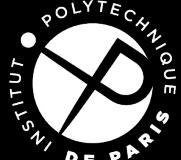


# Thèse de doctorat

NNT : 2025IPPA074



INSTITUT  
POLYTECHNIQUE  
DE PARIS



## Extended reference prior theory for objective and practical inference, application to robust and auditable seismic fragility curves estimation

Thèse de doctorat de l’Institut Polytechnique de Paris  
préparée à l’École polytechnique

École doctorale n°574 École doctorale de mathématiques Hadamard (EDMH)  
Spécialité de doctorat: Mathématiques appliquées

Thèse présentée et soutenue à Palaiseau, le 1er octobre 2025, par

**ANTOINE VAN BIESBROECK**

Composition du Jury :

Mathilde Mougeot Professeure, ENSIIE - ENS Paris-Saclay, Centre Borelli	Présidente du jury
Nicolas Bousquet Professeur associé, Sorbonne université, LPSM	Rapporteur
Daniel Straub Professeur, TU Munich, Engineering Risk Analysis Group	Rapporteur
Sophie Ancelet Ingénierie de recherche, ASNR	Examinateuse
Julyan Arbel Chargé de recherche, Inria Grenoble	Examinateur
Olivier Le Maître Directeur de recherche, École polytechnique, CMAP	Examinateur
Josselin Garnier Professeur, École polytechnique, CMAP	Directeur de thèse
Cyril Feau Ingénieur de recherche, CEA Saclay, SEMT	Encadrant
Clément Gauchy Ingénieur de recherche, CEA Saclay, SGTS	Invité



# Extended reference prior theory for objective and practical inference, application to robust and auditable seismic fragility curves estimation

*By Antoine Van Biesbroeck  
Under the supervision of Josselin Garnier*

*A thesis submitted in fulfillment of the requirements for the degree of  
Doctor of Philosophy in Applied Mathematics*

*Prepared at  
CENTRE DE MATHÉMATIQUES APPLIQUÉES, ÉCOLE POLYTECHNIQUE  
&  
COMMISSARIAT À L'ÉNERGIE ATOMIQUE ET AUX ÉNERGIES ALTERNATIVES*

*Publicly defended on October 1, 2025, in front of a committee composed of*

Mathilde Mougeot	ENS Paris-Saclay	President
Nicolas Bousquet	Sorbonne université	Reviewer
Daniel Straub	TU Munich	Reviewer
Sophie Ancelet	ASNR	Examiner
Julyan Arbel	Inria Grenoble	Examiner
Olivier Le Maître	École polytechnique	Examiner
Josselin Garnier	École polytechnique	Thesis director
Cyril Feau	CEA Saclay	Advisor



*A Pop  
et Maminette*

Heureux sont les coeurs purs, car ils verront Dieu.



# Remerciements

Puisque je ne saurai tous les nommer, je tiens à remercier toutes les personnes qui m'ont accompagné, qui m'ont aidé, qui m'ont soutenu pendant ces trois années. Ainsi que celles et ceux avec qui j'ai ri, j'ai souri et j'ai rêvé.

En particulier et en premier lieu, il est évident et important pour moi de remercier Josselin. Pour avoir dirigé mon travail avec justesse et t'être engagé dans nos discussions parfois à "bâtons rompus". Tu m'as précieusement accordé ta patience et ta bienveillance. Grâce à ta confiance, j'ai éveillé mon attrait pour la recherche. Tu m'as laissé libre de travailler sur ce qui m'a plu tout en m'apportant ton soutien lorsque j'en ai eu besoin. Sous ta direction, j'ai eu la sincère impression de grandement éléver mon recul et la qualité de mon travail.

Toujours en premier lieu (ex æquo), j'adresse un vrai merci à Cyril. En plus d'avoir partagé avec moi ton bureau il y a 4 ans et demi, tu m'as laissé et tu m'as aidé à en faire le siège de l'équipe officielle des mathématiciens d'EMSI, même au travers de nos aventures au LTA ou au LM2S. Présent et à l'écoute, ta bonne humeur et ta passion ont grandement participé à ma motivation quotidienne, passant toujours au travers des obstacles et des déceptions. C'est en discutant qu'on fait de la recherche et cette phrase prend tout son sens avec toi. Refaire le monde et questionner modèles et résultats avec toi ont émulé mon attrait pour le réel. La richesse de ce travail t'est particulièrement due.

Ensuite, j'adresse un grand merci à mes collègues de thèse avec qui j'ai partagé tant de temps. Déjà au CMAP, où il y a trop de noms pour tous les citer, que ce soit les anciennes et anciens que j'ai trop vite connus, ou les nouvelles et nouveaux que j'ai trop peu rencontrés. Alors merci Constantin, Solange, Clément, Madeleine, Leila pour m'avoir accueilli entre vos murs, vos pauses, et vos bureaux. Merci à Richard, Orso, Jean, Mano, Grégoire, Adriano, Anna, Yushan qui ont grandi au moins en partie avec moi. Et merci à Luce, Oskar, Louis, Mattéo, Théo pour avoir pris la relève. Je n'oublie pas d'adresser une mention spéciale aux cooccupantes et cooccupants du bureau 2015 : merci Corentin, Claire, Ignacio, Alexandre, Valentin, pour avoir gardé cette pièce chaleureuse et accueillante, même les jours les moins drôles. Et enfin, il y a celles et ceux avec qui j'ai surtout passé du temps en dehors. Je pense principalement à tous les représentants de la Josselinerie, Baptiste, Paul, Raphaël, Ali, pour ne citer que vous, une vraie secte multi-générationnelle qui partage souvenirs de conférences, karokés, et, pour beaucoup, processus gaussiens.

A toutes celles et tous ceux qui sont encore dans ce laboratoire ou qui y seront, je vous souhaite que ces murs restent les témoins de l'ambiance, la cohésion et l'amicalité que j'ai connue. Je dis bravo en retard à toutes celles, tous ceux qui ont soutenu et bravo en avance à toutes et tous les autres.

Mais mes pensées particulières vont tout de même vers celles et ceux qui, au travers de pauses cafés qui s'éternisent en rires, de confidences sur le canapé, ou de soirées après le travail, sont devenus des amis au-delà d'être d'anciens collègues. Manon, Jessie, Guillaume, Armand, Benjamin, Vanessa, Christoph, ces mots-là sont surtout pour vous. Nul besoin de mille phrases pour vous dire merci puisque je ne saurais comment les écrire. Je ne vous l'exprime sûrement pas assez bien, mais je suis profondément attaché à vous.

Évidemment je n'oublie pas EMSI, où j'ai mis un point d'honneur à venir au moins un jour chaque semaine de ma thèse. Réunions de labo mémorables comme bulles temporelles trop tranquilles caractérisent les journées de ce petit laboratoire, qui vit comme une famille bien atypique qui m'a admis parmi ses membres déjà en 2021. Je pense déjà à Darius qui m'a accueilli tel un électron libre dans son laboratoire. Merci

aussi à Giuseppe, qui vient couper une après-midi de travail seulement pour débattre de sujets essentiels et d'expertise. Tout comme Pierre-Etienne, imitateur hors pair aux mots bien mesurés. Merci également à Fan, Philippe, Thierry, Sandra pour les discussions qui animent autour d'un café des journées qui commencent trop tôt. Mais je pense aussi surtout à la jeunesse, promis Reine tu en fais partie, Bastien, Amal, Breno, Antoine, merci pour avoir apporté de nouvelles têtes dans ce laboratoire. Merci Israe, qui est arrivée en même temps que moi et qui a rendu quelques journées bien vides bien moins moroses. Et bien sûr, bravo Nils pour avoir très bien choisi ton sujet de stage, je te souhaite une très bonne suite de ton parcours.

Aussi, la vie de chercheur ne se limite pas aux murs de ses laboratoires, et je tiens alors à saluer celles et ceux que j'ai croisé et que je continue de croiser, parfois à travers la France et même en dehors. J'adresse un merci particulier à la communauté mascot-num (ou air-tuque) de m'avoir accueilli tel que je suis (c-à-d, bayésien). Je remercie en fait tous ces chercheurs et ces chercheuses intéressés et intéressants, avec qui j'ai apprécié discuter, et que j'ai parfois pris en exemple. Notamment, celles et ceux que j'ai eu la chance de pouvoir choisir pour être membre de mon jury de thèse, et que je remercie sincèrement d'avoir accepté.

Sur cette lancée, ce sont finalement des lieux, et des souvenirs partagés que je remercie, ou ne remercie pas. Par exemple, je ne remercie pas la météo de Chatou, qui éclabousse parfois trop des chercheurs malheureux qui cherchent à rejoindre l'île des impressionnistes le temps d'un séminaire. Par contre, je remercie particulièrement les couloirs souterrains de la station Denfert-Rochereau qui, après les tourniquets, ont abrité interminables discussions, réflexions, débats et confidences, au moins autant que l'angle des via della geppa et roma de Trieste.

Pour finir, merci immensément aux gens qui m'aiment. A mes amis proches, Marc, Philo, Maëva, Raph, Stan à qui je donne parfois je trouve trop peu de nouvelles. Merci de partager secrets, rires, doutes, et projets, et de faire partie de ma vie bien au-delà de ce doctorat. Puis à ma famille : mes cousines et mes cousins, ma sœur et mes frères, mes beaux-frères et mes belles-sœurs, mes parents et mes beaux-parents, mes oncles et mes tantes, ma nièce. Merci de m'entourer, d'être là, ensemble, pour moi.

Je conclus enfin ces pages en m'adressant à celle qui partage ma vie, loin des yeux, mais à toujours près du cœur. Laurilene, mon éternelle alliée, merci pour ton soutien, ta force, et ton amour.

# Abstract

Reference prior theory provides a principled framework for objective Bayesian inference, aiming to minimize subjective input and allow data-based information to drive the estimates distribution. For this reason, the application of this theory to the estimation of seismic fragility curves is particularly relevant. Indeed, these curves are essential elements of seismic probabilistic risk assessment studies; they express the probability of failure of a mechanical structure as a function of indicators that define seismic scenarios. Since they inform critical decisions in infrastructure safety, a complete auditability of the pipeline that leads to the estimates of these curves is required.

This thesis investigates the interplay between reference prior theory and seismic fragility curves estimation, yielding original contributions in these two domains. First, we complement the theoretical foundations of reference priors by developing novel constructions of them. Our goal is to support their objectivity while improving their practical applicability. Our results take the form of theoretical contributions in this domain that are based on a generalized definition of the mutual information. Our approaches tackle the principal issues of reference priors, namely their improper characteristic or that of their posterior, and their complex formulation for practical use.

Second, we revisit the estimation of seismic fragility curves based on the prominent probit-lognormal model in a context where the data are particularly sparse. Our goal is to conduct a Bayesian estimation of seismic fragility curves that leverages the optimization of every sort of information, including the *a priori* one, in order to provide estimates that are robust and auditable. Our results highlight the limitations and irregularities of the model and propose methods that provide accurate and efficient estimates of the curves. The evaluations of our approaches are carried out on different case studies taken from the nuclear industry.

This thesis builds a strong link between these two domains. The application to seismic fragility curves not only motivated theoretical developments but also directly benefited them, ultimately producing a more robust, interpretable, and verifiable estimation framework.



# Résumé

La théorie des priors de référence fournit un cadre approprié à une inférence bayésienne objective, puisqu'elle vise à minimiser la subjectivité introduite et à permettre aux informations issues des données d'orienter la distribution des estimations. Pour cette raison, l'application de cette théorie à l'estimation des courbes de fragilité sismique est particulièrement pertinente. En effet, ces courbes sont des éléments essentiels des études sismiques probabilistes de sûreté ; elles expriment la probabilité de défaillance d'une structure mécanique en fonction d'indicateurs définissant des scénarios sismiques. Puisqu'elles informent des décisions critiques en matière de sécurité des infrastructures, une auditabilité complète de l'approche qui conduit aux estimations de ces courbes est nécessaire.

Cette thèse étudie l'interaction entre la théorie des priors de référence et l'estimation des courbes de fragilité sismique, apportant des contributions originales dans ces deux domaines. Tout d'abord, nous complétons les fondements théoriques des priors de référence en développant de nouvelles constructions de ceux-ci. Notre objectif est de soutenir leur objectivité tout en améliorant leur applicabilité pratique. Nos résultats prennent la forme de contributions théoriques dans ce domaine qui sont basées sur une définition généralisée de l'information mutuelle. Nos approches abordent les principaux problèmes des priors de référence, à savoir le caractère impropre de leur distribution ou de leur distribution *a posteriori*, et leur formulation complexe pour une utilisation pratique.

Ensuite, nous revisitons l'estimation des courbes de fragilité sismique basée sur le modèle probit-lognormal dans un contexte où les données sont particulièrement rares. Notre objectif est de réaliser une estimation bayésienne des courbes de fragilité qui tire parti de l'optimisation de toutes sortes d'informations, y compris l'information *a priori*, afin de fournir des estimations robustes et auditables. Nos résultats mettent en évidence les limites et les irrégularités du modèle et proposent des méthodes qui fournissent des estimations précises et efficaces des courbes. Les évaluations de nos approches sont réalisées sur différents cas d'étude issus de l'industrie nucléaire.

Cette thèse établit un lien fort entre ces deux domaines. L'application aux courbes de fragilité sismique a non seulement motivé les développements théoriques mais leur a aussi directement profité, produisant finalement un cadre d'estimation plus robuste, interprétable et vérifiable.



# Contents

<b>1</b>	<b>Introduction</b>	<b>17</b>
1.1	Motivation and positioning of the thesis . . . . .	17
1.2	Outline of the manuscript and contributions . . . . .	23
<b>2</b>	<b>Introduction en français</b>	<b>27</b>
2.1	Motivation et positionnement de la thèse . . . . .	27
2.2	Esquisse du manuscrit et contributions . . . . .	33
<b>I</b>	<b>Contribution to the reference prior theory</b>	<b>37</b>
<b>3</b>	<b>Review of the reference prior theory</b>	<b>39</b>
3.1	Introduction . . . . .	39
3.2	The standard Bayesian framework . . . . .	41
3.3	Mutual information and reference priors . . . . .	42
3.4	Limits of the classical framework of reference priors and solution . . . . .	47
3.5	Conclusive remarks . . . . .	53
<b>4</b>	<b>Generalized mutual information and their reference priors</b>	<b>55</b>
4.1	Introduction and motivations . . . . .	56
4.2	Generalized mutual information . . . . .	56
4.3	Generalized reference priors . . . . .	58
4.4	Towards $\delta$ -divergences-reference priors . . . . .	60
4.5	Discussions . . . . .	63
4.6	Proofs of the main results . . . . .	68
4.7	Conclusion and prospects . . . . .	76
<b>5</b>	<b>Properly constrained generalized reference priors</b>	<b>77</b>
5.1	Introduction . . . . .	78
5.2	Definitions, notations and motivation . . . . .	78
5.3	Constrained $D_\delta$ -reference priors . . . . .	81
5.4	Discussion . . . . .	83
5.5	An example . . . . .	84
5.6	Detailed proofs . . . . .	85
5.7	Conclusion . . . . .	91

<b>6 Variational approximation of generalized reference priors using neural networks</b>	<b>93</b>
6.1 Introduction . . . . .	93
6.2 Variational approximation of the reference prior (VA-RP) . . . . .	94
6.3 Numerical experiments on different models . . . . .	101
6.4 Conclusion . . . . .	105
<b>II Seismic fragility curve estimation</b>	<b>107</b>
<b>7 Seismic fragility curves in probabilistic risk assessment studies</b>	<b>109</b>
7.1 Introduction . . . . .	109
7.2 Data: from seismic signals to equipments failures . . . . .	110
7.3 Modeling of the fragility curve . . . . .	114
7.4 Example of case studies . . . . .	117
7.5 Conclusion: which case study and which data for a Bayesian estimation of fragility curves? . . . . .	120
<b>8 Bayesian estimation of seismic fragility curves using a reference prior</b>	<b>123</b>
8.1 Introduction . . . . .	124
8.2 Probit-lognormal model and Bayesian framework . . . . .	125
8.3 Jeffreys prior construction for the probit-lognormal model . . . . .	126
8.4 Competing approaches and estimation tools . . . . .	133
8.5 Limits of the estimates given by the three approaches: the curse of degeneracy . . . . .	134
8.6 Performance evaluation metrics . . . . .	137
8.7 Numerical applications . . . . .	138
8.8 A review of the properties of the SK posterior . . . . .	142
8.9 Conclusion . . . . .	145
<b>9 Appropriate constraint incorporation in the probit-lognormal reference prior</b>	<b>147</b>
9.1 Introduction . . . . .	148
9.2 Reminder of the probit-lognormal model and of the Bayesian framework . . . . .	149
9.3 Constrained reference prior for probit-lognormal fragility curves . . . . .	150
9.4 Variational approximation of the constrained reference prior . . . . .	152
9.5 Comparison between the VA-RP and the interpolated reference prior . . . . .	154
9.6 Application of the method on the piping system case study . . . . .	157
9.7 Impacts of the neural network's architecture on the VA-RP . . . . .	159
9.8 Conclusion . . . . .	161
<b>10 Design of experiments with constrained reference priors for robust inference</b>	<b>163</b>
10.1 Introduction . . . . .	164
10.2 Probit-lognormal modeling and constrained reference prior . . . . .	165
10.3 Sequential design of experiments . . . . .	166
10.4 Estimates and benchmarking metrics . . . . .	169
10.5 Application of the method on the piping system . . . . .	170
10.6 Behavior of the method on toy case studies . . . . .	175
10.7 Conclusion . . . . .	180

<b>III Conclusion &amp; perspectives</b>	<b>181</b>
11.1 Summary . . . . .	183
11.2 On the contributions to the reference prior theory . . . . .	184
11.3 On the contributions to the estimation of seismic fragility curves . . . . .	185
11.4 General outlook . . . . .	185
11.5 Perspectives and connections with other fields . . . . .	186
11.6 Final words . . . . .	187
<b>Appendices</b>	<b>189</b>
<b>A Influence of the IM choice on fragility curves estimated using a reference prior</b>	<b>191</b>
A.1 Introduction . . . . .	192
A.2 Bayesian problem . . . . .	193
A.3 Reference prior theory . . . . .	193
A.4 Jeffreys prior construction . . . . .	194
A.5 Estimation tools, competing approaches and benchmarking metrics . . . . .	195
A.6 Numerical application . . . . .	198
A.7 Conclusion . . . . .	202
<b>B Design of experiments for a low fidelity model of seismic fragility curves</b>	<b>203</b>
B.1 Introduction . . . . .	204
B.2 Low fidelity model for fragility curves . . . . .	205
B.3 Sensitivity index for design of experiments . . . . .	207
B.4 Numerical application . . . . .	208
B.5 Discussion . . . . .	211
B.6 Conclusion . . . . .	212
B.7 Details regarding the construction of $\mathbf{z}$ . . . . .	212
<b>C Non-asymptotic reference priors, a brief study</b>	<b>215</b>
C.1 Motivations and context . . . . .	215
C.2 Bayesian framework and mutual information . . . . .	216
C.3 Non-asymptotic reference priors . . . . .	216
C.4 Discussion and link with asymptotic reference priors . . . . .	219
<b>* Ma thèse en quelques pages</b>	<b>221</b>
La thèse en quelques mots, et quelques phrases . . . . .	221
Initiation à la réflexion sur l'a priori Bayésien . . . . .	222
Le lien avec les séismes et les centrales nucléaires dans tout ça . . . . .	222
Concrètement, qu'est ce qu'on a fait, qu'est-ce qu'on en conclut ? . . . . .	223
<b>Bibliography</b>	<b>225</b>



# 1

# Introduction

**Abstract** First of all, we present and motivate the research work carried out as part of this thesis. The aim is to improve methods for seismic probabilistic risk assessment studies, given the lack of response provided by the state of the art regarding the choice of the prior in Bayesian inference. Secondly, we present the organization of the manuscript, which is based on a reformulation of the problematic under in the form of six main questions.

<b>1.1 Motivation and positioning of the thesis . . . . .</b>	<b>17</b>
1.1.1 Probabilistic risk assessment studies . . . . .	17
1.1.2 Uncertainty quantification in probabilistic risk assessment studies . . . . .	19
1.1.3 The choice of the prior in Bayesian studies . . . . .	21
1.1.4 What motivates research on prior elicitation for SPRA studies . . . . .	22
<b>1.2 Outline of the manuscript and contributions . . . . .</b>	<b>23</b>
1.2.1 Problem statement and organization of the thesis . . . . .	23
1.2.2 List of contributions . . . . .	24

## 1.1 Motivation and positioning of the thesis

### 1.1.1 Probabilistic risk assessment studies

#### History

Probabilistic risk assessment (PRA) studies refer to a set of technical analysis methods that allow for the quantification of the risks faced by a facility when exposed to a particular event. This event may be of natural or artificial origin (technological, human, etc.), and may stem from internal or external sources. It can take the form of an earthquake, a flood, a series of internal failures, among others.

Following the early recommendations of F. R. Farmer (a safety expert at the UK Atomic Energy Authority) in the 1960s concerning the reliability of nuclear facilities, these methods share a common feature: the incorporation of the notion of uncertainty into the characterization of the event, its phenomena, and its attributes (thus referred to as the “hazard”). Their framework and concepts were quickly adopted and further developed in the United States (see the report by the Nuclear Regulatory Commission, NRC, 1983). Notably, many studies began integrating seismic reliability analyses since 1968 (Cornell, 1968), thereby laying the groundwork for seismic probabilistic risk assessment (SPRA) studies.

Earthquakes indeed represent a significant risk factor in safety assessments. Firstly, although commonly characterized by their magnitude and distance from the source, seismic signals are far more complex than a simple bivariate function. Two earthquakes with the same magnitude and source can exhibit significantly

## Chapter 1. Introduction

different characteristics (and consequences). Secondly, because an earthquake can simultaneously affect all elements —both external and internal— to a facility, it has the potential to cause severe damage to equipment and structures. The potential cost of the consequences of a seismic hazard can thus be high and critical in the nuclear context, making it an event of major and decisive importance, even in geographical areas where such events are rare.

Nowadays, accounting for seismic hazard within the framework of probabilistic risk assessments is an international recommendation in the context of the nuclear industry. In France, its application within that industry is defined by the nuclear safety and radiation protection authority (ASN in French<sup>1</sup>), as outlined in the fundamental safety rule (ASN, 2002).

### The evolution of PRA in France, at the CEA, and in this Thesis

The evolution of methods and knowledge related to the safety of nuclear installations progresses in parallel with changes in the regulations and standards imposed upon them. Concerning seismic hazard, the relationship between evolving rules and methods (and thus of PRA studies) brings into play differing perspectives, primarily between the main operator (EDF) and the experts of the safety authority (formerly the IRSN, now unified with the nuclear safety authority under the ASN). For the operator, the robustness of the installation should not be fundamentally questioned by advancements in knowledge or methodology, as uncertainties are already accounted for through safety margins incorporated during the equipment's design. The authority's experts, on the other hand, argue that robustness must be continuously reassessed, and that safety margins are not intended to be gradually consumed as knowledge advances (Roger, 2020). In this dialogue, the ASN acts as a kind of referee. Among other responsibilities, it defines the "augmented safety earthquake" (a conservative amplification of the "maximum historically credible earthquake"), which serves as the reference margin in demonstrating an equipment's seismic robustness.

This arbitration is thus both sensitive and critical. The incident that occurred in 2011 at the Fukushima-Daiichi nuclear power plant illustrates this clearly. The seismic hazard (which triggered the tsunami) had been underestimated by the consensus of Japanese experts, and it is ultimately this underestimation that led to the accident (see the report by the international atomic energy agency, IAEA, 2015). In the wake of this event, the ASN decided to amplify the "augmented safety earthquake" by a factor of 1.5 in France<sup>2</sup> for the equipments that belong to the "hard core" (a limited list of systems, structures, and essential components to the upkeep of a nuclear installation). This led to require operators to provide an amplified demonstration of the robustness of their installations.

Within this context, the role of the CEA is ambivalent in the French nuclear landscape. On one hand, it operates research nuclear facilities and is therefore responsible for demonstrating the seismic robustness of its own equipment. On the other hand, it is also involved in collaborative research related to the civil nuclear fleet operated by EDF. A part of the research and expertise on the consequences of seismic hazards is carried out at the CEA's EMSI laboratory, which is equipped with an experimental platform (called Tamaris). This facility enables mechanical testing of equipment under simulated seismic conditions. Advancing methods for the SPRA framework is a focus area for the CEA that falls within the mission of this laboratory. As the CEA is accountable to safety authorities for its installations, it seeks to develop increasingly rigorous methods to justify their seismic performance.

This thesis fits into that process of evolving SPRA methodologies and reinforce safety demonstrations. Funded by the CEA as part of a research training grant, its aim is to enrich the state of the art in seismic fragility assessment for equipment and installations. The goal is to develop methods that are: (i) effective in the face of the complexity of both the studied systems and the hazard itself, (ii) robust over time and resilient to potential reevaluations of safety criteria, and (iii) transparent and auditable by safety authorities.

---

<sup>1</sup>Formerly ASN: the nuclear safety authority (ASN) was unified with the radioprotection and nuclear safety institute (IRSN) in January 1, 2025.

<sup>2</sup>Defining in this respect the "hard core earthquake".

### 1.1.2 Uncertainty quantification in probabilistic risk assessment studies

#### Principles and steps of uncertainty quantification

Probabilistic risk assessment studies rely on the identification and quantification of uncertainties that arise during the evaluation of risk within a physical system. Uncertainty quantification is, in fact, a systematic process that aims at modeling and analyzing uncertainty, its sources, and how it propagates through the modeling of the studied physical system. This process lies at the intersection of physics, engineering, and applied mathematics, and has even emerged as a unique discipline within the field of statistics.

It focuses on three key elements that describe the interaction between the physical system and the hazard: input parameters  $\mathbf{X}$ , the physical response of the system  $Y$ , and a modeling of the latter:  $Y = \mathcal{M}(\mathbf{X})$ . This functional  $\mathcal{M}$  represents the (often complex) modeling of the system and its physical properties. This model may involve solving physical equations, running numerical simulations, or even analyzing the result of a mechanical experiment. The general approach to uncertainty quantification typically consists into several key steps (see, for example, Sudret, 2007; Iooss, 2009) that we detail below:

1. Identification of uncertainty sources: It is common to classify uncertainties into two main categories. On one hand, irreducible uncertainties, which arise from the inherent and “natural” randomness embedded in the hazard and the physical system itself. On the other hand, epistemic uncertainties, which result from a lack of information and are therefore considered reducible, in contrast to the former (Hüllermeier and Waegeman, 2019). This step leads to a probabilistic modeling of the input  $\mathbf{X}$ .
2. Uncertainty propagation: This step involves approximating the distribution of the model output  $\mathcal{M}(\mathbf{X})$ , and evaluating a quantity of interest that depends on it, such as a variance, a failure probability, etc. More generally, this means evaluating a quantity of the form  $\mathbb{E}\phi(Y)$ .
3. Sensitivity analysis: This step aims to provide insight into how uncertainties propagate through the system, by identifying the influence of one or more of the input parameters on the output  $Y$  (Iooss and Lemaître, 2015).

#### Mathematical tools in uncertainty quantification

Numerous mathematical tools support the aforementioned steps of uncertainty quantification. While this manuscript does not aim to provide an exhaustive overview of them, the following highlights several examples that are directly relevant to the work presented in this thesis:

- Surrogate modeling helps to address the complexity involved in evaluating the model  $\mathcal{M}$ , by constructing a surrogate model  $\tilde{\mathcal{M}}$ , trained on a dataset of observations  $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_k, y_k)$ . The surrogate model is, by design, simpler to evaluate than the original model. Its definition is not limited and its form can range from a simple parametric function to the output of a “black-box” neural network. Among the most commonly used approaches we can cite Gaussian processes (or kriging) and polynomial chaos expansions.
- Global sensitivity indices are a core component of the sensitivity analysis step. Since the early work of Sobol' (1993), they have become essential tools for statistically measuring how the system output  $Y$  is influenced by one or several of the input variables  $X_i$  (where  $\mathbf{X} = (X_1, \dots, X_p)$ ). In this context, the impact of input  $X_i$  on  $Y$ , denoted  $S_i$ , is expressed as the expected divergence between the distribution  $P_Y$  of  $Y$  and its conditional distribution given  $X_i$ ,  $P_{Y|X_i}$  (Da Veiga, 2015):

$$S_i = \mathbb{E}_{X_i}[D(P_Y || P_{Y|X_i})], \quad (1.1)$$

where  $D$  is a dissimilarity measure between two probability distributions. The choice of  $D$  is crucial in defining  $S_i$ . For instance, choosing  $D(P||Q) = \|\mathbb{E}_{X \sim P}[X] - \mathbb{E}_{X \sim Q}[X]\|^2$  corresponds to a first-order Sobol' index. The choice of  $D$  can be driven by various considerations, such as the desire to detect independence between  $X_i$  and  $Y$ . The next item explores a range of potential choices for  $D$ .

- Information theory lies at the heart of comparing probability measures, and thus, of uncertainty quantification. By nature, the ways in which two distributions spread information over a space are numerous. In the definition of  $S_i$  above, the selection of  $D$  (and therefore the form of  $S_i$ ) is inherently tied to information-theoretic principles.

A widely used class of divergences, perceived as an extension of Shannon entropy, is the  $f$ -divergences class, introduced by Csiszár (1967). These divergences are used across many fields beyond sensitivity analysis, such as in variational inference (Minka, 2005; Bach, 2023), surrogate model design (Nguyen, Wainwright, and Jordan, 2009), and PAC-Bayesian learning (Picard-Weibel and Guedj, 2022). When  $f$  is convex and satisfies  $f(1) = 0$ , the  $f$ -divergence is defined as:

$$D_f(P||Q) = \int_{\mathcal{X}} f\left(\frac{p(x)}{q(x)}\right) q(x) d\omega(x), \quad (1.2)$$

where  $p$  and  $q$  are densities of  $P$  and  $Q$  with respect to a common base measure  $\omega$  on their domain  $\mathcal{X}$ . Thus,  $f$ -divergences reduce the choice of  $D$  to a choice of the function  $f$ . A commonly used subclass is the  $\delta$ -divergences class (H. Zhu and Rohwer, 1995), defined fixing  $f = f_\delta$  where:

$$f_\delta(x) = \begin{cases} \frac{x^\delta - \delta x - (1-\delta)}{\delta(\delta-1)} & \text{if } \delta \notin \{0, 1\}, \\ x \log x - x + 1 & \text{if } \delta = 1, \\ -\log x + x - 1 & \text{if } \delta = 0. \end{cases} \quad (1.3)$$

Notably, the well-known Kullback-Leibler divergence (denoted KL) can be viewed as a special case of  $\delta$ -divergences. We recall that:

$$\text{KL}(Q||P) = \int \log\left(\frac{q(x)}{p(x)}\right) q(x) d\omega(x) \quad (1.4)$$

(with the same notations as in eq. (1.2)), so that:

$$D_{f_\delta}(P||Q) = \text{KL}(P||Q) \text{ if } \delta = 1, \quad D_{f_\delta}(P||Q) = \text{KL}(Q||P) \text{ if } \delta = 0. \quad (1.5)$$

Of course, many dissimilarity measures are not  $f$ -divergences. First-order Sobol' indices, mentioned above, are one such example. Another notable example is the Maximum Mean Discrepancy (MMD). Let  $P, Q$  be probability measures over a set  $\mathcal{X}$ , and let  $\mathcal{H}$  be a reproducing kernel hilbert space (RKHS) over  $\mathcal{X}$  with reproducing kernel  $k : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}$ . The MMD (Gretton et al., 2012) is defined as:

$$\text{MMD}(\mathcal{H}; P||Q) = \sup_{\substack{f \in \mathcal{H} \\ \|f\|_{\mathcal{H}} \leq 1}} |\mathbb{E}_{X \sim P} f(X) - \mathbb{E}_{X \sim Q} f(X)|, \quad (1.6)$$

or more simply:

$$\text{MMD}^2(\mathcal{H}; P||Q) = \mathbb{E}_{X, X' \sim P \otimes P} [k(X, X')] + \mathbb{E}_{Y, Y' \sim Q \otimes Q} [k(Y, Y')] - 2\mathbb{E}_{X, Y \sim P \otimes Q} [k(X, Y)]. \quad (1.7)$$

- Reproducing kernel Hilbert spaces, although not widely used in this thesis, represent a powerful tool for modeling and representing complex objects. Through the use of their kernel, they enable the manipulation of non-linear objects in a Hilbert linear space. In the case of the MMD defined above, RKHSs are used to represent and compare probability measures.

A RKHS consists in a given Hilbert space  $\mathcal{H}$ , sub-space of a functional space  $\mathcal{F} = \{f : \mathcal{X} \rightarrow \mathbb{C}\}$ , such that for any  $x \in \mathcal{X}$ , the mapping  $f \mapsto f(x)$  is continuous on  $\mathcal{H}$ . Then, the reproducing kernel is the unique mapping  $k : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}$  such that for all  $x \in \mathcal{X}$ ,  $f \in \mathcal{H}$ ,

$$f(x) = \langle f, k(x, \cdot) \rangle_{\mathcal{H}}, \quad (1.8)$$

where  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$  denotes the scalar product on  $\mathcal{H}$ . Among other fundamental properties of the theory, Aronszajn's theorem (Aronszajn, 1950) states the existence of an RKHS from a given kernel that is symmetric and positive definite on  $\mathcal{X}$ . For an in-depth look to the theory, we refer the interested reader to Schölkopf and Smola, 2001.

- Bayesian analysis is a fundamental tool in uncertainty quantification, as it fundamentally introduces a form of uncertainty in one or more parameters of the studied model. The uncertainty encoded in the Bayesian framework typically appears in the form of a prior distribution. Propagation of uncertainties within this framework is achieved by computing the posterior distribution, based on statistical observations  $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_k, y_k)$ . The choice of the prior is critical: it belongs to the step of uncertainty identification and significantly influences the posterior. As such, in the context of probabilistic risk assessments, it directly affects safety evaluations.

Bayesian analysis is whole paradigm. Since it is central to this thesis, the next section is fully dedicated to it.

### 1.1.3 The choice of the prior in Bayesian studies

Bayesian analysis is a branch of statistics built upon an interpretation of probability as a degree of plausibility of events, generally referred to as “credibility”. Like any statistical method, it aims to establish a connection between observed data and their underlying probability distribution. It relies on Bayes’ theorem, treating the unknown parameter that define the data distribution as random. Statistical observations then inform the distribution of these unknown parameters, updating the credibility attributed to one possible distribution of the data over another.

By contrast, the frequentist approach views probability as a frequency of occurrence, observable through repeated experimentations. In this paradigm, the unknown quantity exists as a fixed value, and the frequencies observed through repeated trials are more or less probable depending on which value of the unknown we assume to be true. This leads to what is typically termed “confidence”.

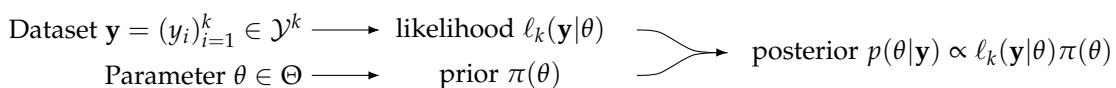
It is important to note that these two approaches are complementary and they often intersect, depending on the specific context. In practical applications, the nature of the case study and its particular constraints may lead to favoring one approach over the other.

This thesis focuses on Bayesian analysis within the context of statistical inference. Here, a variable of interest  $Y \in \mathcal{Y}$  is modeled using a parametric distribution  $P_{Y|\theta}$ , where  $\theta \in \Theta$  is the unknown parameter, now treated as a random variable under the Bayesian framework. Accordingly,  $\theta$  is assigned a distribution known as the **prior**. When we observe successive and conditionally independent realizations of  $Y$ , the **posterior** is defined as the distribution of  $\theta$  conditionally to those observations. A credibility region at level  $r$  is defined as a region within  $\Theta$  that holds at least probability  $r$  under the posterior distribution. We note that various results establish links between the Bayesian and frequentist approaches to estimating  $\theta$ . One such result is the Bernstein–von Mises theorem (see e.g. Vaart, 1992), which shows convergence between the two perspectives under certain conditions.

Let  $\mathbf{y} = (y_1, \dots, y_k)$  represent the observed values from  $k$  realizations of  $Y$ . Assuming all necessary quantities are well-defined, let:  $\pi$  be the density of the prior (with respect to some measure  $\nu$ ),  $\ell(\cdot|\theta)$  be the density of  $P_{Y|\theta}$  (with respect to a measure  $\mu$ ), and  $\ell_k(\mathbf{y}|\theta) = \prod_{i=1}^k \ell(y_i|\theta)$  be the likelihood of the model. With these notations, Bayes’ theorem defines the posterior density (with respect to  $\nu$ ) as:

$$p(\theta|\mathbf{y}) = \frac{\ell_k(\mathbf{y}|\theta)\pi(\theta)}{\int_{\Theta} \ell_k(\mathbf{y}|\theta)\pi(\theta) d\nu(\theta)}. \quad (1.9)$$

This formula formalizes the link between the prior, the model (represented via the likelihood), the data, and the posterior, which ultimately reflects the updated, credible prediction. It illustrates the Bayesian framework as an information transmission chain leading to the *a posteriori*, from two sources: the “real” (the observations) and the *a priori*. This structure is illustrated in the diagram below.



If the information provided by the data is generally not questioned, it is the prior information that captures the most attention. By nature, the prior is the source of uncertainty in the result. Thus, its construction

must be handled with care during the modeling stage and cannot be arbitrary. Yet the Bayesian formulation let it significantly influence the posterior, especially in situations where the dataset is limited in size. An illustration of this phenomenon is proposed fig. 1.1, where different priors —that seem similar— lead to very different posteriors, all based on the same observations. As always, the devil is in the detail: here, it is the tails of the prior distributions —barely visible in the figure— that account for the discrepancy. Still, this example rightly raises concerns about the choice of prior.

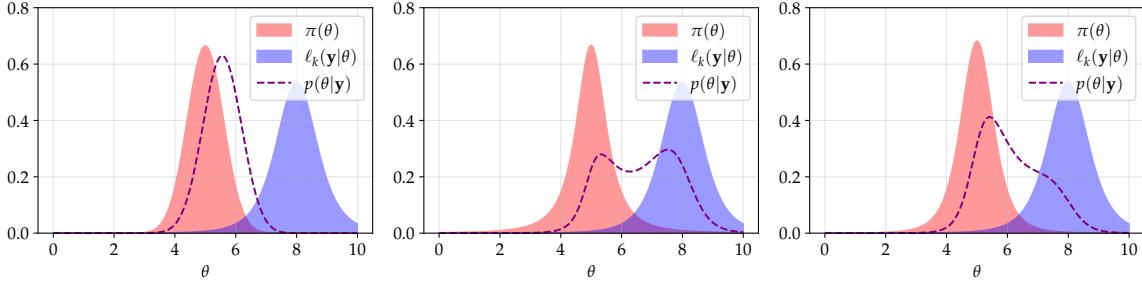


Figure 1.1: Different posteriors (dashed lines) derived from the same likelihood (blue) but different priors (red).

The question of how to construct a prior remains an open issue in the literature. For some, it is an opportunity to incorporate information from external sources, such as expert judgment, historical data, or known properties expected of the result. Such priors are typically described as informative. For others, the prior is instead a way for injecting uncertainty into the model. In other term, a way of expressing a lack of prior knowledge and allowing the data to speak for themselves. These practitioners place their trust in the modeling of the likelihood and their prior will be qualified as non-informative.

What emerges from these introductory paragraphs is that the question is unsolvable: there can be no universally accepted method for selecting a prior. This issue represents one of the “holes” in the Bayesian workflow identified by Gelman and Yao (2020): overly informative priors can undermine confidence in the results, while excessively vague priors can lead to credibility regions so wide they become impractical. That said, this does not diminish the relevance of the Bayesian approach, this section is not intended to discredit it, and the remainder of the manuscript will prove that. But it must leave the one who use it aware of the impact and limitations of their choices. The selection of a prior is as critical as the choice of the model itself.

#### 1.1.4 What motivates research on prior elicitation for SPRA studies

Bayesian analysis has become increasingly popular across many fields, including reliability and safety studies. Indeed, it is widely appreciated for its ability (i) to introduce and propagate uncertainty in an inference problem, and (ii) to exhibit more regular behavior than many frequentist methods, particularly when the number of observed data is limited.

The Bayesian inference framework naturally aligns with the uncertainty quantification process. Identifying uncertainties in the input  $\mathbf{X}$  or within the model  $\mathcal{M}$  itself (or a surrogate, when applicable) can be achieved through their parameterization, resulting in a parametric distribution of the output  $Y$ :  $P_{Y|X,\theta}$ . The variable  $\theta$ , which is treated as random under the Bayesian paradigm, can represent various sources of uncertainty, differing in nature depending on the context (Bousquet, 2024a). This approach has been widely employed across a range of models, including Gaussian processes (e.g., Gu and Berger, 2016) and neural networks (Arbel et al., 2023).

Sometimes, the methodology is adopted specifically for the flexibility it offers in incorporating expert judgment or prior knowledge. Yet, it is this possibility that also questions, as it introduces a degree of subjectivity that is often difficult to justify. In the context of SPRA in the nuclear industry, demonstrating robustness is essential. Given the high complexity of such case studies, one could find as many priors as there are experts if we follow this approach. Such a caricatural situation puts at risk the auditability of any resulting posterior credibility region.

It becomes clear that in an auditable safety study, a “good” credibility region for an estimated parameter is not necessarily a narrow one, even if it seems to suggest that safety is achieved. Conversely, a region that is too wide is not inherently “good” either, as it fails to demonstrate the expected robustness of the equipment. The appropriate balance is complex, and the “right” prior is the one that produces results that inspire trust, while accurately reflecting the equipment’s ability or inability to handle the hazard.

All these reflections highlight the importance of constructing the prior in a methodical and disciplined way within the context of SPRA, using a rigorous framework that seeks, as much as possible, to avoid the introduction of any form of subjectivity into the process.

## 1.2 Outline of the manuscript and contributions

### 1.2.1 Problem statement and organization of the thesis

Several central questions emerge from this introduction and the preceding reflections. They are listed below and represent the major issues that arose during this thesis, which the work presented in this manuscript aims to address.

**Question i** How can one define and support the objectivity of a prior?

**Question ii** What are the limitations of implementing such non-informative priors, and how can their use be reconciled with practical needs?

**Question iii** How can such priors be constructed and derived in practice?

**Question iv** In the context of SPRA, what forms do these objective priors take within a seismic fragility curve model?

**Question v** What are the implications of information scarcity in such models, and how can they be addressed?

**Question vi** How can the different sources of information (*a priori* information and data-based information) be best integrated across the full Bayesian workflow of the model under study?

In this thesis, we aim to address these six questions through a study structured along two axes. The first axis is theoretical in nature and tackles the issues surrounding the construction of so-called “objective” priors. It focuses on the development of a theoretical framework known as reference prior theory. The second axis is more practical, examining the application of the theoretical developments to real-world case studies. It focuses on the estimation of seismic fragility curves, which are a key tool in the framework of seismic probabilistic risk assessment studies. It is important to emphasize that, although distinct, these two major axes of work are deeply interconnected. Practical concerns inspired the theoretical investigations, just as theoretical insights have been applied to serve practical studies.

The part I of this manuscript is devoted to the development of the reference prior theory.

In chapter 3, the theory is introduced and its state of the art is reviewed. This chapter presents the notion of reference prior as it is commonly defined in the literature and situates it within the objective priors. This chapter allows to formalize the mathematical framework that will be used throughout the rest of this part.

In chapter 4, we observe that the definition of reference priors relies on a dissimilarity measure. The chapter seeks to generalize this measure in order to reinforce the objectivity of reference priors. This work provides an answer to question i.

In chapter 5, the theory is extended to address the practical challenges that arise when reference priors, are difficult to use. The chapter introduces the concept of constrained reference priors, whose solutions propose an answer to question ii.

In chapter 6, a numerical method is developed to approximate reference priors. This method is based on

## Chapter 1. Introduction

variational inference, defining the prior as the output of a neural network in order to avoid an explicit calculation. This chapter provides an answer to **question iii**.

The part II of this manuscript address the topic of objective Bayesian estimation of seismic fragility curves. In [chapter 7](#), seismic fragility curves are introduced along with their definition, historical background, and role within the SPRA. A review of the main estimation methods, based on the types of available data, is presented.

In [chapter 8](#), a detailed computation and study of the reference prior is carried out for a classical model used to estimate fragility curves. Bayesian inference using this prior is then compared to other methods. This chapter addresses **question iv**.

In [chapter 9](#), the limitations of the model used for fragility curve estimation are taken into account. A solution is proposed, supported by the theoretical developments from the first part. This chapter proposes an answer to **question v**.

In [chapter 10](#), the study is concluded with a method designed to enhance Bayesian estimation, not only by optimizing the prior selection, but also by improving data selection itself through an experimental design strategy. This chapter provides an answer to **question vi**.

A general conclusion ends the manuscript in the final [part III](#).

### 1.2.2 List of contributions

The research conducted through this thesis led to the production of several contributions to the scientific litterature. Hereafter are listed the published and submitted works.

#### Journal papers

- **A. Van Biesbroeck**, C. Gauchy, C. Feau & J. Garnier (2025). "Robust a posteriori estimation of probit-lognormal seismic fragility curves via sequential design of experiments and constrained reference prior". *arXiv* 2503.07343. DOI: [10.48550/arXiv.2503.07343](https://doi.org/10.48550/arXiv.2503.07343)
- N. Baillie, **A. Van Biesbroeck** & C. Gauchy (2025). "Variational inference for approximate reference priors using neural networks". *arXiv* 2502.02364. DOI: [10.48550/arXiv.2502.02364](https://doi.org/10.48550/arXiv.2502.02364)
- **A. Van Biesbroeck** (2024). "Properly constrained reference priors decay rates for efficient and robust posterior inference", *arXiv* 2409.13041. DOI: [10.48550/arXiv.2409.13041](https://doi.org/10.48550/arXiv.2409.13041)
- **A. Van Biesbroeck**, C. Gauchy, C. Feau & J. Garnier (2025). "Design of experiments based on a low fidelity model for seismic fragility curves estimation". *ESAIM: ProcS* (to appear). HAL: [hal-04719458v1](https://hal.archives-ouvertes.fr/hal-04719458v1)
- **A. Van Biesbroeck** (2023). "Generalized mutual information and their reference priors under Csizar f-divergence". *arXiv* 2310.10530. DOI: [10.48550/arXiv.2310.10530](https://doi.org/10.48550/arXiv.2310.10530)
- **A. Van Biesbroeck**, C. Gauchy, C. Feau & J. Garnier (2024). "Reference prior for Bayesian estimation of seismic fragility curves". *Probabilistic Engineering Mechanics*, 76, pp 103622. DOI: [10.1016/j.probengmech.2024.103622](https://doi.org/10.1016/j.probengmech.2024.103622)

#### Proceedings papers

- **A. Van Biesbroeck**, C. Feau & J. Garnier (2025). "Design of experiments for efficient and conform Bayesian learning of seismic fragility curves". *Proceedings of the 28th conference on Structural Mechanics in Reactor Techology (SMiRT)*.
- N. Baillie, **A. Van Biesbroeck**, C. Feau & C. Gauchy (2025). "Bayesian estimation of seismic fragility curves based on variational reference priors using neural networks". *Proceedings of the 6th Thematic*

Conference on Uncertainty Quantification in Computational Sciences and Engineering (UNCECOMP). URL: [2025.uncecomp.org/proceedings/pdf/21225.pdf](https://2025.uncecomp.org/proceedings/pdf/21225.pdf)

- **A. Van Biesbroeck**, C. Gauchy, J. Garnier & C. Feau (2023). “Connections between reference prior theory and global sensitivity analysis, an illustration with f-divergences”. *Proceedings des 54èmes Journées de Statistiques (JdS)*. HAL: [hal-04171446](https://hal-04171446)
- **A. Van Biesbroeck**, C. Gauchy, C. Feau & J. Garnier (2023). “Influence of the choice of the seismic intensity measure on fragility curves estimation in a Bayesian framework based on reference prior”. *Proceedings of the 5th Thematic Conference on Uncertainty Quantification in Computational Sciences and Engineering (UNCECOMP)*, pp. 94-111. DOI: [10.7712/120223.10327.19899](https://doi.org/10.7712/120223.10327.19899)

Each of these publications, whether it takes the form of a journal paper or a proceedings paper, proposes an original contribution that is tied to this thesis. Naturally, the contributions presented as journal papers are considered as richer.

Apart from two exceptions, each of these contributions lies in the center of a chapter or an appendix of this manuscript. The first exception is the proceedings paper published at JdS in 2023. Indeed, this work constituted an introduction to the study that led to the journal paper entitled “Generalized mutual information and their reference priors under Csizar f-divergence”. In this manuscript, the proceedings paper is incorporated into the chapter tied with the latter ([chapter 4](#)). The second exception is the proceedings paper published at SMiRT in 2025. The original methodology presented in this paper remains conceptual, and its development faced the end of the thesis. The study thus lacks for an in-depth examination that might be conducted in the future.



# 2

## Introduction en français

**Résumé** Ce chapitre ne diffère du [chapitre 1](#) que par sa rédaction en langue française. En premier lieu, nous introduisons et motivons les travaux de recherche qui ont été menés au cours de cette thèse. Ils s'inscrivent dans une problématique qui vise à améliorer les méthodes d'études sismiques probabilistes du fait du manque de réponse qu'apporte l'état de l'art à la question du choix du prior en inférence bayésienne. En second lieu, nous présentons l'organisation du manuscrit qui s'appuie sur une reformulation de la problématique sous la forme de six principales questions.

<b>2.1 Motivation et positionnement de la thèse</b>	27
2.1.1 Etudes probabilistes de sûreté	27
2.1.2 La quantification d'incertitudes dans les études probabilistes de sûreté	29
2.1.3 Le choix du prior dans les études Bayésiennes	31
2.1.4 Ce qui motive la recherche sur l'élicitation de priors pour les études sismiques probabilistes de sûreté	33
<b>2.2 Esquisse du manuscrit et contributions</b>	33
2.2.1 Problématiques et plan	33
2.2.2 Liste des contributions	34

### 2.1 Motivation et positionnement de la thèse

#### 2.1.1 Etudes probabilistes de sûreté

##### Historique

Les études probabilistes de sûreté (EPS) désignent un ensemble de méthodes d'analyses techniques qui permettent de quantifier un risque encouru par une installation lorsqu'elle est sujette à un événement. L'événement peut être d'origine naturelle comme artificielle (technologique, humaine, etc), il peut être de provenance interne comme externe ; il peut s'agir d'un séisme, d'une inondation, d'une combinaison de défaillances internes, entre autres.

Faisant suite aux premières recommandations de F. R. Farmer (expert sûreté à la UK atomic authority) dans les années 1960 pour la fiabilité des installations nucléaires, ces méthodes ont pour caractéristique commune l'introduction de la notion d'incertitude dans la qualification de l'événement, ses phénomènes et ses caractéristiques (on parle alors d'aléa). Leur cadre et leur concept ont été rapidement adoptés et développés aux Etats-Unis (cf. le rapport de la Nuclear Regulatory Commission, NRC, [1983](#)). En particulier, de nombreuses études sont venues dès 1968 (Cornell, [1968](#)) y inscrire les analyses de fiabilité parasismique, définissant alors le cadre des études sismiques probabilistes de sûreté (SPRA en anglais).

Le séisme représente en effet un facteur de risque remarquable des études de sûreté. Premièrement, bien que communément caractérisé par sa magnitude et sa distance à la source, son signal est bien plus riche qu'une fonction bivariée et deux séismes de même magnitude et de même source peuvent avoir des caractéristiques (et des conséquences) significativement différentes. Deuxièmement, puisqu'il touche à la fois tous les éléments aussi bien externes qu'internes à l'installation, il est la potentielle source de conséquences lourdes sur les équipements et structures. Le coût potentiel des conséquences d'un aléa sismique peut alors être élevé et critique dans le contexte nucléaire, ce qui en fait un événement d'intérêt majeur et décisif même dans les zones géographiques où il est rare.

Aujourd'hui, la prise en compte de l'aléa sismique sous le cadre des études probabilistes de sûreté est une recommandation internationale dans le contexte de l'industrie nucléaire. Leur cadre d'application dans l'industrie nucléaire française est précisé par l'autorité de sûreté nucléaire et de radioprotection (ASN)<sup>1</sup> dans la règle fondamentale de sûreté (ASN, 2002).

### L'évolution des EPS en France, au CEA, et dans cette thèse

L'évolution des méthodes et des connaissances relatives à la sûreté des installations nucléaires se fait de manière parallèle à l'évolution de la règle et de la norme imposées à celles-ci. Sur le sujet de l'aléa sismique, le rapport à l'évolution de la règle et de la méthode (et donc des EPS) oppose différents points de vue, principalement entre le principal exploitant (EDF) et les experts de l'autorité de sûreté (anciennement l'IRSN, maintenant unifié avec l'autorité de sûreté nucléaire devenue l'ASN). Pour le premier, la robustesse de l'installation n'est normalement pas remise en question par l'avancée des connaissances et des méthodes puisque l'incertitude sur celles-ci fait parti des marges prises en compte à la construction de l'équipement. Le second pense au contraire que la robustesse est une question perpétuelle et que la marge n'est pas faite pour être mordue au fil des connaissances qui s'ajoutent (Roger, 2020). L'ASN se place alors en arbitre dans ce dialogue, entre autre elle définit le "séisme majoré de sécurité"<sup>2</sup> (il s'agit d'une majoration du spectre du "séisme maximal historiquement vraisemblable" qui est défini par les sismologues) qui produit la marge de référence dans la démonstration de robustesse parasismique d'un équipement.

Cet arbitrage est donc à la fois sensible et critique. L'incident survenu en 2011 à la centrale nucléaire de Fukushima-Daiichi le démontre. L'aléa sismique (qui est la cause du tsunami) de référence a été sous-évalué par le consensus des experts japonais, et c'est cette sous-évaluation qui est au final la cause de l'incident (cf. le rapport de l'agence internationale de l'énergie atomique, IAEA, 2015). Suite à cet événement, l'ASN a pris la décision de majorer d'un facteur 1.5 le séisme majoré de sécurité<sup>3</sup> en France pour les équipements du "noyau dur" (liste limitée de systèmes, structures et composants essentiels au maintien des fonctions de sûreté d'une installation nucléaire). Ceci imposant alors une démonstration amplifiée de la robustesse des différentes installations par leur exploitant.

Dans ce cadre, la place du CEA est ambivalente dans l'échiquier nucléaire français. D'une part, il est exploitant d'installations nucléaires de recherche et joue alors son propre rôle quant à la démonstration de robustesse de ses équipements. Aussi, il participe conjointement à des actions de recherche relatives au parc nucléaire civil exploité par EDF. Une partie de la recherche et l'expertise sur les conséquences de l'aléa sismique se fait au CEA au laboratoire EMSI, qui dispose d'une plateforme expérimentale (Tamaris) qui permet de procéder à divers tests mécaniques sur des équipements sous séismes. L'approfondissement des méthodes des études sismiques probabilistes de sûreté est un enjeu pour le CEA qui s'inscrit dans les missions de ce laboratoire. Le CEA étant responsable devant les autorités de sûreté de ses installations, il cherche à développer des méthodes toujours rigoureuses pour justifier de leur tenue sismique.

Cette thèse s'inscrit dans la démarche d'évolution du SPRA et de démonstration de sûreté. Financée par le CEA dans le cadre d'un contrat de formation par la recherche, elle a pour objectif d'étoffer l'état de l'art en terme d'estimation de la fragilité sismique des équipements et des installations. L'objectif est de développer des méthodes qui doivent le plus possible être (i) efficaces devant la complexité des objets étudiés et de l'aléa, (ii) être robustes dans le temps et devant une potentielle réévaluation des critères de sûreté, et (iii) être

<sup>1</sup> Anciennement ASN: l'autorité de sûreté nucléaire (ASN), a été unifiée avec l'institut de radioprotection et de sûreté nucléaire (IRSN) le 1<sup>er</sup> janvier 2025.

<sup>2</sup> Aujourd'hui plutôt appelé "séisme de dimensionnement".

<sup>3</sup> Définissant alors le "séisme noyau dur".

transparentes et auditables par les autorités de sûreté.

### 2.1.2 La quantification d'incertitudes dans les études probabilistes de sûreté

#### Principe et étapes de la quantification d'incertitudes

Les études probabilistes de sûreté s'appuient sur la prise en compte et la quantification des incertitudes qui apparaissent dans le procédé d'évaluation d'un risque au sein d'un système physique. La quantification d'incertitudes représente en fait une démarche qui consiste à modéliser et analyser méthodiquement l'incertitude, sa source, et sa propagation au sein de la modélisation du système physique étudié. Elle se construit alors à l'intersection de la physique, de l'ingénierie et des mathématiques appliquées, et s'érige même comme une branche des statistiques à part entière.

Elle concentre son étude sur 3 éléments qui décrivent l'interaction entre le système physique et l'aléa : des paramètres d'entrée  $\mathbf{X}$ , une réponse physique du système  $Y$ , et une modélisation de celle-ci selon :  $Y = \mathcal{M}(\mathbf{X})$ . Cette fonction  $\mathcal{M}$  est liée à la modélisation, souvent complexe, du système et de ses propriétés physiques (il peut s'agir de la résolution d'équations physiques, de simulations numériques, ou même du résultat d'une expérience mécanique). La démarche générale de la quantification d'incertitudes est souvent décrite au travers de différentes étapes clés (voir par ex. Sudret, 2007; Iooss, 2009), à savoir :

1. L'identification des sources d'incertitudes : il est usuel de classifier les incertitudes selon deux grandes catégories. D'une part les incertitudes irréductibles, qui proviennent du hasard "naturel" embarqué dans l'aléa et le système physique. D'autre part, les incertitudes épistémiques, qui existent par manque d'information, et qui sont alors considérées comme réductibles en opposition aux précédentes (Hüllermeier & Waegeman, 2019). Il s'en suit une modélisation probabiliste de l'entrée  $\mathbf{X}$ .
2. La propagation des incertitudes : cette étape correspondant à l'approximation de la distribution du modèle  $\mathcal{M}(\mathbf{X})$ , et de l'évaluation d'une quantité d'intérêt qui dépend de celle-ci : comme une variance, une probabilité de défaillance, etc ; plus généralement une quantité de la forme  $\mathbb{E}\phi(Y)$ .
3. L'analyse de sensibilité : son rôle est de comprendre avec plus de recul la manière dont les incertitudes se propagent dans un système, en identifiant l'impact d'un ou de plusieurs des paramètres d'entrées sur la sortie  $Y$  (Iooss & Lemaître, 2015).

#### Des outils mathématiques en quantification d'incertitudes

De nombreux outils mathématiques viennent appuyer les étapes sus-mentionnées. Bien que ce manuscrit n'ait pas pour but d'en décrire un panel exhaustif, ci-dessous sont listés quelques exemples qui reviennent dans les travaux de cette thèse.

- La métamodélisation permet de contrevainir à la complexité d'évaluation du modèle  $\mathcal{M}$ , en construisant un méta-modèle  $\tilde{\mathcal{M}}$ , entraîné à partir d'une base de données d'observations  $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_k, y_k)$ . Le méta-modèle se veut par principe plus simple à évaluer que le modèle réel. Rien ne limite sa définition explicitement qui peut aller d'un simple modèle paramétrique à la sortie d'un réseau de neurones "boîte noire" ; parmi les plus communs on peut évoquer la métamodélisation par processus gaussien (ou krigage) et par polynômes du chaos.
- Les indices globaux de sensibilité s'inscrivent pleinement dans l'étape d'analyse de sensibilité du système. Ils constituent depuis les premiers travaux de Sobol' (1993) des outils essentiels pour mesurer statistiquement comment  $Y$  est impactée par un ou plusieurs des  $X_i$  (où  $(X_1, \dots, X_p) = \mathbf{X}$ ). Dans ce cadre, l'impact, noté  $S_i$ , de l'entrée  $X_i$  sur  $Y$  s'exprime comme une divergence moyenne entre la distribution  $\mathbb{P}_Y$  de  $Y$  et sa distribution conditionnellement à  $\mathbf{X}$ ,  $\mathbb{P}_{Y|\mathbf{X}}$  (Da Veiga, 2015) :

$$S_i = \mathbb{E}_{X_i}[D(\mathbb{P}_Y || \mathbb{P}_{Y|X_i})], \quad (2.1)$$

où  $D$  est une mesure de dissimilarité entre deux mesures de probabilité. Le choix de  $D$  est alors décisif dans l'étude de  $S_i$ . On peut noter, par exemple, que choisir  $D$  définie par  $D(P||Q) = \|\mathbb{E}_{X \sim P} X - \mathbb{E}_{X \sim Q} X\|$

$\mathbb{E}_{X \sim Q} X \|^2$  revient à étudier un indice de Sobol' du premier ordre. Le choix de  $D$  peut être motivé par divers intérêts, par exemple, celui de discriminer l'indépendance entre  $X_i$  et  $Y$ . Le point suivant détaille un large panel de choix possibles pour  $D$ .

- La théorie de l'information est au centre de la comparaison de mesures de probabilité et donc de la quantification d'incertitudes. Par essence, les manières de comparer la façon qu'ont deux mesures de probabilité de distribuer l'information dans un espace sont vastes. Dans la définition de  $S_i$  plus haut, le choix de  $D$  (et donc le choix de définition de  $S_i$ ) relève de la théorie de l'information.

Un choix alors commun, perçu comme une extension de l'entropie de Shannon, est celui des  $f$ -divergences de Csiszár (1967). On peut noter qu'elles sont employées dans des domaines variés, au-delà de l'analyse de sensibilité, tel que l'inférence variationnelle (Minka, 2005; Bach, 2023), le design de métamodèles (Nguyen, Wainwright, & Jordan, 2009), l'apprentissage PAC-bayésien (Picard-Weibel & Guedj, 2022). Lorsque  $f$  est convexe et que  $f(1) = 0$ , elles sont définies selon :

$$D_f(P||Q) = \int_{\mathcal{X}} f\left(\frac{p(x)}{q(x)}\right) q(x) d\omega(x), \quad (2.2)$$

en notant  $p, q$  des densités associées à  $P$  et à  $Q$  par rapport à une mesure commune  $\omega$  sur leur ensemble de définition  $\mathcal{X}$ . Les  $f$ -divergences réduisent le choix de  $D$  à celui de  $f$ , les plus communément employées étant les  $\delta$ -divergences (H. Zhu & Rohwer, 1995) définies en fixant  $f = f_\delta$  où

$$f_\delta(x) = \begin{cases} \frac{x^\delta - \delta x - (1-\delta)}{\delta(\delta-1)} & \text{si } \delta \notin \{0, 1\}, \\ x \log x - x + 1 & \text{si } \delta = 1, \\ -\log x + x - 1 & \text{si } \delta = 0. \end{cases} \quad (2.3)$$

Il est remarquable que la plus connue des mesures de dissimilarité (la divergence de Kullback-Leibler) peut être vue comme un cas particulier des  $\delta$ -divergences. En effet, rappelons que la divergence de Kullback-Leibler, notée KL, est définie par

$$\text{KL}(Q||P) = \int_{\mathcal{X}} \log\left(\frac{q(x)}{p(x)}\right) q(x) d\omega(x) \quad (2.4)$$

(en adoptant les mêmes notations qu'en équation (2.2)), si bien que

$$D_{f_\delta}(P||Q) = \text{KL}(P||Q) \text{ si } \delta = 1, \quad D_{f_\delta}(P||Q) = \text{KL}(Q||P) \text{ si } \delta = 0. \quad (2.5)$$

Bien sûr, il existe aussi de nombreuses mesures de dissimilarité qui ne sont pas des  $f$ -divergences, à ce titre, les indices de Sobol' du premier ordre évoqués plus haut font office d'exemple. Un autre exemple notable présenté ici est celui des normes maximales de discrépances (MMD). Soient  $P, Q$  deux mesures de probabilité définies sur un ensemble  $\mathcal{X}$ , et soit  $\mathcal{H}$  un espace de Hilbert à noyau reproduisant (RKHS ; un rappel sur cette notion et ses définitions est proposé dans le point suivant) sur  $\mathcal{X}$ , dont le noyau reproduisant est noté  $k : \mathcal{X} \times \mathcal{X} \longrightarrow \mathbb{C}$ . On définit la MMD (Gretton et al., 2012) par

$$\text{MMD}(\mathcal{H}; P||Q) = \sup_{\substack{f \in \mathcal{H} \\ \|f\|_{\mathcal{H}} \leq 1}} |\mathbb{E}_{X \sim P} f(X) - \mathbb{E}_{X \sim Q} f(X)|, \quad (2.6)$$

ou, dans une forme plus simple :

$$\text{MMD}^2(\mathcal{H}; P||Q) = \mathbb{E}_{X, X' \sim P \otimes P} [k(X, X')] + \mathbb{E}_{Y, Y' \sim Q \otimes Q} [k(Y, Y')] - 2\mathbb{E}_{X, Y \sim P \otimes Q} [k(X, Y)]. \quad (2.7)$$

- Les espaces de Hilbert à noyau reproduisant, bien que peu utilisés dans cette thèse, représentent un outil performant pour modéliser et représenter des objets complexes. Ils permettent, via l'emploi de leur noyau, de manipuler des objets non linéaires dans un espace linéaire de Hilbert. Dans le cas de la MMD définie précédemment, les RKHS sont employés pour représenter et comparer des mesures de probabilité.

Un RKHS consiste en la donnée d'un espace de hilbert  $\mathcal{H}$ , sous-espace d'un ensemble de fonction  $\mathcal{F} = \{f : \mathcal{X} \rightarrow \mathbb{C}\}$ , et tel que pour tout  $x \in \mathcal{X}$ , l'application  $f \mapsto f(x)$  est continue sur  $\mathcal{H}$ . Le noyau reproduisant est alors l'unique application  $k : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}$  telle que pour tout  $x \in \mathcal{X}$ ,  $f \in \mathcal{H}$ ,

$$f(x) = \langle f, k(\cdot, x) \rangle_{\mathcal{H}}, \quad (2.8)$$

où  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$  exprime le produit scalaire sur  $\mathcal{H}$ . Entre autres propriétés fondamentales de la théorie, le théorème d'Aronszajn (Aronszajn, 1950) assure l'existence d'un RKHS par la donnée seule d'un noyau symétrique et défini positif sur l'ensemble  $\mathcal{X}$ . Pour approfondir la théorie, nous renvoyons le lecteur intéressé vers Schölkopf & Smola, 2001.

Parmis les RKHS populaires, on peut citer celui défini par le noyau Gaussien (appelé aussi RBF) sur  $\mathbb{R}^d$ :  $k(x, y) = \exp(-\tau \|x - y\|_2^2)$  pour un certain  $\tau > 0$ .

- L'analyse bayésienne est un outil incontournable de la quantification d'incertitude, puisqu'elle a pour principe fondamental d'introduire une forme d'incertitude dans un ou plusieurs des paramètres du modèle étudié. L'incertitude incorporée par le cadre bayésien peut prendre diverses formes, et se traduit au travers de la définition de *l'a priori*. La propagation d'incertitudes au travers de ce canevas se quantifie en calculant *l'a posteriori*, à partir d'observations statistiques  $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_k, y_k)$ . La définition de *l'a priori* est critique, elle prend place dans l'étape d'identification des incertitudes, et se révèle impactante sur *l'a posteriori*, et donc, dans un contexte d'étude probabiliste de sûreté, sur le risque.

Puisque ce dernier outil, qui est en fait un paradigme à lui tout seul, est au centre de cette thèse, la section qui suit lui est pleinement consacrée.

### 2.1.3 Le choix du prior dans les études Bayésiennes

L'analyse bayésienne est une théorie des statistiques qui se construit selon une interprétation des probabilités qui sont vues comme des degrés de plausibilité des événements, on parle généralement de "crédibilité". Comme toute étude statistique, elle cherche à faire le lien entre des observations statistiques et leur distribution probabiliste. Elle s'appuie sur le théorème de Bayes, en considérant l'inconnue qui permet de définir la distribution des données comme étant aléatoire. Les observations statistiques informent alors la distribution de cette inconnue, ce qui met à jour le degré de crédibilité attribué à une distribution que suivraient les données par rapport à une autre.

Le raisonnement contraire, dit fréquentiste, considère la probabilité comme une fréquence d'occurrence observable au prix de reproduire l'expérience. Sous ce paradigme, l'inconnue existe, et les fréquences données par la répétition de l'expérience sont plus probables selon que l'on considère certaines valeurs de l'inconnue plutôt que d'autres, on parle généralement de "confiance".

Au final, il est important de noter que ces deux raisonnements se complètent, voire parfois se confondent suivant les cas, et ils sont loin d'être antinomiques. Dans un contexte pratique, le cas d'étude et ses spécificités peuvent néanmoins orienter le choix de favoriser l'un par rapport à l'autre.

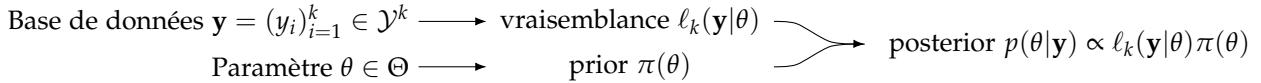
Dans cette thèse, on se concentre sur l'analyse bayésienne dans le cadre de l'inférence statistique. Dans ce contexte, une variable d'intérêt  $Y \in \mathcal{Y}$  est modélisée selon une distribution paramétrique  $\mathbb{P}_{Y|\theta}$ ,  $\theta \in \Theta$  étant l'inconnue, alors considérée comme étant une variable aléatoire dans le canevas Bayésien. En toute logique,  $\theta$  admet une distribution, que l'on appelle le **prior**. Lorsque des réalisations de  $Y$  successives et indépendantes conditionnellement à  $\theta$  sont observées, le **posterior** se définit comme la distribution de  $\theta$  conditionnellement à ces observations. Une région de crédibilité à un niveau  $r$  se définit comme une région de  $\Theta$  dont la probabilité selon le posterior est au moins égale à  $r$ . On notera que dans ce contexte, divers résultats lient les approches bayésienne et fréquentiste d'estimation de  $\theta$ . Le théorème de Berstein Von-Mises (voir par ex. Vaart, 1992) en est un exemple.

Notons  $\mathbf{y} = (y_1, \dots, y_k)$  les valeurs observées de  $k$  réalisations de  $Y$ . En supposant que tous ces objets existent et sont bien définis, notons  $\pi$  la densité du prior (selon une mesure  $\nu$ ),  $\ell(\cdot|\theta)$  les densités de  $\mathbb{P}_{Y|\theta}$  (selon une mesure  $\mu$ ) et  $\ell_k(\mathbf{y}|\theta) = \prod_{i=1}^k \ell(y_i|\theta)$ , appelée la vraisemblance du modèle. Alors avec ces notations

le théorème de Bayes définit la densité du posterior (par rapport à  $\nu$ ) selon :

$$p(\theta|y) = \frac{\ell_k(y|\theta)\pi(\theta)}{\int_{\Theta} \ell_k(y|\theta)\pi(\theta)d\nu(\theta)}. \quad (2.9)$$

Cette formule fait le lien entre le prior, le modèle (représenté par la vraisemblance), les données, et le posterior qui traduit la prédiction crédible. Elle permet de schématiser le canevas bayésien comme une chaîne de transmission d'information vers l'*a posteriori*, depuis deux sources : le "réel" (les observations) et l'*a priori*. Ce schéma est celui présenté ci-dessous.



Si l'information issue des données ne pose normalement pas question, c'est l'information *a priori* qui concentre l'attention. Par essence, elle est la source de l'incertitude dans le résultat, sa construction doit se faire de manière attentive lors de l'étape de modélisation et ne peut être arbitraire. Cependant, la formulation bayésienne lui attribue un impact significatif sur le posterior, particulièrement dans des cas où la taille de la base de données est limitée. Une illustration de ce phénomène est proposée en figure 2.1, où différents priors, en apparence proches, donnent lieu à des posteriors très différents, à partir des mêmes observations. Bien sûr, le diable est dans les détails, et ici, ce sont les queues de distribution des priors, peu discernables sur l'illustration, qui produisent cette différence. Malgré tout cet exemple questionne, à juste titre, sur le choix du prior.

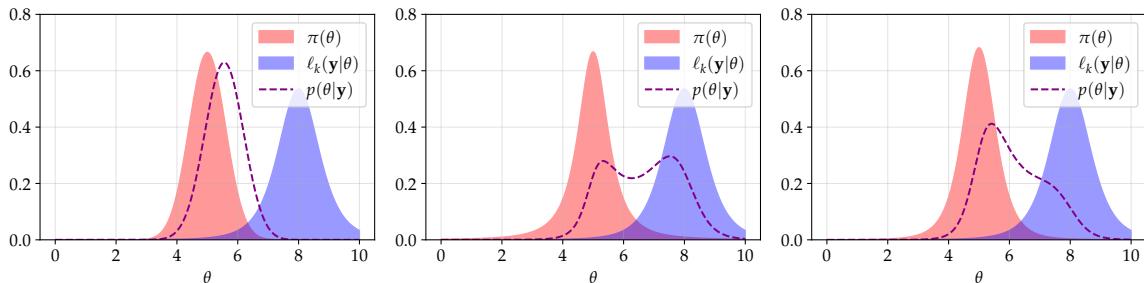


Figure 2.1: Différents calculs de posteriors (en pointillé) à partir d'une même vraisemblance (bleu) mais de différents priors (rouge).

La question de la construction du prior reste une question ouverte dans la littérature. Pour certains, elle est l'opportunité d'ajouter de l'information provenant d'une source extérieure, il peut s'agir d'un jugement d'expert, de données historiques, ou de connaissances sur certaines propriétés attendues du résultat. Leurs priors sont plutôt qualifiés d'informatifs. Pour d'autres, elle est au contraire le moyen d'introduire une incertitude dans le modèle, l'*a priori* est alors dans un tel cas une absence de connaissance, qui vient laisser l'information provenir essentiellement des données. Ces derniers ont confiance en la modélisation de la vraisemblance, et leur prior sera qualifié de non-informatif.

On comprend au travers de ces paragraphes introductifs que cette question est insoluble : il ne peut pas exister de méthode unanime de sélection de prior. Elle illustre même l'un des "trous" du workflow bayésien selon Gelman & Yao (2020) : les priors trop informés ne permettent pas d'avoir confiance en le moindre résultat tandis que les priors trop peu informés donnent lieu à des zones de crédibilité parfois trop larges et inutilisables. Ce n'est pas pour autant que l'approche bayésienne perde de son intérêt et cette section n'est pas vouée à discréditer la démarche, la suite du manuscrit le prouvera. Mais elle doit laisser celui ou celle qui l'emploie conscient ou consciente des impacts et des limites de ses choix. Le choix d'un prior est aussi critique que le choix du modèle lui-même.

### 2.1.4 Ce qui motive la recherche sur l'élicitation de priors pour les études sismiques probabilistes de sûreté

L'analyse bayésienne a gagné en popularité dans de nombreux domaines, y compris dans des études de fiabilité ou de sûreté. En effet, elle est plébiscitée pour sa capacité (i) à introduire et propager une incertitude dans un problème d'inférence, et (ii) à être plus régulière que beaucoup de méthodes fréquentistes, en particulier lorsque le nombre de données observées est limité.

Le cadre de l'inférence bayésienne s'introduit en effet bien dans la démarche de quantification d'incertitudes. L'identification des incertitudes sur l'entrée  $\mathbf{X}$  ou sur le modèle même  $\mathcal{M}$  (ou bien s'il y a lieu le métamodèle) peut se faire au travers d'une paramétrisation de celles-ci, en donnant lieu à une distribution paramétrique de  $Y$  :  $\mathbb{P}_{Y|\mathbf{X},\theta}$ . La variable  $\theta$ , aléatoire sous le paradigme bayésien, peut alors représenter diverses sources d'incertitudes, de différentes natures selon les cas (Bousquet, 2024a). Cette approche a été employée à pléthore dans divers types de modèles, on peut citer par exemple les cas des processus gaussiens (par ex. Gu & Berger, 2016) ou des réseaux de neurones (Arbel et al., 2023).

Parfois, certaines études adoptent la démarche pour la possibilité qu'elle offre d'introduire un jugement ou une connaissance *a priori*. C'est pourtant ce même point qui fait aussi débat puisqu'il est la potentielle source d'une subjectivité difficile à justifier. Dans le cadre des études sismiques probabilistes de sûreté en industrie nucléaire, la démonstration de robustesse est centrale. Les cas d'études y étant souvent très complexes, on trouvera alors parfois autant de priors qu'il y a d'experts en suivant cette démarche. De tels cas caricaturaux mettent en péril l'auditabilité de la moindre région de crédibilité *a posteriori*.

On comprend en fait que dans une étude de sûreté qui se veut auditable, une "bonne" région de crédibilité sur un paramètre estimé n'est pas une région nécessairement étroite, même si elle suggère que la sûreté est satisfaite. A l'inverse, une région trop large n'est pas pour autant "bonne" car elle ne permet pas de démontrer la robustesse attendue de l'équipement. L'équilibre est complexe, et le "bon" prior est celui qui, bien que démontrant au mieux la capacité ou l'incapacité de l'équipement à résister à l'aléa, produit un résultat inspirant confiance.

Toutes ces réflexions mettent en lumière l'intérêt d'une construction méthodique du prior dans le cadre des études sismiques probabilistes de sûreté, au travers d'un cadre rigoureux qui vise à éviter, dans la mesure du possible, l'introduction de la moindre subjectivité dans la démarche.

## 2.2 Esquisse du manuscrit et contributions

### 2.2.1 Problématiques et plan

Plusieurs questions principales se détachent de cette introduction et de ces premières réflexions. Elles sont listées ci-dessous, et décrivent les problématiques majeures qui se sont manifestées au cours de cette thèse et auxquelles les travaux présentés dans ce manuscrit cherchent à répondre.

**Question i** Comment peut-on définir et appuyer l'objectivité d'un prior ?

**Question ii** Quelles sont les limites de l'emploi d'un tel prior, peu informatif par essence, et comment concilier son usage avec les besoins pratiques ?

**Question iii** Comment construire et implémenter de tels priors dans un cas pratique ?

**Question iv** Dans le contexte des études sismiques probabilistes de sûreté, comment s'implémentent ces priors objectifs dans un modèle concret d'estimation de fiabilité sismique ?

**Question v** Quelles conséquences peut avoir le manque d'information *a priori* ou provenant des données dans ce modèle et comment les solutionner ?

**Question vi** Comment peut-on bénéficier au mieux des différentes sources d'information (celle *a priori* et celle issue des données) dans le workflow bayésien du modèle étudié dans son ensemble ?

Dans cette thèse, nous tentons de répondre à ces six problématiques au travers d'une étude selon deux axes. Le premier axe a une dimension qualifiée de théorique, en adressant les problématiques liées à la construction de prior dits "objectifs" dans son ensemble. Il se concentre sur le développement d'une théorie appelée la théorie des priors de référence. Le second axe est alors plutôt qualifié de pratique, en étudiant la mise en œuvre des développements théoriques sur des cas d'études réels, en se concentrant sur l'estimation de courbes de fragilité sismique, qui représentent un outil important du cadre des études sismiques probabilistes de sûreté. Il reste important de rappeler que, bien qu'indépendants, les deux grands axes de travail sus-mentionnés s'alimentent l'un l'autre. Ce sont autant les problématiques pratiques qui ont motivé les recherches théoriques, que les découvertes théoriques qui ont été mises au service des études pratiques.

La **partie I** de ce manuscrit est alors dévouée au développement de la théorie des priors de référence.

Dans le [chapitre 3](#), on introduit la théorie et on développe son état de l'art. Il permet d'introduire la notion de prior de référence telle qu'elle est communément définie dans la littérature, et sa place parmi les priors objectifs. Ce chapitre permet aussi de formaliser le cadre mathématique de travail pour la suite de la partie.

Dans le [chapitre 4](#), on observe que la définition des priors de référence s'appuie sur une mesure de dissimilarité. On cherche alors dans ce chapitre à généraliser la définition de cette dernière afin d'appuyer le caractère objectif des priors de référence. Ce chapitre apporte une réponse à la **question i**.

Dans le [chapitre 5](#), on cherche une ouverture au cadre des priors de référence, lorsqu'il donne lieu à des priors difficilement utilisables en pratique. On propose alors une définition de ce qu'on appelle priors de référence contraints. Les solutions de cette définition apportent une réponse à la **question ii**.

Dans le [chapitre 6](#), on développe une méthode numérique d'approximation des priors de référence. Cette méthode s'appuie sur l'inférence variationnelle en définissant le prior comme la sortie d'un réseau de neurones, afin d'éviter un calcul explicite de celui-ci. Ce chapitre apporte une réponse à la **question iii**.

La **partie II** de ce manuscrit aborde le sujet de l'estimation bayésienne dite objective des courbes de fragilité sismique.

Dans le [chapitre 7](#), on introduit les courbes de fragilité sismique, leur définition, leur historique et leur intégration dans le cadre des études sismiques probabilistes de sûreté. Un état de l'art des méthodes d'estimation de ces courbes suivant les différents types de données à disposition y est proposé.

Dans le [chapitre 8](#), on propose un calcul et une étude en profondeur du prior de référence pour un modèle classique employé pour l'estimation de ces courbes de fragilité. L'inférence bayésienne via l'emploi de ce prior y est comparée à d'autres méthodes. Ce chapitre apporte une réponse à la **question iv**.

Dans le [chapitre 9](#), on s'intéresse en profondeur aux limites du modèle considéré pour l'estimation des courbes de fragilité. On propose une solution qui est appuyée par les développements théoriques de la première partie. Ce chapitre apporte une réponse à la **question v**.

Dans le [chapitre 10](#), on termine notre étude en proposant une méthode qui vient appuyer l'estimation bayésienne en optimisant l'information issue des données non plus seulement lors du choix du prior mais également lors de la sélection de celles-ci, au travers d'une méthode de planification d'expériences. Ce chapitre apporte une réponse à la **question vi**.

Une conclusion générale viendra clore le manuscrit dans une **partie III** finale.

### 2.2.2 Liste des contributions

La recherche conduite au cours de cette thèse a donné lieu à plusieurs contributions dans la littérature scientifique. Ci-après sont listés les travaux publiés et soumis pendant celle-ci.

#### Articles de revues

- A. Van Biesbroeck, C. Gauchy, C. Feau & J. Garnier (2025). "Robust a posteriori estimation of probit-lognormal seismic fragility curves via sequential design of experiments and constrained reference

prior". *arXiv* 2503.07343. DOI: [10.48550/arXiv.2503.07343](https://doi.org/10.48550/arXiv.2503.07343).

- N. Baillie, A. Van Biesbroeck & C. Gauchy (2025). "Variational inference for approximate reference priors using neural networks". *arXiv* 2502.02364. DOI: [10.48550/arXiv.2502.02364](https://doi.org/10.48550/arXiv.2502.02364)
- A. Van Biesbroeck (2024). "Properly constrained reference priors decay rates for efficient and robust posterior inference", *arXiv* 2409.13041. DOI: [10.48550/arXiv.2409.13041](https://doi.org/10.48550/arXiv.2409.13041)
- A. Van Biesbroeck, C. Gauchy, C. Feau & J. Garnier (2025). "Design of experiments based on a low fidelity model for seismic fragility curves estimation". *ESAIM: ProcS* (to appear). HAL: [hal-04719458v1](https://hal.archives-ouvertes.fr/hal-04719458v1)
- A. Van Biesbroeck (2023). "Generalized mutual information and their reference priors under Csizar f-divergence". *arXiv* 2310.10530. DOI: [10.48550/arXiv.2310.10530](https://doi.org/10.48550/arXiv.2310.10530)
- A. Van Biesbroeck, C. Gauchy, C. Feau & J. Garnier (2024). "Reference prior for Bayesian estimation of seismic fragility curves". *Probabilistic Engineering Mechanics*, 76, pp 103622. DOI: [10.1016/j.probengmech.2024.103622](https://doi.org/10.1016/j.probengmech.2024.103622)

### Articles de conférences

- A. Van Biesbroeck, C. Feau & J. Garnier (2025). "Design of experiments for efficient and conform Bayesian learning of seismic fragility curves". *Proceedings of the 28th conference on Structural Mechanics in Reactor Technology (SMiRT)*.
- N. Baillie, A. Van Biesbroeck, C. Feau & C. Gauchy (2025). "Bayesian estimation of seismic fragility curves based on variational reference priors using neural networks". *Proceedings of the 6th Thematic Conference on Uncertainty Quantification in Computational Sciences and Engineering (UNCECOMP)*. URL: [2025.uncecomp.org/proceedings/pdf/21225.pdf](https://2025.uncecomp.org/proceedings/pdf/21225.pdf)
- A. Van Biesbroeck, C. Gauchy, J. Garnier & C. Feau (2023). "Connections between reference prior theory and global sensitivity analysis, an illustration with f-divergences". *Proceedings des 54èmes Journées de Statistique (JdS)*. HAL: [hal-04171446](https://hal.archives-ouvertes.fr/hal-04171446)
- A. Van Biesbroeck, C. Gauchy, C. Feau & J. Garnier (2023). "Influence of the choice of the seismic intensity measure on fragility curves estimation in a Bayesian framework based on reference prior". *Proceedings of the 5th Thematic Conference on Uncertainty Quantification in Computational Sciences and Engineering (UNCECOMP)*, pp. 94-111. DOI: [10.7712/120223.10327.19899](https://doi.org/10.7712/120223.10327.19899)

Chacune de ces publications, qu'elle prenne la forme d'un article de journal ou de conférence propose une contribution originale et liée à cette thèse. Naturellement, les contributions présentées dans des articles de revues sont considérées comme plus percutantes.

Mis à part deux exceptions, chacune de ces contributions se voit au centre d'un chapitre ou d'une annexe de ce manuscrit. La première exception est l'article de conférence publié aux JdS de 2023. En effet ce travail a constitué une introduction à l'étude qui a donné lieu plus tard à l'article de revue intitulé "Generalized mutual information and their reference priors under Csizar f-divergence". Dans cette thèse, l'article de conférence est donc incorporé au chapitre lié à ce dernier papier de revue ([chapitre 4](#)). La seconde exception est l'article de conférence publié à SMiRT en 2025. La méthodologie originale proposée dans cet article reste conceptuelle et s'est heurtée au calendrier de fin de thèse. Elle manque alors d'un approfondissement qui pourrait lui être apporté dans le futur.



# I

## CONTRIBUTION TO THE REFERENCE PRIOR THEORY



# 3

## Review of the reference prior theory

**Abstract** The reference prior theory is a prominent tool when one researches for a prior embedded with an objective construction. It is designed to ensure the information in the posterior comes promptly from the observations rather than from the prior itself. In this chapter, the theory is reviewed, and its formalism is rigorously introduced. We recall the main results about the reference priors that are proven in the literature. Additionally, because we find the usual Bayesian framework limiting for the study of reference priors, we propose a novel Bayesian formalism that provides a rigorous incorporation of improper priors. The latter fixes the notations we use in rest of the manuscript.

3.1	Introduction	39
3.2	The standard Bayesian framework	41
3.2.1	A minimal framework	41
3.2.2	Likelihood and densities	41
3.2.3	Improper priors	42
3.3	Mutual information and reference priors	42
3.3.1	Mutual information	42
3.3.2	Reference priors	43
3.3.3	Properties and Jeffreys prior	45
3.4	Limits of the classical framework of reference priors and solution	47
3.4.1	Novel framework for Bayesian inference with improper priors	47
3.4.2	Notations and compatibility with the usual framework	48
3.4.3	Proof of proposition 3.1	51
3.5	Conclusive remarks	53

### 3.1 Introduction

Bayesian analysis offers a coherent framework for integrating prior information and updating beliefs with data. The incorporation of prior knowledge can be motivated by various needs, such as enhancing interpretability, introducing uncertainty, or embedding existing knowledge into the analysis. Actually, the selection of this a priori centralizes a lot of attention in Bayesian inference and can be a critical aspect of the method. As a matter of fact, it can significantly influence the posterior distribution and, consequently, the conclusions drawn from the analysis. Thus, an inconsistent or insufficiently justified choice of prior can jeopardize the validity of the entire study, a concern now well identified in the modern Bayesian workflow described in Gelman, Vehtari, et al., 2020, and echoed across applied domains.

A plethora of approaches exist in the literature to address the prior elicitation problem. They offer different strategies for different purposes. Comprehensive reviews of such methods are proposed in Mikkola et al., 2023 and in Consonni et al., 2018, for instance. Among the different challenges that are commonly addressed by these methods, we can emphasize (i) the incorporation of information, (ii) the ease of posterior inference, or (iii) the research for interpretability.

The former refers to the so-called “informative priors”, because their distributions tend to concentrate the information. They aim to incorporate genuine prior knowledge about parameters, often arising from past data, domain expertise, physical constraints, or theoretical understanding. The difficulty lies in doing so transparently, reproducibly, and justifiably, especially when seeking to avoid subjective bias.

Regarding the second, we can recall conjugate priors (see e.g. Robert, 2007) that are designed to provide a simple formulation of the posterior. Penalized Complexity (PC) priors (Simpson et al., 2017) and sparse priors (Castillo, Schmidt-Hieber, and Vaart, 2015) are made to be suitable for high-dimensional problems; and g-priors (Liang et al., 2008; Galharret and Philippe, 2023) are used for variable selection in normal regression models.

Thirdly, in order to provide a better interpretability of the information transmission process, several studies favor a hierarchical design, in which the prior is the result of a latent prior modeling. This approach allows the use of historical data, as in power priors (Chen and Ibrahim, 2006) and Information Matrix priors (M. Gupta and Ibrahim, 2009); or even imaginary data (Pérez and Berger, 2002; Spitzner, 2011). Posterior priors, developed for sensitivity analysis studies (Bousquet, 2024b), are also part of this approach. These hierarchical constructions are appreciated for their ability to tackle the improper aspect (they integrate to infinity) of most low-informative priors (priors whose distribution is more diffuse through the parameter’s space, in opposition with informative priors).

Many of these methods involve subjective choices that remain open to criticism. Often, they are based on the selection of a class and hyperparameters that remain to be tuned.

The research for so-called “objective priors” arose in contexts where the incorporation of uncritical beliefs is impossible (either by lack of available beliefs, or by lack of confidence within the existing ones). In those contexts, there is consensus that an appropriate prior belongs within the non-informative ones. In this respect, Lindley (1956) already suggested the maximization of the Shannon’s entropy, as a way to choose a prior that would provide the least possible information. The same philosophy has been the source of several rules, which lead to different definitions of objective priors. A review of the existing rules and their application on simple examples is proposed by Kass and Wasserman, 1996 and by Berger, Bernardo, and Sun (2015). We also mention the work of Datta and M. Ghosh (1996), in which the question of invariance by re-parametrization of several of these rules is thoroughly focused.

In this chapter and in this thesis, we focus on the reference prior theory, first introduced by Bernardo (1979a). Based on an “expected utility maximization”, the idea was to shift from minimizing the information in the prior to maximizing the knowledge brought by the observations over prior distribution itself. This notion of expected information, is defined with the support of the mutual information which allows for a formal construction of priors that are designed to let the data dominate the inference.

The reference prior theory has been extensively studied since its deepen formalization by Berger, Bernardo, and Sun (2009). The derivation of reference priors and the study of their properties in different contexts has been extensively focused. We can mention the works in Berger and Sun, 2008 for bivariate normal models, Berger and Bernardo, 1992b for multinomial models, and Gu and Berger, 2016 for Gaussian process models, among others. More recent studies on the theory are scarce, yet we can cite the work of Bodnar and Elster (2014), who suggest the derivation of reference priors using sequential maximization of Shannon’s entropy, and the work of Gao, Ramesh, and Chaudhari (2022) who consider reference priors to the training of deep neural networks. We mention that Muré (2018) provides a comprehensive review of the theory, and introduces original results about the uniqueness of the reference priors.

This chapter proposes a novel brief review of the reference prior theory. We fix the framework and define important main tools for the theory in section 3.2. The reference priors are formally defined in section 3.3, and we review their different properties that are proven in the literature. Eventually, this chapter is the occasion to discuss the limitations of the standard framework of the reference priors, and to suggest a solution based

on an original formalization of the Bayesian inference tools. The latter is done in section 3.4 and fixes the mathematical elements and notations that we will manipulate throughout the whole manuscript. A short conclusion terminates the chapter in section 3.5.

## 3.2 The standard Bayesian framework

### 3.2.1 A minimal framework

The Bayesian framework is commonly defined from the given of a statistical model. Let  $(\Omega, \mathcal{P}, \mathbb{P})$  be a probability space, and define  $Y$ , a random variable defined on  $\Omega$  and taking values on a measurable space  $(\mathcal{Y}, \mathcal{Y})$ . A statistical model is characterized by a collection of parameterized probability measures  $(\mathbb{P}_{Y|\theta})_{\theta \in \Theta}$  on  $(\mathcal{Y}, \mathcal{Y})$ . The Bayesian viewpoint considers a random variable  $T : \Omega \rightarrow \Theta$  taking values in a measurable space  $(\Theta, \mathcal{T})$ , following a **prior distribution** denoted  $\Pi$ , and which is such that the conditional distribution of  $Y$  to  $T = \theta$  is  $\mathbb{P}_{Y|\theta}$ . In other terms:

$$\forall A \in \mathcal{Y}, B \in \mathcal{T}, \mathbb{P}(Y \in A, T \in B) = \int_B \mathbb{P}_{Y|\theta}(A) d\Pi(\theta). \quad (3.1)$$

In practice,  $k$  realizations  $\mathbf{y} = (y_1, \dots, y_k)$  of the r.v.  $Y$  are observed, they are supposed to be identically distributed and independent conditionally to  $T$ , meaning they are the realization of a random vector  $\mathbf{Y}_k = (Y_1, \dots, Y_k)$ , whose distribution  $\mathbb{P}_{\mathbf{Y}_k}$  is defined on  $(\mathcal{Y}^k, \mathcal{Y}^{\otimes k})$  by:

$$\forall A \in \mathcal{Y}^{\otimes k}, \mathbb{P}_{\mathbf{Y}_k}(A) = \int_{\Theta} \mathbb{P}_{Y_k|\theta}(A) d\Pi(\theta), \quad (3.2)$$

where  $\mathbb{P}_{Y_k|\theta} := \mathbb{P}_{Y|\theta}^{\otimes k}$ . The distribution  $\mathbb{P}_{\mathbf{Y}_k}$  is called the **marginal distribution**.

Given the observations  $\mathbf{y}$ , the **posterior distribution**, denoted  $\mathbb{P}_{T|\mathbf{y}}$ , is then defined as the  $\mathbb{P}_{\mathbf{Y}_k}$ -a.s. unique one verifying

$$\forall A \in \mathcal{Y}^{\otimes k}, \forall B \in \mathcal{T}, \int_A \mathbb{P}_{T|\mathbf{y}}(B) d\mathbb{P}_{\mathbf{Y}_k}(\mathbf{y}) = \int_B \mathbb{P}_{Y_k|\theta}(A) d\Pi(\theta). \quad (3.3)$$

We can notice that the posterior distribution is always absolutely continuous w.r.t. the prior distribution. Its existence and  $\mathbb{P}_{\mathbf{Y}_k}$ -a.s. uniqueness are ensured by the disintegration theorem (see e.g. Chang and Pollard, 1997), while  $(\mathcal{Y}, \mathcal{Y})$  and  $(\Theta, \mathcal{T})$  are Borel spaces.

### 3.2.2 Likelihood and densities

It is common to assume that the statistical model admits a likelihood: there exists a collection of density functions  $(\ell(\cdot|\theta))_{\theta \in \Theta}$  w.r.t. a common  $\sigma$ -finite measure  $\mu$  on  $\mathcal{Y}$  such that for  $\Pi$ -a.e.  $\theta \in \Theta$ :

$$\forall A \in \mathcal{Y}, \mathbb{P}_{Y|\theta}(A) = \int_A \ell(y|\theta) d\mu(y). \quad (3.4)$$

**Example 3.1** (Multinomial model). The multinomial model is an example of statistical model where  $\theta \in (0, 1)^d$ ,  $\mathcal{Y} = \{y \in \{0, \dots, n\}^d, \sum_{i=1}^d y_i = n\}$  for a fixed  $n \in \mathbb{N}^*$ , and with  $\ell(y|\theta) = \frac{n!}{y_1! \dots y_d!} \prod_{i=1}^d \theta_i^{y_i}$ .

This way,  $\mathbb{P}_{\mathbf{Y}_k}$  admits a **marginal density**  $p_{\mathbf{Y}_k}$  w.r.t.  $\mu^{\otimes k}$  defined as

$$\forall \mathbf{y} \in \mathcal{Y}^k, p_{\mathbf{Y}_k}(\mathbf{y}) = \int_{\Theta} \prod_{i=1}^k \ell(y_i|\theta) d\Pi(\theta) = \int_{\Theta} \ell_k(\mathbf{y}|\theta) d\Pi(\theta), \quad (3.5)$$

where  $\ell_k(\mathbf{y}|\theta)$  denotes  $\prod_{i=1}^k \ell(y_i|\theta)$  and is called the likelihood.

Also, we can suppose that  $\Pi$  admits a **prior density**  $\pi$  w.r.t. a  $\sigma$ -finite measure  $\nu$  on  $\Theta$  (usually, it is simply assumed that  $\Theta \subset \mathbb{R}^d$  and  $\nu$  is the Lebesgue measure). The posterior distribution admits a **posterior**

density  $p(\cdot|\mathbf{y})$  w.r.t.  $\nu$  defined by

$$\forall \theta \in \Theta, p(\theta|\mathbf{y}) = \frac{\ell_k(\mathbf{y}|\theta)\pi(\theta)}{p_{\mathbf{Y}_k}(\mathbf{y})} \text{ if } p_{\mathbf{Y}_k}(\mathbf{y}) \neq 0, \quad p(\theta|\mathbf{y}) = 1 \text{ otherwise.} \quad (3.6)$$

To conclude this section, we acknowledge the usual confusion between the notation for the random variable  $T$  and notation for the values it takes  $\theta$ . Thus, we might refer as “the distribution of  $\theta$ ” or the “distribution conditionally to  $\theta$ ” to respectively mention the prior distribution or the conditional distributions to  $T = \theta$ .

### 3.2.3 Improper priors

It is common when dealing with non-informative priors in Bayesian analysis to define improper ones. Improper priors are defined as distributions admitting a density  $\pi$  that integrates to infinity:  $\int_{\Theta} \pi(\theta) d\nu(\theta) = \infty$ .

Of course, such a prior is not a probability distribution anymore, so that the framework elucidated in sections 3.2.1 and 3.2.2 does not stand. A way-around to be able to use improper priors is to focus on restrictions of the model, as formalized in the definition below.

**Definition 3.1.** As  $\nu$  is a  $\sigma$ -finite measure, there exists an increasing sequence  $(B_n)$  of elements in  $\Theta$  such that  $\Theta = \bigcup_{n \in \mathbb{N}} B_n$  and  $\nu(B_n) < \infty$ . An improper prior is said to be admissible if for any  $n$ ,  $\int_{B_n} \pi(\theta) d\nu(\theta) < \infty$ , and if there exists  $n_0$  such that  $\int_{B_{n_0}} \pi(\theta) d\nu(\theta) > 0$ .

In this case, on any  $\tilde{\Theta} \in \mathcal{T}$  such that there exists  $N \in \mathbb{N}$  verifying  $\tilde{\Theta} \subset \bigcup_{n \geq n_0}^N B_n$  and such that  $B_{n_0} \subset \tilde{\Theta}$ , one can define the probability space  $(\tilde{\Theta}, \tilde{\mathcal{T}}, \tilde{\Pi})$  with:

$$\tilde{\mathcal{T}} = \{B \cap \tilde{\Theta}, B \in \mathcal{T}\}, \quad \forall \tilde{B} \in \tilde{\mathcal{T}}, \tilde{\Pi}(\tilde{B}) = \frac{\int_{\tilde{B}} \pi(\theta) d\nu(\theta)}{\int_{\tilde{\Theta}} \pi(\theta) d\nu(\theta)}. \quad (3.7)$$

That modeling is the restricted modeling driven by  $\tilde{\Theta}$ . The probability  $\tilde{\Pi}$  (respectively its density) is referred as the renormalized restriction of the prior  $\Pi$  (resp. of the prior density  $\pi$ ) to  $\tilde{\Theta}$ .

On any restricted model, the posterior and the marginal distributions exist. Given two restrictions driven by  $\tilde{\Theta}_1$  and  $\tilde{\Theta}_2$ , the renormalized restricted prior densities  $\tilde{\pi}_1$  and  $\tilde{\pi}_2$  are equal on  $\tilde{\Theta}_1 \cap \tilde{\Theta}_2$  up to a constant:

$$\tilde{\pi}_1(\theta) = K\tilde{\pi}_2(\theta), \quad \text{so that} \quad \forall \mathbf{y} \in \mathcal{Y}^k, \tilde{p}_{1,\mathbf{Y}_k}(\mathbf{y}) = K\tilde{p}_{2,\mathbf{Y}_k}(\mathbf{y}), \quad (3.8)$$

where  $\tilde{p}_{1,\mathbf{Y}_k}$  (reps.  $\tilde{p}_{2,\mathbf{Y}_k}$ ) denotes the marginal density in the restricted model driven by  $\tilde{\Theta}_1$  (resp.  $\tilde{\Theta}_2$ ). Thus, calling  $\tilde{p}_1(\cdot|\mathbf{y})$  (resp.  $\tilde{p}_2(\cdot|\mathbf{y})$ ) the posterior density given the observations  $\mathbf{y}$  and under the restricted model driven by  $\tilde{\Theta}_1$  (resp.  $\tilde{\Theta}_2$ ), we have

$$\forall \theta \in \tilde{\Theta}_1 \cap \tilde{\Theta}_2, \tilde{p}_1(\theta|\mathbf{y}) = \tilde{p}_2(\theta|\mathbf{y}). \quad (3.9)$$

We conclude that the posterior density is always well-defined on  $\Theta$  when the improper prior is admissible. If the posterior density integrates to  $\infty$ , the posterior is said to be improper.

This construction is called a way-around because the framework remains incomplete:  $T$  is not embedded with a probability distribution anymore, and the marginal distribution  $\mathbb{P}_{\mathbf{Y}_k}$  is not appropriately defined. Later in this chapter, we will introduce a novel construction that authorizes the definition of improper priors in a more suitable way. Before that, the current construction is enough to define the reference priors among the admissible priors: priors that are either proper, either admissible improper priors.

## 3.3 Mutual information and reference priors

### 3.3.1 Mutual information

The principle of the theory is to minimize the role of the prior within the posterior, in order to maximize the information gain from the data. Thus, it amounts to maximize how “away” is the prior to the posterior.

When it resorts to compare information between probability distributions, we generally consider dissimilarity measures. The most common one is the Kullback-Leibler divergence, it is the one that is considered in the historical settings of the reference prior theory. We recall below the expression of the Kullback-Leibler divergence between two distributions  $P$  and  $Q$  on  $\mathcal{X}$ , which admit densities  $p, q$  w.r.t. a common measure  $\omega$ :

$$\text{KL}(Q||P) = \int_{\mathcal{X}} \log \left( \frac{q(x)}{p(x)} \right) q(x) d\omega(x). \quad (3.10)$$

We extend the definition of the Kullback-Leibler divergence when  $P$  and  $Q$  are not probability distributions, but are such that  $P(\mathcal{X}), Q(\mathcal{X}) < \infty$  by writing  $\text{KL}(Q||P) := \text{KL}(\tilde{Q}||\tilde{P})$  where  $\tilde{P}, \tilde{Q}$  respectively equal  $P/P(\mathcal{X})$  and  $Q/Q(\mathcal{X})$ .

Thus, the idea is to maximize the divergence  $\text{KL}(\mathbb{P}_{T|y}||\Pi)$ . As the observed sample  $y$  is unknown, the average value of the divergence is considered:

**Definition 3.2** (Mutual information). Given a Bayesian framework defined as in section 3.2 with a proper prior  $\Pi$ , and given a number of observations  $k$ , the mutual information is defined as the quantity:

$$\mathcal{I}^k(\Pi) := \mathbb{E}_{Y_k \sim \mathbb{P}_{Y_k}} [\text{KL}(\mathbb{P}_{T|Y_k}||\Pi)] = \int_{\mathcal{Y}^k} \text{KL}(\mathbb{P}_{T|y}||\Pi) p_{Y_k}(y) d\mu^{\otimes k}(y). \quad (3.11)$$

We recall that the definition of the mutual information is not limited to the scope of Bayesian analysis. In information theory it is common to refer to the mutual information  $\mathcal{I}(X, Z)$  between a random variable  $X$  and a random variable  $Z$  (see e.g. MacKay, 2003). The quantity in definition 3.2 aligns with their definition of the mutual information  $\mathcal{I}(T, Y_k)$  between the random parameter  $T$  and the random data  $Y_k$ .

One can apply Fubini-Lebesgue's theorem to write the mutual information as

$$\mathcal{I}^k(\Pi) = \mathbb{E}_{Y_k \sim \mathbb{P}_{Y_k}} [\text{KL}(\mathbb{P}_{T|Y_k}||\Pi)] = \mathbb{E}_{T \sim \Pi} [\text{KL}(\mathbb{P}_{Y_k|T}||\mathbb{P}_{Y_k})] \quad (3.12)$$

This second expression allows an interpretation of the mutual information from another viewpoint: it measures how the parameter  $T$  impacts the distribution of  $Y_k$ .

### 3.3.2 Reference priors

Reference priors are defined as the priors that are expected to maximize the role of the observed data within the posterior distribution. Consequently, they are historically sought as maximizers of the mutual information.

However, formally maximizing the mutual information  $\mathcal{I}^k(\Pi)$  w.r.t.  $\Pi$  presents three main issues:

1. First, its expression depends on the value of the number of observations  $k$ , and  $k$  alters the form of its maximal argument. This problem is tackled arguing that  $\mathcal{I}^k$  only measures a limited quantity of information brought by  $T$  onto the distribution of the data  $Y$ . The Bayesian framework does considerate the data  $(Y_1, \dots, Y_k)$  to be non-independent, so that the distribution of  $Y$  is enriched as  $k$  rises. According to Bernardo and Smith (1994), the limit  $\mathcal{I}^\infty$  (if it exists) measures the knowledge missing from the prior. Thus, the formal definition tends to maximize  $\mathcal{I}^k$  as  $k$  diverges to  $\infty$ .
2. The second issue comes with the fact that the quantity  $\mathcal{I}^k(\Pi)$  generally does not admit a finite limit as  $k$  diverges to  $\infty$ .
3. The third issue is that the mutual information is only correctly defined when the prior is proper, while many non-informative priors used in Bayesian analysis are improper.

The commonly admitted definition for the reference priors, which takes into account the three issues aforementioned is the following, adapted from Bernardo, 2005.

**Definition 3.3** (Reference priors). Let  $\mathcal{P}$  be a set of priors on  $\Theta$ , a prior  $\Pi \in \mathcal{P}$  is a reference prior over  $\mathcal{P}$  if there exists an appropriate increasing sequence  $(\Theta_i)_{i \in \mathbb{N}}$  such that  $\bigcup_{i \in \mathbb{N}} \Theta_i = \Theta$  with for any  $i$ :  $0 < \Pi(\Theta_i) < \infty$  and

$$\lim_{k \rightarrow \infty} \mathcal{I}_i^k(\Pi_i) - \mathcal{I}_i^k(P_i) \geq 0 \text{ for all } P \in \mathcal{P}_i \quad (3.13)$$

with  $\Pi_i$  (resp.  $P_i$ ) being the renormalized restriction of  $\Pi$  (resp.  $P$ ) on  $\Theta_i$ ,  $\mathcal{I}_i^k$  denoting the mutual information on the restricted modeling driven by  $\Theta_i$ , and  $\mathcal{P}_i = \{P \in \mathcal{P}, 0 < P(\Theta_i) < \infty\}$ .

This first definition considers a set  $\mathcal{P}$  called a set of priors. We provide a thorough definition of such a class in section 3.4. Until then, it refers to a set of priors that are admissible as described in section 3.2.3. This definition also requires the sequence  $(\Theta_i)$  to be “appropriate”. Actually, the original definition does not restrict the choices of that sequence, so that this qualifying adjective can be ignored in this section. However, we will see in the sequel that some studies on reference priors require to impose some assumptions on  $(\Theta_i)_i$ .

**Example 3.2.** If  $\Theta$  is finite with  $\mathcal{T} = \mathcal{P}(\Theta)$ , and if  $\mathbb{P}_{Y|\theta} \neq \mathbb{P}_{Y|\theta'}$  for any  $\theta' \neq \theta$  then the reference prior  $\Pi$  over the set of all priors is the uniform prior:  $\Pi(B) = \Pi(B')$  for all  $B, B' \subset \Theta$ . A proof of this result is proposed in Muré, 2018.

We must note that other definitions of reference priors exist. In the genesis of the reference prior theory, Bernardo (1979b) suggested maximizing the mutual information  $\mathcal{I}^k$ , and derive the limit of the maximizers when  $k \rightarrow \infty$  to define reference priors. In Berger, Bernardo, and Sun, 2009, the authors prove that in the real and one parameter case (i.e.  $\Theta \subset \mathbb{R}$ ), this strategy is consistent with definition 3.3. Their result is the following, sometimes taken as a definition.

**Theorem 3.1** (Explicit form of the reference prior). Assume  $\Theta \subset \mathbb{R}$  with  $v$  being the Lebesgue measure.

Call  $\mathcal{P}_s$  the set of admissible priors that are positive and continuous, and that issue proper posterior distributions.

Let  $\Pi^* \in \mathcal{P}_s$  we call  $p^*(\cdot|\cdot)$  its posterior and define for any interior point  $\theta_0 \in \Theta$ ,

$$f_k(\theta) = \exp \left( \int_{\mathcal{Y}_k} \ell_k(\mathbf{y}|\theta) \log(p^*(\theta|\mathbf{y})) d\mu^{\otimes k}(\mathbf{y}) \right) \quad \text{and} \quad f(\theta) = \lim_{k \rightarrow \infty} \frac{f_k(\theta)}{f_k(\theta_0)}. \quad (3.14)$$

If: (i)  $\forall \theta, \varepsilon > 0$ ,  $\int_{|\tilde{\theta} - \theta| < \varepsilon} p^*(\tilde{\theta}|\mathbf{y}) d\tilde{\theta} \xrightarrow[k \rightarrow \infty]{\mathbb{P}_{Y|\theta}} 1$ ,

(ii) each  $f_k$  is continuous with  $\forall k, \theta, \frac{f_k(\theta)}{f_k(\theta_0)} < h(\theta)$  such that  $h$  is integrable on any compact set, and

(iii)  $f$  is the density of  $F \in \mathcal{P}_s$  and is such that for any increasing sequence of compact sets  $(\Theta_i)_{i \in \mathbb{N}}$  that covers  $\Theta$ , calling  $g_i$  the restricted posterior density of  $F$  on  $\Theta_i$ , we have  $\forall \mathbf{y}, \int_{\Theta_i} g_i(\theta|\mathbf{y}) \log \frac{g_i(\theta|\mathbf{y})}{g(\theta|\mathbf{y})} d\theta \xrightarrow[i \rightarrow \infty]{} 0$  where  $g$  is  $F$ 's posterior,

then  $F$  is a reference prior over  $\mathcal{P}_s$ .

The form of the functions  $f_k$  in the above theorem comes from maximizing the mutual information (expressed in definition 3.2) when  $k$  is fixed. Indeed, it is possible to research for a maximal argument of the mutual information from a constrained optimization viewpoint. A Lagrange multipliers theorem should give an implicit expression of the density whose prior maximizes the mutual information for a fixed  $k$ : it is proportional to  $\exp \left( \int_{\mathcal{Y}_k} \ell_k(\mathbf{y}|\theta) \log(p(\theta|\mathbf{y})) d\mu^{\otimes k}(\mathbf{y}) \right)$ . This result is rigorously proven in appendix C for a particular case. It was intuited by Bernardo and Smith (1994), with the idea that, asymptotically, the posterior becomes independent of the chosen prior. The theorem formalizes the intuition, as the posterior in  $f_k$ 's expression comes from any initial prior  $\Pi^* \in \mathcal{P}_s$ .

### Hierarchical construction in high dimensional settings

Definition 3.3 and theorem 3.1 both define reference priors as maximizers of the mutual information. While there is consensus that such an approach leads to appropriate objective priors in low-dimensional cases, it is not always the case in high-dimensional settings. Indeed, the prior maximizing the mutual information has often non-desirable properties in such scenarios, according to Berger, Bernardo, and Sun (2015). In

Berger and Bernardo, 1992a, the authors detail their suggestion to define the reference priors hierarchically, by considering an ordering of the parameters:

$$\theta = (\theta_1, \dots, \theta_r) \in \Theta = \Theta_1 \times \dots \times \Theta_r. \quad (3.15)$$

Typically, it is recommended to assume  $\Theta_j \subset \mathbb{R}^{d_j}$  with small dimensions  $d_j$  (e.g., lower or equal than 2) for any  $j \in \{1, \dots, r\}$ , and to sequentially build a reference prior on the  $\Theta_j$ ,  $j \in \{1, \dots, r\}$ :

1. initially fix  $\ell_k^1 = \ell_k$ ;
2. for any values of  $\theta_{j+1}, \dots, \theta_r \in \Theta_{j+1} \times \dots \times \Theta_r$ , compute a reference prior (in the sense of definition 3.3) under the model with likelihood  $\theta_j \mapsto \ell_k^j(\mathbf{y}|\theta_j, \dots, \theta_r)$ , denote  $\pi_j(\cdot|\theta_{j+1}, \dots, \theta_r)$  its normalized density;
3. derive  $\ell_k^{j+1}$  such as

$$\ell_k^{j+1}(\mathbf{y}|\theta_{j+1}, \dots, \theta_r) = \int_{\Theta_j} \ell_k^j(\mathbf{y}|\theta_j, \dots, \theta_r) d\pi_j(\theta_j|\theta_{j+1}, \dots, \theta_r). \quad (3.16)$$

This construction is often considered in the literature when it resorts to define reference priors in multidimensional context. However, it relies on the hierarchization of the parameter's coordinates fixed in eq. (3.15), and the recommended one depends on the model of interest. There are a few models, such as the multinomial (defined in example 3.1) one, for which the different reference priors resulting from this construction have been extensively studied. In the case of the multinomial model, it is claimed that this construction with an ordering where  $r = d$  when  $\Theta \subset \mathbb{R}^d$  must be favored (Berger and Bernardo, 1992b; Berger, Bernardo, and Sun, 2015). In this case, the authors show that the density  $\pi^*$  of this reference prior given by this construction is given by

$$\pi^*(\theta) = K \left[ \prod_{j=1}^r \prod_{i=1}^{d_j} \theta_j^{(i)-1/2} \right] \left[ \prod_{j=1}^{r-1} \left( 1 - \sum_{j'=1}^j \sum_{i=1}^{d_j} \theta_{j'}^{(i)} \right)^{-d_{j+1}/2} \right] \left[ 1 - \sum_{j=1}^r \sum_{i=1}^{d_j} \theta_j^{(i)} \right]^{-1/2}, \quad (3.17)$$

where  $K$  is a normalization constant.

### 3.3.3 Properties and Jeffreys prior

The elucidation of the reference priors remains an open question for some restricted sets of priors  $\mathcal{P}$  or when we lack of regularity assumptions on the model. In most studies,  $\mathcal{P}$  remains as large as possible and is restricted only by some mild regularity assumptions, such as that the priors admitting continuous densities.

In the general case, Muré (2018) studied the uniqueness of reference priors. His study requires to define the sequence  $(\Theta_i)_i$  in definition 3.3 “appropriate” if and only if the  $\Theta_i$  are open subsets of  $\Theta$ . We summarize his main result as follows:

**Theorem 3.2.** Let  $\mathcal{P}$  be the set of admissible priors. Call an “appropriate” increasing sequence of subsets  $(\Theta_i)_i$ , an increasing sequence such that each for any  $i$ ,  $\Theta_i$  is open. Under the assumption that there exists a weakly estimator of  $\theta$ , the reference prior over  $\mathcal{P}$ , if it exists, is unique up to a multiplicative constant.

This result allows referring to “the reference prior” instead to “reference priors” in some cases. Also, there does not exist to our knowledge a result about the existence of the reference prior in the general case.

Actually, the reference priors are mostly studied in what are called regular cases. The following assumption is a regularity assumption on the statistical model, it ensures useful tools to be well-defined.

**Assumption 3.1.**  $\Theta$  is a subset of  $\mathbb{R}^d$  with non-empty interior, and the likelihood is such that

1. for  $\mu$ -a.e.  $y$ ,  $\theta \mapsto \ell(y|\theta)$  is twice differentiable, with  $\nabla_\theta^2 \ell(y|\theta)$  being negative definite,

2. for any  $\theta \in \Theta$ , there exists  $\tau > 0$  such that

$$\mathbb{E}_{Y \sim \mathbb{P}_{Y|\theta}} \left[ \sup_{\theta', \|\theta - \theta'\| < \tau} \|\nabla_\theta^2 \log \ell(Y|\theta')\|^2 \right] \quad \text{and} \quad \mathbb{E}_{Y \sim \mathbb{P}_{Y|\theta}} \left[ \sup_{\theta', \|\theta - \theta'\| < \tau} \|\nabla_\theta \log \ell(Y|\theta')\|^2 \right], \quad (3.18)$$

are continuous functions of  $\theta$  on the interior of  $\Theta$ .

Under assumption 3.1, the Fisher information matrix and the Jeffreys prior are well-defined (see Lehmann, 1999). The Fisher information matrix  $\mathcal{I}$  is expressed as

$$\mathcal{I}(\theta) = (\mathcal{I}(\theta)_{i,j})_{i,j=1}^d \quad \text{with} \quad \mathcal{I}(\theta)_{i,j} = \int_{\mathcal{Y}} [\partial_{\theta_i} \log \ell(y|\theta)] [\partial_{\theta_j} \log \ell(y|\theta)] \ell(y|\theta) d\mu(y), \quad (3.19)$$

under assumption 3.1 its coefficients can be expressed as follows

$$\mathcal{I}(\theta)_{i,j} = - \int_{\mathcal{Y}} [\partial_{\theta_i \theta_j}^2 \log \ell(y|\theta)] \ell(y|\theta) d\mu(y). \quad (3.20)$$

The Jeffreys prior refers to the prior whose density is  $J$ , defined by

$$J(\theta) = \sqrt{\det \mathcal{I}(\theta)}. \quad (3.21)$$

Under assumption 3.1,  $J$  is a continuous and positive function on  $\Theta$ . Here are a few examples of Jeffreys priors in different modelings.

**Example 3.3** (A proper Jeffreys prior in the multinomial case). The multinomial model is defined as  $\theta \in (0, 1)^d$ ,  $\mathcal{Y} = \{y \in \{0, \dots, n\}^d, \sum_{i=1}^d y^i = n\}$  for a fixed  $n \in \mathbb{N}^*$ . The likelihood is

$$\ell(y|\theta) = \frac{n!}{y^1! \dots y^d!} \prod_{i=1}^d \theta_i^{y^i}. \quad (3.22)$$

In this model, the Jeffreys prior corresponds to a Dirichlet( $1/2, \dots, 1/2$ ) distribution:

$$J(\theta) \propto \prod_{i=1}^d \theta_i^{-1/2} (1 - \theta_i)^{-1/2}. \quad (3.23)$$

It is a proper prior and corresponds to the prior given in eq. (3.17) when  $r = d$ .

**Example 3.4** (Jeffreys prior in the Gaussian case). The Gaussian model is defined by  $\mathcal{Y} = \mathbb{R}$ ,  $\ell(y|\theta) = \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}$ , with  $\theta = \mu \in \mathbb{R}$  in the known variance case, or  $\theta = (\mu, \sigma^2) \in \mathbb{R} \times (0, \infty)$  in the unknown variance case. In both cases, the Jeffreys prior has a density:

$$J(\theta) \propto \frac{1}{\sigma}. \quad (3.24)$$

In the known variance case the prior for  $\mu$  corresponds to a uniform distribution on  $\mathbb{R}$ . In both cases, it is improper. However, its posteriors are always proper.

Clarke and Barron (1994) have shown that the Jeffreys prior is the reference prior over the priors with continuous and positive densities on  $\Theta$ . Their result can be expressed as follows.

**Theorem 3.3.** Under assumption 3.1. Call an “appropriate” increasing sequence of set  $(\Theta_i)_i$ , an increasing sequence such that for any  $i$   $\Theta_i$  is compact. Assume  $\nu$  is the Lebesgue measure and let  $\mathcal{P}$  be the set of priors admitting a continuous and positive density. Then the Jeffreys prior is the unique reference prior over  $\mathcal{P}$ .

Muré (2018) has shown that the above theorem still stands when “appropriate” refers to open subsets of  $\Theta$ . This result can be anticipated by reminding that the form of the density maximizing the mutual information for a fixed  $k$  is  $\exp \left( \int_{\mathcal{Y}_k} \ell_k(\mathbf{y}|\theta) \log(p(\theta|\mathbf{y})) d\mu^{\otimes k}(\mathbf{y}) \right)$ . Indeed, according to the Bernstein von Mises theorem

(see e.g. Vaart, 1992), under appropriate assumptions, the posterior density  $p(\theta|y)$  is asymptotically close to a centered Gaussian density whose variance is  $\mathcal{I}(\theta)^{-1/2}/\sqrt{k}$  when  $y$  are distributed w.r.t.  $\mathbb{P}_{Y_k|\theta}$ . This convergence does not depend on the prior. We thus asymptotically have

$$\log p(\theta|y) \simeq \log \frac{k}{2\pi} + \frac{1}{2} \log \det \mathcal{I}(\theta), \quad (3.25)$$

and we approximate

$$\int_{\mathcal{Y}^k} \ell_k(y|\theta) \log p(\theta|y) d\mu^{\otimes k}(y) \simeq \log \frac{k}{2\pi} + \frac{1}{2} \log \det \mathcal{I}(\theta). \quad (3.26)$$

Calling  $f_k(\theta) = \exp \left( \int_{\mathcal{Y}^k} \ell_k(y|\theta) \log(p(\theta|y)) d\mu^{\otimes k}(y) \right)$ , the above heuristic gives that  $f_k$  is asymptotically proportional to  $\exp \left( \frac{1}{2} \log \det \mathcal{I}(\theta) \right)$ , which is the density of the Jeffreys prior.

We conclude this section by noticing that the Jeffreys prior arises as a cornerstone in the reference prior theory. We recall that this prior has been introduced by Jeffreys (1946). It is a prior that is already praised in Bayesian statistics for its main property to be invariant by re-parametrization of the model: define  $\phi = g(\theta)$  with  $g$  being a diffeomorphism, and define the model induced considering  $\phi$  instead of  $\theta$ . Then the Jeffreys prior for this re-parameterized model admits the density  $J_g$ :

$$J_g(\phi) = J(g^{-1}(\phi)) |\det d_\phi g^{-1}(\phi)|. \quad (3.27)$$

This ensures that inference is not affected by the choice of parameterization, preserving coherence. We also recall that the Jeffreys prior is known to be improper in most cases.

## 3.4 Limits of the classical framework of reference priors and solution

During this brief review, we mentioned some limits in the framework of the reference priors definition. They are related to the incorporation of improper priors in the admissible priors. As a matter of fact, when dealing with non-informative distributions, it is common to find one of the main assumptions on probability distributions (that is, their total mass is 1) limiting. The fact is that in some cases, improper distributions seem really “close” to proper ones. As an example, we could mention that the improper uniform distribution on  $\mathbb{R}$  can be seen as a Gaussian distribution with a variance that diverges to infinity. A logarithmic re-parametrization of the latter gives the improper distribution of density proportional to  $x \mapsto 1/x$ .

Additionally, when dealing with densities, nothing strongly forbid the definition of the posterior density when the prior is improper (see section 3.2.3). It became therefore common to normalize the use of improper priors, and to carry out *a posteriori* inference from them when the issued posterior is proper.

Even though the way-around works, we lose the existence of a latent probability space with random variables expressing formally the existence of a path that links the “chaos” of a universe to the phenomena observed and modeled. This section is dedicated to the elucidation of such construction, in a framework that allows priors to be improper.

### 3.4.1 Novel framework for Bayesian inference with improper priors

Since the early work of Kolmogorov (1933), it is appreciated to be reminded that there exists an abstract, underlying structure from which result the probabilistic description and tools that are used to perform inference. This structure, which takes the form of a probability space in the axiomatic of Kolmogorov, represents the indescribable phenomena that induce randomness in the observations.

However, the usual Kolmogorov’s axiomatic does not take into account improper distributions to construct the underlying space that provides the usual statistical modeling. In this section, we construct from scratch a framework for Bayesian statistics, in a way that allows priors to be improper. We re-define the objects defined in section 3.2.

Our suggestion is based on the work of Taraldsen and Lindqvist (2016), who adapted the propositions of Renyi (1970). Their idea is to extend the notion of a probability space when the probability is only a  $\sigma$ -finite

measure. Up to a multiplicative constant, their framework is shown to coincide with standard probability spaces when the measure has a finite total mass, or when conditioning on a set with finite mass.

We let  $(\Theta, \mathcal{T})$  be a measurable space. If  $\Pi_1$  and  $\Pi_2$  are two non-null, non-negative and  $\sigma$ -finite measures we can define the relation  $\simeq$  by

$$\Pi_1 \simeq \Pi_2 \iff \exists t > 0, \Pi_1 = t\Pi_2. \quad (3.28)$$

Define by  $\mathcal{M}$  the space of non-null, non-negative and  $\sigma$ -finite measures on  $\Theta$ . The following properties are verified:

- The relation  $\simeq$  is an equivalence relation on  $\mathcal{M}$ . We note  $[\Pi]$  the class of a  $\Pi \in \mathcal{M}$ .
- Denote by  $\mathcal{M}^\nu$  the space of absolutely continuous measures w.r.t.  $\nu \in \mathcal{M}$ . It is stable by  $\simeq$ .
- Define by  $\mathcal{R}$  the space of non-negative non-null measurable functions from  $\Theta$  to  $\mathbb{R}$ . The relation

$$f \propto g \iff \exists t > 0, f = tg \quad (3.29)$$

is an equivalence relation on  $\mathcal{R}$ . Let  $\nu \in \mathcal{M}$ , the Radon-Nikodym Theorem defines a natural surjection from the quotient space  $\mathcal{R}/\propto$  to  $\mathcal{M}^\nu/\simeq$ . A class in  $\mathcal{R}/\propto$  is called a density class of the unique class in  $\mathcal{M}^\nu/\simeq$  that it induces by this mapping. If two elements  $[f], [g]$  of  $\mathcal{R}/\propto$  are density classes of a same element in  $\mathcal{M}^\nu/\simeq$ , they are equal  $\nu$ -a.e. in the sense  $\exists t > 0, f = tg$   $\nu$ -a.e.

Bioche and Druilhet (2016) consider such a definition of  $\sigma$ -finite measures equal up-to-a constant to formalize the notion of prior and posterior approximations, whether they are proper or improper. Notably, they define a topology induced by a convergence in  $\mathcal{M}/\simeq$ : the Q-vague convergence.

A prior is then generally defined as a class  $[\Pi]$  in  $\mathcal{M}/\simeq$ . The tuple  $(\Theta, \mathcal{T}, [\Pi])$  corresponds to what Taraldsen and Lindqvist (2016) call a conditional probability space for any  $[\Pi] \in \mathcal{M}/\simeq$ . For any  $U \in \mathcal{T}$  such that  $\Pi(U) \in (0, +\infty)$ , that space issues a unique probability space on  $U$ , with the probability  $\Pi(\cdot|U) := \Pi(\cdot \cap U)/\Pi(U)$ , it is independent of the choice of the representative of  $[\Pi]$ . This framework gives what is necessary to construct a general conditional probability space that models our problem, as we express below.

**Proposition 3.1.** *We assume that the observations take values in a Polish space  $(\mathcal{Y}, \mathcal{Y})$  and are statistically modeled by the collection of conditional probabilities  $(\mathbb{P}_{Y|\theta})_{\theta \in \Theta}$  with  $(\Theta, \mathcal{T})$  being a Polish space and such that  $\forall A \in \mathcal{Y}$ ,  $\theta \mapsto \mathbb{P}_{Y|\theta}(A)$  is measurable. We consider  $\Pi \in \mathcal{M}$ .*

*Then there exist a conditional probability space  $(\Omega, \mathcal{E}, [\Pi])$ , a measurable process  $\bar{Y} = (Y_i)_{i \in \mathbb{N}}$ ,  $Y_i : \Omega \rightarrow \mathcal{Y}$ , and a measurable function  $T : \Omega \rightarrow \Theta$  such that*

1. for any  $U \in \mathcal{T}$ ,  $\Pi(T \in U) = \Pi(U)$ ;
2. if  $\Pi(U) \in (0, +\infty)$  then the unique probability space conditioned on  $\{T \in U\}$  is such that the conditional probability of any  $(Y_{j_l})_{l=1}^k$  to  $T = \theta$  is  $\mathbb{P}_{Y|\theta}^{\otimes k}$ . More explicitly, if  $A_{j_1}, \dots, A_{j_k} \in \mathcal{Y}$  then

$$\Pi\left(\bigcap_{l=1}^k Y_{j_l} \in A_{j_l} | T \in U\right) = \int_{\Theta} \mathbb{P}_{Y|\theta}^{\otimes k}(A_{j_1} \times \dots \times A_{j_k}) d\Pi(\theta|U). \quad (3.30)$$

Proposition 3.1 is proven in section 3.4.3. It ensures the existence of an underlying abstract space  $\Omega$  from only the given of the statistical model (the collection of probability measures  $(\mathbb{P}_{Y|\theta})_{\theta \in \Theta}$ ) and the prior (which is a measure  $\Pi \in \mathcal{M}$ ). When  $\Pi$  is improper, the tuple  $(\Omega, \mathcal{E}, [\Pi])$  is not a probability space anymore, but they still describe the randomness that goes beyond the observations and the parameter. In the following section, we develop the compatibility of this framework with the usual one.

### 3.4.2 Notations and compatibility with the usual framework

Given the construction proposed in section 3.4.1, the Bayesian framework is defined on the basis of probability measures  $(\mathbb{P}_{Y|\theta})_{\theta \in \Theta}$  and a prior  $\Pi$  which is a representative of an element  $[\Pi] \in \mathcal{M}/\simeq$ , under the

assumptions in proposition 3.1. The consideration of another representative  $\tilde{\Pi}$  of  $[\Pi]$  does not alter the construction.

That framework allows the construction of a marginal distribution  $\mathbb{P}_{Y_k}$ , defined as the pushforward measure of  $\Pi$  by  $Y_k$ . In the same way as  $\Pi$ , it is unique up to a constant. The posterior distribution  $\mathbb{P}_{T|y}$  is well-defined for  $\mathbb{P}_{Y_k}$ -a.e.  $y$  by

$$\forall A \in \mathcal{Y}^{\otimes k}, \forall B \in \mathcal{T}, \int_A \mathbb{P}_{T|y}(B) d\mathbb{P}_{Y_k}(y) = \int_B \mathbb{P}_{Y_k|\theta}(A) d\Pi(\theta). \quad (3.31)$$

It belongs to  $\mathcal{M}$ , and its class  $[\mathbb{P}_{T|y}]$  in  $\mathcal{M}/\simeq$  is uniquely defined.

## Densities

From now on, we assume that the model admits a likelihood: there exists a collection of density functions  $(\ell(\cdot|\theta))_{\theta \in \Theta}$  w.r.t. a common  $\sigma$ -finite measure  $\mu$  on  $\mathcal{Y}$  such that for  $\Pi$ -a.e.  $\theta$ :

$$\forall A \in \mathcal{Y}, \mathbb{P}_{Y|\theta}(A) = \int_A \ell(y|\theta) d\mu(y). \quad (3.32)$$

We also assume that  $\Pi$  belongs to  $\mathcal{M}^\nu$  (it is absolutely continuous w.r.t.  $\nu$ ), with  $\nu \in \mathcal{M}$ . We denote by  $\pi \in \mathcal{R}$  a density of  $\Pi$  w.r.t.  $\nu$ . For  $y \in \mathcal{Y}^k$  there exists a posterior density w.r.t.  $\nu$  defined in the class  $[p(\cdot|y)] \in \mathcal{R}/\propto$  where

$$p(\theta|y) \propto \ell_k(y|\theta)\pi(\theta) = \prod_{i=1}^k \ell(y_i|\theta)\pi(\theta) \quad (3.33)$$

If the posterior is proper for  $\mathbb{P}_{Y_k}$ -a.e.  $y$ , a marginal density can be defined  $\mathbb{P}_{Y_k}$ -a.e. by

$$p_{Y_k}(y) = \int_{\Theta} \ell_k(y|\theta)\pi(\theta) d\theta. \quad (3.34)$$

## Impact on the framework of reference priors

The elucidating of the densities in this novel Bayesian construction is consistent with the usual construction. When restricting the framework to a  $\tilde{\Theta} \in \mathcal{T}$  such that  $0 < \Pi(\tilde{\Theta}) < \infty$ , the restricted model driven by  $\tilde{\Theta}$  (in the sense given in section 3.2.3) corresponds to the model on  $\tilde{\Theta}$  considering the prior  $\Pi(\cdot|\tilde{\Theta})$ .

Therefore, the notations and definitions written in the previous sections still stand. The modeling remains the same for any prior that equals  $\Pi$  up to a positive multiplicative constant, so that if  $\Pi \in \mathcal{M}$  is a reference prior, any prior in  $[\Pi]$  also is. For this reason it becomes appropriate to talk about a reference prior class. We will authorize however the slight abuse of language referring to “the reference prior” instead of “reference prior in the reference prior class”, when the latter is unique. We formalize this terminology in the following definitions.

**Definition 3.4.** A set of priors refers to a subset of  $\mathcal{M}^\nu$ .

When  $\nu$  is the Lebesgue measure, a set of continuous prior refers to a subset  $\mathcal{P} \subset \mathcal{M}^\nu$  where for any  $\Pi \in \mathcal{P}$ ,  $\Pi$  admits a  $\nu$ -a.e. continuous and everywhere locally bounded density  $\pi$ . We call  $\mathcal{M}_C^\nu$  the set of continuous priors.

The subset of  $\mathcal{R}$  composed by  $\nu$ -a.e. continuous and everywhere locally bounded functions is denoted  $\mathcal{R}_{Cb}$ .

**Proposition 3.2.**  $\mathcal{M}_C^\nu$  is stable by  $\simeq$  and  $\mathcal{R}_{Cb}$  is stable by  $\propto$ .

If  $\Theta \subset \mathbb{R}^d$ , and if  $f$  is  $\nu$ -a.e. continuous,  $f$  is  $\nu$ -a.e. locally bounded.

Any  $f$  in  $\mathcal{R}_{Cb}$  is bounded on every compact.

*Proof.* 1. The stability of  $\mathcal{R}_{\mathcal{C}^b}$  to  $\infty$  comes since multiplying a function by a constant does not affect its continuity nor its locally bounded characteristic.

We consider  $M$  the mapping assigning a density in  $\mathcal{R}$  to its associated measure in  $\mathcal{M}^\nu$  as

$$M : f \in \mathcal{R} \mapsto \left( B \in \mathcal{T} \mapsto \int_B f(\theta) d\nu(\theta) \right). \quad (3.35)$$

We have  $f \propto g \implies M(f) \simeq M(g)$ , so that  $M(\mathcal{R}_{\mathcal{C}^b})$  is stable by  $\simeq$ . The second statement of the proposition comes since  $M(\mathcal{R}_{\mathcal{C}^b}) = \mathcal{M}_{\mathcal{C}}^\nu$  by definition.

2. For the second statement, we consider  $f$   $\nu$ -a.e. continuous. We note  $E = \{\theta, \forall \tau > 0, f \text{ is not bounded on } B(\theta, \tau)\}$ . Let  $\theta \in E$ , for any  $n > 0$ ,  $f$  is not bounded on  $B(\theta, 1/n)$ , so that there exists  $\theta_n \in B(\theta, 1/n)$  such that  $f(\theta_n) > n$ . That leading to  $f(\theta_n) \xrightarrow{n \rightarrow \infty} \infty$  while  $\theta_n \xrightarrow{n \rightarrow \infty} \theta$ ,  $f$  is discontinuous on  $\theta$ . We deduce that  $\nu(E) = 0$ .

3. If  $\tilde{\Theta}$  is a compact subset of  $\Theta$ , we can cover  $\tilde{\Theta}$  by open balls on which  $f$  is bounded (consequently to statement 2.). Since  $\tilde{\Theta}$  is compact, the open cover admits a finite subcover.  $f$  being bounded on this finite subcover, we deduce that it is bounded on  $\tilde{\Theta}$ .  $\square$

**Definition 3.5** (Advanced definition of reference priors). Let  $\mathcal{P}$  be a set of priors. A prior  $\Pi \in \mathcal{P}$  is a reference prior over  $\mathcal{P}$  if there exists an openly increasing sequence of compact sets  $(\Theta_i)_{i \in \mathbb{N}}$  such that  $\bigcup_{i \in \mathbb{N}} \Theta_i = \Theta$  with for any  $i$ :  $0 < \Pi(\Theta_i) < \infty$  and

$$\lim_{k \rightarrow \infty} \mathcal{I}^k(\Pi(\cdot | \Theta_i)) - \mathcal{I}^k(P(\cdot | \Theta_i)) \geq 0 \text{ for all } P \in \mathcal{P} \text{ verifying } 0 < P(\Theta_i) < \infty. \quad (3.36)$$

Let  $\mathcal{Q}$  be a subset of  $\mathcal{M}^\nu / \simeq$ . And denote by  $q$  the surjection  $q : \Pi \in \mathcal{M}^\nu \mapsto [\Pi]$ . A class  $[\Pi] \in \mathcal{Q}$  is called a reference prior class over  $\mathcal{Q}$  if  $\Pi$  is a reference prior over  $q^{-1}(\mathcal{Q})$ .

The above definition is congruous with definition 3.3, apart from the necessary condition on the sequence  $(\Theta_i)_i$ . The definition of an openly increasing sequence of compacts is stated below. It is, according to ourselves, an appropriate condition on this sequence, as it allows the consistency between the compact case and the general case.

**Definition 3.6.** A sequence  $(\Theta_i)_{i \in \mathbb{N}}$  of subsets of  $\Theta$  is said to be openly increasing if there exist  $i_0 \geq 0$  and a sequence  $(V_i)_{i \geq i_0}$  of open subsets of  $\Theta$  such that for any  $i \geq i_0$ :

$$\Theta_i \subset V_i \subset \Theta_{i+1}. \quad (3.37)$$

Such sequence does not necessarily exist if  $\Theta$  has an “unusual” form. In most cases, one would work on sets  $\Theta$  that are “almost” open or “almost” closed, i.e.  $\nu(\partial\Theta) = 0$ . In this case, one can restrict the study to  $\overline{\Theta}$  or to  $\mathring{\Theta}$ . If  $\Theta$  is moreover a subset of  $\mathbb{R}^d$  then there always exist openly increasing sequence of compact subsets  $(\Theta_i)_{i \in \mathbb{N}}, (\Theta'_i)_{i \in \mathbb{N}}$  with  $\bigcup_{i \in \mathbb{N}} \Theta_i = \mathring{\Theta}$  and  $\bigcup_{i \in \mathbb{N}} \Theta'_i = \overline{\Theta}$ . We express that statement in the proposition below.

**Proposition 3.3.** 1. If  $\Theta$  is an open or a closed subset of  $\mathbb{R}^d$  there exist an openly increasing sequence of compact subsets  $(\Theta_i)_{i \in \mathbb{N}}$  such that  $\bigcup_{i \in \mathbb{N}} \Theta_i = \Theta$ .

2. If  $\Theta$  is a countable intersection of open subsets of  $\mathbb{R}^d$ , there exist an openly increasing sequence of compact subsets  $(\Theta_i)_{i \in \mathbb{N}}$  such that  $\bigcup_{i \in \mathbb{N}} \Theta_i = \Theta$ .

*Proof.* If  $\Theta$  is open, then we can write  $\Theta$  as an increasing countable union of open balls:

$$\Theta = \bigcup_{n \in \mathbb{N}^*} \bigcup_{\theta \in \Theta \cap \mathbb{Q}^d} B(\theta, \min(R(\theta) - \frac{1}{n}, n)) \quad \text{with} \quad R(\theta) = \sup\{\tau > 0, B(\theta, \tau) \subset \Theta\}. \quad (3.38)$$

We write  $\Theta \cap \mathbb{Q}^d = \{\theta_i, i \in \mathbb{N}^*\}$ , and write  $\Theta = \bigcup_{n \in \mathbb{N}^*} V_n$  with  $V_n = \bigcup_{i=1}^n B(\theta_i, \hat{R}_n^i)$ , where  $\hat{R}_n^i =$

$\min(R(\theta_i) - \frac{1}{n}, n)$ . This way, each  $V_n$  is a finite union of closed balls, so that  $\overline{V_n}$  is compact. Also,  $V_{n+1}$  contains balls with the same center as the one that constitute  $V_n$ , yet with higher radius. Thus,  $\overline{V_n} \subset V_{n+1}$ . Therefore, setting  $\Theta_i = \overline{V_i}$  for all  $i$  makes  $(\Theta_i)_{i \in \mathbb{N}^*}$  being an openly increasing sequence of compact subsets that cover  $\Theta$ .

If  $\Theta$  is closed, we recall that  $\mathcal{F} = \{\Theta \cap B(\theta, \tau), \theta \in \mathbb{Q}^d, \tau \in \mathbb{Q}_+^*\}$  is a subbase of the induced topology of  $\mathbb{R}^d$  on  $\Theta$ . We can thus cover  $\Theta$  by elements in  $\mathcal{F}$ :  $\Theta = \bigcup_{i \in I} U_i$ . It is possible to index this cover by  $\mathbb{N}$ . We now construct the sequences  $(\Theta_i)_{i \in \mathbb{N}}$  and  $(V_i)_{i \in \mathbb{N}}$  by induction. Starting from  $V_0 = U_0$  and  $\Theta_0 = \overline{V_0}$ , we then take for all  $i$ ,  $V_{i+1} = \bigcup_{j=1}^{m(i)} U_j$ , where  $m(i) = \min\{j' \mid \bigcup_{j=0}^{j'} U_j \supset \Theta_i\}$ . By induction,  $\Theta_i$  is compact, so that  $m(i) < \infty$  and  $\Theta_{i+1} = \overline{V_{i+1}}$  also is. That makes  $(\Theta_i)_{i \in \mathbb{N}}$  being an openly increasing sequence of subsets that covers  $\Theta$ .

To prove the second statement, we write  $\Theta = \bigcup_{n \in \mathbb{N}} U_n$  where the  $(U_n)_{n \in \mathbb{N}}$  are open subsets of  $\mathbb{R}^d$ . Using the first statement, there exists an openly increasing cover for each of the  $(U_n)_{n \in \mathbb{N}}$ :  $U_n = \bigcup_{j \in \mathbb{N}} \Theta_j^n$ , with  $\Theta_j^n \subset V_j^n \subset \Theta_{j+1}^n$  for some open subsets  $(V_j^n)_{j \in \mathbb{N}}$  of  $\Theta$ . Let us define  $\Theta_i = \bigcup_{j, n \leq i} \Theta_j^n$ . We verify that  $\Theta_i$  is compact and that  $\Theta_i \subset \bigcup_{j, n \leq i} V_j^n$ . Since for all  $n, j$   $V_j^n$  is an open subset of  $U_n$ , it is an open subset of  $\Theta$ . This concludes the proof since  $\Theta = \bigcup_{i \in \mathbb{N}} \Theta_i$ .  $\square$

**Corollary 3.1.** *If  $\Theta$  is an open subset of  $\mathbb{R}^d$  and  $f : \Theta \rightarrow \mathbb{R}$  is a v-a.e. continuous function of  $\Theta$ , there exists an openly increasing sequence of compact subsets  $(\Theta_i)_{i \in \mathbb{N}}$  such that  $\bigcup_{i \in \mathbb{N}} \Theta_i = \Theta \setminus E$ , where  $E$  is the negligible set of points on which  $f$  is discontinuous.*

*Proof.* The corollary follows from the previous proposition recalling that  $E$  is a countable intersection of closed subsets of  $\Theta$ .  $\square$

**Proposition 3.4.** *Assume  $\Theta$  is compact. Then if  $\Pi$  is a reference prior over a set of priors  $\mathcal{P}$ , it verifies*

$$\lim_{k \rightarrow \infty} \mathcal{I}^k(\Pi) - \mathcal{I}^k(P) \geq 0 \text{ for all } P \in \mathcal{P}. \quad (3.39)$$

*Proof.* We consider an openly increasing sequence of compact sets  $(\Theta_i)_{i \in \mathbb{N}}$  such that  $\bigcup_{i \in \mathbb{N}} \Theta_i = \Theta$ . By definition there exist  $i_0 \geq 0$  and a sequence  $(V_i)_{i \geq i_0}$  of open subsets of  $\Theta$  such that for any  $i \geq i_0$

$$\Theta_i \subset V_i \subset \Theta_{i+1}. \quad (3.40)$$

This way,  $\bigcup_i V_i = \Theta$  and the compacity of  $\Theta$  imposes it to be a finite union, so that  $\Theta_i = \Theta$  for any  $i \geq i_1$  for some  $i_1 \geq 0$ . Since  $\Pi$  is a reference prior over  $\mathcal{P}$  its renormalized restriction to  $\Theta_{i_1}$  verifies eq. (3.39). Hence the result.  $\square$

### 3.4.3 Proof of proposition 3.1

Let  $1 \leq \mathbf{i} \leq \infty$ ,  $I = \{i \in \mathbb{N}, i < \mathbf{i}\}$ , and  $(U_i)_{i \in I}$  be a sequence of disjoint non-empty sets in  $\mathcal{T}$  that covers  $\Theta$  such that  $\Pi(U_i) < \infty \forall i$ . When  $\Pi(U_i) = 0$ , we denote by  $\Pi(\cdot | U_i)$  an arbitrary probability distribution on  $U_i$  (for instance a Dirac distribution); otherwise, it is defined by  $\Pi(\cdot | U_i) = \Pi(\cdot \cap U_i) / \Pi(U_i)$ . Fix firstly  $i \in I$  and for any  $k \geq 1$  let us construct  $\mathbb{P}_{Y_k, T}^i$  on  $\mathcal{Y}^{\otimes k} \otimes \mathcal{T}$  such that

$$\mathbb{P}_{Y_k, T}^i(A_1 \times \cdots \times A_k \times B) = \int_B \mathbb{P}_{Y| \theta}^{\otimes k}(A_1 \times \cdots \times A_k) d\Pi(\theta | U_i), \quad (3.41)$$

so that  $\mathbb{P}_{Y_k, T}^i$  is a probability distribution on  $\mathcal{Y}^k \times U_i$ . Therefore, using Kolmogorov extension theorem, there exist a probability space  $(\Omega^i, \mathcal{X}^i, \mathbb{P}^i)$ , a random process  $\overline{Y^i} = (Y_j^i)_{j \geq 1}$ ,  $Y_j^i : \Omega^i \rightarrow \mathcal{Y}$ , and a random variable

### Chapter 3. Review of the reference prior theory

$T^i : \Omega^i \rightarrow \Theta$  such that their joint distribution is uniquely defined by the distribution  $\mathbb{P}_{Y_k, T}^i$  of  $(T^i, Y_k^i)$  for any  $k$ , with  $Y_k^i = (Y_j^i)_{j=1}^k$ .

Denote  $\Omega = I \times \prod_{i \in I} \Omega^i$ . A natural  $\sigma$ -algebra on  $\Omega$  is the one generated by the cylinder sets:

$$\Xi := \sigma(N \times E_0 \times \cdots \times E_{k-1} \times \prod_{k \leq i < \mathbf{i}} \Omega^i, N \in \mathcal{P}(I), E_0, \dots, E_{k-1} \in \mathcal{X}^i, k < \mathbf{i}). \quad (3.42)$$

We can define  $\bar{\Pi}$  by

$$\bar{\Pi}\left(N \times \prod_{i \in I} E_i\right) = \sum_{n \in N} \mathbb{P}^n(E_n)\Pi(U_n), \quad (3.43)$$

for any  $N \in \mathcal{P}(I)$ ,  $(E_i)_{i \in I} \in \prod_{i \in I} \mathcal{X}^i$  with  $E_i = \Omega^i$  for any  $i \geq k$  for some  $k < \mathbf{i}$ . The existence of a measure  $\Pi$  on  $\Xi$ , that coincides with  $\bar{\Pi}$  on the cylinder sets is guaranteed by Carathéodory's extension Theorem.

Now, we call  $p^n : (m, (\omega^i)_i) \in \Omega \mapsto \omega^n$  for any  $n \in I$ ,  $c : (m, (\omega^i)_i) \in \Omega \mapsto m$ , and  $p : \mathbf{w} \in \Omega \mapsto p^c(\mathbf{w})(\mathbf{w})$ . This way, we can define  $T : \Omega \rightarrow \Theta$  by  $T(\mathbf{w}) = T^c(\mathbf{w})(p(\mathbf{w}))$  and  $\bar{Y} = (Y_j)_{j \geq 1}$  with for any  $j \geq 1$ ,  $Y_j : \Omega \rightarrow \mathcal{Y}$  by  $Y_j(\mathbf{w}) = Y_j^c(\mathbf{w})(p(\mathbf{w}))$ .

We can verify that  $T$  is a measurable: if  $B \in \mathcal{T}$ ,

$$T^{-1}(B) = \bigcup_{i \in I} \{\mathbf{w} \in \Omega, c(\mathbf{w}) = i, T^i(p(\mathbf{w})) \in B\} = \bigcup_{i \in I} c^{-1}(\{i\}) \cap (T^i(p^i))^{-1}(B) \quad (3.44)$$

which is measurable. The same arguments stand for the measurability of  $Y_j$  for any  $j$ .

We have, for any  $V \in \mathcal{T}$ :

$$\begin{aligned} \Pi(T \in V) &= \bar{\Pi}\left(\bigcup_{i \in I} \{i\} \times \Omega^1 \times \cdots \times \{T^i \in V\} \times \Omega^{i+1} \times \cdots\right) \\ &= \sum_{i \in I} \mathbb{P}^i(T^i \in V)\Pi(U_i) = \sum_{i \in I} \Pi(V|U_i)\Pi(U_i) = \Pi(V); \end{aligned} \quad (3.45)$$

and if  $\Pi(V) < \infty$  then for any  $(A_{jl})_{l=1}^k \in \mathcal{Y}^k$ :

$$\begin{aligned} \Pi\left((Y_{jl})_{l=1}^k \in (A_{jl})_{l=1}^k | T \in V\right) \Pi(T \in V) \\ = \Pi\left(\bigcup_{i \in I} \{i\} \times \Omega^1 \times \cdots \{(Y_{jl})_{l=1}^k \in (A_{jl})_{l=1}^k\} \cap \{T^i \in V\} \times \Omega^{i+1} \cdots\right), \end{aligned} \quad (3.46)$$

so that

$$\begin{aligned} \Pi\left((Y_{jl})_{l=1}^k \in (A_{jl})_{l=1}^k | T \in V\right) &= \sum_{i \in I} \mathbb{P}^i\left(\{(Y_{jl})_{l=1}^k \in (A_{jl})_{l=1}^k\} \cap \{T^i \in V\}\right) \frac{\Pi(U_i)}{\Pi(V)} \\ &= \sum_{i \in I} \int_{V \cap U_i} \mathbb{P}_{Y|\theta}^{\otimes k}((A_{jl})_{l=1}^k) \frac{d\Pi(\theta)}{\Pi(V)} = \int_{\Theta} \mathbb{P}_{Y|\theta}^{\otimes k}((A_{jl})_{l=1}^k) d\Pi(\theta|V). \end{aligned} \quad (3.47)$$

We precise that the above expressions rely on the fact that  $\Pi(I \times \prod_{i \in I} E_i) = \sum_{i \in I} \mathbb{P}^i(E_i)\Pi(U_i)$ , for any  $E_i \in \mathcal{X}^i$ , for every  $i \in I$ . To state that equality, we write  $I \times \prod_{i \in I} E_i = \bigcup_{n \geq 1} \bigcap_{k < \mathbf{i}} (I^n \times \prod_{i \in I} E'_{i,k})$  with  $I^n = \{i \in I, i < n\}$  and  $E'_{i,k} = E_i$  if  $i \leq k$ ,  $E'_{i,k} = \Omega^i$  otherwise. This way for any  $n \geq 1$ :  $I^n \times \prod_{i \in I} E'_{i,k} \subset I^{n+1} \times \prod_{i \in I} E'_{i,k}$  for any  $k < \mathbf{i}$ , and  $\bigcap_{k < \mathbf{i}} I^n \times \prod_{i \in I} E'_{i,k} \subset \bigcap_{k < \mathbf{i}} I^{n+1} \times \prod_{i \in I} E'_{i,k}$ . Thus,

$$\Pi\left(I \times \prod_{i \in I} E_i\right) = \lim_{n \rightarrow \infty} \Pi\left(\bigcap_{k < \mathbf{i}} (I^n \times \prod_{i \in I} E'_{i,k})\right). \quad (3.48)$$

For any  $n \geq 1$ ,  $(I^n \times \prod_{i \in I} E'_{i,k})$  is a decreasing sequence, because  $E'_{i,k+1} \subset E'_{i,k}$  for any  $k < \mathbf{i}-1$ . Also,  $\Pi(I^n \times \prod_{i \in I} E'_{i,0}) = \sum_{i=0}^{n-1} \Pi(U_i) < \infty$ . Therefore, we can write:

$$\begin{aligned} \Pi\left(\bigcap_{k < \mathbf{i}} (I^n \times \prod_{i \in I} E'_{i,k})\right) &= \lim_{k \rightarrow \infty} \Pi\left(I^n \times \prod_{i < k, \mathbf{i}} E_i \times \prod_{k \leq i < \mathbf{i}} \Omega^i\right) \\ &= \lim_{k \rightarrow \infty} \left[ \sum_{i < k, \mathbf{i}} \mathbb{P}^i(E_i)\Pi(U_i) + \sum_{k \leq i < n, \mathbf{i}} \Pi(U_i) \right] = \sum_{i < n, \mathbf{i}} \mathbb{P}^i(E_i)\Pi(U_i), \end{aligned} \quad (3.49)$$

because the sums are all finite. Eventually, the limit in eq. (3.48) is equal to  $\sum_{i < i} \mathbb{P}^i(E_i)\Pi(U_i)$  as expected, hence the result.

## 3.5 Conclusive remarks

Reference priors are designed as a solution to the search for objective priors. To do so, they are constructed on informatic-theoretic foundations, under a framework though to ensure they maximize the role of the data in the definition of the posterior distribution.

The theory is provided with a formal definition and several works propose rigorous studies of reference priors, in different contexts. That let them be elected for their lack of subjectivity in various practical studies (e.g. Chen, Ibrahim, and S. Kim, 2008; Gu and Berger, 2016; D'Andrea et al., 2021) and various statistical models, such as Gaussian process-based models (Paulo, 2005; Gu and Berger, 2016), generalized linear models (Natarajan and Kass, 2000) or general exponential family models (Clarke and Ghosal, 2010).

Moreover, the framework that we have introduced allows a rigorous incorporation of improper priors in Bayesian modelings. This provides a better interpretability of the reference priors framework when it leads to an improper prior distribution.

However, some limits remain. First, the theory is still built on some choices, such as the choice of the dissimilarity measure to quantify how different are the prior and the posterior. Second, the incorporation of improper priors is appealing yet can represent an issue in cases where they issue improper posteriors. Actually, reference priors are known to lead to improper posteriors in some cases. Last, existing theoretical developments limit their studies to sets of priors that are as general as possible, and effects of its restriction are rarely focused.



# 4

## Generalized mutual information and their reference priors

This chapter is mainly based on the submitted work: Antoine Van Biesbroeck (2024a). “Generalized mutual information and their reference priors under Csizar f-divergence”. arXiv:2310.10530. DOI: [10.48550/arXiv.2310.10530](https://doi.org/10.48550/arXiv.2310.10530)

**Abstract** This chapter complements the reference prior theory. We propose an interpretation of the theory from a sensitivity analysis viewpoint. This leads us to propose a new way of defining mutual information, using different dissimilarity measures between probability distributions. Our construction creates a new framework for reference priors, which are rigorously studied. Our main result gives a limit of the generalized mutual information when the dissimilarity measure considered resembles to a  $\delta$ -divergence. It makes it easier to derive reference priors under constraints or within specific sets. In the absence of constraints, we prove that the Jeffreys prior maximizes the generalized mutual information, reinforcing its objective characteristic.

<b>4.1</b>	<b>Introduction and motivations</b>	56
<b>4.2</b>	<b>Generalized mutual information</b>	56
4.2.1	Mutual information and definitions	56
4.2.2	Mutual information as a sensitivity index	57
4.2.3	Generalized mutual information definition	57
<b>4.3</b>	<b>Generalized reference priors</b>	58
<b>4.4</b>	<b>Towards <math>\delta</math>-divergences-reference priors</b>	60
4.4.1	A first result	61
4.4.2	Results when $\delta < 0$	61
4.4.3	Results when $\delta > 0$	62
<b>4.5</b>	<b>Discussions</b>	63
4.5.1	About the results	63
4.5.2	About the introduction of simple constraints	63
4.5.3	About the assumptions and their limitations	64
4.5.4	About the robustness of Jeffreys prior with different divergences	65
<b>4.6</b>	<b>Proofs of the main results</b>	68
4.6.1	Proof of the convergence in probability	68
4.6.2	Proof of the results when $\delta < 0$	70
4.6.3	Proof of the results when $\delta > 0$	72
<b>4.7</b>	<b>Conclusion and prospects</b>	76

## 4.1 Introduction and motivations

The reference prior theory aims at defining priors that are the most objective possible. A review of the theory can be found in [chapter 3](#). It is built on the maximization of the mutual information, which is designed to measure the information brought by the data in the posterior distribution.

Although there are infinite ways to compare probability distributions, the mutual information is classically defined as an expected Kullback-Leibler divergence between the prior and the posterior. The settings under which the reference priors are usually defined result from the aforementioned choice. Since the aim is to define priors that are not impacted by potentially subjective choice, we ask how the results about reference priors change when the definition of the mutual information is modified.

Extensions of the mutual information in the reference prior definition have already been explored in the literature (e.g. Chen, Dey, et al., [2010](#); Liu et al., [2014](#); Le, [2014](#)), and some do not necessarily lead to the Jeffreys prior (Hashimoto, [2021](#); Clarke and Sun, [1997](#); M. Ghosh, Mergel, and Liu, [2011](#)). They are all based on the expression of mutual information as an average divergence between the posterior and the prior distributions.

In this chapter, we contribute to the reference prior theory with an original derivation of the mutual information from a Global Sensitivity Analysis (GSA) based viewpoint. GSA, whose principle is to measure how the uncertainty of an output is impacted by that of some of its input (Iooss and Lemaître, [2015](#)), allows us indeed to provide an interpretation of the reference prior as a maximizer of such sensitivity influence that the observations get from the parameter of interest. This suggestion leads to defining what we call generalized mutual information, by analogy with global sensitivity indices (which are introduced for instance in Da Veiga et al., [2021](#)). It relies on the wide range of existing dissimilarity measures between two probability distributions. An example is the  $f$ -divergence subclass (Csiszár, [1967](#)), commonly employed as an extension of Shannon's entropy for various purposes in statistics, such as, variational inference (Minka, [2005](#); Bach, [2023](#)), surrogate model design (Nguyen, Wainwright, and Jordan, [2009](#)), PAC Bayesian learning (Picard-Weibel and Guedj, [2022](#)) and Differential Privacy (Mironov, [2017](#)). A study of those divergences within our generalized mutual information is also a main contribution of this chapter, with the goal of deriving what one is invited to call generalized reference priors. We provide an accomplished formalization for the generalized reference prior settings, and results based on classical  $f$ -divergences for sensitivity analysis such as  $\delta$ -divergences. Our main result takes the form of a limit w.r.t. the number of data of the mutual information. Its analytical expression as a function of the prior permits a simple expression of our reference priors: they are its maximal arguments. It opens the path of easier theoretical derivation of reference priors among the ones that satisfy different kinds of constraints, or that belong to some particular sets of priors.

In the next section, we elucidate our motivation from a Global Sensitivity Analysis viewpoint for an enrichment of the mutual information. The generalized mutual information that we define is constructed from any possible dissimilarity measure. That definition is used in section [4.3](#) to define the generalized reference priors. We show in that section that the usual properties satisfied by reference priors still stand under our framework. A deeper study is afterward conducted in section [4.4](#) in the case where the dissimilarity measure is an  $f$ -divergence resembling to a  $\delta$ -divergence. These results, alongside open thoughts about their robustness when the dissimilarity measure is not limited to such  $f$ -divergences, are discussed in section [4.5](#). After proving our statements in section [4.6](#), we conclude this chapter in section [4.7](#).

## 4.2 Generalized mutual information

### 4.2.1 Mutual information and definitions

We consider a statistical modeling characterized by a collection of distributions  $(\mathbb{P}_{Y|\theta})_{\theta \in \Theta}$ . The Bayesian framework is constructed as described in [chapter 3](#) (section [3.4](#)). We suppose that  $\Theta \subset \mathbb{R}^d$  and denote by  $\nu$  the Lebesgue measure on  $\mathbb{R}^d$ . The priors considered are absolutely continuous distributions w.r.t.  $\nu$ .

We consider the same notations as in [chapter 3](#):  $Y_k$  denotes a random vector of  $k$  data items whose distribution conditionally to  $T = \theta$  is  $\mathbb{P}_{Y_k|\theta} := \mathbb{P}_{Y|\theta}^{\otimes k}$ , where  $T$  is an r.v. whose distribution is the prior  $\Pi$

(which is not necessarily a probability distribution, see section 3.4.1). The density of  $\Pi$  is denoted by  $\pi$ .

We suppose that the modeling admits a likelihood, denoted by  $\ell$  with  $\forall \theta \in \Theta, \mathbf{y} \in \mathcal{Y}^k, \ell_k(\mathbf{y}|\theta) = \prod_{i=1}^k \ell(y_i|\theta)$ . For all  $\theta$ ,  $\ell(\cdot|\theta)$  is the density of  $\mathbb{P}_{T|\theta}$  w.r.t. a common measure  $\mu$  on  $\mathcal{Y}$ . We denote by  $p_{\mathbf{Y}_k}$  the marginal density of  $\mathbf{Y}_k$ , and by  $p(\cdot|\mathbf{y})$  the posterior density given the observations  $\mathbf{y} \in \mathcal{Y}^k$ . Furthermore, we suppose the modeling to be regular (i.e. assumption 3.1 in chapter 3 is verified), so that the Fisher information matrix (that we denote  $\mathcal{I}$ ) and the Jeffreys prior exist. A density of the Jeffreys prior is denoted by  $J$ .

Under this framework, when the prior  $\Pi$  is proper, we recall that the mutual information is classically defined as the following quantity:

$$\mathcal{I}^k(\Pi) = \mathbb{E}_{\mathbf{Y}_k \sim \mathbb{P}_{\mathbf{Y}_k}} [\text{KL}(\mathbb{P}_{T|\mathbf{Y}_k} || \Pi)]. \quad (4.1)$$

We recall that in the above expression, the posterior and marginal distributions depend on the prior considered. Applying Fubini-Lebesgue's theorem allows to also write the mutual information as follows

$$\mathcal{I}^k(\Pi) = \mathbb{E}_{T \sim \Pi} [\text{KL}(\mathbb{P}_{\mathbf{Y}_k|T} || \mathbb{P}_{\mathbf{Y}_k})]. \quad (4.2)$$

That last quantity expresses the mutual information as a measure of the influence that the stochastic parameter  $T$  has on the data  $\mathbf{Y}_k$ . That interpretation aligns with a sensitivity analysis viewpoint. Actually, the expression in eq. (4.2) can be seen as a sensitivity index, where the impact of the input  $T$  on the output  $\mathbf{Y}_k$  is studied. We develop this interpretation in the following section.

*Proof that eq. (4.1) and eq. (4.2) are equal.* One can remark that for any  $x \in (1, \infty)$ ,  $|\log x| \leq x - \log x$ . That statement allows applying proposition 4.3, which is stated later on, with  $f = -\log$ . It provides the desired equality whether the involved quantities are finite or infinite.  $\square$

## 4.2.2 Mutual information as a sensitivity index

The integration of the Bayesian formalism in the settings of uncertainty quantification resorts in modeling  $\mathbf{Y}_k$  as the output of a stochastic system whose inputs include  $T$ . In other terms, we write  $\mathbf{Y}_k = \mathcal{M}(T, \epsilon)$ . Here,  $T$  embeds uncertainty on the input parameters and  $\epsilon$  represents an unknown variable embedding inherent uncertainty.

A well-designed Bayesian modeling parameterized by  $T$  would be such that the input between  $\epsilon$  and  $T$  whose impact on the output is the highest is  $T$ . Such a modeling should reduce the impact of the irreducible uncertainty embedded in  $\epsilon$ .

This impact is measured by a sensitivity index as follows, considering a dissimilarity measure  $D$ ,

$$S = \mathbb{E}_{T \sim \Pi} [D(\mathbb{P}_{\mathbf{Y}_k} || \mathbb{P}_{\mathbf{Y}_k|T})]. \quad (4.3)$$

It corresponds to the mutual information  $\mathcal{I}^k(\Pi)$  when  $D$  is defined as  $D(P||Q) = \text{KL}(Q||P)$ .

Actually, different choices for dissimilarity measures  $D$  can be done to define sensitivity indices (Da Veiga, 2015). For instance, setting  $D(P||Q) = |\mathbb{E}_{X \sim P} X - \mathbb{E}_{X \sim Q} X|^2$  gives the un-normalized Sobol' index (Sobol', 1993). Among the most common ones, we recall the  $f$ -divergences, defined from a real and measurable function  $f$  that is convex and maps 1 to 0. The  $f$ -divergence is denoted  $D_f$  and defined as

$$D_f(P||Q) = \int_{\mathcal{X}} f\left(\frac{p(x)}{q(x)}\right) q(x) d\omega(x) \quad (4.4)$$

where  $p, q$  respectively are densities of  $P$  and  $Q$  w.r.t. a common measure  $\omega$  on  $\mathcal{X}$ .

## 4.2.3 Generalized mutual information definition

The same way as sensitivity indices are supported by various dissimilarity measures, our suggestion is to define the mutual information using other dissimilarity measures than the Kullback-Leibler divergence. Our novel definition is the following.

**Definition 4.1** ( $D$ -mutual information). Let  $D$  be a dissimilarity measure, for a fixed  $k$ , the  $D$ -mutual information of a proper prior  $\Pi$  with total mass equal to 1 under  $k$  observations is defined as

$$\mathcal{I}_D^k(\Pi) := \mathbb{E}_{T \sim \Pi}[D(\mathbb{P}_{Y_k} || \mathbb{P}_{Y_k|T})]. \quad (4.5)$$

In the case where  $D = D_f$  is an  $f$ -divergence, the mutual information equals

$$\mathcal{I}_{D_f}^k(\Pi) = \mathbb{E}_{T \sim \pi}[D_f(\mathbb{P}_{Y_k} || \mathbb{P}_{Y_k|T})] = \int_{\Theta} \int_{\mathcal{Y}^k} f\left(\frac{p_{Y_k}(\mathbf{y})}{\ell_k(\mathbf{y}|\theta)}\right) \ell_k(\mathbf{y}|\theta) d\mu^{\otimes k}(\mathbf{y}) \pi(\theta) d\theta. \quad (4.6)$$

It corresponds to the classical definition in the case where  $f = -\log$ . Actually, the classical definition is a subcase of the study of a particular class of the  $f$ -divergences: the  $\delta$ -divergences. They are defined setting  $f = f_\delta$  where

$$f_\delta(x) = \begin{cases} \frac{x^\delta - \delta x - (1-\delta)}{\delta(\delta-1)} & \text{if } \delta \notin \{0, 1\}, \\ x \log x - x + 1 & \text{if } \delta = 1, \\ -\log x + x - 1 & \text{if } \delta = 0. \end{cases} \quad (4.7)$$

When  $\delta = 0$ , we retrieve the original form (KL form) of the mutual information.

### 4.3 Generalized reference priors

Reference priors are defined as asymptotic maximizers of the mutual information. Using the same formalism as in their original definition (that is expressed in chapter 3), we suggest the following definition of generalized reference priors using our generalized mutual information.

**Definition 4.2** (Generalized reference prior). Let  $D$  be a dissimilarity measure and  $\mathcal{P}$  a set of priors on  $\Theta$ . A prior  $\Pi \in \mathcal{P}$  is called a  $D$ -reference prior over  $\mathcal{P}$  with rate  $\varphi(k)$  if there exists an openly increasing sequence of compact subsets  $(\Theta_i)_{i \in \mathbb{N}}$  such that  $\bigcup_{i \in \mathbb{N}} \Theta_i = \Theta$  and for any  $i$ :  $0 < \Pi(\Theta_i) < \infty$  and

$$\lim_{k \rightarrow \infty} \varphi(k)[\mathcal{I}_D^k(\Pi(\cdot|\Theta_i)) - \mathcal{I}_D^k(P(\cdot|\Theta_i))] \geq 0 \text{ for all } P \in \mathcal{P} \text{ verifying } 0 < P(\Theta_i) < \infty; \quad (4.8)$$

where  $\varphi(k)$  is a positive and monotonous function of  $k$ .

This definition matches with the original one when the dissimilarity measure is the Kullback-Leibler divergence. Nevertheless, it introduces the definition of an associated rate proposed by ourselves. In fact, our work provides elements showing that such rate exists and may vary as a function of the dissimilarity measure considered. The idea behind  $\varphi$  is to look for an asymptotic expansion of  $\mathcal{I}_D^k(P_1) - \mathcal{I}_D^k(P_2)$  that stands for any pair of probability distributions  $(P_1, P_2)$ :

$$\mathcal{I}_D^k(P_1) - \mathcal{I}_D^k(P_2) \underset{k \rightarrow \infty}{=} \varphi(k)(L(P_1) - L(P_2)) + o(\varphi(k)), \quad (4.9)$$

where  $L$  is a mapping that does not depend on  $P_1$  or  $P_2$ . When it resorts to the original mutual information, the rate  $\varphi$  is constant.

The results provided below ensure that the generalized reference priors verify similar properties as the original reference priors.

**Proposition 4.1** (Invariance by reparametrization). Consider a reparametrization  $\phi = g(\theta)$  with  $g$  being a diffeomorphism and  $\mathcal{P}$  a set of priors on  $\Theta$ . Then if  $\Pi$  is a reference prior over  $\mathcal{P}$  for the model parameterized by  $\theta$ ,  $\tilde{\Pi}$  is a reference prior over  $\tilde{\mathcal{P}}$  for the model parameterized by  $\phi$ , where:

$$\tilde{\mathcal{P}} = \{M(\pi \cdot |\det dg_\phi|^{-1}), M(\pi) \in \mathcal{P}\}, \quad (4.10)$$

with  $M : f \in \mathcal{R} \mapsto (B \mapsto \int_B f(\theta) d\theta) \in \mathcal{M}^v$ .

*Proof.* The proposition comes from the fact that  $\mathcal{I}_D^k$  is clearly invariant by reparametrization: for any prior  $\Pi$  having a density  $\pi$  on  $\Theta$ , and any  $\tilde{\Theta} \subset \Theta$  with  $\Pi(\tilde{\Theta}) \in (0, \infty)$ , calling  $\tilde{\Xi} = g(\tilde{\Theta})$ ,

$$\mathcal{I}_D^k(\Pi(\cdot|\tilde{\Theta})) = \mathbb{E}_{T \sim \Pi(\cdot|\tilde{\Theta})}[D(\mathbb{P}_{Y_k}||\mathbb{P}_{Y_k|T})] = \mathbb{E}_{\Phi \sim \tilde{\Pi}(\cdot|\tilde{\Xi})}[D(\mathbb{P}_{Y_k}||\mathbb{P}_{Y_k|g^{-1}(\Phi)})] =: \tilde{\mathcal{I}}_D^k(\tilde{\Pi}(\cdot|\tilde{\Xi})) \quad (4.11)$$

where  $\tilde{\Pi}$  is defined by its density  $\tilde{\pi} = \pi \cdot |\det d_\phi g|^{-1}$ , and  $\tilde{\mathcal{I}}_D^k$  denoted the  $D$ -mutual information on the reparameterized model.  $\square$

**Proposition 4.2** (Consistency with compact cases). *Assume  $\Theta$  is compact. Let  $\mathcal{P}$  be a set of continuous priors. If  $\Pi$  is a reference prior over  $\mathcal{P}$  with rate  $\varphi$ , then*

$$\lim_{k \rightarrow \infty} \varphi(k)[\mathcal{I}_D^k(\Pi) - \mathcal{I}_D^k(P)] \geq 0 \text{ for all } P \in \mathcal{P}. \quad (4.12)$$

*Proof.* As the priors are continuous, their restrictions are proper on any compact. Thus, the quantities involved in the proposition are well-defined. As  $\Pi$  is a reference prior over  $\mathcal{P}$  we can consider an associated openly increasing sequence of compacts  $(\Theta_i)_{i \in \mathbb{N}}$  that covers  $\Theta$ . Since  $\Theta$  is compact there exists  $i$  such that  $\Theta_i = \Theta$ . Hence the result since the renormalized restriction  $\Pi(\cdot|\Theta_i)$  verifies the statement of the proposition by definition.  $\square$

**Proposition 4.3** (Invertibility). *Consider that  $\Pi$  is a continuous prior and let  $f$  be a convex function.*

1. *If for a compact set  $\tilde{\Theta}$ ,*

$$\sup_{\theta \in \tilde{\Theta}} \sup_{\theta' \in \tilde{\Theta}} \int_{\mathcal{Y}^k} f\left(\frac{\ell_k(\mathbf{y}|\theta')}{\ell_k(\mathbf{y}|\theta)}\right) \ell_k(\mathbf{y}|\theta) d\mu^{\otimes k}(\mathbf{y}) < \infty \quad (4.13)$$

*Then the  $D_f$ -mutual information value of  $\Pi(\cdot|\tilde{\Theta})$  is finite.*

2. *If there exists  $K \geq 0$  such that for any  $x > 0$ ,  $|f(x)\mathbb{1}_{f(x)<0}| \leq Kx$  then the  $D_f$ -mutual information restricted to compact expression can be inverted back:*

$$\mathbb{E}_{T \sim \Pi(\cdot|\tilde{\Theta})}[D_f(\mathbb{P}_{Y_k}||\mathbb{P}_{Y_k|T})] = \mathbb{E}_{Y_k \sim \mathbb{P}_{Y_k}}[D(\Pi(\cdot|\tilde{\Theta})||\mathbb{P}_{T|Y_k})]. \quad (4.14)$$

*Proof.* We denote  $\tilde{\Pi} = \Pi(\cdot|\tilde{\Theta})$  and  $\tilde{\pi}$  the latter's density. We have

$$\begin{aligned} \mathcal{I}_{D_f}^k(\tilde{\Pi}) &= \int_{\tilde{\Theta}} \int_{\mathcal{Y}^k} f\left(\frac{p_{Y_k}(\mathbf{y})}{\ell_k(\mathbf{y}|\theta)}\right) \ell_k(\mathbf{y}|\theta) d\mu^{\otimes k}(\mathbf{y}) \tilde{\pi}(\theta) d\theta \\ &\leq \sup_{\theta \in \tilde{\Theta}} \int_{\mathcal{Y}^k} \int_{\tilde{\Theta}} f\left(\frac{\ell_k(\mathbf{y}|\theta')}{\ell_k(\mathbf{y}|\theta)}\right) \ell_k(\mathbf{y}|\theta) \tilde{\pi}(\theta') d\theta' d\mu^{\otimes k}(\mathbf{y}), \end{aligned} \quad (4.15)$$

using the convexity of  $f$ . Since  $\tilde{\Theta}$  is compact,  $\tilde{\pi}$  is bounded. Thus, we obtain  $\mathcal{I}_{D_f}^k(\tilde{\Pi}) < \infty$ .

For inverting the integrals, we notice that the assumption on  $f$  allows writing

$$\int_{\tilde{\Theta}} \int_{\mathcal{Y}^k} \left| f\left(\frac{p_{Y_k}(\mathbf{y})}{\ell_k(\mathbf{y}|\theta)}\right) \right| \ell_k(\mathbf{y}|\theta) d\mu^{\otimes k}(\mathbf{y}) \tilde{\pi}(\theta) d\theta \leq 2K + \int_{\tilde{\Theta}} \int_{\mathcal{Y}^k} f\left(\frac{p_{Y_k}(\mathbf{y})}{\ell_k(\mathbf{y}|\theta)}\right) \ell_k(\mathbf{y}|\theta) d\mu^{\otimes k}(\mathbf{y}) \tilde{\pi}(\theta) d\theta \quad (4.16)$$

using  $|f(x)| \leq 2Kx\mathbb{1}_{f(x)<0} + f(x)$ . We deduce that, if  $\mathbb{E}_{T \sim \tilde{\Pi}}[D_f(\mathbb{P}_{Y_k}||\mathbb{P}_{Y_k|T})]$  is finite, then Fubini-Lebesgue's theorem can be applied, leading to the equality (4.14). The same way, we have

$$\int_{\mathcal{Y}^k} \int_{\tilde{\Theta}} \left| f\left(\frac{\tilde{\pi}(\theta)}{p(\theta|\mathbf{y})}\right) \right| p(\theta|\mathbf{y}) d\theta p_{Y_k}(\mathbf{y}) d\mu^{\otimes k}(\mathbf{y}) \leq 2K + \int_{\mathcal{Y}^k} \int_{\tilde{\Theta}} f\left(\frac{\tilde{\pi}(\theta)}{p(\theta|\mathbf{y})}\right) p(\theta|\mathbf{y}) d\theta p_{Y_k}(\mathbf{y}) d\mu^{\otimes k}(\mathbf{y}), \quad (4.17)$$

so that if  $\mathbb{E}_{\mathbf{Y}_k \sim \mathbb{P}_{\mathbf{Y}_k}}[D_f(\tilde{\Pi} || \mathbb{P}_{T|\mathbf{Y}_k})]$  is finite, the integrals can also be inverted and equality (4.14) still stands. We conclude by deducing that, if one of the quantity in eq. (4.14) is infinite, the other cannot be finite, letting the equality to stand as well.  $\square$

**Proposition 4.4** (Compatibility with sufficient statistics). *Suppose  $D$  is an  $f$ -divergence such that there exists  $K \geq 0$  with for all  $x > 0$ ,  $|f(x)\mathbb{1}_{f(x)<0}| \leq Kx$ . Consider  $Z := z(\mathbf{Y}_k) \in \mathcal{Z}$  a sufficient statistic of the model built on  $(\mathbb{P}_{Y|\theta})_{\theta \in \Theta}$ . The  $D_f$ -mutual information remains stable by considering  $Z$  instead of  $\mathbf{Y}_k$ :*

$$\mathcal{I}_{D_f}^k(\Pi(\cdot|\tilde{\Theta})) = \mathbb{E}_{T \sim \Pi(\cdot|\tilde{\Theta})}[D(\mathbb{P}_{\mathbf{Y}_k} || \mathbb{P}_{\mathbf{Y}_k|T})] = \mathbb{E}_{T \sim \Pi(\cdot|\tilde{\Theta})}[D(\mathbb{P}_Z || \mathbb{P}_{Z|T})] =: \tilde{\mathcal{I}}_{D_f}^k(\Pi(\cdot|\tilde{\Theta})) \quad (4.18)$$

for all  $\Pi \in \mathcal{M}^\nu$ ; where  $\tilde{\mathcal{I}}^k$  denotes the mutual information considering the second model.

*Proof.* If  $Z$  is a sufficient statistic, the posterior verifies  $p(\theta|z(\mathbf{y})) = p(\theta|\mathbf{y})$ . Thus, the terms  $D_f(\Pi(\cdot|\tilde{\Theta}) || \mathbb{P}_{T|\mathbf{Y}_k})$  and  $D_f(\Pi(\cdot|\tilde{\Theta}) || \mathbb{P}_{T|Z})$  are equal. Hence the result since

$$\mathcal{I}_{D_f}^k(\Pi(\cdot|\tilde{\Theta})) = \mathbb{E}_{\mathbf{Y}_k \sim \mathbb{P}_{\mathbf{Y}_k}}[D_f(\Pi(\cdot|\tilde{\Theta}) || \mathbb{P}_{T|\mathbf{Y}_k})] \quad \text{and} \quad \tilde{\mathcal{I}}_{D_f}^k(\Pi(\cdot|\tilde{\Theta})) = \mathbb{E}_{Z \sim \mathbb{P}_Z}[D_f(\Pi(\cdot|\tilde{\Theta}) || \mathbb{P}_{T|Z})]. \quad (4.19)$$

$\square$

To conclude this section, we mention the work of Le (2014). The author proposes a formulation of  $D_{f_\delta}$ -reference priors as limit of priors when  $\Theta \subset \mathbb{R}$ , in the same fashion as in Berger, Bernardo, and Sun, 2009 (see chapter 3, theorem 3.1).

**Theorem 4.1** (Le, 2014). *Consider a  $\delta$ -divergence  $D_{f_\delta}$  with  $\delta \in [0, 1]$ . Assume  $\Theta \subset \mathbb{R}$ . Let  $\mathcal{P}_s$  be the set of continuous positive priors admitting proper posteriors. Let  $\Pi^* \in \mathcal{P}_s$  we call  $p^*(\cdot|\cdot)$  its posterior density and define for any interior point  $\theta_0 \in \Theta$ ,*

$$f_k(\theta) = \exp \left( \int_{\mathcal{Y}^k} \ell_k(\mathbf{y}|\theta) \log(p^*(\theta|\mathbf{y})) d\mu^{\otimes k}(\mathbf{y}) \right) \quad \text{and} \quad f(\theta) = \lim_{k \rightarrow \infty} \frac{f_k(\theta)}{f_k(\theta_0)}. \quad (4.20)$$

*Under mild assumptions that are similar to the ones of Berger, Bernardo, and Sun, 2009, the prior admitting  $f$  (if it exists) as density is a reference prior over  $\mathcal{P}_s$ .*

## 4.4 Towards $\delta$ -divergences-reference priors

The objective of this section is to study the  $D_f$ -reference priors for certain classes of functions  $f$  that include  $\delta$ -divergence. We precise that from now on, the priors will always be considered continuous (they admit  $\nu$ -a.e. continuous and locally bounded density functions). We recall that the set of such priors is denoted by  $\mathcal{M}_C^\nu$ , and they all admit a density in the set  $\mathcal{R}_{C^b}$ .

We recall that the random vectors are given by a process  $\bar{\mathbf{Y}}$  defined on a conditional probability space  $(\Omega, \Xi, [\Pi])$  (see proposition 3.1). If  $\tilde{\Theta}$  is a compact subset of  $\Theta$ , the restriction  $(\tilde{\Omega}, \tilde{\Xi}, \Pi(\cdot|T \in \tilde{\Theta}))$  of the conditional probability space to  $\{T \in \tilde{\Theta}\}$  is a probability space. For some  $\theta \in \tilde{\Theta}$  we will adopt the notation  $\mathbb{P}_\theta$  for the conditional probability on  $\tilde{\Omega}$  to  $\{T = \theta\}$ . That means, for any  $k$ , any  $A \in \mathcal{Y}^{\otimes k}$ ,

$$\mathbb{P}_\theta(\mathbf{Y}_k \in A) = \mathbb{P}_{\mathbf{Y}_k|\theta}(A). \quad (4.21)$$

We denote by  $\mathbb{E}_\theta$  the associated expectation. The proofs of the results that are stated in this section are given in section 4.6.

#### 4.4.1 A first result

A first useful result comes when we consider an asymptotic expansion of  $f$  in the neighborhood of 0:

$$f(x) \underset{x \rightarrow 0}{=} g_1(x) + o(g_2(x)) \quad (4.22)$$

for some  $g_1 : (0, \infty) \rightarrow \mathbb{R}$ ,  $g_2 : (0, \infty) \rightarrow (0, \infty)$ .

**Proposition 4.5.** Let  $\tilde{\Theta}$  be a compact subset of  $\Theta$ . Consider a prior  $\Pi \in \mathcal{M}_C^\vee$  and denote by  $\pi \in \mathcal{R}_{C^b}$  a density of  $\Pi(\cdot | \tilde{\Theta})$ . For almost every  $\theta$  in the interior of  $\text{Supp } \pi$  there exists  $c = c_\theta$ ,  $C = C_\theta$  positive constants such that the following convergence in probability holds:

$$\tilde{g}_2^{C,c}(k^{-d/2})^{-1} \left| f\left(\frac{p_{Y_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)}\right) - g_1\left(k^{-d/2}\pi(\theta)(2\pi)^{d/2}|\mathcal{I}(\theta)|^{-1/2}\exp\left(\frac{1}{2}S_k^T\mathcal{I}(\theta)^{-1}S_k\right)\right) \right| \xrightarrow[k \rightarrow \infty]{\mathbb{P}_\theta} 0, \quad (4.23)$$

where  $\tilde{g}_2^{C,c}(k^{-d/2}) = \sup_{x \in [c, C]} g_2(xk^{-d/2})$ , and  $S_k$  denotes  $\frac{1}{\sqrt{k}} \sum_{i=1}^k \nabla_\theta \log \ell(Y_i|\theta)$ .

There are three things we can say about this first asymptotic result. First, it emphasizes the asymptotic link between two ratios: (i) the marginal over the likelihood on the one hand, (ii) the chosen prior density over Jeffreys one on the other hand. The intuition comes from noticing that  $S_k^T\mathcal{I}(\theta)^{-1}S_k$  converges in distribution to a standard Gaussian, which does not depend on  $\theta$ . Second, this result also highlights that it is the behavior of  $f$  in the neighborhood of 0 that is decisive. Third, this result is two steps away from giving an asymptotic expansion of the  $D_f$ -mutual information: (i) stating that this limit stands in  $L^1(\mathbb{P}_\theta)$ , (ii) verifying the limit can be switched with the expectation w.r.t.  $T \sim \Pi(\cdot | \tilde{\Theta})$ .

One can notice that this proposition can be applied with  $f = -\log$ ,  $g_1 = f$  and  $g_2 = 1$ . In that case we obtain the convergence in  $\mathbb{P}_\theta$ -probability already stated in Clarke and Barron, 1990:

$$\log \frac{\ell_k(\mathbf{Y}_k|\theta)}{p_{Y_k}(\mathbf{Y}_k)} - \frac{d}{2} \log k + \log \pi(\theta) - \frac{1}{2} \log |\mathcal{I}(\theta)| + \frac{1}{2} S_k^T \mathcal{I}(\theta)^{-1} S_k \xrightarrow[k \rightarrow \infty]{\mathbb{P}_\theta} 0. \quad (4.24)$$

#### 4.4.2 Results when $\delta < 0$

In what follows, we precise the asymptotic form of  $f$  in the neighborhood of 0:

$$f(x) \underset{x \rightarrow 0}{=} ax^\delta + o(x^\delta). \quad (4.25)$$

We also require  $f$  to be locally bounded and to be controlled as  $x \rightarrow \infty$ :

$$f \underset{x \rightarrow \infty}{=} O(x). \quad (4.26)$$

We invite the reader to remark that such function  $f$  aligns with translated  $\delta$ -divergences, among others.

In this section we suppose  $\delta \in (-1, 0)$ . An asymptotic limit of the  $D_f$ -mutual information can be given once the following assumption is satisfied.

**Assumption 4.1.** For every compact subset  $\tilde{\Theta}$  of  $\Theta$ , there exist  $K_0$  such that the quantity

$$\mathbb{E}_\theta \left[ \sup_{\theta \in \tilde{\Theta}} \|\nabla_\theta^2 \log \ell(Y_1|\theta)\| \right] \leq K_0. \quad (4.27)$$

**Theorem 4.2.** Under assumption 4.1, let  $\tilde{\Theta} \subset \Theta$  be a compact set and  $\Pi \in \mathcal{M}_C$  such that  $\Pi(\cdot | \tilde{\Theta})$  admits a positive density  $\pi$ . The quantity  $k^{d\delta/2} \mathcal{J}_{D_f}^k(\Pi(\cdot | \tilde{\Theta}))$  has a positive limit when  $k \rightarrow \infty$ :

$$\lim_{k \rightarrow \infty} k^{d\delta/2} \mathcal{J}_{D_f}^k(\Pi(\cdot | \tilde{\Theta})) = a C_\delta \int_{\tilde{\Theta}} \pi(\theta)^{1+\delta} |\mathcal{I}(\theta)|^{-\delta/2} d\theta, \quad (4.28)$$

where  $C_\delta = (2\pi)^{d\delta/2}(1-\delta)^{-d/2}$ .

**Proposition 4.6.** *With the assumptions of the above theorem, if  $a(\delta+1) > 0$ , then*

$$\lim_{k \rightarrow \infty} k^{d\delta/2} (\mathcal{J}_{D_f}^k(\mathbf{J}(\cdot|\tilde{\Theta})) - \mathcal{J}_{D_f}^k(\Pi(\cdot|\tilde{\Theta}))) \geq 0, \quad (4.29)$$

where  $\mathbf{J}(\cdot|\tilde{\Theta})$  is the restricted Jeffreys prior on  $\tilde{\Theta}$ , its density is  $J(\theta) = |\mathcal{I}(\theta)|^{1/2} / \int_{\tilde{\Theta}} |\mathcal{I}(\tilde{\theta})|^{1/2} d\tilde{\theta}$ . The equality in eq. (4.29) stands if and only if  $\pi = J$  a.e.

The above proposition makes the Jeffreys prior class being the unique  $D_f$ -reference prior class over  $\mathcal{M}_{C^*}^V$ , where  $\mathcal{M}_{C^*}^V$  represents the subset of  $\mathcal{M}_C^V$  of priors admitting positive densities on  $\Theta$ . We formalize this statement in the following theorem.

**Theorem 4.3.** *Let  $b \in \mathbb{R}$ , define  $\bar{f} = f + b$  and assume  $a(\delta+1) > 0$ . Under assumption 4.1, the Jeffreys prior class is the unique  $D_{\bar{f}}$ -reference prior class over  $\mathcal{M}_{C^*}^V / \simeq$ . That means the Jeffreys prior  $\mathbf{J}$  is a  $D_{\bar{f}}$ -reference prior over  $\mathcal{M}_{C^*}^V$  and if  $\Pi$  is another reference prior over  $\mathcal{M}_{C^*}^V$ , then  $\Pi \simeq \mathbf{J}$ .*

#### 4.4.3 Results when $\delta > 0$

As in the previous section, we suppose that  $f$  takes the following form in the neighborhood of 0:

$$f(x) \underset{x \rightarrow 0}{=} ax^\delta + o(x^\delta). \quad (4.30)$$

And we require as well  $f$  to be locally bounded and to be controlled as  $x \rightarrow \infty$ :

$$f \underset{x \rightarrow \infty}{=} O(x). \quad (4.31)$$

In comparison with the preceding section, we suppose here that  $\delta \in (0, 1)$ . The elucidating of a similar result than the one in the previous section requires the satisfaction of stronger assumptions.

**Assumption 4.2.** For any compact subset  $\tilde{\Theta} \subset \Theta$ , there exists  $m > 0$  such that for all  $\theta \in \tilde{\Theta}$ ,

$$\mathbb{P}_\theta \left( \forall x \in \mathbb{R}^d, \inf_{\tilde{\theta} \in \tilde{\Theta}} -x^T \nabla_{\tilde{\theta}}^2 \log \ell(Y_1|\tilde{\theta}) x > m \|x\|^2 \right) = 1. \quad (4.32)$$

**Assumption 4.3.** For any compact subset  $\tilde{\Theta} \subset \Theta$ , the random variables  $S_k = \frac{1}{\sqrt{k}} \sum_{i=1}^k \nabla_\theta \log \ell(y_i|\theta)$  are sub-Gaussians: there exist  $\xi > \delta/2$  and  $K_1 > 0$  such that for all  $\theta \in \tilde{\Theta}$ , and all  $k$ ,  $\mathbb{E}_\theta e^{\frac{\xi}{m} \|S_k\|^2} < K_1$ .

**Theorem 4.4.** *Suppose assumptions 4.2 and 4.3. Let  $\tilde{\Theta} \subset \Theta$  be a compact set and  $\Pi \in \mathcal{M}_C^V$ . We denote  $\pi \in \mathcal{R}_{C^b}$  a density of  $\Pi(\cdot|\tilde{\Theta})$ . Then  $k^{d\delta/2} \mathcal{J}_{D_f}^k(\Pi(\cdot|\tilde{\Theta}))$  admits a finite limit when  $k \rightarrow \infty$ :*

$$\lim_{k \rightarrow \infty} k^{d\delta/2} \mathcal{J}_{D_f}^k(\Pi(\cdot|\tilde{\Theta})) = l(\pi) = a C_\delta \int_{\tilde{\Theta}} \pi(\theta)^{1+\delta} |\mathcal{I}(\theta)|^{-\delta/2} d\theta, \quad (4.33)$$

where  $C_\delta = (2\pi)^{d\delta/2}(1-\delta)^{-d/2}$ .

**Proposition 4.7.** *With the assumptions of the above theorem, if  $a(\delta+1) > 0$ , then*

$$\lim_{k \rightarrow \infty} k^{d\delta/2} (\mathcal{J}_{D_f}^k(\mathbf{J}(\cdot|\tilde{\Theta})) - \mathcal{J}_{D_f}^k(\Pi(\cdot|\tilde{\Theta}))) \geq 0, \quad (4.34)$$

where  $\mathbf{J}(\cdot|\tilde{\Theta})$  is the restricted Jeffreys prior on  $\tilde{\Theta}$ , its density is  $J(\theta) = |\mathcal{I}(\theta)|^{1/2} / \int_{\tilde{\Theta}} |\mathcal{I}(\tilde{\theta})|^{1/2} d\tilde{\theta}$ . The equality in eq. (4.34) stands if and only if  $\pi = J$  a.e.

In the same way as in the previous subsection, this result implies that the Jeffreys prior is the up-to-a-constant-unique  $D_f$ -reference prior over  $\mathcal{M}_C^V$ .

**Theorem 4.5.** Let  $b \in \mathbb{R}$ , define  $\bar{f} = f + b$  and assume  $a(\delta + 1) > 0$ . Under assumptions 4.2 and 4.3, the Jeffreys prior class is the unique  $D_{\bar{f}}$ -reference prior class over  $\mathcal{M}_C^V / \simeq$ .

## 4.5 Discussions

### 4.5.1 About the results

Regarding the optimal characteristic of the Jeffreys prior, we argue that it actually supports the relevance of our construction, in continuity with the original reference prior framework which we rely on. Indeed, we remind that Jeffreys prior is optimal for KL-mutual information (as shown in Clarke and Barron, 1994). Although its multivariate derivation is criticized, its expression remains robust and appreciated in low-dimensional cases.

For a practical implementation when the dimension is high, some authors would recommend referring to hierarchical strategies such as the one proposed by Berger, Bernardo, and Sun (2015) and evoked in chapter 3 (section 3.3.2). As these strategies are based on the sequential derivation of reference priors on sequentially conditioned models, they take the form of hierarchically derived low dimensional Jeffreys priors. Consequently, since our result supports the consideration of Jeffreys prior as a universal maximizer of the mutual information, they left that methodology unchanged.

Additionally, we draw attention to the open-ended nature of the results we express. Indeed, theorems 4.2 and 4.4 provide an analytical limit of the mutual information that takes the form of a functional  $l$  of the prior. One can notice that a reference prior is a prior that maximizes  $l$ . Thus, a maximization of our function  $l$  over a well-chosen set of priors constitutes an easy path to express reference priors that could differ from Jeffreys, as expressed in the theorem below.

**Theorem 4.6.** Let  $f \underset{x \rightarrow 0}{=} ax^\delta + o(x^\delta)$ ,  $f \underset{x \rightarrow \infty}{=} O(x)$ . Assume  $\delta \in (-1, 1) \setminus \{0\}$ ,  $a(\delta + 1) > 0$ . Assume assumption 4.1 if  $\delta < 0$  or assumptions 4.2 and 4.3 if  $\delta > 0$ . For some  $\pi \in \mathcal{R}_{C^b}$ , define

$$l(\pi) = aC_\delta \int_{\Theta} \pi(\theta)^{1+\delta} |\mathcal{I}(\theta)|^{-\delta/2} d\theta \quad (4.35)$$

with  $C_\delta = (2\pi)^{d\delta/2}(1 - \delta)^{-d/2}$ .

Let  $\mathcal{P}$  be a set of continuous priors. Denote by  $\mathcal{R} \subset \mathcal{R}_{C^b}$  the set of their associated densities. Let  $\Pi \in \mathcal{P}$  whose density is denoted by  $\pi \in \mathcal{R}$ . If there exists an openly increasing sequence  $(\Theta_i)_{i \in \mathbb{N}}$  such that  $\bigcup_{i \in \mathbb{N}} \Theta_i = \Theta$  with for any  $i$ :  $\Pi \in (0, \infty)$  and  $\pi \cdot \mathbf{1}_{\Theta_i} \propto \pi_i$  where:

$$\pi_i \in \arg \max_{p_i \in \mathcal{R}_i} l(p_i), \text{ with } \mathcal{R}_i = \left\{ p : \theta \mapsto \frac{\tilde{p}(\theta)}{\int_{\Theta_i} \tilde{p}(\tilde{\theta}) d\tilde{\theta}}, \tilde{p} \in \mathcal{R}, \int_{\Theta_i} \tilde{p} \in (0, \infty) \right\}; \quad (4.36)$$

then  $\Pi$  is a  $D_f$ -reference prior over  $\mathcal{P}$ .

One is invited to notice that when  $\Theta$  is compact, maximizing  $l$  over the normalized densities on  $\Theta$  is sufficient. The following subsection suggests an example of a constrained problem.

### 4.5.2 About the introduction of simple constraints

Here we illustrate with a simple example an application of theorem 4.6 to the derivation of constrained reference priors. Consider a Bernoulli modeling:  $Y|(T = \theta) \sim \mathcal{B}(\theta)$ ,  $\mathcal{B}(\theta)$  being the Bernoulli distribution

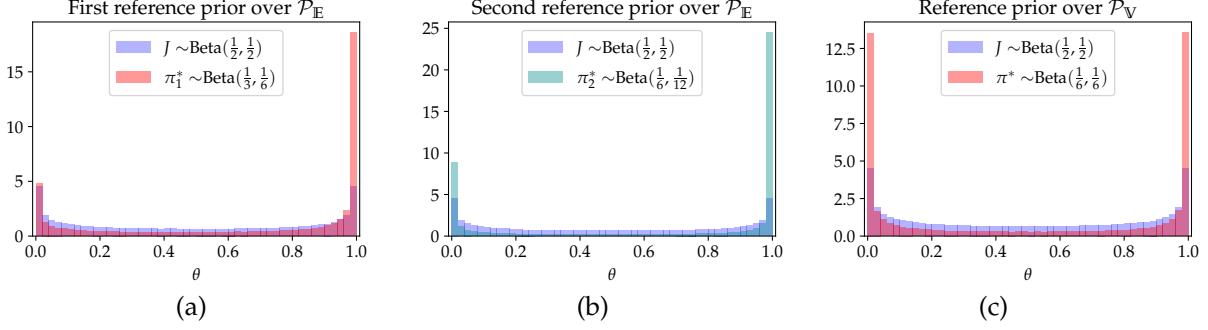


Figure 4.1: Comparison of histograms of samples from Jeffreys prior and from the  $D_f$ -reference priors over a set  $\mathcal{P}_{\mathbb{E}} = \{\pi = \text{Beta}, \mathbb{E}_{T \sim \pi} T = \frac{2}{3}\}$  (fig. (a) and (b)), and over a set  $\mathcal{P}_V = \{\pi = \text{Beta}, \mathbb{V}_{T \sim \pi} T = \frac{3}{16}\}$  (fig. (c)). In this example, there exist 2 solutions  $\pi_1^*$  (in fig. (a)) and  $\pi_2^*$  (in fig. (b)) of reference priors over  $\mathcal{P}_{\mathbb{E}}$ .  $\delta$  has been set to 1/2 in this example.

with parameter  $\theta$ . We take  $\Theta = [0, 1]$  and write  $\ell(y|\theta) = \theta^y(1-\theta)^{1-y}$ . Within this modeling, the Jeffreys prior takes the form of a Beta distribution:  $\text{Beta}(1/2, 1/2)$  (see e.g. Robert, 2007).

A possible type of constraints is the restriction of the moments of the prior. For instance, we denote  $\mathcal{P}_{\mathbb{E}}$  the class of priors  $\mathcal{P}_{\mathbb{E}} = \{\pi = \text{Beta}(\lambda_1, \lambda_2), \mathbb{E}_{T \sim \pi} [T] = 1/c\}$ . A prior in  $\mathcal{P}_{\mathbb{E}}$  can be parameterized as  $\pi_{\lambda} = \text{Beta}(\lambda, \lambda(c-1))$ ,  $\lambda \in (0, 1)$ . Then, solving  $\lambda^* = \arg \max_{\lambda} l(\pi_{\lambda})$  can be carried out numerically or analytically to issue the reference prior  $\pi_{\lambda^*}$  over  $\mathcal{P}_{\mathbb{E}}$ . The same way, some class  $\mathcal{P}_V$  could be considered this time to fix the variance of the prior.

In fig. 4.1 are plotted the  $D_f$ -reference priors over different sets of Beta priors with constrained expectations first, and with constrained variances second. We notice that, when it concerns the class  $\mathcal{P}_{\mathbb{E}}$ , two priors are reference priors. Actually, the considered class is not convex in this case, which makes the uniqueness not ensured. An additional constraint on the researched reference prior should be set to make it unique. For instance, with the idea of constructing non-informative priors, the one with maximal Shanon's entropy could be chosen. In our example,  $\pi_1^*$  (fig. 4.1.(a)) has a higher entropy than  $\pi_2^*$  (fig. 4.1.(b)).

### 4.5.3 About the assumptions and their limitations

The sub-Gaussian assumption over the likelihood, which is required for the statement of the result in the case  $\delta < 0$  remains somehow restrictive. For instance, the univariate Gaussian model with unknown variance does not satisfy this assumption. However, it is still verified by other regular statistical models, such as the univariate normal model with known variance, or continuous compact models ( $\mathcal{Y}$  compact and  $(y, \theta) \mapsto \ell(y|\theta)$  continuously differentiable). Note that due to the asymptotic normality of the r.v.  $\mathcal{I}(\theta)^{-1/2} S_k$  induced by the central limit theorem, the parameter  $\xi$  of its assumed Gaussian tail cannot exceed 1/2. This restricts  $\delta$  to be smaller than 1.

Furthermore, in opposition with the classical settings for  $f$ -divergences which tend to suppose the convexity of  $f$ , the Jeffreys prior optimally within our results is actually implied by the concavity of  $x \mapsto f(1/x)$  in the neighborhood of 0. This statement is showcased by the heuristic developed in the next subsection.

That gives some insight about the potential use of other functions  $f$  than the ones we have considered. Note that Clarke and Sun (1997) proved under some assumptions that another prior than Jeffreys prior is the  $D$ -reference prior when  $D$  is a chi-square distance. This dissimilarity measure corresponds to the case where  $f(x) = x^{-1}$ , making  $x \mapsto f(1/x)$  both convex and concave. A conclusion about the optimum of such generalized mutual information thus requires stronger regularity assumptions on the likelihood to push the asymptotic analysis further.

#### 4.5.4 About the robustness of Jeffreys prior with different divergences

When  $D_f$  is a  $\delta = 0$ -divergence

The results elucidated in the previous section explicitly address  $f$ -divergences which are  $\delta$ -divergences. It is important to recall that when  $\delta = 0$ , the  $\delta$ -divergence is usually defined as a KL divergence, more precisely it is an  $f$ -divergence with  $f = -\log$ .

This way, the  $\delta = 0$ -divergence leads to the original definition of the mutual information. In this case, results similar to those of theorems 4.3 and 4.5 exist (see chapter 3, section 3.3.2).

#### A heuristic for other $f$ -divergences

We provide a heuristic for the derivation of the limit of a  $D_f$ -mutual information when  $k \rightarrow \infty$ . It is similar to discussions conducted by Muré (2018) and Xie and Barron (1997).

The result of proposition 4.5 takes the form of asymptotic connection between  $p_{Y_k}(Y_k)/\ell_k(Y_k|\theta)$  and  $e^{\frac{1}{2}S_k^T\mathcal{I}(\theta)^{-1}S_k}$ . This connection can be anticipated from the second order Taylor expansion written as follows: for every  $y, \theta, \tilde{\theta}$ , there exists a  $\theta'$  in the segment between  $\theta$  and  $\tilde{\theta}$  such that

$$\log \left( \frac{\prod_{i=1}^k \ell(y_i|\tilde{\theta})}{\prod_{i=1}^k \ell(y_i|\theta)} \right) = (\tilde{\theta} - \theta)^T \sum_{i=1}^k \nabla_\theta \log \ell(y_i|\theta) + \frac{1}{2} (\tilde{\theta} - \theta)^T \sum_{i=1}^k \nabla_\theta^2 \log \ell(y_i|\theta') (\tilde{\theta} - \theta). \quad (4.37)$$

Indeed,  $p_{Y_k}(Y_k)/\ell_k(Y_k|\theta) = \int_{\Theta} \frac{\prod_{i=1}^k \ell(Y_i|\tilde{\theta})}{\prod_{i=1}^k \ell(Y_i|\theta)} \pi(\tilde{\theta}) d\tilde{\theta}$ . The terms  $\nabla_\theta^2 \log \ell_k(y_i|\theta)$  are definite negative, so that when we exponentialize the Taylor expansion, they vanish quickly when  $\tilde{\theta}$  is not close to  $\theta$ .

We deduce that the mass is concentrated around  $\theta$  in the integrals that define  $p_{Y_k}(Y_k)/\ell_k(Y_k|\theta)$ , so that

$$\begin{aligned} p_{Y_k}(Y_k)/\ell_k(Y_k|\theta) &\approx \pi(\theta) \int_{\Theta} \exp \left( \sqrt{k}(\tilde{\theta} - \theta)^T S_k - \frac{k}{2} (\tilde{\theta} - \theta)^T \mathcal{I}(\theta) (\tilde{\theta} - \theta) \right) d\tilde{\theta} \\ &= \pi(\theta) (2\pi)^{d/2} |k\mathcal{I}(\theta)|^{-1/2} \exp \left( \frac{1}{2} S_k^T \mathcal{I}(\theta)^{-1} S_k \right). \end{aligned} \quad (4.38)$$

To go toward an asymptotic expression of the mutual information, we must evaluate the last term with  $f$  and compute its expectation w.r.t.  $S_k$ . Conditionally to  $\theta$ , the terms  $\nabla_\theta \log \ell(Y_i|\theta)$  are i.i.d. and a central limit theorem gives that  $(S_k)_k$  converges in distribution to a Gaussian with mean  $\mathbb{E}_\theta \nabla_\theta \log \ell(Y|\theta) = 0$  and covariance matrix  $\mathcal{I}(\theta)$ .

That leads to approximating the  $D_f$ -mutual information as

$$\mathcal{I}_{D_f}^k(\Pi) \approx \int_{\Theta} \int_{\mathbb{R}^k} f \left( \pi(\theta) (2\pi)^{d/2} |k\mathcal{I}(\theta)|^{-1/2} e^{-\frac{\|x\|^2}{2}} \right) e^{-\frac{\|x\|^2}{2}} (2\pi)^{-d/2} dx \pi(\theta) d\theta. \quad (4.39)$$

We notice that if  $x \mapsto f(1/x)$  is concave, the right-hand term in the above equation is smaller than

$$\int_{\mathbb{R}^k} f \left( \left( \int_{\Theta} (2\pi)^{-d/2} |k\mathcal{I}(\theta)|^{1/2} e^{-\frac{\|x\|^2}{2}} d\theta \right)^{-1} \right) e^{-\frac{\|x\|^2}{2}} (2\pi)^{-d/2} dx, \quad (4.40)$$

which equals the right-hand term in eq. (4.39) evaluated with  $\pi(\theta) = |k\mathcal{I}(\theta)|^{1/2} / \int_{\Theta} |k\mathcal{I}(\tilde{\theta})|^{1/2} d\tilde{\theta}$ .

This simple heuristic emphasizes that the decisive assumption on  $f$  is that  $x \mapsto f(1/x)$  is concave, rather than  $f$  being convex. Regarding our developments in section 4.4, the final results on the optimality of Jeffreys stood while  $a(\delta + 1) > 0$ . In our cases, this condition implies the convexity of  $f$  in the neighborhood of 0, that is equivalent in these cases to  $x \mapsto f(1/x)$  being concave in the neighborhood of 0.

#### A simple development with a Maximum Mean Discrepancy divergence

We also investigate briefly the  $D$ -mutual information in a case where  $D$  is not an  $f$ -divergence. The dissimilarity measure we have chosen to study is a squared maximum mean discrepancy (MMD):

$$\text{MMD}^2(\mathcal{H}; P||Q) = \mathbb{E}_{X, X' \sim P \otimes P} [K(X, X')] + \mathbb{E}_{Y, Y' \sim Q \otimes Q} [K(Y, Y')] - 2\mathbb{E}_{X, Y \sim P \otimes Q} [K(X, Y)]; \quad (4.41)$$

where  $\mathcal{H}$  is a reproducing kernel Hilbert space (RKHS), and  $K$  is a reproducing kernel.

We study the simple Gaussian model with unknown mean:  $\mathcal{Y} = \mathbb{R}$ ,  $\Theta \subset \mathbb{R}$ , and  $\ell(y|\theta) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\theta)^2}{2\sigma^2}}$ .

We consider the kernels  $K_k : \mathbf{y}, \mathbf{y}' \in \mathcal{Y}^k$ ,  $K_k(\mathbf{y}, \mathbf{y}') = \frac{1}{(\sqrt{2\pi}h)^k} e^{-\frac{\|\mathbf{y}-\mathbf{y}'\|^2}{2h^2}}$ , and  $\mathcal{H}_k$  their associated RKHS. For any  $k$  it is thus possible to define the MMD between  $\mathbb{P}_{\mathbf{Y}_k}$  and  $\mathbb{P}_{\mathbf{Y}_k|\theta}$  by considering the RKHS  $\mathcal{H}_k$ . Let  $\tilde{\Theta} \subset \Theta$  be compact. For some prior  $\Pi \in \mathcal{M}_C^\nu$ , denoting  $\pi \in \mathcal{R}_{C^b}$  a density of  $\Pi(\cdot|\tilde{\Theta})$  we then have

$$\begin{aligned} \mathcal{J}_{\text{MMD}^2}^k(\Pi(\cdot|\tilde{\Theta})) &= \frac{1}{(2\pi)^{3k/2} h^k \sigma^{2k}} \int_{\tilde{\Theta}} \int_{\mathbb{R}^k \times \mathbb{R}^k} \prod_{i=1}^k e^{-\frac{(y_i-y'_i)^2}{2h^2}} e^{-\frac{(y_i-\theta)^2}{2\sigma^2}} e^{-\frac{(y'_i-\theta)^2}{2\sigma^2}} d\mathbf{y} d\mathbf{y}' \pi(\theta) d\theta \\ &\quad - \frac{1}{(2\pi)^{3k/2} h^k \sigma^{2k}} \int_{\tilde{\Theta} \times \tilde{\Theta}} \int_{\mathbb{R}^k \times \mathbb{R}^k} \prod_{i=1}^k e^{-\frac{(y_i-y'_i)^2}{2h^2}} e^{-\frac{(y_i-\theta)^2}{2\sigma^2}} e^{-\frac{(y'_i-\theta)^2}{2\sigma^2}} d\mathbf{y} d\mathbf{y}' \pi(\theta) \pi(\theta') d\theta d\theta'. \end{aligned} \quad (4.42)$$

In this very specific case, it is possible to derive and study the  $\text{MMD}^2$ -reference priors. The proposition below summarizes our results.

**Proposition 4.8.** *Let  $\tilde{\Theta} \subset \Theta$  be a compact set.*

1. *Let  $\Pi, \tilde{\Pi} \in \mathcal{M}_C^\nu$ , we denote respectively  $\pi$  and  $\tilde{\pi}$  the densities of  $\Pi(\cdot|\tilde{\Theta})$  and  $\tilde{\Pi}(\cdot|\tilde{\Theta})$ . We have:*

$$\lim_{k \rightarrow \infty} C_1^{-k+1} \sqrt{k} (\mathcal{J}_{\text{MMD}^2}^k(\Pi(\cdot|\tilde{\Theta})) - \mathcal{J}_{\text{MMD}^2}^k(\tilde{\Pi}(\cdot|\tilde{\Theta}))) = - \int_{\tilde{\Theta}} \pi(\theta)^2 d\theta + \int_{\tilde{\Theta}} \tilde{\pi}(\theta)^2 d\theta, \quad (4.43)$$

*with  $C_1 = \frac{(2\sigma^2+h^2)^{-1/2}}{\sqrt{2\pi}}$ . And the Jeffreys prior class is the unique  $\text{MMD}^2$ -reference prior class over  $\mathcal{M}_C^\nu$ .*

2. *For a density  $\pi$  we denote by  $M(\pi)$  its associated normalized prior:  $M(\pi) = B \mapsto \int_B \pi / \int_{\tilde{\Theta}} \pi$ . Let  $\pi_k^*$  be continuous densities such that for any  $k$ ,  $\mathcal{J}_{\text{MMD}^2}^k(M(\pi_k^*)) \geq \mathcal{J}_{\text{MMD}^2}^k(M(\pi))$  for any other  $\pi \in \mathcal{M}_C^\nu$ . If  $(\pi_k^*)_k$  converges uniformly to a function  $\pi^*$  on  $\tilde{\Theta}$ , then  $\pi^*$  is the Jeffreys prior density:  $\pi^* \propto 1$ .*

*Proof.* First we derive for any  $\theta, \theta' \in \tilde{\Theta}$

$$\begin{aligned} &\int_{\mathbb{R} \times \mathbb{R}} e^{-\frac{(y-y')^2}{2h^2}} e^{-\frac{(y-\theta)^2}{2\sigma^2}} e^{-\frac{(y'-\theta')^2}{2\sigma^2}} d\mathbf{y} d\mathbf{y}' \\ &= \sqrt{2\pi\tilde{\sigma}^2} \int_{\mathbb{R}} e^{-\frac{y'^2}{2h^2}} e^{-\frac{\theta'^2}{2\sigma^2}} e^{\frac{1}{2\tilde{\sigma}^2} \left( \frac{\tilde{\sigma}^2}{h^2} y' + \frac{\tilde{\sigma}^2}{\sigma^2} \theta \right)^2} e^{-\frac{(y'-\theta')^2}{2\sigma^2}} d\mathbf{y}' \quad \text{with } \tilde{\sigma}^{-2} = \frac{1}{h^2} + \frac{1}{\sigma^2} \\ &= 2\pi\tilde{\sigma}\bar{\sigma} e^{-\frac{\theta'^2}{2\sigma^2}} e^{-\frac{\theta'^2}{2\sigma^2}} e^{\frac{1}{2} \frac{\tilde{\sigma}^2}{\sigma^2} \theta^2} e^{\frac{1}{2\tilde{\sigma}^2} \left( \frac{\tilde{\sigma}^2}{\sigma^2} \theta' - \frac{\tilde{\sigma}^2\bar{\sigma}^2}{h^2\sigma^2} \theta \right)^2} \quad \text{with } \bar{\sigma}^{-2} = \frac{1}{h^2} - \frac{\tilde{\sigma}^2}{h^4} + \frac{1}{\sigma^2} = \frac{1}{\sigma^2 + h^2} + \frac{1}{\sigma^2} \\ &= 2\pi\tilde{\sigma}\bar{\sigma} e^{-\frac{1}{2} \frac{1}{2\sigma^2+h^2} (\theta'-\theta)^2}. \end{aligned} \quad (4.44)$$

That allows us to write

$$\mathcal{J}_{\text{MMD}^2}^k(\Pi(\cdot|\tilde{\Theta})) = -C_1^k \int_{\tilde{\Theta} \times \tilde{\Theta}} e^{-\frac{k}{2} \frac{1}{2\sigma^2+h^2} (\theta'-\theta)^2} \pi(\theta) \pi(\theta') d\theta d\theta' + C_1^k, \quad (4.45)$$

with  $C_1 = \frac{\tilde{\sigma}\bar{\sigma}}{\sqrt{2\pi}h\sigma^2} = \frac{(2\sigma^2+h^2)^{-1/2}}{\sqrt{2\pi}}$ .

1. To prove the first statement of the theorem, we start by computing the limit as  $k \rightarrow \infty$  of the integrals  $C_1 \sqrt{k} \int_{\tilde{\Theta}} e^{-\frac{k}{2} \frac{1}{2\sigma^2+h^2} (\theta'-\theta)^2} \pi(\theta) d\theta$  for a fixed  $\theta'$  in which  $\pi$  is continuous. The scheme of this proof is standard as this integral is the convolution of  $\pi$  with an approximate Dirac distribution. Let  $\varepsilon > 0$ , by continuity of  $\pi$  there exists a neighborhood  $V$  of  $\theta'$  such that for any  $\theta \in V$ ,  $|\pi(\theta) - \pi(\theta')| <$

$\varepsilon/2$ . Additionally, it is possible to choose a  $k_0$  such that for all  $k \geq k_0$ ,  $C_1\sqrt{k}\int_{\mathbb{R}\setminus V} e^{-\frac{k}{2}\frac{(\theta'-\theta)^2}{2\sigma^2+h^2}}d\theta < \varepsilon/(4\sup_{\tilde{\Theta}}\pi)$ . This way, for all  $k \geq k_0$ ,

$$\left|C_1\sqrt{k}\int_{\tilde{\Theta}} e^{-\frac{k}{2}\frac{1}{2\sigma^2+h^2}(\theta'-\theta)^2}\pi(\theta)d\theta - \pi(\theta')\right| \leq C_1\sqrt{k}\int_{\mathbb{R}} e^{-\frac{k}{2}\frac{1}{2\sigma^2+h^2}(\theta'-\theta)^2}|\pi(\theta)\mathbb{1}_{\theta\in\tilde{\Theta}} - \pi(\theta')|d\theta < \varepsilon. \quad (4.46)$$

That being stated, we can apply a dominated convergence theorem since for  $\nu$ -a.e.  $\theta'$  and for all  $k$ , the terms  $|C_1\sqrt{k}\int_{\tilde{\Theta}} e^{-\frac{k}{2}\frac{(\theta-\theta')^2}{2\sigma^2+h^2}}\pi(\theta)\pi(\theta')d\theta| \leq (\sup_{\tilde{\Theta}}\pi)^2$ , which is summable on  $\tilde{\Theta}$ . We obtain

$$\lim_{k\rightarrow\infty} C_1^{-k+1}\sqrt{k}(\mathcal{J}_{\text{MMD}^2}^k(\Pi(\cdot|\tilde{\Theta})) - \mathcal{J}_{\text{MMD}^2}^k(\tilde{\Pi}(\cdot|\tilde{\Theta}))) = -\int_{\tilde{\Theta}}\pi(\theta)^2d\theta + \int_{\tilde{\Theta}}\tilde{\pi}(\theta)^2d\theta. \quad (4.47)$$

We recall that the Jeffreys prior  $\mathbf{J}$  for this model is the uniform prior:  $J \propto 1$ . Since Chauchy-Schwarz inequality implies  $\int_{\tilde{\Theta}}\tilde{\pi}(\theta)d\theta \leq \sqrt{\nu(\tilde{\Theta})}(\int_{\tilde{\Theta}}\pi(\theta)^2d\theta)^{1/2}$ , with equality if and only if  $\tilde{\pi} \propto 1$ , we deduce

$$\lim_{k\rightarrow\infty} C_1^{-k+1}\sqrt{k}(\mathcal{J}_{\text{MMD}^2}^k(\mathbf{J}(\cdot|\tilde{\Theta})) - \mathcal{J}_{\text{MMD}^2}^k(\tilde{\Pi}(\cdot|\tilde{\Theta}))) \geq 0 \text{ with equality iff } \mathbf{J}(\cdot|\tilde{\Theta}) = \tilde{\Pi}(\cdot|\tilde{\Theta}). \quad (4.48)$$

2. We consider the vectorial space  $\mathcal{C}(\tilde{\Theta})$  of continuous real functions defined on  $\tilde{\Theta}$ . It is a Banach space when equipped with the norm  $\|p\| = \sup_{\tilde{\Theta}}|p|$ . The expression of  $\mathcal{J}_{\text{MMD}^2}^k(\Pi(\cdot|\tilde{\Theta}))$  elucidated in eq. (4.45) can be seen as a functional of  $\pi$  of the form  $\pi \mapsto B(\pi, \pi)$ ,  $B$  being a continuous bilinear form on  $E^2$ . This function is differentiable w.r.t.  $\pi$ . We thus apply a Lagrange multipliers theorem to write that if  $\pi_k^* \in E$  maximizes  $\pi \mapsto \mathcal{J}_{\text{MMD}^2}^k(\Pi(\cdot|\tilde{\Theta}))$  on the constrained space  $E \cap \{p \in E, \int_{\tilde{\Theta}} p = 1\}$ , then there exists  $\lambda \in \mathbb{R}$  such that for all  $p \in E$ :

$$-2C_1^k \int_{\tilde{\Theta} \times \tilde{\Theta}} e^{-\frac{k}{2}\frac{1}{2\sigma^2+h^2}(\theta-\theta')^2} p(\theta)\pi_k^*(\theta')d\theta d\theta' - \lambda \int_{\tilde{\Theta}} p(\theta)d\theta = 0. \quad (4.49)$$

This being true for any  $p \in E$ , we deduce for all  $\theta' \in \tilde{\Theta}$ ,

$$-2C_1^k \int_{\tilde{\Theta}} e^{-\frac{k}{2}\frac{(\theta-\theta')^2}{2\sigma^2+h^2}} \pi_k^*(\theta)d\theta = \lambda \quad \text{and} \quad -2C_1^k \int_{\tilde{\Theta} \times \tilde{\Theta}} e^{-\frac{k}{2}\frac{(\theta-\theta')^2}{2\sigma^2+h^2}} \pi_k^*(\theta)\pi_k^*(\theta')d\theta d\theta' = \lambda. \quad (4.50)$$

We now apply a similar reasoning as in 1. to derive the limits of the above integrals. The  $\pi_k^*$  are assumed to form a uniformly convergent sequence of functions. That ensures its limit  $\pi^*$  is continuous on  $\tilde{\Theta}$ . We fix  $\theta' \in \tilde{\Theta}$ . Let  $\varepsilon > 0$ , we fix a neighborhood  $V$  of  $\theta'$  such that for all  $\theta \in V$ ,  $|\pi^*(\theta) - \pi^*(\theta')| < \varepsilon/4$ . By uniform convergence of  $(\pi_k^*)$ , we can fix  $k_0$  such that for all  $k \geq k_0$ ,  $\|\pi_k^* - \pi^*\| < \varepsilon/4$ . It is also possible to fix  $k_1$  such that for all  $k \geq k_1$ ,  $C_1\sqrt{k}\int_{\mathbb{R}\setminus V} e^{-\frac{k}{2}\frac{(\theta-\theta')^2}{2\sigma^2+h^2}}d\theta < \varepsilon/(8\sup_{k'}\|\pi_{k'}^*\|)$ . We thus have for all  $k \geq \max(k_1, k_0)$

$$\begin{aligned} \left|C_1\sqrt{k}\int_{\tilde{\Theta}} e^{-\frac{k}{2}\frac{(\theta-\theta')^2}{2\sigma^2+h^2}}\pi_k^*(\theta)d\theta - \pi^*(\theta')\right| &\leq C_1\sqrt{k}\int_{\mathbb{R}} e^{-\frac{k}{2}\frac{(\theta-\theta')^2}{2\sigma^2+h^2}}|\pi_k^*(\theta)\mathbb{1}_{\theta\in\tilde{\Theta}} - \pi^*(\theta')|d\theta \\ &\leq C_1\sqrt{k}\int_{\tilde{\Theta}} e^{-\frac{k}{2}\frac{(\theta-\theta')^2}{2\sigma^2+h^2}}|\pi_k^*(\theta)\mathbb{1}_{\theta\in\tilde{\Theta}} - \pi^*(\theta)\mathbb{1}_{\theta\in\tilde{\Theta}}|d\theta + C_1\sqrt{k}\int_{\mathbb{R}} e^{-\frac{k}{2}\frac{(\theta-\theta')^2}{2\sigma^2+h^2}}|\pi^*(\theta)\mathbb{1}_{\theta\in\tilde{\Theta}} - \pi^*(\theta')|d\theta < \varepsilon. \end{aligned} \quad (4.51)$$

As the integrated terms are all bounded, a dominated converge theorem gives:

$$\begin{aligned} \lim_{k\rightarrow\infty} C_1\sqrt{k}\int_{\tilde{\Theta}} e^{-\frac{k}{2}\frac{(\theta-\theta')^2}{2\sigma^2+h^2}}\pi_k^*(\theta)d\theta &= \pi^*(\theta') \\ \text{and } \lim_{k\rightarrow\infty} C_1\sqrt{k}\int_{\tilde{\Theta} \times \tilde{\Theta}} e^{-\frac{k}{2}\frac{(\theta-\theta')^2}{2\sigma^2+h^2}}\pi_k^*(\theta)\pi_k^*(\theta')d\theta d\theta' &= \int_{\tilde{\Theta}}\pi^*(\theta)^2d\theta. \end{aligned} \quad (4.52)$$

Since the above terms are equal, we conclude that  $\pi^*$  is constant. □

## 4.6 Proofs of the main results

### 4.6.1 Proof of the convergence in probability

*Proof of proposition 4.5.* We fix  $\theta$  in the interior of  $\text{Supp } \pi$  such that  $\pi$  is continuous in  $\theta$ . Let  $\varepsilon, \tilde{\varepsilon} > 0$ . We want to show that there exist  $k_0$ , such that for all  $k \geq k_0$

$$\mathbb{P}_\theta \left( \tilde{g}_2^{C,c}(k^{-d/2})^{-1} \left| f \left( \frac{p_{Y_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)} \right) - g_1 \left( k^{-d/2} \pi(\theta) (2\pi)^{d/2} |\mathcal{I}(\theta)|^{-1/2} \exp \left( \frac{1}{2} S_k^T \mathcal{I}(\theta)^{-1} S_k \right) \right) \right| < \tilde{\varepsilon} \right) > 1 - \varepsilon, \quad (4.53)$$

for some  $C, c = C_\theta, c_\theta$ . We start by defining some events that will be asymptotically negligible:

1. Remind the definition of the r.v.  $S_k = \frac{1}{\sqrt{k}} \sum_{i=1}^k \nabla_\theta \log \ell(Y_i|\theta)$ . The central limit theorem ensures that when  $\mathbf{Y}_k \sim \mathbb{P}_{\mathbf{Y}_k|\theta}$ ,  $S_k$  converges in distribution to  $\mathcal{N}(0, \mathcal{I}(\theta))$  as  $k \rightarrow \infty$ . Therefore, one can choose  $M_\theta$  such that the events  $\underline{\mathbf{A}}_k(\theta)$  defined as  $\underline{\mathbf{A}}_k(\theta) = \{\|S_k\| > M_\theta\}$  verify

$$\mathbb{P}_\theta(\underline{\mathbf{A}}_k(\theta)) \xrightarrow[k \rightarrow \infty]{} \varepsilon/5 \quad (4.54)$$

2. Denote for any  $i = 1, \dots, k$ , and for any  $\mathbf{y} \in \mathcal{Y}^k$ ,

$$\kappa_i(\mathbf{y}) = \{-x^T \nabla_\theta^2 \log \ell(y_i|\theta') x, \theta' \in \tilde{\Theta}, \|x\| = 1\}. \quad (4.55)$$

They are compact subsets of  $(0, +\infty)$ . The law of large numbers gives

$$\frac{1}{k} \sum_{i=1}^k \inf \kappa_i(\mathbf{Y}_k) \xrightarrow[k \rightarrow \infty]{} \mathbb{E}_\theta \left[ \inf_{\theta' \in \tilde{\Theta}, \|x\|=1} -x^T \nabla_\theta^2 \log \ell(Y_1|\theta') x \right], \quad (4.56)$$

which is positive as the infima are positive  $\mathbb{P}_\theta$ -a.s. We choose an  $\hat{m}_\theta > 0$  smaller than this limit, and we denote

$$\underline{\mathbf{B}}_k(\theta) = \left\{ \frac{1}{k} \sum_{i=1}^k \inf \kappa_i(\mathbf{Y}_k) < \hat{m} \right\}. \quad (4.57)$$

Then  $\mathbb{P}_\theta(\underline{\mathbf{B}}_k(\theta)) \xrightarrow[k \rightarrow \infty]{} 0$ .

3. Beforehand, denote

$$C_\theta = \exp \left( M_\theta^2 / (2\hat{m}_\theta) \right) \quad \text{and} \quad c_\theta = \pi(\theta) (2\pi)^{d/2} |\mathcal{I}(\theta)|^{-1/2}, \quad (4.58)$$

which are finite and positive. Using the assumed local behavior of  $f$  around 0, one can choose  $w, \eta > 0$  such that for any  $c_\theta < |x| < C_\theta$ ,  $|l| < w$ , and  $|t| < \eta$ ,

$$|f(xt + lt) - g_1(xt)| \leq \tilde{\varepsilon} g_2(xt) \leq \tilde{\varepsilon} \tilde{g}_2^{C_\theta, c_\theta}(t). \quad (4.59)$$

Using the  $\mathbb{P}_\theta$ -a.s. convergence of  $-\frac{1}{k} \sum_{i=1}^k \nabla_\theta^2 \log \ell(Y_i|\theta)$  to  $\mathcal{I}(\theta)$ , we have  $\mathbb{P}_\theta(\underline{\mathbf{C}}_k(\theta)) \xrightarrow[k \rightarrow \infty]{} 0$  where

$$\begin{aligned} \underline{\mathbf{C}}_k(\theta) = & \left\{ \forall \|x\| \leq M, \left| \pi(\theta) (2\pi)^{d/2} |k\mathcal{I}(\theta)|^{-1/2} \exp \left( \frac{1}{2} x^T \mathcal{I}(\theta)^{-1} x \right) \right. \right. \\ & \left. \left. - \pi(\theta) (2\pi)^{d/2} |k\hat{\mathcal{I}}_k(\theta)|^{-1/2} \exp \left( \frac{1}{2} x^T \hat{\mathcal{I}}_k(\theta)^{-1} x \right) \right| \geq wk^{-d/2} \right\}, \end{aligned} \quad (4.60)$$

with  $\hat{\mathcal{I}}_k(\theta')$  denoting  $-\frac{1}{k} \sum_{i=1}^k \nabla_\theta^2 \log \ell(Y_i|\theta')$ .

4. Let us choose  $R_\theta > 0$  large enough to have

$$\left( \sup_{\tilde{\Theta}} \pi \right) \int_{\|x\| > R_\theta} e^{M_\theta x} e^{-\frac{\hat{m}_\theta}{2} x^2} dx < w, \quad (4.61)$$

and  $u > 0$  such that  $\forall |x| < uR_\theta^2$ ,  $|e^x - 1| < w/(2C_\theta R_\theta^d \sup_{\tilde{\Theta}} \pi)$ .

From assumption 3.1, the function  $\theta' \mapsto \|\nabla_\theta^2 \log \ell(y|\theta') - \nabla_\theta^2 \log \ell(y|\theta)\|$  is dominated on the Euclidean ball  $B(\theta, \tau_1)$  by a variable that has a finite expectation  $\mathbb{E}_\theta$ . Therefore, a dominated convergence theorem gives that the function

$$\theta' \mapsto \mathbb{E}_\theta \left[ \|\nabla_\theta^2 \log \ell(Y_1|\theta') - \nabla_\theta^2 \log \ell(Y_1|\theta)\| \right] \quad (4.62)$$

converges to 0 when  $\theta' \rightarrow \theta$  as a consequence of the continuity of the second order partial derivatives of the log-likelihood.

Thus, if we choose  $\tilde{\tau}_\theta$  such that the above quantity is smaller than  $u\varepsilon$  and if we define  $\underline{D}_k(\theta)$  as the event

$$\underline{D}_k(\theta) = \left\{ \sup_{\theta' \in \tilde{\Theta}, \|\theta' - \theta\| < \tilde{\tau}_\theta} \|\hat{\mathcal{I}}_k(\theta') - \hat{\mathcal{I}}_k(\theta)\| > u \right\}, \quad (4.63)$$

one gets by applying Markov's inequality that  $\mathbb{P}_\theta(\underline{D}_k(\theta)) < \varepsilon/4$ .

Now, notice that for any  $\tilde{\theta} \in \tilde{\Theta}$ ,  $\mathbf{y} \in \mathcal{Y}^k$ , there exists  $\hat{\theta}$  on the segment between  $\theta$  and  $\tilde{\theta}$  such that

$$\log \left( \frac{\prod_{i=1}^k \ell(y_i|\tilde{\theta})}{\prod_{i=1}^k \ell(y_i|\theta)} \right) = (\tilde{\theta} - \theta)^T \sum_{i=1}^k \nabla_\theta \log \ell(y_i|\theta) + \frac{1}{2} (\tilde{\theta} - \theta)^T \sum_{i=1}^k \nabla_\theta^2 \log \ell(y_i|\hat{\theta})(\tilde{\theta} - \theta). \quad (4.64)$$

Let us write  $p_{\mathbf{Y}_k}(\mathbf{y})/\ell_k(\mathbf{y}|\theta) = \int_{\tilde{\Theta}} h_{\mathbf{y}}(\tilde{\theta}, \theta) \pi(\tilde{\theta}) d\tilde{\theta}$  where

$$h_{\mathbf{y}}(\tilde{\theta}, \theta) = \exp \left( (\tilde{\theta} - \theta)^T \nabla_\theta \log \ell_k(\mathbf{y}|\theta) \right) \exp \left( \frac{1}{2} (\tilde{\theta} - \theta)^T \nabla_\theta^2 \log \ell_k(\mathbf{y}|\hat{\theta})(\tilde{\theta} - \theta) \right). \quad (4.65)$$

We fix  $k$  and work under the event  $\underline{\mathbf{A}}_k(\theta)^c \cap \underline{\mathbf{B}}_k(\theta)^c \cap \underline{\mathbf{C}}_k(\theta)^c \cap \underline{\mathbf{D}}_k(\theta)^c$ . i.e. we consider  $\omega \in \underline{\mathbf{A}}_k(\theta)^c \cap \underline{\mathbf{B}}_k(\theta)^c \cap \underline{\mathbf{C}}_k(\theta)^c \cap \underline{\mathbf{D}}_k(\theta)^c$  and denote  $\mathbf{y} = \mathbf{Y}_k(\omega)$ ,  $S_k^{\mathbf{y}} = \frac{1}{\sqrt{k}} \nabla_\theta \log \ell_k(\mathbf{y}|\theta)$ ,  $\hat{\mathcal{I}}_k^{\mathbf{y}}(\theta) = -\frac{1}{k} \nabla_\theta^2 \log \ell_k(\mathbf{y}|\theta)$ . We can thus bound  $h_{\mathbf{y}}(\tilde{\theta}, \theta)$  as

$$h_{\mathbf{y}}(\tilde{\theta}, \theta) \leq \exp \left( M_\theta \sqrt{k} \|\tilde{\theta} - \theta\| \right) \exp \left( -\frac{k}{2} \hat{m}_\theta \|\tilde{\theta} - \theta\|^2 \right) \leq C_\theta. \quad (4.66)$$

The rest of the proof only consists in bounding four quantities by  $wk^{-d/2}$ .

- Define  $\tilde{R}_k = R_\theta k^{-1/2}$ . Inequality in eq. (4.61) implies

$$\mathbf{I}_1 := \int_{\|\tilde{\theta} - \theta\| > \tilde{R}_k} \exp \left( (\tilde{\theta} - \theta)^T \sqrt{k} S_k^{\mathbf{y}} \right) \exp \left( -\frac{k}{2} (\tilde{\theta} - \theta)^T \hat{\mathcal{I}}_k^{\mathbf{y}}(\theta)(\tilde{\theta} - \theta) \right) \pi(\theta) d\tilde{\theta} < wk^{-d/2}. \quad (4.67)$$

- Recalling the upper bound in eq. (4.66), we have the same way

$$\mathbf{I}_2 := \int_{\|\tilde{\theta} - \theta\| > \tilde{R}_k} h_{\mathbf{y}}(\tilde{\theta}, \theta) \pi(\tilde{\theta}) d\tilde{\theta} < wk^{-d/2}. \quad (4.68)$$

- As  $\pi$  is continuous in  $\theta$ , there exist  $k_1$  such that if  $k \geq k_1$  then for all  $\|\tilde{\theta} - \theta\| < \tilde{R}_k$ ,  $|\pi(\tilde{\theta}) - \pi(\theta)| \leq w/(2C_\theta R_\theta^d)$ . We write:

$$\begin{aligned} \mathbf{I}_3 &:= \int_{\|\tilde{\theta} - \theta\| < \tilde{R}_k} h_{\mathbf{y}}(\tilde{\theta}, \theta) \pi(\tilde{\theta}) d\tilde{\theta} - \int_{\|\tilde{\theta} - \theta\| < \tilde{R}_k} \exp \left( (\tilde{\theta} - \theta)^T \sqrt{k} S_k^{\mathbf{y}} \right) \exp \left( -\frac{k}{2} (\tilde{\theta} - \theta)^T \hat{\mathcal{I}}_k^{\mathbf{y}}(\theta)(\tilde{\theta} - \theta) \right) \pi(\theta) d\tilde{\theta} \\ |\mathbf{I}_3| &\leq \int_{\|\tilde{\theta} - \theta\| < \tilde{R}_k} |\hat{h}_{\mathbf{y}}(\tilde{\theta}, \theta, \hat{\theta}) - \hat{h}_{\mathbf{y}}(\tilde{\theta}, \theta, \theta)| \pi(\tilde{\theta}) d\tilde{\theta} + \int_{\|\tilde{\theta} - \theta\| < \tilde{R}_k} \hat{h}_{\mathbf{y}}(\tilde{\theta}, \theta, \theta) |\pi(\tilde{\theta}) - \pi(\theta)| d\tilde{\theta} \end{aligned} \quad (4.69)$$

where  $\hat{h}_y(\tilde{\theta}, \theta, \hat{\theta}) = \exp((\tilde{\theta} - \theta)^T \sqrt{k} S_k^y) \exp\left(-\frac{k}{2}(\tilde{\theta} - \theta)^T \hat{\mathcal{I}}_k^y(\hat{\theta})(\tilde{\theta} - \theta)\right)$ . Notice that for any  $\theta', \tilde{\theta}$  in the Euclidean ball  $B(\theta, \tilde{R}_k)$ ,

$$|\hat{h}_y(\tilde{\theta}, \theta, \theta') - \hat{h}_y(\tilde{\theta}, \theta, \theta)| \leq \hat{h}_y(\tilde{\theta}, \theta, \theta') \left| \exp\left(-\frac{k}{2}(\tilde{\theta} - \theta)^T (\hat{\mathcal{I}}_k^y(\theta') - \hat{\mathcal{I}}_k^y(\theta))(\tilde{\theta} - \theta)\right) - 1 \right|. \quad (4.70)$$

Thus, if  $k \geq k_2$  where  $k_2$  is such that  $\tilde{R}_{k_2} < \tilde{\tau}$ , we obtain  $(\tilde{\theta} - \theta)^T (\hat{\mathcal{I}}_k^y(\theta') - \hat{\mathcal{I}}_k^y(\theta))(\tilde{\theta} - \theta) < u$  consequently to  $\mathbf{y} = \mathbf{Y}_k(\omega)$  with  $\omega \in \underline{\mathbf{D}}_k(\theta)^c$ . It then comes  $|\mathbf{I}_3| < wk^{-d/2}$ .

- By recognizing a Gaussian density in the integral below, we get the equality

$$\begin{aligned} & \int_{\mathbb{R}^d} \exp((\tilde{\theta} - \theta)^T \sqrt{k} S_k^y) \exp\left(\frac{1}{2}(\tilde{\theta} - \theta)^T \nabla_\theta^2 \ell_k(\mathbf{y}|\theta)(\tilde{\theta} - \theta)\right) \pi(\theta) d\tilde{\theta} \\ &= \pi(\theta) (2\pi)^{d/2} |k \hat{\mathcal{I}}_k^y(\theta)|^{-1/2} \exp\left(\frac{1}{2} S_k^{yT} \hat{\mathcal{I}}_k^y(\theta)^{-1} S_k^y\right). \end{aligned} \quad (4.71)$$

Defining  $\mathbf{I}_4 = (4.71) - \pi(\theta) (2\pi)^{d/2} |\mathcal{I}(\theta)|^{-1/2} \exp\left(\frac{1}{2} S_k^{yT} \mathcal{I}(\theta)^{-1} S_k^y\right)$ , we obtain, using that  $\omega \in \underline{\mathbf{C}}_k(\theta)^c$ ,

$$|\mathbf{I}_4| \leq wk^{-d/2}. \quad (4.72)$$

If we define  $k_3$  such that  $k^{-d/2} < \eta/4$ , and if  $k \geq \max(k_1, k_2, k_3)$ , then the four inequalities elucidated above allows evaluating eq. (4.59) with  $t = k^{-d/2}$ ,

$$\begin{aligned} x(\mathbf{y}) &= \int_{\tilde{\Theta}} \exp\left((\tilde{\theta} - \theta)^T \sqrt{k} S_k^y\right) \exp\left(-\frac{k}{2}(\tilde{\theta} - \theta)^T \hat{\mathcal{I}}_k^y(\theta)(\tilde{\theta} - \theta)\right) \pi(\theta) d\tilde{\theta}, \\ \text{and } l(\mathbf{y}) &= x(\mathbf{y}) - \frac{p_y(\mathbf{y})}{\ell_k(\mathbf{y}|\theta)} k^{d/2} = \mathbf{I}_1 + \mathbf{I}_2 + \mathbf{I}_3 + \mathbf{I}_4. \end{aligned} \quad (4.73)$$

All in all, we have proven that for  $k \geq \max(k_1, k_2, k_3)$  the inclusion

$$\underline{\mathbf{A}}_k(\theta)^c \cap \underline{\mathbf{B}}_k(\theta)^c \cap \underline{\mathbf{C}}_k(\theta)^c \cap \underline{\mathbf{D}}_k(\theta)^c \subset \left\{ \left| f\left(\frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)}\right) - g_1(k^{-d/2} x(\mathbf{Y}_k)) \right| < \tilde{g}_2^{C_\theta, c_\theta}(k^{-d/2}) \tilde{\varepsilon} \right\} \quad (4.74)$$

stands. To conclude we denote  $k_4$  such that for any  $k \geq k_4$ ,

$$\mathbb{P}_\theta(\underline{\mathbf{A}}_k(\theta)) < \varepsilon/4, \quad \mathbb{P}_\theta(\underline{\mathbf{B}}_k(\theta)) < \varepsilon/4, \quad \mathbb{P}_\theta(\underline{\mathbf{C}}_k(\theta)) \varepsilon/4, \quad \mathbb{P}_\theta(\underline{\mathbf{D}}_k(\theta)) < \varepsilon/4. \quad (4.75)$$

This way, defining  $k_0 = \max(k_1, k_2, k_3, k_4)$ , the following inequality is true for any  $k \geq k_0$ :

$$\mathbb{P}_\theta\left(\tilde{g}_2^{C_\theta, c_\theta}(k^{-d/2})^{-1} \left| f\left(\frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)}\right) - g_1\left(k^{-d/2} \pi(\theta) (2\pi)^{d/2} |\mathcal{I}(\theta)|^{-1/2} \exp\left(\frac{1}{2} S_k^{yT} \mathcal{I}(\theta)^{-1} S_k^y\right)\right)\right| < \tilde{\varepsilon}\right) > 1 - \varepsilon. \quad (4.76)$$

## 4.6.2 Proof of the results when $\delta < 0$

*Proof of theorems 4.2 and 4.3 and proposition 4.6.* We suppose that assumption 4.1 is verified. We recall that in this proof  $\delta < 0$ . The following lemma is required for the proof. It is proven later on.

**Lemma 4.1.** Let  $\zeta < 0$ . There exists a constant  $K$  independent of  $\theta$  such that

$$k^{d\zeta/2} \mathbb{E}_\theta \left[ \left( \frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)} \right)^\zeta \right] \leq K. \quad (4.77)$$

We recall that proposition 4.5 states the following convergence in probability for a.e.  $\theta$ :

$$k^{d\delta/2} \left| f \left( \frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)} \right) - ak^{-d\delta/2} \pi(\theta)^\delta (2\pi)^{d\delta/2} |\mathcal{I}(\theta)|^{-\delta/2} \exp \left( \frac{\delta}{2} S_k^T \mathcal{I}(\theta)^{-1} S_k \right) \right| \xrightarrow[k \rightarrow \infty]{\mathbb{P}_\theta} 0. \quad (4.78)$$

The first goal in this prove will be to show that this convergence is a convergence in  $L^1(\mathbb{P}_\theta)$ . This is shown by proving that the sequence of r.v. is uniformly integrable. Let  $\rho > 0$  and choose an even  $\underline{\mathbf{A}}$  such that  $\mathbb{P}_\theta(\underline{\mathbf{A}}) = \rho$ . Firstly, we have

$$\mathbb{E}_\theta \left[ \mathbb{1}_{\underline{\mathbf{A}}} \pi(\theta)^\delta (2\pi)^{d\delta/2} e^{\frac{\delta}{2} S_k^T \mathcal{I}(\theta)^{-1} S_k} \right] \leq \pi(\theta)^\delta (2\pi)^{d\delta/2} \rho, \quad (4.79)$$

and secondly, considering  $t, \hat{C}$  and  $\hat{C}'$  such that  $|f(x)| \leq \hat{C}x$  if  $x > t$  and  $|f(x)| \leq \hat{C}'x^\delta$  otherwise; we obtain using  $\underline{\mathbf{G}}_k = \underline{\mathbf{G}}_k(\theta)\{p_{\mathbf{Y}}(\mathbf{y}) > \gamma \ell_k(\mathbf{y}|\theta)\}$  and Hölder's inequality:

$$\begin{aligned} \mathbb{E}_\theta \left[ \mathbb{1}_{\underline{\mathbf{A}}} k^{d\delta/2} f \left( \frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)} \right) \right] &\leq \mathbb{E}_\theta \left[ \mathbb{1}_{\underline{\mathbf{G}}_k} \mathbb{1}_{\underline{\mathbf{A}}} k^{d\delta/2} \hat{C} \frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)} \right] + \mathbb{E}_\theta \left[ \mathbb{1}_{\underline{\mathbf{G}}_k^c} \mathbb{1}_{\underline{\mathbf{A}}} k^{d\delta/2} \hat{C}' \left( \frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)} \right)^\delta \right] \\ &\leq \hat{C}\rho^{1/2} + \hat{C}'\rho^{1/p} K^{1/q}, \end{aligned} \quad (4.80)$$

with  $1/p + 1/q = 1$ . The above ensures the required uniform-integrability. That allows us to write:

$$\mathbb{E}_\theta k^{d\delta/2} \left| f \left( \frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)} \right) - ak^{-d\delta/2} \pi(\theta)^\delta (2\pi)^{d\delta/2} |\mathcal{I}(\theta)|^{-\delta/2} \exp \left( \frac{\delta}{2} S_k^T \mathcal{I}(\theta)^{-1} S_k \right) \right| \xrightarrow[k \rightarrow \infty]{} 0. \quad (4.81)$$

This limit stands for every  $\theta$  in the interior of  $\text{Supp } \pi$  within which  $\pi$  is continuous. For such a  $\theta$  it comes as  $k \rightarrow \infty$

$$k^{d\delta/2} D_f(\mathbb{P}_{\mathbf{Y}_k} || \mathbb{P}_{\mathbf{Y}_k|\theta}) = a\pi(\theta)^\delta (2\pi)^{d\delta/2} |\mathcal{I}(\theta)|^{-\delta/2} \mathbb{E}_\theta \left[ e^{\frac{\delta}{2} \|\mathcal{I}(\theta)^{-1/2} S_k\|^2} \right] + o(1). \quad (4.82)$$

The sequence of r.v.  $\mathcal{I}(\theta)^{-1/2} S_k$  converges in distribution to a standard normal r.v. denoted  $X$ . Thus, since  $\delta < 0$ ,  $\mathbb{E}_\theta \left[ e^{\frac{\delta}{2} \|\mathcal{I}(\theta)^{-1/2} S_k\|^2} \right] \xrightarrow[k \rightarrow \infty]{} \mathbb{E}_\theta \left[ e^{\frac{\delta}{2} \|X\|^2} \right]$ . Therefore,

$$k^{d\delta/2} D_f(\mathbb{P}_{\mathbf{Y}_k} || \mathbb{P}_{\mathbf{Y}_k|\theta}) \xrightarrow[k \rightarrow \infty]{} a\pi(\theta)^\delta (2\pi)^{d\delta/2} |\mathcal{I}(\theta)|^{-\delta/2} \mathbb{E}_\theta \left[ e^{\frac{\delta}{2} \|X\|^2} \right] + o(1). \quad (4.83)$$

The second step of the proof it to integrate this limit w.r.t.  $\pi$ , using a dominated convergence theorem. It is possible since the following domination stands:

$$\begin{aligned} k^{d\delta/2} D_f(\mathbb{P}_{\mathbf{Y}_k} || \mathbb{P}_{\mathbf{Y}_k|\theta}) &= k^{d\delta/2} \mathbb{E}_\theta \left[ \mathbb{1}_{\underline{\mathbf{G}}_k} f \left( \frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)} \right) \right] + k^{d\delta/2} \mathbb{E}_\theta \left[ \mathbb{1}_{\underline{\mathbf{G}}_k^c} f \left( \frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)} \right) \right] \\ &\leq \hat{C} + \hat{C}'K. \end{aligned} \quad (4.84)$$

The domination above makes the following licit:

$$\lim_{k \rightarrow \infty} k^{d\delta/2} \mathcal{J}_{D_f}^k(\Pi(\cdot | \tilde{\Theta})) = a(2\pi)^{d\delta/2} (1-\delta)^{-d/2} \int_{\tilde{\Theta}} \pi(\delta)^{\delta+1} |\mathcal{I}(\theta)|^{-\delta/2} d\theta. \quad (4.85)$$

This concludes the proof of theorem 4.2.

The results of proposition 4.6 and theorem 4.3 comes by concavity of  $x \mapsto ax^{-\delta}$ , which results from  $a(\delta+1) > 0$  and gives

$$\begin{aligned} a \int_{\tilde{\Theta}} \pi(\theta)^\delta |\mathcal{I}(\theta)|^{-\delta/2} \pi(\theta) d\theta \\ \leq a \left( \int_{\tilde{\Theta}} \pi(\theta)^{-1} |\mathcal{I}(\theta)|^{1/2} \pi(\theta) d\theta \right)^{-\delta} = a \left( \int_{\tilde{\Theta}} |\mathcal{I}(\theta)|^{1/2} d\theta \right)^{-\delta}. \end{aligned} \quad (4.86)$$

The concavity of  $x \mapsto ax^{-\delta}$  being strict, the above inequality is an equality if and only if  $\pi(\theta)^\delta |\mathcal{I}(\theta)|^{-\delta/2}$  is constant with respect to  $\theta$ , i.e.  $\pi \propto J$ .

□

*Proof of lemma 4.1.* Set  $\theta$  that belongs in the interior of  $\text{Supp } \pi$ . We write,

$$\mathbb{E}_\theta \left[ \left( \frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)} \right)^\zeta \right] \leq \mathbb{E}_\theta \left[ \left( \int_{\|\tilde{\theta}-\theta\| \leq \tilde{R}_k} \exp((\tilde{\theta}-\theta)^T \sqrt{k} S_k) \exp \left( -\frac{k}{2} (\tilde{\theta}-\theta)^T \hat{\mathcal{I}}_k(\hat{\theta})(\tilde{\theta}-\theta) \right) \pi(\tilde{\theta}) d\tilde{\theta} \right)^\zeta \right], \quad (4.87)$$

where  $\hat{\theta}$  belongs to the segment between  $\theta$  and  $\tilde{\theta}$ , and with  $\hat{\mathcal{I}}_k(\theta')$  denoting  $-\frac{1}{k} \sum_{i=1}^k \nabla_{\theta'}^2 \log \ell(y_i|\theta')$ . Denote  $H_k = \sup_{\tilde{\theta} \in \tilde{\Theta}, \|x\|=1} x^T \hat{\mathcal{I}}_k(\tilde{\theta}) x$ , and  $m_\pi = \inf_{\tilde{\Theta}} \pi > 0$ . We have

$$\begin{aligned} \mathbb{E}_\theta \left[ \left( \frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)} \right)^\zeta \right] &\leq \mathbb{E}_\theta \left[ m_\pi^\zeta \left( k^{-d/2} \int_{\mathbb{R}^d} e^{x^T S_k} e^{-\frac{H_k}{2} \|x\|^2} dx \right)^\zeta \right] \\ &= \mathbb{E}_\theta \left[ m_\pi^\zeta k^{-d\zeta/2} (2\pi)^{d\zeta/2} H_k^{-\zeta/2} e^{\frac{\zeta}{2} H_k \|S_k\|^2} \right]. \end{aligned} \quad (4.88)$$

We now bound  $H_k$  as below:

$$H_k = \sup_{\tilde{\Theta} \in \tilde{\Theta}, \|x\|=1} -\frac{1}{k} \sum_{i=1}^k x^T \nabla_{\theta}^2 \log \ell(y_i|\theta) x \leq \frac{1}{k} \sum_{i=1}^k \sup_{\tilde{\Theta}} \|\nabla_{\theta}^2 \ell(y_i|\theta)\|. \quad (4.89)$$

That allows to conclude using Jensen's inequality:

$$k^{d\zeta/2} \mathbb{E}_\theta \left[ \left( \frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)} \right)^\zeta \right] \leq m_\pi^\zeta (2\pi)^{d\zeta/2} K_0^{-\zeta/2}. \quad (4.90)$$

□

### 4.6.3 Proof of the results when $\delta > 0$

*Proof of theorems 4.4 and 4.5 and proposition 4.7.* We suppose that assumptions 4.2 and 4.3 are verified, we recall that in this proof  $\delta > 0$ . The following lemmas are required for the proof. They are proven later on.

**Lemma 4.2.** For any  $\zeta < 2\xi$ , there exists a constant  $K_2 > 0$  independent of  $\theta \in \tilde{\Theta}$  such that for any  $k > 0$ ,

$$k^{d\zeta/2} \mathbb{E}_\theta \left[ \left( \frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)} \right)^\zeta \right] \leq K_2. \quad (4.91)$$

**Lemma 4.3.** There exists a constant  $K_3 > 0$  independent of  $\theta \in \tilde{\Theta}$  such that for any  $k > 0$  and any  $\rho > 0$ ,

$$k^{d\delta/2} \mathbb{E}_\theta \left[ \mathbb{1}_{\underline{\mathbf{E}}_k} \frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)} \right] \leq \rho^{1/2} K_3 \quad (4.92)$$

with  $\underline{\mathbf{E}}_k = \underline{\mathbf{E}}_k(\theta) = \{ \frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)} > \rho^{-1/\delta} \}$ .

We start by proving the following claims about the consequences of the sub-Gaussianity of  $S_k$  (assumption 4.3).

- **Claim 1.** Let  $\zeta < \xi$ , then if  $X_k = \mathcal{I}(\theta)^{-1/2} S_k$ , then  $\mathbb{E}_\theta e^{\zeta \|X_k\|^2} < K_1$ .

*Proof.* Assumption 4.2 makes  $\|\mathcal{I}(\theta)^{-1/2}\| \leq m^{-1/2}$ , i.e.  $\|X_k\|^2 \leq \|S_k\|^2/m$ . Hence the result since  $\mathbb{E}_\theta e^{\frac{\zeta}{m} \|S_k\|^2} < K_1$ .

- **Claim 2.** Let  $t \in \mathbb{R}$  and  $v \in \mathbb{R}^d$  with  $\|v\| = 1$ . We have  $\mathbb{E}_\theta e^{tv^T S_k} \leq K_1 e^{\frac{t^2 \chi^2}{2}}$  with  $\chi^2 = \xi/(2m)$ .

*Proof.* For any  $s > 0$ , we have  $|tv^T S_k| \leq \frac{s}{2}t^2 + \frac{1}{2s}|v^T S_k|^2 \leq \frac{s}{2}t^2 + \frac{1}{2s}\|S_k\|^2$ . Choosing  $s = m/(2\xi)$ , we get

$$\mathbb{E}_\theta e^{tv^T S_k} \leq e^{\frac{mt^2}{4\xi}} \mathbb{E}_\theta e^{\frac{\xi}{m}\|S_k\|^2} \leq K_1 e^{\frac{t^2 \chi^2}{2}}. \quad (4.93)$$

We recall that proposition 4.5 states the following convergence in probability for a.e.  $\theta$ :

$$k^{d\delta/2} \left| f\left(\frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)}\right) - ak^{-d\delta/2} \pi(\theta)^\delta (2\pi)^{d\delta/2} |\mathcal{I}(\theta)|^{-\delta/2} \exp\left(\frac{\delta}{2} S_k^T \mathcal{I}(\theta)^{-1} S_k\right) \right| \xrightarrow[k \rightarrow \infty]{\mathbb{P}_\theta} 0. \quad (4.94)$$

The first goal in this prove will be to show that this convergence is a convergence in  $L^1(\mathbb{P}_\theta)$ . This is shown by proving that the sequence of r.v. is uniformly integrable. Let  $\rho > 0$  and choose an even  $\underline{\mathbf{A}}$  such that  $\mathbb{P}_\theta(\underline{\mathbf{A}}) = \rho$ . Denoting  $X_k = \mathcal{I}(\theta)^{-1/2} S_k$ , using Hölder's inequality it comes

$$\begin{aligned} \mathbb{E}_\theta \left[ \mathbb{1}_{\underline{\mathbf{A}}} \pi(\theta)^\delta (2\pi)^{d\delta/2} e^{\frac{\delta}{2} S_k^T \mathcal{I}(\theta)^{-1} S_k} \right] &\leq \pi(\theta)^\delta (2\pi)^{d\delta/2} \rho^{1/p} \mathbb{E}_\theta \left[ e^{\frac{\delta}{2} q \|X_k\|^2} \right]^{1/q} \\ &\leq \rho^{1/p} (\sup_\Theta \pi)^\delta (2\pi)^{d\delta/2} K_1^{1/q} \end{aligned} \quad (4.95)$$

for any  $p, q$  such that  $1/p + 1/q = 1$  and  $q\delta/2 < \xi$ . Second, consider  $\hat{C}$  and  $\hat{C}'$  such that  $|f(x)| \leq \hat{C}x$  if  $x > \rho^{-1/\delta}$  and  $|f(x)| \leq \hat{C}'x^\delta$  otherwise. Using  $\underline{\mathbb{E}}_k = \underline{\mathbb{E}}_k(\theta) = \{p_{\mathbf{Y}_k}(\mathbf{Y}_k)/\ell_k(\mathbf{Y}_k|\theta) > \rho^{-1/\delta}\}$  and Hölder's inequality, we compute:

$$\begin{aligned} \mathbb{E}_\theta \left[ \mathbb{1}_{\underline{\mathbf{A}}} k^{d\delta/2} f\left(\frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)}\right) \right] &\leq \mathbb{E}_\theta \left[ \mathbb{1}_{\underline{\mathbb{E}}_k} \mathbb{1}_{\underline{\mathbf{A}}} k^{d\delta/2} \hat{C} \frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)} \right] + \mathbb{E}_\theta \left[ \mathbb{1}_{\underline{\mathbb{E}}_k^c} \mathbb{1}_{\underline{\mathbf{A}}} k^{d\delta/2} \hat{C}' \left(\frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)}\right)^\delta \right] \\ &\leq \hat{C} \rho^{1/2} K_3 + \hat{C}' \rho^{1/p} K_2^{1/q}, \end{aligned} \quad (4.96)$$

where  $1/p = 1 - 1/q$  with  $q\delta/2 < \xi$ . Equations (4.95) and (4.96) ensure the required uniform-integrability. That allows us to write:

$$\mathbb{E}_\theta k^{d\delta/2} \left| f\left(\frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)}\right) - ak^{-d\delta/2} \pi(\theta)^\delta (2\pi)^{d\delta/2} |\mathcal{I}(\theta)|^{-\delta/2} \exp\left(\frac{\delta}{2} S_k^T \mathcal{I}(\theta)^{-1} S_k\right) \right| \xrightarrow[k \rightarrow \infty]{} 0. \quad (4.97)$$

This limit stands for every  $\theta$  in the interior of  $\text{Supp } \pi$  within which  $\pi$  is continuous. For such a  $\theta$  it comes as  $k \rightarrow \infty$

$$k^{d\delta/2} D_f(\mathbb{P}_{\mathbf{Y}_k} || \mathbb{P}_{\mathbf{Y}_k|\theta}) = a \pi(\theta)^\delta (2\pi)^{d\delta/2} |\mathcal{I}(\theta)|^{-\delta/2} \mathbb{E}_\theta \left[ e^{\frac{\delta}{2} \|X_k\|^2} \right] + o(1). \quad (4.98)$$

The second step of the proof is to precise the above asymptotic expansion by studying the limit of  $\mathbb{E}_\theta e^{\frac{\delta}{2} \|X_k\|^2}$ . The sequence of r.v.  $X_k$  converges in distribution to a standard normal r.v. denoted  $X$ . This implies that for any  $A > 0$ , the expectation  $\mathbb{E}_\theta \left[ e^{\frac{\delta}{2} \|X_k\|^2} \mathbb{1}_{\|X_k\| \leq A} \right]$  converges to  $\mathbb{E}_\theta \left[ e^{\frac{\delta}{2} \|X\|^2} \mathbb{1}_{\|X\| \leq A} \right]$ . Moreover, under Assumption 4.3 there exists a constant  $K'$  independent of  $k$  such that for any  $k$

$$\begin{aligned} \mathbb{E}_\theta \left[ e^{\frac{\delta}{2} \|X_k\|^2} \left| 1 - \mathbb{1}_{\|X_k\| \leq A} \right| \right] &\leq \mathbb{E}_\theta \left[ e^{\frac{\delta q}{2} \|X_k\|^2} \right]^{1/q} \mathbb{P}_\theta(\|X_k\| \geq A)^{1/p} \\ &\leq K_1^{1/q} \mathbb{P}_\theta(\exp(t^2 \|X_k\|^2) \geq \exp(t^2 A)) \\ &\leq K' e^{-t^2 A^2/p} \end{aligned} \quad (4.99)$$

using Hölder inequality with  $1/p + 1/q = 1$  such as  $\delta q/2 \leq \xi$  first, and Markov inequality with  $t^2 < \xi$  second. Thus, for any  $\varepsilon > 0$ , one can choose  $A$  such that (4.99)  $\leq \varepsilon/3$  and  $\mathbb{E}_\theta \left[ e^{\frac{\delta}{2} \|X\|^2} \left| 1 - \mathbb{1}_{\|X\| \leq A} \right| \right] < \varepsilon/3$  as well to get

$$\left| \mathbb{E}_\theta \left[ e^{\frac{\delta}{2} \|X_k\|^2} \right] - \mathbb{E}_\theta \left[ e^{\frac{\delta}{2} \|X\|^2} \right] \right| < 2\varepsilon/3 + \left| \mathbb{E}_\theta \left[ e^{\frac{\delta}{2} \|X_k\|^2} \mathbb{1}_{\|X_k\| \leq A} \right] - \mathbb{E}_\theta \left[ e^{\frac{\delta}{2} \|X\|^2} \mathbb{1}_{\|X\| \leq A} \right] \right|. \quad (4.100)$$

Eventually,  $k$  can be chosen large enough such that the right-hand side of above equation is smaller than  $\varepsilon$ , which states the convergence  $\mathbb{E}_\theta e^{\frac{\delta}{2}\|X_k\|^2} \xrightarrow[k \rightarrow \infty]{} \mathbb{E}_\theta e^{\frac{\delta}{2}\|X\|^2}$ . This lets us write

$$k^{d\delta/2} D_f(\mathbb{P}_{Y_k} || \mathbb{P}_{Y_k|\theta}) \xrightarrow{k \rightarrow \infty} a\pi(\theta)^\delta (2\pi)^{d\delta/2} |\mathcal{I}(\theta)|^{-\delta/2} (1-\delta)^{-d/2} + o(1). \quad (4.101)$$

The third and last step of the proof it to carry out a dominated convergence theorem to elucidate the limit of  $\mathcal{J}_{D_f}^k(\Pi(\cdot|\tilde{\Theta}))$ . We denote  $\underline{F}_k = \underline{F}_k(\theta) = \{p_{Y_k}(Y_k)/\ell_k(Y_k|\theta) > 1\}$ , we write

$$\begin{aligned} k^{d\delta/2} |D_f(\mathbb{P}_Y || \mathbb{P}_{Y|\theta})| (a(2\pi)^{d/2} \sup_{\tilde{\Theta}} \pi|\mathcal{I}|^{-1/2})^{-\delta} &\leq k^{d\delta/2} \mathbb{E}_\theta \left[ \mathbb{1}_{\underline{F}_k} \hat{C} \frac{p_{Y_k}(Y_k)}{\ell_k(Y_k|\theta)} \right] + \mathbb{E}_\theta \left[ \mathbb{1}_{\underline{F}_k^c} k^{d\delta/2} \hat{C}' \left( \frac{p_{Y_k}(Y_k)}{\ell_k(Y_k|\theta)} \right)^\delta \right] \\ &\leq \hat{C} K_3 + \hat{C}' K_2. \end{aligned} \quad (4.102)$$

The domination above makes the following licit:

$$\lim_{k \rightarrow \infty} k^{d\delta/2} \mathcal{J}_{D_f}^k(\Pi(\cdot|\tilde{\Theta})) = a(2\pi)^{d\delta/2} (1-\delta)^{-d/2} \int_{\tilde{\Theta}} \pi(\delta)^{\delta+1} |\mathcal{I}(\theta)|^{-\delta/2} d\theta. \quad (4.103)$$

This concludes the proof of theorem 4.4. The result of proposition 4.7 and theorem 4.5 comes by concavity of  $x \mapsto ax^{-\delta}$ , which results from  $a(\delta+1) > 0$  and gives

$$\begin{aligned} a \int_{\tilde{\Theta}} \pi(\theta)^\delta |\mathcal{I}(\theta)|^{-\delta/2} \pi(\theta) d\theta \\ \leq a \left( \int_{\tilde{\Theta}} \pi(\theta)^{-1} |\mathcal{I}(\theta)|^{1/2} \pi(\theta) d\theta \right)^{-\delta} = a \left( \int_{\tilde{\Theta}} |\mathcal{I}(\theta)|^{1/2} d\theta \right)^{-\delta}. \end{aligned} \quad (4.104)$$

The concavity of  $x \mapsto ax^{-\delta}$  being strict, the above inequality is an equality if and only if  $\pi(\theta)^\delta |\mathcal{I}(\theta)|^{-\delta/2}$  is constant with respect to  $\theta$ , i.e.  $\pi \propto J$ .

□

*Proof of lemma 4.2.* We work under the assumptions of theorem 4.4. We write

$$\begin{aligned} \mathbb{E}_\theta \left[ \left( \frac{p_{Y_k}(Y_k)}{\ell_k(Y_k|\theta)} \right)^\zeta \right] &= \mathbb{E}_\theta \left( \int_{\tilde{\Theta}} e^{(\tilde{\theta}-\theta)^T \sqrt{k} S_k} e^{-\frac{k}{2}(\tilde{\theta}-\theta)^T \hat{I}_k(\hat{\theta})(\tilde{\theta}-\theta)} \pi(\tilde{\theta}) d\tilde{\theta} \right)^\zeta \\ &\leq \mathbb{E}_\theta k^{-d\zeta/2} \left( \int_{\mathbb{R}^d} e^{u^T S_k} e^{-\frac{1}{2} u^T \hat{I}_k(\hat{\theta}) u} du \sup_{\tilde{\Theta}} |\pi| \right)^\zeta \\ &\leq \mathbb{E}_\theta k^{-d\zeta/2} \left( \int_{\mathbb{R}^d} e^{u^T S_k} e^{-\frac{\|u\|^2}{2} \frac{1}{k} \sum_{i=1}^k \inf_{\tilde{\Theta}} L^{Y_i}} du \sup_{\tilde{\Theta}} |\pi| \right)^\zeta \end{aligned} \quad (4.105)$$

with  $\hat{\theta}$  being a point in  $\mathbb{R}^d$  within the segment between  $\theta$  and  $\tilde{\theta} = u/\sqrt{k} + \theta$ ; with  $L^y(\theta)$  being the smallest eigenvalue of  $(\nabla_\theta^2 \log \ell(y|\theta))^{1/2}$ . Therefore,  $\inf_{\tilde{\Theta}} L^{Y_i} > m$   $\mathbb{P}_\theta$  a.s. and

$$\mathbb{E}_\theta \left[ \left( \frac{p_{Y_k}(Y_k)}{\ell_k(Y_k|\theta)} \right)^\zeta \right] \leq k^{-d\zeta/2} K'' \mathbb{E}_\theta \left[ m^{-d\zeta/2} e^{\frac{\zeta}{2m} \|S_k\|^2} \right] \quad (4.106)$$

for some constant  $K'' > 0$ . By considering the sub-Gaussianity of  $S_k$ , the proof of the Lemma is complete.

□

*Proof of lemma 4.3.* We work under the assumptions of theorem 4.4. Notice

$$\mathbb{E}_\theta \left[ \mathbb{1}_{\underline{E}_k} \frac{p_{Y_k}(Y_k)}{\ell_k(Y_k|\theta)} \right] = \int_{\mathcal{Y}^k} \mathbb{1}_{\underline{E}_k} d\mathbb{P}_{Y_k}(\mathbf{y}). \quad (4.107)$$

According to Markov's inequality,

$$\begin{aligned} \mathbb{E}_\theta \left[ \mathbb{1}_{\mathbf{E}_k} \frac{p_{\mathbf{Y}_k}(\mathbf{y})}{\ell_k(\mathbf{y}|\theta)} \right] &\leq \rho^{\tilde{\delta}/\delta} \int_{\mathcal{Y}^k} \left( \frac{p_{\mathbf{Y}_k}(\mathbf{y})}{\ell_k(\mathbf{y}|\theta)} \right)^{\tilde{\delta}} p_{\mathbf{Y}_k}(\mathbf{y}) d\mu^{\otimes k}(\mathbf{y}) \\ &\leq \rho^{\tilde{\delta}/\delta} \int_{\mathcal{Y}^k} \left( \frac{p_{\mathbf{Y}_k}(\mathbf{y})}{\ell_k(\mathbf{y}|\theta)} \right)^{\tilde{\delta}} \int_{\tilde{\Theta}} \ell_k(\mathbf{y}|\tilde{\theta}) \pi(\tilde{\theta}) d\tilde{\theta} d\mu^{\otimes k}(\mathbf{y}) \\ &\leq \rho^{\tilde{\delta}/\delta} \int_{\tilde{\Theta}} \mathbb{E}_{\tilde{\theta}} \left[ \left( \frac{p_{\mathbf{Y}_k}(\mathbf{y})}{\ell_k(\mathbf{y}|\tilde{\theta})} \right)^{\tilde{\delta}} \left( \frac{\ell_k(\mathbf{y}|\tilde{\theta})}{\ell_k(\mathbf{y}|\theta)} \right)^{\tilde{\delta}} \right] \pi(\tilde{\theta}) d\tilde{\theta}, \end{aligned} \quad (4.108)$$

for some  $\tilde{\delta} > 0$  that we determine later. Using the expansion of the likelihood quotient expressed in eq. (4.64) we can write, for some  $\hat{\theta}$  in the segment between  $\theta$  and  $\tilde{\theta}$ :

$$\begin{aligned} \mathbb{E}_\theta \left[ \mathbb{1}_{\mathbf{E}_k} \frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)} \right] &\leq \rho^{\tilde{\delta}/\delta} \int_{\tilde{\Theta}} \mathbb{E}_{\tilde{\theta}} \left[ \left( \frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\tilde{\theta})} \right)^{\tilde{\delta}} e^{\tilde{\delta}(\tilde{\theta}-\theta)^T \sqrt{k} S_k} e^{-\frac{k}{2}\tilde{\delta}(\tilde{\theta}-\theta)^T \hat{\mathcal{I}}_k(\hat{\theta})(\tilde{\theta}-\theta)} \right] \pi(\tilde{\theta}) d\tilde{\theta} \\ &\leq \rho^{\tilde{\delta}/\delta} \left( \int_{\tilde{\Theta}} \mathbb{E}_{\tilde{\theta}} \left[ \left( \frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\tilde{\theta})} \right)^{p\tilde{\delta}} \right] \pi(\tilde{\theta}) d\tilde{\theta} \right)^{1/p} \\ &\quad \times \left( \int_{\tilde{\Theta}} \mathbb{E}_{\tilde{\theta}} \left[ e^{q\tilde{\delta}(\tilde{\theta}-\theta)^T \sqrt{k} S_k} e^{-q\frac{k}{2}\tilde{\delta}(\tilde{\theta}-\theta)^T \hat{\mathcal{I}}_k(\hat{\theta})(\tilde{\theta}-\theta)} \right] \pi(\tilde{\theta}) d\tilde{\theta} \right)^{1/q}, \end{aligned} \quad (4.109)$$

the second inequality being a consequence of Hölder's inequality (necessarily  $1/p + 1/q = 1$ ). The first integral in the right-hand side of the equation verifies

$$\int_{\tilde{\Theta}} \mathbb{E}_{\tilde{\theta}} \left[ \left( \frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\tilde{\theta})} \right)^{p\tilde{\delta}} \right] \pi(\tilde{\theta}) d\tilde{\theta} \leq K_2 k^{-p\tilde{\delta}d/2} \quad (4.110)$$

when  $\tilde{\delta}p/2 < \xi$  according to Lemma 4.2.

Now, we look for an upper-bound of the second integral:

$$I = \int_{\tilde{\Theta}} \mathbb{E}_{\tilde{\theta}} \left[ e^{q\tilde{\delta}(\tilde{\theta}-\theta)^T \sqrt{k} S_k} e^{-q\frac{k}{2}\tilde{\delta}(\tilde{\theta}-\theta)^T \hat{\mathcal{I}}_k(\hat{\theta})(\tilde{\theta}-\theta)} \right] \pi(\tilde{\theta}) d\tilde{\theta}. \quad (4.111)$$

Using the claim 2. on the consequence of the sub-Gaussianity of  $S_k$ , we get for any  $\tilde{\theta} \in \tilde{\Theta}$ :

$$\mathbb{E}_{\tilde{\theta}} \left[ e^{q\tilde{\delta}(\tilde{\theta}-\theta)^T \sqrt{k} S_k} e^{-q\frac{k}{2}\tilde{\delta}(\tilde{\theta}-\theta)^T \hat{\mathcal{I}}_k(\hat{\theta})(\tilde{\theta}-\theta)} \right] \leq K_1 e^{\frac{\xi^2}{2} q^2 \tilde{\delta}^2 k \|\tilde{\theta}-\theta\|^2} e^{-q\frac{k}{2}\tilde{\delta}m \|\tilde{\theta}-\theta\|^2}. \quad (4.112)$$

Therefore,

$$\begin{aligned} I &\leq K_1 k^{-d/2} \int_{\mathbb{R}^d} e^{\frac{\xi^2}{2} q^2 \tilde{\delta}^2 \|x\|^2} e^{-q\frac{k}{2}\tilde{\delta}m \|x\|^2} \sup_{\tilde{\Theta}} \pi dx \\ &\leq K_1 k^{-d/2} \int_{\mathbb{R}^d} e^{\frac{m}{4\xi} q^2 \tilde{\delta}^2 \|x\|^2} e^{-q\frac{k}{2}\tilde{\delta}m \|x\|^2} \sup_{\tilde{\Theta}} \pi dx. \end{aligned} \quad (4.113)$$

Let us suppose the integral of the right hand term above to be a finite constant that we denote  $L$ ; for this to be true, the inequality  $q\tilde{\delta} < 2\xi$  must be satisfied. In that case, one gets

$$\mathbb{E}_\theta \left[ \mathbb{1}_{\mathbf{E}_k} \frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)} \right] \leq \rho^{\tilde{\delta}/\delta} K_2^{1/p} L^{1/q} K_1^{1/q} k^{-d(\tilde{\delta}+1/q)/2}. \quad (4.114)$$

Choose now  $\tilde{\delta} = \delta/2$ , and  $q = p = 2$ . Thus  $p\tilde{\delta}/2 < \xi$  and  $q\tilde{\delta}/2 < \xi$  as required. Moreover,  $\tilde{\delta} + 1/q > \delta$  given  $1 > \delta$ . This way,

$$\mathbb{E}_\theta \left[ \mathbb{1}_{\mathbf{E}_k} \frac{p_{\mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k|\theta)} \right] \leq \rho^{1/2} K_3 k^{-d\delta/2}, \quad (4.115)$$

with  $K_3 = K_2^{1/2} L^{1/2} K_1^{1/2}$ .

□

## 4.7 Conclusion and prospects

This chapter contributes to the reference prior theory on multiple fronts.

First, we interpreted reference priors as maximizers of sensitivity indices. This interpretation reinforces the trustworthiness of their selection in practice: they are priors that minimize the impact of irreducible uncertainty on the studied systems.

Second, building on the aforementioned interpretation we provided a formalized framework for a generalization of the mutual information. We rigorously proved that usual properties of reference priors are still verified under that novel formalism. These generalized reference priors have been elucidated, and our main results show that they are Jeffreys priors in all the different cases that we study, namely when the mutual information is defined using a  $\delta$ -divergence with  $\delta \in (-1, 1)$ .

These results push further the robustness of reference priors that are constructed from Jeffreys—or hierarchical versions of Jeffreys—and support the objectivity of their choice in practice.

We conclude by emphasizing that our results provide an analytical criterion that aids in the definition of some generalized reference priors: they maximize a concave functional. That criterion constitutes an open path for the derivation of generalized reference priors over less general sets of priors. While this thought is slightly explored in this chapter, we note that it is furthered studied in the following chapter.

# 5

## Properly constrained generalized reference priors

This chapter is mainly based on the submitted work: Antoine Van Biesbroeck (2024b). “Properly constrained reference priors decay rates for efficient and robust posterior inference”. arXiv:2409.13041. DOI: [10.48550/arXiv.2409.13041](https://doi.org/10.48550/arXiv.2409.13041)

**Abstract** Reference priors are widely recognized for their objective nature. Yet, they often lead to intractable and improper priors, which complicates their application. Besides, informed prior elicitation methods are penalized by the subjectivity of the choices they require. In this chapter, we aim at proposing a reconciliation of the aforementioned aspects. Leveraging the objective aspect of the reference prior theory, we introduce two strategies of constraint incorporation to build tractable reference priors. One provides a simple and easy-to-compute solution when the improper aspect is not questioned, and the other introduces constraints to ensure the reference prior is proper, or it provides proper posterior. Our methodology emphasizes the central role of Jeffreys prior decay rates in this process, and the practical applicability of our results is demonstrated using an example taken from the literature.

<b>5.1</b>	<b>Introduction</b>	78
<b>5.2</b>	<b>Definitions, notations and motivation</b>	78
5.2.1	Notations	78
5.2.2	Objective and motivation	79
5.2.3	A useful definition: quasi-reference priors	80
<b>5.3</b>	<b>Constrained <math>D_\delta</math>-reference priors</b>	81
5.3.1	Constrained $D_\delta$ -reference priors based on Jeffreys' asymptotic decay rates	81
5.3.2	Properly constrained $D_\delta$ -reference priors	82
<b>5.4</b>	<b>Discussion</b>	83
<b>5.5</b>	<b>An example</b>	84
<b>5.6</b>	<b>Detailed proofs</b>	85
5.6.1	Proof of theorem 5.2	85
5.6.2	Proof of proposition 5.2	89
5.6.3	Proof of theorem 5.3	89
5.6.4	Proof of theorem 5.4	90
<b>5.7</b>	<b>Conclusion</b>	91

## 5.1 Introduction

The reference prior theory represents a widely elected theory for constructing priors that are qualified as “objective”. The theory has been thoroughly introduced in [chapter 3](#). The theory provides a formal mechanism to incorporate prior information in a way that maximizes the information gained from the data within the issued *a posteriori* quantities. In opposition with a plethora of existing methods for building a prior (see e.g. Mikkola et al., [2023](#)), this process is tuned to prevent the incorporation of subjective beliefs in the workflow.

However, despite their objective nature, their implementation is often cumbersome and not always recommended in high dimensions (Berger, Bernardo, and Sun, [2015](#)). Moreover, the low-informative nature of these priors is associated with their common improper aspect, necessitating careful handling to ensure valid statistical inference. Thus, the construction of priors is expected to strike a balance between several criteria. While many works restrict their sets of priors to ones that are tractable, proper, or suitable for high dimensions, others seek to minimize any source of subjectivity.

This chapter aims to reconcile all these viewpoints to improve the prior elicitation. Our contribution takes the form of an enrichment of the reference prior theory to leverage the objective aspect that it provides to reference priors. Building on the developments presented in [chapter 4](#), we restrict priors to the ones that belong to well-chosen—and not too restrictive—sets, and we introduce two strategies to define convenient reference priors. Our first strategy provides a simple, tractable solution for constraining reference priors when the improper aspect is not questioned. The second, by contrast, introduces constraints that lead to reference priors that are proper, or lead to proper posteriors. For both strategies, we try to examine the potential loss of objectivity induced by the constraints, and we discuss their limits. Our results emphasize the central role of Jeffreys prior decay rates when they are improper. Additionally, we draw attention to the fact that our methodology opens a way to define various reference priors on the basis of constraints that could result from any other motivation.

This chapter is organized as follows. In section [5.2](#), after reviewing briefly the notations for the generalized reference priors framework, we develop our motivation and the objective of this work. This section is also the occasion to introduce a novel definition that is useful for the rest of the study: the quasi  $D$ -reference priors. Our main results on constrained reference priors are presented in section [5.3](#), and discussed in section [5.4](#). Then, the practical aspect of our work is studied by the application of our method to an example taken from the literature in section [5.5](#). Detailed mathematical proofs are compiled in section [5.6](#). Section [5.7](#) terminates the chapter with a conclusion.

## 5.2 Definitions, notations and motivation

### 5.2.1 Notations

In this work we consider a statistic model characterized by a collection of probability distributions  $(\mathbb{P}_{Y|\theta})_{\theta \in \Theta}$  on a measurable set  $(\mathcal{Y}, \mathcal{Y})$ . We consider the same construction of the Bayesian framework as in [chapter 3](#) (section [3.4](#)): considering any prior  $\Pi$  (that is a  $\sigma$ -finite measure on  $\Theta$ ) we denote  $\mathbf{Y}_k$  a random vector of  $k$  observations whose distribution conditionally to  $T = \theta$  is  $\mathbb{P}_{Y^k|\theta} = \mathbb{P}_{Y|\theta}^{\otimes k}$ , where  $T$  is a r.v. whose distribution is the prior  $\Pi$  (see section [3.4.1](#)).

The modeling is supposed to be regular: we assume  $\Theta \subset \mathbb{R}^d$  with  $\nu$  being the Lebesgue measure on  $\mathbb{R}^d$  and every prior  $\Pi$  is supposed to admit a density  $\pi$  w.r.t.  $\nu$  (i.e.  $\Pi \in \mathcal{M}^\nu$ ). We also assume that the model admits a likelihood, denoted by  $\ell$  with for any  $\theta \in \Theta$  and  $\mathbf{y} \in \mathcal{Y}^k$ ,  $\ell_k(\mathbf{y}|\theta) := \prod_{i=1}^k \ell_i(\mathbf{y}_i|\theta)$ . We suppose that it verifies assumption [3.1](#) in [chapter 3](#), making the Fisher information matrix (denoted  $\mathcal{I}$ ) and the Jeffreys prior (whose density is denoted  $J$ ) being well-defined. The marginal distribution (resp. density) is denoted by  $\mathbb{P}_{\mathbf{Y}_k}$  (resp.  $p_{\mathbf{Y}_k}$ ) and the posterior distribution (resp. density) given the observations  $\mathbf{y} \in \mathcal{Y}^k$  is denoted by  $\mathbb{P}_{T|\mathbf{y}}$  (resp.  $p(\cdot|\mathbf{y})$ ).

Given these notations, we recall the expression of the generalized mutual information, defined in [chapter 4](#) when  $\Pi$  is proper:

$$\mathcal{I}_D^k(\Pi) := \mathbb{E}_{T \sim \Pi}[D(\mathbb{P}_{\mathbf{Y}_k} || \mathbb{P}_{\mathbf{Y}_k|T})], \quad (5.1)$$

with  $D$  being a dissimilarity measure. In this chapter, we mostly focus on such generalized mutual information when  $D$  is a  $\delta$  divergence with  $\delta \in (0, 1)$ . For some  $\delta \in (0, 1)$ , the notation  $D_\delta$  will refer to the  $\delta$ -divergence whose expression is reminded below:

$$D_\delta(P||Q) = \int_{\mathcal{X}} f_\delta \left( \frac{p(x)}{q(x)} \right) q(x) d\omega(x) \quad \text{with} \quad f_\delta(x) = \frac{x^\delta - \delta x - (1-\delta)}{\delta(1-\delta)}, \quad (5.2)$$

where  $p, q$  respectively are densities of  $P$  and  $Q$  w.r.t. a common measure  $\omega$  on  $\mathcal{X}$ . When evoking reference priors in a general way, we will refer to generalized reference priors, as proposed in [chapter 4](#). We remind below their definition:

**Definition 5.1** (Generalized reference prior). Let  $D$  be a dissimilarity measure and  $\mathcal{P}$  a set of priors on  $\Theta$ . A prior  $\Pi \in \mathcal{P}$  is called a  $D$ -reference prior over  $\mathcal{P}$  with rate  $\varphi(k)$  if there exists an openly increasing sequence of compact subsets  $(\Theta_i)_{i \in \mathbb{N}}$  such that  $\bigcup_{i \in \mathbb{N}} \Theta_i = \Theta$  and for any  $i$ :  $0 < \Pi(\Theta_i) < \infty$  and

$$\lim_{k \rightarrow \infty} \varphi(k) [\mathcal{I}_D^k(\Pi(\cdot|\Theta_i)) - \mathcal{I}_D^k(P(\cdot|\Theta_i))] \geq 0 \text{ for all } P \in \mathcal{P} \text{ verifying } 0 < P(\Theta_i) < \infty; \quad (5.3)$$

where  $\varphi(k)$  is a positive and monotonous function of  $k$ . It is said to be unique if for any other  $D$ -reference prior  $\Pi'$ ,  $\Pi \simeq \Pi'$ .

We also remind the following result on the  $D_\delta$ -mutual information and their reference priors (see [chapter 4](#)).

**Theorem 5.1.** Suppose  $\Theta$  to be compact and  $\Pi \in \mathcal{M}_{\mathcal{C}}^V$  be a prior with  $\Pi(\Theta) = 1$ . The  $D_\delta$ -mutual information admits a limit:

$$\lim_{k \rightarrow \infty} k^{d\delta/2} \mathcal{I}_{D_\delta}(\Pi) = l(\pi) - (\delta(1-\delta))^{-1}, \quad l(\pi) = C_\delta \int_{\Theta} \pi(\theta)^{1+\delta} |\mathcal{I}(\theta)|^{-\delta/2} d\theta, \quad (5.4)$$

where  $\pi$  is the density of  $\Pi$ , and with  $C_\delta = (2\pi)^{d\delta/2} (1-\delta)^{-d/2} / (\delta(\delta-1))$ .

Call  $\mathcal{R} \subset \mathcal{R}_{\mathcal{C}^b}$  a set of densities such that  $M(\mathcal{R}) = \mathcal{P}$ , with  $M$  mapping a density to its associated prior. Then  $\Pi$  is a  $D_\delta$ -reference prior over  $\mathcal{P}$  iff  $\pi$  maximizes  $l$  over  $\mathcal{R}$ .

## 5.2.2 Objective and motivation

We already know that the definition of the reference prior and the  $D_\delta$ -reference prior is satisfied by the Jeffreys prior over the large set of priors  $\mathcal{M}_{\mathcal{C}}^V$  in most cases (see chapters [3](#) and [4](#)).

This result is, however, limiting and disappointing in some cases. The reasons are the following ones: (i) the Jeffreys prior is not recommended in high-dimensional problems as it is known to be “either too diffuse or too concentrated” (Berger, Bernardo, and Sun, [2015](#)); moreover (ii) when the expression of the likelihood is itself complex, the computation of the Jeffreys prior can become intractable; also (iii) the Jeffreys prior is known to often lead to an improper prior, which does not necessarily issue a proper posterior distribution, essential for practical *a posteriori* inference and sampling.

To tackle these limitations, we propose in this work to restrict the set of priors over which we derive the reference priors. Indeed, the reference prior definition is usually considered with very large sets of priors, which are constrained only by some regularity assumptions imposed to the priors (such as continuity, positivity). These regularity assumptions do not generally discriminate the Jeffreys prior from the studied set of priors. In this chapter, different restricted sets of priors will be suggested, they are sets that are though to counter the limitations (ii) and (iii) aforementioned. In most cases, they will not include the Jeffreys prior.

The tackling of limitation (i) mentioned above is not a purpose of this work. We recall that it is actually frequently tackled by a sequential construction of the reference prior as suggested by Bernardo ([1979b](#)). On the condition that an ordering of the parameters is set:

$$\theta = (\theta_1, \dots, \theta_r) \in \Theta = \Theta_1 \times \dots \times \Theta_r, \quad (5.5)$$

this construction considers a hierarchical construction of the reference prior. It is already described in [chapter 3](#) (section [3.3.2](#)). We remind below the steps of the sequential construction:

1. initially fix  $\ell_k^1 = \ell_k$ ;
2. for any values of  $\theta_{j+1}, \dots, \theta_r \in \Theta_{j+1} \times \dots \times \Theta_r$ , compute a reference prior (in the sense of definition 3.3) under the model with likelihood  $\theta_j \mapsto \ell_k^j(\mathbf{y}|\theta_j, \dots, \theta_r)$ , denote  $\pi_j(\cdot|\theta_{j+1}, \dots, \theta_r)$  its normalized density;
3. derive  $\ell_k^{j+1}$  such as

$$\ell_k^{j+1}(\mathbf{y}|\theta_{j+1}, \dots, \theta_r) = \int_{\Theta_j} \ell_k^j(\mathbf{y}|\theta_j, \dots, \theta_r) d\pi_j(\theta_j|\theta_{j+1}, \dots, \theta_r). \quad (5.6)$$

In this work, the reference priors will be derived only given their formal definition (definition 5.1). Yet, our results can be incorporated in this sequential construction. Indeed, step 2 of the method depicted above consists of the derivation of a reference prior w.r.t. the variable  $\theta_j$ . Additionally, we invite to note that this construction does not solve the limitations (ii) and (iii) previously evoked. Actually, it makes them essential. Indeed, step 2 requires, firstly, a derivation of a reference prior, so that it would lead to a low-dimensional Jeffreys prior if the set of priors is not constrained. Also step 3 necessitates, secondly, that the latter leads to a proper posterior so that the integral involved does not diverge.

We note that this last issue is taken into account by Berger and Bernardo (1992a) with the suggestion of such construction on an increasing sequence of compact subsets of  $\Theta$ :  $\bigcup_{i \in \mathbb{N}} \Theta_i = \Theta$ . The hierarchical reference prior can then be chosen as a limit of the ones obtained under  $\Theta_i$  when  $i \rightarrow \infty$ . However, this limit can be cumbersome to derive in practice. Another solution suggested by Muré (2018) is to restrict the  $\sigma$ -algebra  $\mathcal{Y}$  until the reference prior derived in step 2 leads to a proper posterior. It is still imperfect, as there is no guarantee that such a restricted  $\sigma$ -algebra exists outside the trivial one.

### 5.2.3 A useful definition: quasi-reference priors

While we aim at restricting the set of priors  $\mathcal{P}$  in definition 5.1, we must notice that such a restriction leaves really unsure the existence of a reference prior. Indeed, the definition is itself restrictive, as to admit a reference prior, the set  $\mathcal{P}$  must contain a prior whose restrictions are optimal on any compact subsets of  $\Theta$ . In this section, we suggest an extension of the definition of reference priors in the case where in the set  $\mathcal{P}$ , the optimal priors on compact subsets of  $\Theta$  are not renormalization of each other, but converge to a prior in  $\mathcal{P}$ . Such convergence is considered in the sense of the Q-vague convergence (Bioche and Druilhet, 2016) on  $\mathcal{M}_C^\vee$ . The Q-vague convergence of a sequence  $(\Pi_n)_n$  to a limit  $\Pi$  is equivalent to the convergence of  $([\Pi_n])_n$  to  $[\Pi]$  in  $\mathcal{M}_C^\vee / \simeq$  for the quotient topology of the vague convergence on  $\mathcal{M}_C^\vee$ .

**Definition 5.2** (Quasi reference prior). Let  $\mathcal{P}$  be a set of priors. We call  $\Pi \in \mathcal{P}$  a quasi  $D$ -reference prior if it exists an openly increasing sequence  $(\Theta_i)_{i \in \mathbb{N}}$  of compact sets with  $\bigcup_{i \in \mathbb{N}} \Theta_i = \Theta$  such that

- (i) for any  $i \in \mathbb{N}$ , there exists a  $D$ -reference prior  $\Pi_i$  over  $\mathcal{P}_i = \{P(\cdot|\Theta_i), P \in \mathcal{P}, P(\Theta_i) \in (0, \infty)\}$ ,
- (ii)  $\Pi$  is the Q-vague limit of the sequence  $(\Pi_i)_{i \in \mathbb{N}}$ .

It is said to be unique if for any other quasi  $D$ -reference prior  $\Pi'$ ,  $\Pi \simeq \Pi'$ .

Proposition below ensures that this definition properly extends definition 5.1 in the case of  $\delta$ -divergences.

- Proposition 5.1.**
- If  $\Pi$  is a  $D_\delta$ -reference prior over a set  $\mathcal{P}$ , then it is a quasi  $D_\delta$ -reference prior.
  - If  $\mathcal{P}$  is a set of priors convex and stable by multiplication by indicator functions over measurable sets, then the quasi  $D_\delta$ -reference prior over  $\mathcal{P}$  is the unique  $D_\delta$ -reference prior over  $\mathcal{P}$ .
  - If  $\mathcal{P}$  is a convex set of priors and if the sequence of subsets  $(\Theta_i)_i$  in definition 5.1 is fixed, then the quasi-reference prior over  $\mathcal{P}$  is unique.

*Proof.* The first statement of the proposition is clear given the definition of a  $D_\delta$ -reference prior. For the second, let us adopt the notations of theorem 5.1 and notice that if  $\Theta$  is compact and if  $\pi^*$  is the maximal argument of  $l$  over  $\mathcal{R}$ , then its renormalized restriction  $\pi_1^*$  on a compact subset  $U$  maximizes  $l$  over the set of all renormalized densities  $\mathcal{R}_U$ . Indeed, if we suppose that  $\pi_1 \in \mathcal{R}_U$  maximizes  $l$  then, denoting  $\pi_0^*$  the renormalized restriction of  $\pi^*$  to  $\Theta \setminus U$ ,  $t = \int_U \pi^*$ , and  $\pi = t\pi_1 + (1-t)\pi_0^*$ ,  $\pi \in \mathcal{R}$  and

$$l(\pi) = t^{\delta+1}l(\pi_1) + (1-t)^{\delta+1}l(\pi_0^*) > t^{\delta+1}l(\pi_1^*) + (1-t)^{\delta+1}l(\pi_0^*) = l(\pi^*). \quad (5.7)$$

Hence  $\pi^*$  does not maximize  $l$  over  $\mathcal{R}$ , which is absurd.

Therefore, in our problem, considering two sequences  $(\pi_i^{(1)})_i$  and  $(\pi_i^{(2)})_i$  respectively defined on  $(\Theta_i^{(1)})_i$  and  $(\Theta_i^{(2)})_i$ , we will get that for any  $i$ ,  $\pi_i^{(1)}(\theta) = \pi_i^{(2)}(\theta)$  for all  $\theta \in \Theta_i^{(1)} \cap \Theta_i^{(2)}$ . Eventually, they are identical on every compact subsets of  $\Theta$ , and equal to their Q-vague limits which are the same.

Finally, the third statement of the proposition results from the strict concavity of  $l$ . Indeed, for any  $i$ , the set  $\mathcal{R}_i$  of renormalized restricted densities on  $\Theta_i$  is convex so that the maximal argument of  $l$  over  $\mathcal{R}_i$  is unique. Hence the uniqueness of the quasi-reference prior over  $\mathcal{P}$ .  $\square$

## 5.3 Constrained $D_\delta$ -reference priors

In this section, we present results of two different constraint incorporations on the reference priors. In section 5.3.1, the constraint limits the set of priors to exponentiation of coordinates of  $\theta$ . The result suggests a simpler reference priors in a case where the improper Jeffreys prior's decay rates are not an issue. In section 5.3.2, we introduce a well-chosen linear constraint on the set of priors that ensures the constraint reference prior is proper, or leads to proper posteriors.

### 5.3.1 Constrained $D_\delta$ -reference priors based on Jeffreys' asymptotic decay rates

In this section, we tackle the computational cost of the reference prior. As mentioned in section 5.2.2, the Jeffreys prior expression is often complex to derive even in low dimensional models. This is even more a problem in practical studies where the prior must be evaluated a numerous number of times, when it resorts to MCMC simulations to provide posterior samples of  $\theta$  for instance.

In some works of the literature, the Jeffreys prior is replaced by its decay rates at the boundary of the domain. For instance, the reference prior for Gaussian processes suggested by Gu (2019) is built on the basis of the decay rates of a Jeffreys prior sequentially computed on the different variables that compose  $\theta$  (following the construction presented in section 5.2.2). Their idea is that, in particular when it is improper, the prior provides the most information from its asymptotic rates, and variations of them are noticed to have a strong influence on the posterior distribution. The result that follows provides a formalization of this intuition, focusing on the case where the Jeffreys prior asymptotically behaves like exponentiation of coordinates of  $\theta$ .

**Theorem 5.2.** Suppose  $\Theta \subset \mathbb{R}$  is an interval of the form  $[c, b)$  (or  $(b, c]$ ). Call  $M : \pi \in \mathcal{R}_{\mathcal{C}^b} \mapsto (B \mapsto \int_B \pi d\nu) \in \mathcal{M}_{\mathcal{C}}^V$ .

- If  $b \in \mathbb{R}$  and  $J(\theta) \underset{\theta \rightarrow b}{\sim} C|\theta - b|^a$  for constants  $C \in \mathbb{R}$  and  $a \leq -1$ , then  $M(\pi^*)$  where  $\pi^*(\theta) \propto |\theta - b|^a$  is the unique quasi  $D_\delta$ -reference prior over  $M(\hat{\mathcal{R}})$  where  $\hat{\mathcal{R}} = \{\pi(\theta) \propto |\theta - b|^u, u \in \mathbb{R}\}$ .
- If  $|b| = \infty$  and  $J(\theta) \underset{\theta \rightarrow b}{\sim} C\theta^a$  for constants  $C \in \mathbb{R}$  and  $a \geq -1$ , then  $M(\pi^*)$  where  $\pi^*(\theta) \propto \theta^a$  is the unique quasi  $D_\delta$ -reference prior over  $M(\hat{\mathcal{R}})$  where  $\hat{\mathcal{R}} = \{\pi(\theta) \propto \theta^u, u \in \mathbb{R}\}$ .

*Proof.* The proof is technical and detailed in section 5.6. The idea is that  $l(\pi)$  can be seen as a negative divergence between  $\pi$  and  $J$ . However, when  $J$  is improper at the boundary of the domain, the maximization

of  $l(\pi)$  gets closer to the minimization of a divergence between  $\pi$  and the improper decay rate of  $J$ .  $\square$

**Remark 5.1.** Theorem 5.2 still stands when  $\Theta = (b, c)$  (or  $(c, b)$ ) if  $c \neq \infty$  and if  $J(\theta)$  admits a non-null and finite limit when  $\theta \rightarrow b$ .

This theorem serves the statement of two conclusions: (i) it emphasizes that when Jeffreys prior is improper, its improper decay rates contain the most relevant information, and (ii) it proposes to choose this asymptotic expansion of Jeffreys as a quasi  $D_\delta$ -reference prior when we look for an easy prior to compute.

### 5.3.2 Properly constrained $D_\delta$ -reference priors

In section 5.3.1, we have provided some elements to construct a tractable reference prior on the coordinates over which Jeffreys prior is improper. The reference prior proposed by our theorem keeps the improper characteristic of Jeffreys prior on the same coordinates. This improper aspect can, however, remain an issue in some cases, especially when the resulting posterior is improper as well.

For this reason, it might happen that some asymptotic rates in some directions still have to be tackled. The work in this section is concluded by results that allow defining a  $D_\delta$ -reference prior (or quasi  $D_\delta$ -reference prior), which benefits from adjusted decay rates from Jeffreys prior. The proposition below constitutes a preliminary result that gives the form of a  $D_\delta$ -reference prior over a set of priors with linear constraints.

**Assumption 5.1.** A family of functions from  $\Theta$  to  $\mathbb{R}$   $(g_j)_{j=1}^p$  is said to satisfy assumption 5.1 if  $g_0, \dots, g_p$  are linearly independent in the space of a.e. continuous functions from  $\Theta$  to  $\mathbb{R}$ , where  $g_0 := \theta \mapsto 1$ .

**Proposition 5.2.** Suppose  $\Theta$  to be a compact subset of  $\mathbb{R}^d$ . Let  $g_1, \dots, g_p$  be functions in  $\mathcal{R}_{C^b}$  that satisfy assumption 5.1. Define  $\tilde{\mathcal{P}}$  the set of priors  $\Pi$  on  $\Theta$  such that  $\forall 1, \dots, p, \int_{\Theta} g_j d\Pi = c_j$ , for some  $c_j \in \mathbb{R}$ . If there exists a  $D_\delta$ -reference prior over  $\tilde{\mathcal{P}}$ , it is unique. If it is positive, its density  $\pi$  verifies

$$\pi(\theta) = J(\theta) \left( \lambda_0 + \sum_{j=1}^p \lambda_j g_j(\theta) \right)^{1/\delta}, \quad (5.8)$$

for some  $\lambda_j \in \mathbb{R}$ . Reciprocally, if there exists a prior whose density verifies the above equation for some  $\lambda_j \in \mathbb{R}$ , it is the  $D_\delta$ -reference prior over  $\tilde{\mathcal{P}}$ .

*Proof.* This proposition results from a Lagrange multipliers theorem. A detailed proof is proposed in section 5.6.  $\square$

**Remark 5.2.** While it is not the subject of this work, we let the reader notice that this proposition opens the way to the introduction of constraints based on expert judgments in prior elicitation. They can take the form of moment constraints or predictive constraints.

**Remark 5.3.** The expression of the reference prior given by proposition 5.2 depends on the chosen  $\delta$ -divergence. While this work considers only the framework of reference priors under  $\delta$ -divergences as a dissimilarity measure, a version of this theorem could be written in the original framework of the reference prior theory that uses the Kullback-Leibler divergence. The expression of the resulting reference prior would be impacted. In appendix C we prove that the expression using the Kullback-Leibler divergence would take the form:

$$\pi^* \propto J \cdot \exp \left( \sum_{j=1}^p \lambda_j g_j \right), \quad (5.9)$$

for some  $\lambda_j$  that remain to be determined. This expression was already intuited by Bernardo and Smith (1994).

Below, given a function  $g$  that is selected to adjust the asymptotics of Jeffreys prior, is stated the expression of a proper  $D_\delta$ -reference prior.

**Theorem 5.3.** Let  $g : \Theta \rightarrow (0, \infty)$  be a function in  $\mathcal{R}_{C^b}$  such that

$$\int_{\Theta} J(\theta) g^{1/\delta}(\theta) d\theta < \infty \quad \text{and} \quad \int_{\Theta} J(\theta) g^{1/\delta+1}(\theta) d\theta < \infty, \quad (5.10)$$

and suppose that  $g$  is bounded in the neighborhood of  $b$  for an element  $b \in \partial\Theta$ . We denote by  $\bar{\mathcal{P}}$  the set of positive priors  $\Pi$  on  $\Theta$  such that  $\int_{\Theta} g d\Pi < \infty$ , and we define  $\Pi \in \bar{\mathcal{P}}$  as the prior whose density  $\pi$  verifies

$$\pi(\theta) \propto J(\theta) g(\theta)^{1/\delta}. \quad (5.11)$$

If  $Jg$  is non-integrable in the neighborhood of  $b$ , then  $\Pi$  is a  $D_{\delta}$ -reference prior over  $\bar{\mathcal{P}}$ . Otherwise, and if  $J$  is improper in the neighborhood of  $b$ ,  $\Pi$  is a  $D_{\delta}$ -reference prior over the set of proper priors in  $\bar{\mathcal{P}}$ .

*Proof.* The statement of this theorem results from the sequential use of proposition 5.2 on an increasing sequence of compact subsets of  $\Theta$ . A detailed proof is written in section 5.6.  $\square$

To improve the above theorem, one would like to relax the first assumption in eq. (5.10), i.e., to let  $\int_{\Theta} Jg^{1/\delta}$  be infinite. Indeed, in this way, the result would provide a reference prior  $\Pi$ —non-necessarily proper—but such that  $\pi g \in L^1$ , where  $\pi$  is a density of  $\Pi$ . With a good choice of  $g$ ,  $\pi^*$  could be built as a prior that provides a proper posterior. It is the purpose of the next theorem. The cost of this relaxation is the provision of a quasi-reference prior instead of a reference prior.

**Theorem 5.4.** Let  $g : \Theta \rightarrow (0, \infty)$  be in  $\mathcal{R}_{C^b}$  such that

$$\int_{\Theta} J(\theta) g(\theta) d\theta = \infty \quad \text{and} \quad \int_{\Theta} J(\theta) g^{1/\delta+1}(\theta) d\theta < \infty, \quad (5.12)$$

and suppose that  $g(\theta) \xrightarrow[\theta \rightarrow b]{} 0$  for an element  $b \in \partial\Theta$  such that  $J$  is non-integrable in the neighborhood of  $b$ .

Let  $(\Theta_i)_{i \in \mathbb{N}}$  be an openly increasing sequence of compact sets that covers  $\Theta$  and  $(c_i)_i$  be a bounded sequence in  $(0, \infty)$ . Define the set of priors  $\bar{\mathcal{P}}' = \{\Pi, \forall i, \int_{\Theta_i} g d\Pi = c_i \int_{\Theta_i} d\Pi\}$ .

Denote for any  $i$   $\bar{\mathcal{P}}'_i$  the set of renormalized restrictions to  $\Theta_i$  of priors in  $\bar{\mathcal{P}}'$ . If for any  $i$  there exists a positive maximum of  $l$  over  $\bar{\mathcal{P}}'_i$ , then  $\Pi$  whose density is denoted  $\pi$  is a quasi  $D_{\delta}$ -reference prior over  $\bar{\mathcal{P}}'$  with

$$\pi(\theta) \propto J(\theta) g(\theta)^{1/\delta}. \quad (5.13)$$

This prior is such that  $\int_{\Theta} \pi g d\nu < \infty$ .

## 5.4 Discussion

The knowledge of Jeffreys prior's decay rates is central in the results presented in this work. These results indicate that in common scenarios where Jeffreys prior is improper, these rates must be explicitly considered in order to construct a reference prior.

We let the reader note that theorems 5.3 and 5.4 introduce results that also depend on the chosen dissimilarity measure. Therefore, a balance must be found between the subjective influence of the constraint and the quest for an informed prior to facilitate possible sampling from the posterior. This is illustrated in the example we address in the following section. However, it is important to observe that using the KL-divergence instead of a  $\delta$ -divergence would result in a stronger influence of the constraint on the final prior. As noted in remark 5.3, the exponentialization of the function  $g$  could lead to a prior with distribution tails that are significantly negligible beyond those of Jeffreys, thereby jeopardizing its objective nature.

Generally, the results we propose in sections 5.3.1 and 5.3.2 address different problems and are thus fundamentally different in nature. In one case, the improper aspect of Jeffreys prior is not necessarily challenged, and an efficient construction of the latter is proposed. In the other case, the goal is to significantly attenuate

its improper aspect while maintaining as much objectivity as possible. In this latter case, however, the expressions of the proposed reference priors still depend on the expression of Jeffreys prior. Nevertheless, when Jeffreys prior is proper, there is no guarantee that a straightforward construction inspired by its convergence rates at the domain boundaries will be relevant. Indeed, although improper tails concentrate an infinite mass that constitutes all the information at the boundaries, when they are proper, the information of interest may need to be sought elsewhere. In this case, a calculation or approximation of the ‘proper’ Jeffreys prior remains to be considered.

Finally, regarding theorem 5.2, although the result is limited to parameter power distribution tails, it is observed that, in practice, these include a wide range of improper Jeffreys priors. For example, this includes Jeffreys priors derived from various Gaussian models, such as those introduced by Neyman and Scott (1948); Jeffreys priors related to specific parameters within Gaussian process models (Gu and Berger, 2016); and those arising in more specialized contexts, like the one that we develop in the part II of this manuscript. Moreover, the invariance of Jeffreys priors under re-parameterization can sometimes allow us to return to this case. Specifically, if  $J$  can be asymptotically written as a power of a function  $f$ , where  $f$  is differentiable, monotone and with bounded derivative (from above and from below), then the re-parameterization  $\vartheta = f(\theta)$  should allow us to recover the reference prior among those expressible as powers of  $f$ .

In the following section, we illustrate an application of our work with an example taken from the literature.

## 5.5 An example

In their work, Rubio and Steel (2014) prove that the two piece location-scale model they proposed has an improper Jeffreys prior, which issues an improper posterior. The model is parameterized by  $\theta = (\mu, \sigma_1, \sigma_2) \in \mathbb{R} \times (0, \infty)^2$ , inferred over observations in  $\mathcal{Y} = \mathbb{R}$ . It has the following likelihood:

$$\ell(y|\theta) = \frac{2}{\sigma_1 + \sigma_2} \left[ f\left(\frac{y - \mu}{\sigma_1}\right) \mathbb{1}_{(-\infty, \mu)}(y) + f\left(\frac{y - \mu}{\sigma_2}\right) \mathbb{1}_{(\mu, \infty)}(y) \right], \quad (5.14)$$

where  $f$  is a density function with support on  $\mathbb{R}$ , assumed to be symmetric with a single mode at zero, and with a few integrability assumptions that are detailed in Rubio and Steel, 2014. The choice of  $f$  is open and we can take, for instance, the standard Gaussian density function. Under this construction, the Fisher information matrix of this model takes the form:

$$\mathcal{I}(\theta) = \begin{pmatrix} \frac{\alpha_1}{\sigma_1 \sigma_2} & -\frac{2\alpha_3}{\sigma_1(\sigma_1 + \sigma_2)} & \frac{2\alpha_3}{\sigma_2(\sigma_1 + \sigma_2)} \\ * & \frac{\alpha_2}{\sigma_1(\sigma_1 + \sigma_2)} + \frac{\sigma_2}{\sigma_1(\sigma_1 + \sigma_2)^2} & -\frac{1}{(\sigma_1 + \sigma_2)^2} \\ * & * & \frac{\alpha_2}{\sigma_2(\sigma_1 + \sigma_2)} + \frac{\sigma_1}{\sigma_2(\sigma_1 + \sigma_2)^2} \end{pmatrix} \quad (5.15)$$

for some positive constants  $\alpha_1, \alpha_2$  and  $\alpha_3$ .

The full Jeffreys prior can be computed as

$$J(\theta) \propto \frac{1}{\sigma_1 \sigma_2 (\sigma_1 + \sigma_2)}, \quad (5.16)$$

it is improper and leads to an improper posterior. In the following, we construct different priors based on the suggestions developed in this chapter.

**Proper priors based on a moment constraint** Considering results in section 5.3.2 and the decay rates of the Jeffreys prior above, a simple correction can be done to issue a proper reference prior w.r.t.  $\sigma_1, \sigma_2$ , which results in a proper posterior. We consider  $\delta \in (0, 1)$ ; given  $\varepsilon \in (0, \frac{1}{1+\delta})$  we have

$$\int J(\theta) (\sigma_1 \sigma_2)^{\varepsilon/\delta+1} d\sigma_1 d\sigma_2 < \infty, \quad (5.17)$$

so that the associate proper  $D_\delta$ -reference prior density  $\pi$ , which is such that  $\pi(\mu, \sigma_1, \sigma_2) \sigma_1^\varepsilon \sigma_2^\varepsilon$  is integrable w.r.t.  $\sigma_1, \sigma_2$ , is

$$\pi(\mu, \sigma_1, \sigma_2) \propto \frac{(\sigma_1 \sigma_2)^{\varepsilon/\delta-1}}{\sigma_1 + \sigma_2}. \quad (5.18)$$

**Overview and sensitivity on the parameters** Our prior densities are compared with the Jeffreys prior one in fig. 5.1.(a), w.r.t. the parameter  $\sigma_2$  the others being fixed to 1. For this comparison, a multiplicative constant had to be chosen on  $J$ , we have chosen the one such that  $J(1, 1, 1) = 2$ . Our priors differ as a function of  $\gamma = \varepsilon/\delta \in (0, \frac{1}{1+\delta}) \subset (0, 1)$ . On the one hand, when  $\gamma$  becomes close to 0, the prior—which we denote by  $\pi_\gamma$  from now on—becomes close to the Jeffreys prior, i.e., the most objective prior w.r.t. the mutual information criterion. However, in this case,  $\pi_\gamma$  becomes close to an improper prior, and its posterior becomes close to an improper posterior. On the other hand, setting  $\gamma$  away from 0 rearranges the quantity of information in the prior. Its referential nature decreases in favor of an increase in its entropy. Therefore, a trade-off has to be made between suitability for inference and objectivity. Finally, note that in fig. 5.1.(a) is also drawn the *independent Jeffreys prior* (IJ) proposed by the authors in Rubio and Steel, 2014 for this model as an alternative to Jeffreys:

$$\text{IJ}(\mu, \sigma_1, \sigma_2) \propto \frac{\sqrt{\sigma_1 + \alpha_2(\sigma_1 + \sigma_2)} \sqrt{\sigma_2 + \alpha_2(\sigma_1 + \sigma_2)}}{\sqrt{\sigma_1 \sigma_2} (\sigma_1 + \sigma_2)^2}. \quad (5.19)$$

Its decay rates are actually the same as the ones of  $\pi_\gamma$  when  $\gamma = 1/2$ , and the two priors are hard to distinguish. Note that this latter prior  $\pi_{1/2}$  equals the hierarchical reference prior (as described in section 5.2.2) constructed from the ordering  $\pi(\theta) = \pi_1(\sigma_1, \sigma_2 | \mu) \pi_2(\mu)$ .

To evaluate a bit further our method, we propose a visualization of the posterior sensitivity to the priors, i.e., to  $\gamma$ . Such influence quantification constitutes a critical step of the *Bayesian workflow*, as expressed by Gelman, Vehtari, et al. (2020). Several methods exist in the literature for this purpose (e.g., Berger, 1990; Nott et al., 2020). An approach is to compare the variations of an *a posteriori* quantity as a function of the parameter (Kallioinen, Paananen, and Bürkner, 2023). In this example, this methodology is yet limited by the improper aspect of Jeffreys posterior. It cannot be considered for any comparison. In fig. 5.1.(b), (c) and (d) are plotted, for numerous  $\gamma$  and several data set sizes  $k$ , the posterior densities that result from a sample of data and from the prior densities  $\pi_\gamma$ . As expected, the influence of the prior appears for small values of  $k$ . Indeed, we can notice that for high values of  $\gamma$ , the posterior is slightly more shifted to the right and seems to be a little flatter. It is remarkable that this observation becomes limited when  $k$  increases. When  $k = 50$ , the difference between the posterior densities is hard to distinguish. This indicates that the little losses of objectivity should induce small variations in the resulting inference in this example.

For practical details, the chosen  $f$  is the standard Gaussian density, and the different data sets have been generated according to the likelihood in eq. (5.14) conditionally to  $\theta^* = (2, 2, 2)$ .

## 5.6 Detailed proofs

### 5.6.1 Proof of theorem 5.2

To prove this theorem, we consider to simplify the derivations that  $b = 0$  with  $\Theta = (0, 1]$ . We will show later how to extend the result to the other cases. As the Jeffreys prior can be defined up to a positive multiplicative constant with no incidence on the definition of a  $D_\delta$ -reference prior, we will simplify its decay rate assuming  $J(\theta) \underset{\theta \rightarrow 0}{\sim} \theta^\alpha$ .

Let us define the increasing sequence of compact subsets  $\Theta$ :  $\Theta_i = [\theta_i, 1]$ ,  $i \geq 0$ , with  $\theta_i \xrightarrow{i \rightarrow \infty} 0$ . We denote by  $\psi_i$  and  $\tilde{\psi}_i$  the functions defined as follow:

$$\psi_i(u) = -\frac{\int_{\Theta_i} J(\theta)^{-\delta} \theta^{u(1+\delta)} d\theta}{\left(\int_{\Theta_i} \theta^u d\theta\right)^{1+\delta}}, \quad \tilde{\psi}_i(u) = -\frac{\int_{\Theta_i} \theta^{-\alpha\delta} \theta^{u(1+\delta)} d\theta}{\left(\int_{\Theta_i} \theta^u d\theta\right)^{1+\delta}}. \quad (5.20)$$

The quantity  $\psi_i(u)$  corresponds—up to a positive constant—to the parametrization w.r.t.  $u \in \mathbb{R}$  of  $l(\pi_i)$ ; where  $\pi(\theta) \propto \theta^u$ ,  $\pi_i$  is re-normalized restriction of  $\pi$  to  $\Theta_i$  and where  $l$  is the function defined in theorem 5.1 that we seek to maximize. Therefore, if we call  $u_i \in \arg \max_{u \in \mathbb{R}} \psi_i(u)$  for any  $i$ , and if the associated sequence of priors  $(M(\pi_i))_i$  converge Q-vaguely to a prior  $\Pi^*$  over  $\Theta$ , that would prove that  $\Pi^*$  is a quasi  $D_\delta$ -reference prior.

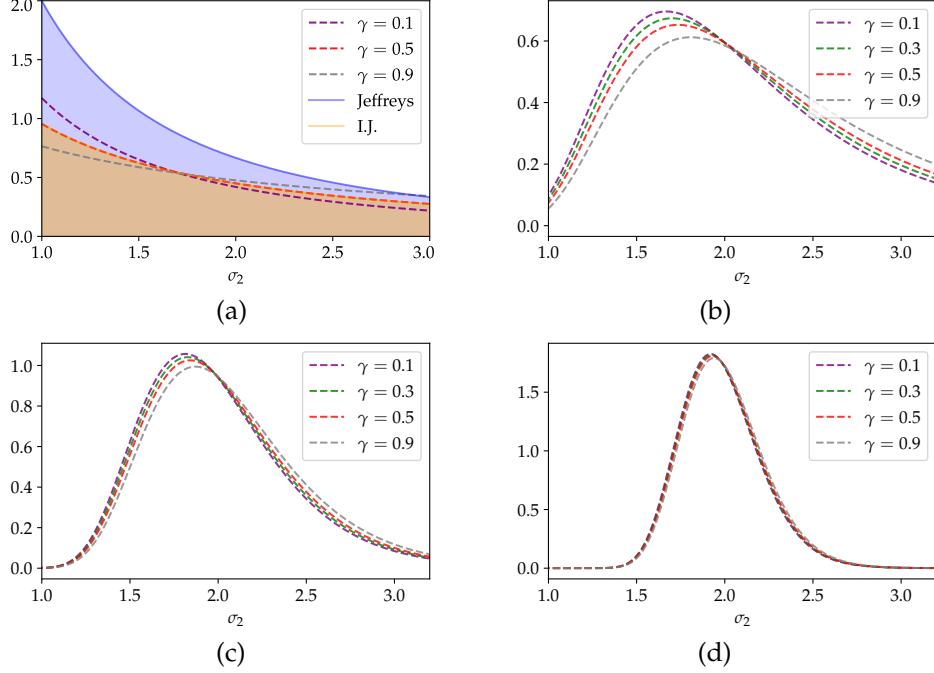


Figure 5.1: In (a), different prior densities  $\pi_\gamma$  (in dashed line) w.r.t.  $\sigma_2$  along with a Jeffreys prior density (that delimits the blue area) and the *Independent Jeffreys* of Rubio and Steel, 2014 (that delimits the orange area). In (b), (c) and (d), posterior densities for different values of  $\gamma$  are plotted. One observed sample was used to calculate all the posteriors in a figure. The observed sample has a size of  $k = 5$  in (b),  $k = 15$  in (c), and  $k = 50$  in (d).

**The case  $a < -1$**  Firstly, we assume that  $a < -1$ . Denote  $U = (-\infty, u_m := \frac{\delta a - 1}{1 + \delta})$ , we want to derive an asymptotic equivalent when  $i \rightarrow \infty$  of  $\psi_i(u)$ , uniformly w.r.t.  $u \in U$ . More precisely, we are about to show that for any  $\varepsilon > 0$  there exists  $i_1$  such that for any  $u \in U$  and  $i > i_1$ ,  $|\psi_i(u) - \tilde{\psi}_i(u)| < \varepsilon |\tilde{\psi}_i(u)|$ .

Let  $\varepsilon > 0$ , using that  $J(\theta)^{-\delta} \underset{\theta \rightarrow 0}{\sim} \theta^{-\delta a}$ , there exists  $i_0$  such that for any  $\theta < \theta_{i_0}$ ,  $|J(\theta)^{-\delta} - \theta^{-\delta a}| < \varepsilon \theta^{-\delta a}$ . Now, we consider an  $i_1$  such that:

$$\frac{\int_{\Theta_{i_0}} J(\theta)^{-\delta} d\theta}{\int_{\theta_{i_1}/\theta_{i_0}}^1 J(\theta_{i_0}\theta)^{-\delta} \theta^{u_m(1+\delta)} d\theta} < \theta_{i_0} \varepsilon / 2 \quad \text{and} \quad \frac{\int_{\Theta_{i_0}} \theta^{-\delta a} d\theta}{\int_{\theta_{i_1}/\theta_{i_0}}^1 (\theta_{i_0}\theta)^{-\delta a} \theta^{u_m(1+\delta)} d\theta} < \theta_{i_0} \varepsilon / 2. \quad (5.21)$$

Thus, for any  $i > i_1$  and any  $u \in U$ :

$$\frac{\int_{\Theta_{i_0}} J(\theta)^{-\delta} \theta^{u(1+\delta)} d\theta}{\int_{\Theta_i \setminus \Theta_{i_0}} J(\theta)^{-\delta} \theta^{u(1+\delta)} d\theta} = \frac{\theta_{i_0}^{-1} \int_{\theta_{i_0}}^1 J(\theta)^{-\delta} (\frac{\theta}{\theta_{i_0}})^{u(1+\delta)} d\theta}{\int_{\theta_i/\theta_{i_0}}^1 J(\theta_{i_0}\theta)^{-\delta} \theta^{u(1+\delta)} d\theta} < \frac{\theta_{i_0}^{-1} \int_{\Theta_{i_0}} J(\theta)^{-\delta} d\theta}{\int_{\theta_{i_1}/\theta_{i_0}}^1 J(\theta_{i_0}\theta)^{-\delta} \theta^{u_m(1+\delta)} d\theta} < \varepsilon / 2 \quad (5.22)$$

and

$$\frac{\int_{\Theta_{i_0}} \theta^{-\delta} \theta^{u(1+\delta)} d\theta}{\int_{\Theta_i \setminus \Theta_{i_0}} \theta^{-\delta} \theta^{u(1+\delta)} d\theta} = \frac{\theta_{i_0}^{-1} \int_{\theta_{i_0}}^1 \theta^{-\delta} (\frac{\theta}{\theta_{i_0}})^{u(1+\delta)} d\theta}{\int_{\theta_i/\theta_{i_0}}^1 (\theta_{i_0}\theta)^{-\delta} \theta^{u(1+\delta)} d\theta} < \frac{\theta_{i_0}^{-1} \int_{\Theta_{i_0}} \theta^{-\delta} d\theta}{\int_{\theta_{i_1}/\theta_{i_0}}^1 (\theta_{i_0}\theta)^{-\delta} \theta^{u_m(1+\delta)} d\theta} < \varepsilon / 2, \quad (5.23)$$

so that  $|\psi_i(u) - \tilde{\psi}_i(u)| < \varepsilon |\tilde{\psi}_i(u)|$  as expected.

Now we want to use this asymptotic equivalence to bound the difference  $|a - u_i|$ , where  $u_i$  is defined as a maximal argument of  $\psi_i$ . The next step is thus to show that such  $(u_i)_i$  exists.

There exist  $\tilde{K}, \tilde{K}'$  such that  $\tilde{K}\theta^{-\delta} \leq J(\theta)^{-\delta} \leq \tilde{K}'\theta^{-\delta}$ . Let  $i \geq 0$ , we can write

$$\tilde{K}|\tilde{\psi}_i(u)| \leq |\psi_i(u)| \leq \tilde{K}'|\tilde{\psi}_i(u)| \quad (5.24)$$

with

$$|\tilde{\psi}_i(u)| = \theta_i^{-\delta(a+1)} \frac{|u+1|^{1+\delta}}{|\delta(u-a) + u+1|} \frac{1 - \theta_i^{-\delta(u-a)-u-1}}{(1 - \theta_i^{-u-1})^{1+\delta}} \xrightarrow[u \rightarrow -\infty]{} +\infty. \quad (5.25)$$

That makes  $|\psi_i|$  being a coercive and continuous function on  $U$ , so that it admits minimal arguments in  $U$ . We denote by  $u_i$  one of them:

$$u_i \in \arg \max_{u \in U} \psi_i(u). \quad (5.26)$$

We recall that, by concavity of  $x \mapsto -x^{-\delta}$ , we find that  $a$  is the only maximal argument of  $\tilde{\psi}_i$  for any  $i$ . This way, for  $i > i_1$ , we write

$$\begin{aligned} |\tilde{\psi}_i(u_i) - \tilde{\psi}_i(a)| &\leq |\psi_i(u_i) - \tilde{\psi}_i(u_i)| + |\psi_i(a) - \tilde{\psi}_i(a)| + |\psi_i(u_i) - \psi_i(a)| \\ \tilde{\psi}_i(a) - \tilde{\psi}_i(u_i) &\leq \varepsilon(|\tilde{\psi}_i(u_i)| + |\tilde{\psi}_i(a)|) + \psi_i(u_i) - \psi_i(a), \end{aligned} \quad (5.27)$$

which leads to

$$\begin{aligned} 2(\tilde{\psi}_i(a) - \tilde{\psi}_i(u_i)) &\leq \varepsilon(|\tilde{\psi}_i(u_i)| + |\tilde{\psi}_i(a)|) + \psi_i(u_i) - \tilde{\psi}_i(u_i) + \tilde{\psi}_i(a) - \psi_i(a) \\ \tilde{\psi}_i(a) - \tilde{\psi}_i(u_i) &\leq \varepsilon(|\tilde{\psi}_i(u_i)| + |\tilde{\psi}_i(a)|). \end{aligned} \quad (5.28)$$

Consequently to the convergence of  $(\theta_i^{\delta(a+1)} \tilde{\psi}_i(a))_i$  toward a positive limit when  $i \rightarrow \infty$ , we deduce that  $\theta_i^{\delta(a+1)} (\tilde{\psi}_i(a) - \tilde{\psi}_i(u_i)) / |\theta_i^{\delta(a+1)} \tilde{\psi}_i(u_i)|$  is asymptotically null. This prevents the sequence  $(\theta_i^{\delta(a+1)} \tilde{\psi}_i(u_i))_i$  to admit a non finite subsequential limit, meaning it has to be bounded and to converges to the same limit as  $(\theta_i^{\delta(a+1)} \tilde{\psi}_i(a))_i$ , i.e.  $-|a+1|^\delta$ .

On another hand, we notice that for any  $M > 0$ , there exist a  $M'$  such that for any  $u < M'$ ,

$$\frac{|u+1|^{1+\delta}}{|\delta(u-a) + u+1|} > M \quad (5.29)$$

and  $|\theta^{\delta(a+1)} \tilde{\psi}_i(u)| > M$  for any  $i \geq 0$ . Thus, as  $(\theta^{\delta(a+1)} \tilde{\psi}_i(u_i))_i$  has been proven to be bounded, so must be  $(u_i)_i$ .

To conclude on that sequence, we denote by  $\rho$  a finite subsequential limit of  $(u_i)_i$ , if  $\rho \neq u_m$  then deriving the limit of  $\theta^{\delta(a+1)} \tilde{\psi}_i(u_i)$  leads to

$$\begin{aligned} -\frac{|\rho+1|^{1+\delta}}{\delta(\rho-a) + \rho+1} &= |a+1|^\delta \\ \text{i.e. } -|\rho+1|(|\rho+1|^\delta - |a+1|^\delta) &= |a+1|^\delta \delta(\rho-a); \end{aligned} \quad (5.30)$$

necessarily,  $\rho = a$ . It remains to prove that  $\rho = u_m$  is absurd. Indeed, in this case the integrals  $\int_{\Theta_i} \theta^{-\delta a + u_i(1+\delta)} d\theta$  converge either to 0, either to  $+\infty$ . Therefore, that would make  $(\theta_i^{\delta(a+1)} \tilde{\psi}_i(u_i))_i$  converging either to  $-\infty$ , either to 0, which in both case is different to  $-|a+1|^\delta$ .

Let us now work beyond the subset  $U$  of  $\mathbb{R}$ . First, if  $u \in (-1, +\infty)$ , the integrals that compose  $\psi_i(u)$  both admit finite and positive limits when  $i \rightarrow \infty$ . The limit of  $(|\psi_i(u)|)_i$  is moreover bounded from below as a consequence of eq. (5.24):

$$|\psi_i(u)| \geq \tilde{K} \frac{(u+1)^{1+\delta}}{\delta(u-a) + u+1} \frac{1 - \theta_i^{\delta(u-a)+u+1}}{(1 - \theta_i^{u+1})^{1+\delta}} \geq \tilde{K}' |\log \theta_i|^{-1-\delta}. \quad (5.31)$$

Thus, there exists  $i_2 \geq 0$  such that for any  $i > i_2$   $|\psi_i(a)| < \tilde{K}' |\log \theta_i|^{-1-\delta}$ , consequently to  $\psi_i(a) \underset{i \rightarrow \infty}{\sim} |a+1|^\delta \theta_i^{-\delta(a+1)}$   $\underset{i \rightarrow \infty}{=} o(|\log \theta_i|^{-\delta-1})$ . As a result, for any  $i > i_2$ :

$$\sup_{u \in (-1, +\infty)} \psi_i(u) < \psi_i(a) \leq \psi_i(u_i). \quad (5.32)$$

Finally, if  $u \in (u_m, -1)$ , analogously than in eq. (5.31), we can write

$$|\psi_i(u)| \geq \tilde{K}(u+1)^{1+\delta} \frac{|\log \theta_i|}{(1-\theta_i^{u+1})^{1+\delta}} \geq \tilde{K}'' \theta_i^{-(u_m+1)(1+\delta)} |\log \theta_i|^{-\delta}. \quad (5.33)$$

Once again, we have  $\psi_i(a) \underset{i \rightarrow \infty}{=} o(\theta_i^{-(u_m+1)(1+\delta)} |\log \theta_i|^{-\delta})$  and we can consider  $i_3 \geq 0$  such that for any  $i > i_3$ :

$$\sup_{u \in (u_m, -1)} \psi_i(u) < \psi_i(a) \leq \psi_i(u_i). \quad (5.34)$$

All the work that precedes proves that any sequence  $(v_i)_i$  defined by  $v_i \in \arg \max_{\mathbb{R}} \psi_i$  converges to  $a$ .

**The case  $a = -1$**  In this case, we easily get that  $\psi_i(a) \underset{i \rightarrow \infty}{\sim} -|\log \theta_i|^{-\delta}$ . When  $u < \eta < a$ ,

$$|\psi_i(u)| \geq \tilde{K} \frac{|u+1|^\delta}{\delta+1} \frac{1-\theta_i^{-(\delta+1)(u+1)}}{(1-\theta_i^{u-1})^{1+\delta}} \geq \hat{K} |\eta+1|^\delta (1-\theta_i^{-(\eta-1)(1+\delta)}) \quad (5.35)$$

and when  $u > \tilde{\eta} > a$ ,

$$|\psi_i(u)| \geq \tilde{K} \frac{|u+1|^\delta}{\delta+1} \frac{1-\theta_i^{(\delta+1)(u+1)}}{(1-\theta_i^{u+1})^{1+\delta}} \geq \hat{K}' |\tilde{\eta}+1|^\delta (1-\theta_i^{(\tilde{\eta}-1)(1+\delta)}). \quad (5.36)$$

Thus, for any  $\varepsilon > 0$ , the equations above with  $\eta = a - \varepsilon$  and  $\tilde{\eta} = a + \varepsilon$  let state that there exists an  $i_1 \geq 0$  such that for any  $i > i_1$ :

$$\sup_{u \in (-\infty, a-\varepsilon)} \psi_i(u) < \psi_i(a) \quad \text{and} \quad \sup_{u \in (a+\varepsilon, \infty)} \psi_i(u) < \psi_i(a) \quad (5.37)$$

so that  $\arg \max_{\mathbb{R}} \psi_i \subset (a - \varepsilon, a + \varepsilon)$ , which let the definition of a sequence  $(u_i)_i$  of maximal arguments of  $\psi_i$  which converges to  $a$ .

**Q-vague convergence** The conclusion concerning the Q-vague convergence of the  $M(\pi_i^*)$  defined from the  $u_i$  constructed in the work above:  $\pi_i^*(\theta) \propto \theta^{u_i}$  is a direct result of Bioche and Druilhet, 2016, Proposition 2.16. Indeed, the convergence of sequence  $(u_i)_i$  toward  $a$ , implies that the sequence of our priors converges Q-vaguely to  $\Pi^*$  such that  $\Pi^* = M(\pi^*)$  with  $\pi^*(\theta) \propto \theta^a$ .

**Extension to other set  $\Theta$**  We shall now demonstrate that the proven result extends itself to the general case:  $\Theta = [c, b]$  or  $(b, c)$ ,  $b \in \mathbb{R} \cup \{-\infty, \infty\}$ .

We first consider  $\Theta = (b, c)$  with  $b \in \mathbb{R}$ . We denote by  $\mathcal{Q}$  the set of priors densities after the substitution  $\vartheta = (\theta - b)/(c - b) \in T = (0, 1)$ :  $\mathcal{Q} = \{\tilde{\pi}(\vartheta) = (b - c)\pi(\vartheta(b - c) + b), \pi \in \hat{\mathcal{P}}\}$ . Thus,  $\mathcal{Q} = \{\pi(\theta) \propto \theta^u, u \in \mathbb{R}\}$ . We define the increasing sequence of compact sets  $(\Theta_i)_i$  by  $\Theta_i = [b + t_i, c]$  with  $t_i \xrightarrow{i \rightarrow \infty} 0$ . Therefore, for  $\pi \in \hat{\mathcal{P}}$ , calling  $\pi_i$  the renormalized restriction of  $\pi$  to  $\Theta_i$  gives

$$l(\pi_i) = \tilde{l}(\tilde{\pi}_i) = C_\delta \int_{T_i} \tilde{\pi}_i(\vartheta)^{1+\delta} \tilde{J}(\vartheta)^{-\delta} d\vartheta \quad (5.38)$$

with  $T_i = [t_i, 1]$ ,  $\tilde{\pi}_i(\vartheta) = (b - c)\pi_i(\vartheta(b - c) + b) \in \mathcal{Q}$  and  $\tilde{J}(\vartheta) = (b - c)J(\vartheta(b - c) + b) \underset{\vartheta \rightarrow 0}{\sim} \vartheta^a$ . Thus, the work done above states that the family of maximal arguments  $\tilde{\pi}_i^*$  of  $\tilde{l}$  over  $\mathcal{Q}_i$  provides a sequence of priors  $(M(\tilde{\pi}_i^*))_i$  that converges Q-vaguely toward  $M(\tilde{\pi}^*)$  where  $\tilde{\pi}^*(\vartheta) \propto \vartheta^a$ . Thus, the associate densities  $\pi_i^*$  maximize  $l$  over  $\mathcal{P}_i$  and issues a sequence of priors  $(M(\pi_i^*))_i$  that converges Q-vaguely toward  $M(\pi^*)$  where  $\pi^*(\theta) \propto |\theta - b|^a$ .

To treat the other cases, other substitutions with analogous work permit to conclude: (i) the substitution  $\vartheta = (b - \theta)/(b - c)$  when  $\Theta = (c, b)$ ,  $b \in \mathbb{R}$ ; (ii) the substitution  $\vartheta = 1/(|\theta - c| + 1)$  when  $\Theta = [c, \infty)$  or  $(-\infty, c]$ .

### 5.6.2 Proof of proposition 5.2

Let us start by the uniqueness. A prior  $\Pi^* \in \mathcal{P}$  is a  $D_\delta$ -reference prior if and only if its density  $\pi^*$  maximizes  $l$  over  $\mathcal{R}$ , where  $M(\mathcal{R}) = \mathcal{P}$ . As the set  $\mathcal{P}$  is convex,  $\mathcal{R}$  is convex as well. Also, we notice that the function  $\pi \mapsto l(\pi)$  is strictly concave, so that its maximizer is unique in the sense that two maximizer are equal  $\nu$ -a.e. That ensures the  $D_\delta$ -reference prior is unique.

Regarding the expression of  $\pi^*$ , let us call  $E$  the space of bounded a.e. continuous functions from  $\Theta$  to  $\mathbb{R}$  and we equip  $E$  with the supremum norm over  $E$ :  $\|f\| = \sup_{\Theta} |f|$ . The set  $\Theta$  being supposed compact, the pair  $(E, \|\cdot\|)$  constitutes a Banach vector space whose restriction  $U$  composed by the positive functions of  $E$  is an open and convex subset. It is possible to see  $l$  as being a concave function defined on  $U$ .

Let us compute the differentiate of  $l$  over  $U$ . One can write  $l = \phi_2 \circ \phi_1$  with

$$\phi_1 : \pi \in E \longmapsto \pi^{1+\delta} \in E; \quad \phi_2 : \pi \in E \longmapsto C_\delta \int_{\Theta} \pi(\theta) |\mathcal{I}(\theta)|^{-\delta/2} d\theta. \quad (5.39)$$

As  $\phi_2$  is a continuous linear mapping from  $E$  to  $\mathbb{R}$ ,  $l$  is differentiable while  $\phi_1$  is, with for any  $h \in E$ :

$$dl(\pi) = \phi_2 \circ d\phi_1(\pi). \quad (5.40)$$

Fix  $\pi \in U$ . For an  $\varepsilon > 0$  there exists  $\tilde{\varepsilon} > 0$  such that while  $|x| < \|\pi\|$  and  $|u| < \tilde{\varepsilon}$  then  $|(x+u)^{1+\delta} - x^\delta - (1+\delta)x^\delta u| < \varepsilon|u|$ . Thus for any  $h \in E$  such that  $\|h\| < \tilde{\varepsilon}$ , we have

$$\|\phi_1(\pi+h) - \phi_1(\pi) - (1+\delta)\pi^\delta h\| < \varepsilon\|h\|. \quad (5.41)$$

We conclude that  $\phi_1$  differentiable on  $U$  with  $d\phi_1(\pi)h = (1+\delta)\pi^\delta h$ , for any  $\pi \in U, h \in E$ . This differentiate is additionally continuous, which makes  $l$  continuously differentiable as well.

Thus, considering the additional constraint  $\int_{\Theta} \pi(\theta) d\theta = 1$ , this problem can be treated applying the Lagrange multipliers theorem (see e.g. (Zeidler, 2012)) to state that there exist  $\lambda_0, \dots, \lambda_p \in \mathbb{R}^{p+1}$  such that

$$dl(\pi^*)h - \lambda \int_{\Theta} h(\theta) d\theta - \sum_{i=1}^p \lambda_i \int_{\Theta} h(\theta) g_i(\theta) d\theta = 0 \quad (5.42)$$

for any  $h \in E$ . Eventually, as  $dl(\pi)h = C_\delta(1+\delta) \int_{\Theta} \pi(\theta)^\delta |\mathcal{I}(\theta)|^{-\delta/2} h(\theta) d\theta$ , we get

$$\pi^*(\theta) = J(\theta) \left( \lambda_0 + \sum_{i=1}^p \lambda_i g_i(\theta) \right)^{1/\delta}. \quad (5.43)$$

Moreover, as  $l$  is strictly concave and the constraints are linear, the second order condition for the Lagrangian states the reciprocal aspect of this result: calling  $C$  the space of functions satisfying the constraints, if  $\pi^* \in U \cap C$  satisfies the last equation above, it maximizes  $l$ .

To conclude, the elements in  $\tilde{\mathcal{P}}$  are associated to densities which belong to  $E$  because  $\Theta$  is compact. They are non necessarily everywhere positive, but only supposed to be non-negative. This set of bounded a.e. continuous and non-negative functions satisfying the constraints is actually the closure of  $U \cap C$  over which  $l$  is continuous. Therefore, calling  $\tilde{\mathcal{P}}^*$  the set of positive priors in  $\tilde{\mathcal{P}}$ , we have  $\arg \max_{M(\pi) \in \tilde{\mathcal{P}}} l(\pi) = \arg \max_{M(\pi) \in \tilde{\mathcal{P}}^*} l(\pi)$  which is  $\pi^*$ .

### 5.6.3 Proof of theorem 5.3

Assuming that it exists, denote by  $\pi^*$  the density of the  $D_\delta$ -reference prior  $\Pi^*$  over  $\overline{\mathcal{P}}$ , denote as well

$$c = \int_{\Theta} \pi^*(\theta) g(\theta) d\theta, \quad Z_i = \int_{\Theta_i} \pi^*(\theta) d\theta \quad (5.44)$$

for any  $i \in \mathbb{N}$ , considering an openly increasing sequence  $(\Theta_i)_{i \in \mathbb{N}}$  of compact sets that cover  $\Theta$  over which  $1 > \Pi^*(\Theta_i) > 0$  for any  $i$  such that  $\Theta_i \neq \Theta$  (we show later that they exist). In particular,  $\Pi^*$  must be the  $D_\delta$ -reference prior over the set  $\mathcal{P}^c = \{\Pi \in \mathcal{M}_C^V, \int_{\Theta} g d\Pi = c\}$ .

Let  $i \in \mathbb{N}$ , if  $\pi_i$  is a density on  $\Theta_i$  such that

$$\int_{\Theta_i} \pi_i g + \frac{1}{Z_i} \int_{\Theta \setminus \Theta_i} \pi^* g = \frac{c}{Z_i}, \quad (5.45)$$

then the prior whose density  $\pi$  on  $\Theta$  defined by  $\pi = Z_i \pi_i + \pi^* \mathbb{1}_{\Theta \setminus \Theta_i}$  belongs to  $\mathcal{P}^c$ . Therefore, denoting  $\pi_i^*$ , the renormalized restriction of  $\pi^*$  to  $\Theta_i$ ,  $l(\pi_i^*)$  must be larger than  $l(\pi_i)$  on  $\Theta_i$  by definition of the reference prior.

Thus,  $\pi_i^*$  is a maximal argument of  $l$  under the constraints

$$\int_{\Theta_i} \pi_i = 1, \quad \int_{\Theta_i} \pi_i g = \frac{c}{Z_i} - \frac{1}{Z_i} \int_{\Theta \setminus \Theta_i} \pi^* g. \quad (5.46)$$

Given the result of proposition 5.2,  $\pi_i^*$  takes the form of:

$$\pi_i^* = J \cdot (\lambda_i^{(1)} + \lambda_i^{(2)} g)^{1/\delta} \quad (5.47)$$

for some  $\lambda_i^{(1)}$  and  $\lambda_i^{(2)}$ .

Then, denoting as well  $\pi_{i+1}^* = J \cdot (\lambda_{i+1}^{(1)} + \lambda_{i+1}^{(2)} g)^{1/\delta}$ , and reminding that  $\pi_{i+1}^* \propto \pi_i^*$  on  $\Theta_i$ , we deduce that  $\lambda_{i+1}^{(1)} = \lambda_i^{(1)}$  as well as  $\lambda_{i+1}^{(2)} = \lambda_i^{(2)}$ . Eventually,  $\pi^* \propto J \cdot (\lambda_1 + \lambda_2 g)^{1/\delta}$ . Thus, in the neighborhood of  $b$ ,  $\pi^*(\theta) \underset{\theta \rightarrow b}{\sim} K \lambda_1^{1/\delta}$  for some  $K \neq 0$  if  $\lambda_1$  is non-null, which is discordant with the satisfaction of the constraint  $\int_{\Theta} \pi^* g < \infty$  in the case where  $\int_{\Theta} Jg = \infty$  or with the constraint  $\int_{\Theta} \pi^* = 1$  otherwise.

Finally,  $\pi^*$  is proportional to  $Jg^{1/\delta}$  and the value of  $c$  must be fixed by

$$c = \left( \int_{\Theta} J \cdot g^{1+1/\delta} \right) \cdot \left( \int_{\Theta} J \cdot g^{1/\delta} \right)^{-1}. \quad (5.48)$$

To finish the proof, we still have to show that  $\Pi^*$  is not null on any of the  $\Theta_i$ , or on any of the  $\Theta \setminus \Theta_i$ . First, there must exist  $i_0$  such that  $\Pi^*(\Theta_{i_0}) > 0$ , so that  $\Pi^*(\Theta_i) > 0$  for any  $i \geq i_0$  and the sequence  $(\Theta_i)_{i \geq i_0}$  can be considered instead of the initial one. Second, for any  $i$ , if  $\Theta \setminus \Theta_i$  is non-empty then it has a non-empty interior, as a consequence of the definition of openly increasing sequences of compact sets. Thus, as  $\Pi^*$  is assumed to be positive,  $\Pi^*(\Theta \setminus \Theta_i)$  is non-null. Hence the result.

#### 5.6.4 Proof of theorem 5.4

Consider an openly increasing sequence of compact sets  $(\Theta_i)_{i \in \mathbb{N}}$  that covers  $\Theta$ . Let  $i \in \mathbb{N}$ , for  $\Pi$  to be a  $D_\delta$ -reference prior over  $\mathcal{P}_i$ , its normalized density  $\pi_i$  must maximize the function  $l$  under the constraints  $\int_{\Theta_i} \pi_i = 1$  and  $\int_{\Theta_i} \pi_i g = c_i$ . Thus, according to proposition 5.2,  $\pi_i$  can be written as:

$$\pi_i(\theta) \propto J(\theta)(\lambda_i^{(1)} + \lambda_i^{(2)} g(\theta))^{1/\delta}, \quad (5.49)$$

for some  $\lambda_i^{(1)}$  and  $\lambda_i^{(2)}$ . Considering, if required, a subsequence of  $(\pi_i)_{i \in \mathbb{N}}$ , we can assume that  $\lambda_i^{(1)}$  and  $\lambda_i^{(2)}$  have constant signs. Also, by convexity we write:

$$\begin{aligned} c_i^\delta &\geq \int_{\Theta_i} J(\theta)(\lambda_i^{(1)} + \lambda_i^{(2)} g(\theta))g(\theta)d\theta \left( \int_{\Theta_i} Jg \right)^{\delta-1} \\ &\geq \lambda_i^{(1)} \left( \int_{\Theta_i} Jg \right)^\delta + \lambda_i^{(2)} \frac{\int_{\Theta_i} Jg^2}{\left( \int_{\Theta_i} Jg \right)^{1-\delta}}, \end{aligned} \quad (5.50)$$

and

$$\begin{aligned} 1 &\geq \int_{\Theta_i} J(\theta)(\lambda_i^{(1)} + \lambda_i^{(2)} g(\theta))d\theta \left( \int_{\Theta_i} J \right)^{\delta-1} \\ &\geq \lambda_i^{(1)} \left( \int_{\Theta_i} J \right)^\delta + \lambda_i^{(2)} \frac{\int_{\Theta_i} Jg}{\left( \int_{\Theta_i} J \right)^{1-\delta}}. \end{aligned} \quad (5.51)$$

We want to identify the possible subsequential limits of  $\lambda_i^{(1)}$ . We assume that one exists in  $(0, \infty]$ , we can consider if required a subsequence to assume that it is the actual the limit of  $\lambda_i^{(1)}$ . This way, eq. (5.51) does not allow  $\lambda_i^{(2)}$  to be non-negative. Thus,  $\lambda_i^{(2)}$  is negative and by convexity,

$$c_i^\delta \geq \lambda_i^{(1)} \left( \int_{\Theta_i} Jg \right)^\delta + \lambda_i^{(2)} \left( \int_{\Theta_i} Jg^{1+1/\delta} \right)^\delta. \quad (5.52)$$

Given that  $(c_i)_i$  is bounded, and that  $\int_{\Theta_i} Jg \xrightarrow{i \rightarrow \infty} \infty$  while  $\int_{\Theta_i} Jg^{1+1/\delta}$  has a finite limit, we deduce that  $\lambda_i^{(2)}$  diverges to  $-\infty$ . However, to ensure the expression of  $\pi_i$  to be well defined, we must have  $\lambda_i^{(1)} + \lambda_i^{(2)} g(\theta) \geq 0$  for any  $i$  and any  $\theta$ . As  $g$  is non-null, it necessary comes that  $(\lambda_i^{(2)} / \lambda_i^{(1)})_i$  is a bounded sequence. Therefore,  $\lambda_i^{(2)} \xrightarrow{i \rightarrow \infty} o\left(\lambda_i^{(1)} \left(\int_{\Theta_i} Jg\right)^\delta\right)$  and the right hand side in eq. (5.52) diverges to  $\infty$ , which is a contradiction with  $(c_i)_i$  being bounded.

As a consequence, any subsequential limit of  $(\lambda_i^{(1)})_i$  must belong to  $[-\infty, 0]$ . Without changing the notations, we can now assume that  $(\lambda_i^{(1)})_i$  has a limit, which we suppose to be strictly negative firstly, also,  $(\lambda_i^{(1)})_i$  can then be assumed to be always negative. Thus, still to ensure that  $\lambda_i^{(1)} + \lambda_i^{(2)} g(\theta) \geq 0$  for any  $i, \theta$ , and reminding that  $g(\theta) \xrightarrow{\theta \rightarrow b} 0$  we deduce  $\frac{\lambda_i^{(2)}}{|\lambda_i^{(1)}|} \xrightarrow{i \rightarrow \infty} \infty$ . Then, the sequence of functions  $(\pi_i / \lambda_i^{(2)})_i$  converges pointwisely to  $\pi^*$ , defined by  $\pi^*(\theta) \propto J(\theta)g(\theta)^{1/\delta}$  with  $|\frac{1}{\lambda_i^{(2)}} \pi_i(\theta)| \leq C + g(\theta)$  for some  $C$ , with  $g$  being bounded on all compact sets. According to Bioche and Druilhet, 2016, Proposition 2.15, what precedes implies that the sequence of priors  $(M(\pi_i))_i$  converges Q-vaguely to  $M(\pi^*)$ .

Eventually, if  $(\lambda_i^{(1)})_i$  admits 0 as a subsequential limit, we can assume that it is its actual limit. Consequently, the sequence of functions  $(\pi_i / \lambda_i^{(2)})_i$  converges pointwisely to the density  $\pi^*$  while being bounded by  $g$ . Therefore, the sequence of priors  $(M(\pi_i))_i$  converges Q-vaguely to  $M(\pi^*)$ . In any cases,  $M(\pi^*)$  is a quasi-reference prior over  $\bar{\mathcal{P}}'$ .

## 5.7 Conclusion

Prior elicitation remains an open topic in Bayesian analysis, with no solution satisfying all criteria (objectivity, computational feasibility, property, ...) simultaneously. In this work, we have focused on the objectivity and exploitability of priors. Specifically, we have provided some solutions for practitioners seeking non-subjective priors, addressing the practical challenges of traditional reference priors, particularly in terms of complexity and proper aspect. Through the example we presented, we demonstrated that simple criteria can lead to the construction of reference priors that are either straightforward to formulate or proper, depending on the case.

Moreover, our results show that one cannot completely dispense with a consideration of the asymptotic properties of Jeffreys prior. This emphasizes the importance of studying this prior for a complete construction of objective priors. Indeed, although one of our results offers a way to define in a simpler way the reference prior when it has improper decay rates, Jeffreys' rate still have to be elucidated. Our other result requires to derive the whole Jeffreys prior.

The approximation of the latter thus becomes a problem of a prior importance. Of course, one can thoroughly focus its explicit expression, yet it can be cumbersome to derive. In chapter 6 we propose a numerical solution for the approximation of the Jeffreys prior.



# 6

## Variational approximation of generalized reference priors using neural networks

This chapter develops the work done by Nils Baillie, during his end-of-study internship that I supervised at CEA Saclay. It is mainly based on the submitted paper: Nils Baillie, Antoine Van Biesbroeck, and Clément Gauchy (2025). “Variational inference for approximate objective priors using neural networks”. arXiv:2502.02364. DOI: [10.48550/arXiv.2502.02364](https://doi.org/10.48550/arXiv.2502.02364)

**Abstract** The computation of priors yielded by the framework of the reference prior theory remains a difficult task in various statistical models. In this chapter, we develop a flexible algorithm based on variational inference to compute approximations of reference priors from a set of distributions parameterized using neural networks. The algorithm is implemented to approximate the maximizer of generalized mutual information. It allows the introduction of constraints in the optimization problem that ensure the resulting prior is proper or issues proper posteriors. Also, our method incorporates the possibility to approximate the posterior distribution using Markov Chain Monte Carlo methods. Numerical experiments on two statistical models are presented, and we evaluate the performance of the algorithm by comparing both prior and posterior distributions. Our results show the usefulness of this approach to recover the target distribution.

<b>6.1</b>	<b>Introduction</b>	93
<b>6.2</b>	<b>Variational approximation of the reference prior (VA-RP)</b>	94
6.2.1	Reference priors, notations	94
6.2.2	Implicit expression of the prior using neural networks	95
6.2.3	Objective functions and their gradient	96
6.2.4	Adaptation for the constrained VA-RP	99
6.2.5	Posterior sampling from the implicit definition	100
<b>6.3</b>	<b>Numerical experiments on different models</b>	101
6.3.1	Multinomial model	101
6.3.2	Normal model	104
<b>6.4</b>	<b>Conclusion</b>	105

### 6.1 Introduction

The reference prior theory has been used in various statistical models, and the reference priors are recognized for their objective nature in several practical studies. However, while the theoretical expression of the reference priors is known in most cases (see [chapter 3](#) for a comprehensive review of the reference prior theory),

they suffer from their low computational feasibility. Indeed, they generally take the form of a Jeffreys prior or hierarchical Jeffreys priors, whose expression necessitate a heavy numerical cost to be derived, and that becomes even more cumbersome as the dimensionality of the problem increases. The developments done in chapters 4 and 5 make the tackling of this problem essential since the expression of the generalized reference prior suggested lead (with or without constrained) to a prior that is derived from Jeffreys'. It is an issue that jeopardizes their applicability since, in many applications, *a posteriori* estimates are obtained using Markov Chain Monte Carlo (MCMC) methods, which require many prior evaluations, further compounding the computational burden.

In general, when we look for sampling or approximating a probability distribution, several approaches arise and may be used within a Bayesian framework. In this work, we focus on variational inference methods. Variational inference seeks to approximate a complex target distribution  $p$ , —e.g. *a posteriori*— by optimizing over a family of simpler parameterized distributions  $q_\lambda$ . The goal then is to find the distribution  $q_{\lambda^*}$  that is the best approximation of  $p$  by minimizing an objective function, such as a dissimilarity measure. Variational inference methods have been widely adopted in various contexts, including popular models such as Variational Autoencoders (VAEs; Kingma and Welling, 2019), which are a class of generative models where one wants to learn the underlying distribution of data samples. We can also mention normalizing flows (see e.g. Papamakarios et al., 2021; Kobyzhev, Prince, and Brubaker, 2021), which consider diffeomorphism transformations to recover the density of the approximated distribution from the simpler one taken as input.

When it resorts to approximate reference priors, it is possible to leverage that they maximize the mutual information, instead of directly maximizing a divergence between a target and an output. Indeed, the mutual information does not depend on the target distribution that we want to reach, so that iterative derivations of the theoretical solution are not necessary. In Nalisnick and Smyth, 2017, the authors propose a variational inference procedure to approximate reference priors using a lower bound of the mutual information as an optimization criterion. By building on these foundations, this chapter proposes a novel variational inference algorithm to compute reference priors.

In this work, the reference priors are approximated from a parametric family of probability distributions implicitly defined by the push-forward probability distribution through a nonlinear function that takes the form of a neural network. The algorithm is thought to handle the maximization of mutual information that is defined using  $f$ -divergence (as suggested in chapter 4), instead of the traditional Kullback-Leibler divergence. Additionally, building on the developments conducted in chapter 5, we extend the framework to allow the integration of constraints on the prior. That last feature permits handling situations where the Jeffreys prior is improper. Improper priors represent a particular challenge in this work because they cannot be sampled from, yet the prior constructed from our methodology is not known explicitly and can only be used for sampling. In comparison with the previous works, we benchmark extensively our algorithm on statistical models of different complexity and nature to ensure its robustness.

In the following, we start by explaining the algorithmic construction of the variational approximation of the reference prior in section 6.2. We define in that section the optimization problem, its adaptation for incorporating constraints, and how the output prior is built and used for posterior sampling. The methodology is afterward implemented in section 6.3 on two different models: the multinomial model and the normal model. A conclusion terminates the chapter in section 6.4.

## 6.2 Variational approximation of the reference prior (VA-RP)

### 6.2.1 Reference priors, notations

In this study, we consider classical and regular statistic models defined by a collection of probability distributions  $(\mathbb{P}_{Y|\theta})_{\theta \in \Theta}$  on a measure set  $\mathcal{Y}$ , which admit likelihoods  $(\ell(\cdot|\theta))_{\theta \in \Theta}$  w.r.t. a common measure  $\mu$  on  $\mathcal{Y}$ . Given a prior  $\Pi$  on  $\Theta$ , one can construct the Bayesian framework as in chapter 3 (section 3.4) to define the  $D_f$ -mutual information when  $k$  data are observed:

$$\mathcal{I}_{D_f}^k(\Pi) = \mathbb{E}_{T \sim \Pi}[D_f(\mathbb{P}_{Y_k} || \mathbb{P}_{Y_k|T})]; \quad (6.1)$$

where  $\mathbb{P}_{Y_k|\theta} := \mathbb{P}_{Y|\theta}^{\otimes k}$ ,  $\mathbb{P}_{Y_k}$  being the marginal distribution and  $D_f$  denoting an  $f$ -divergence. This definition of the mutual information using  $f$ -divergence is an extension of the original definition that uses a Kullback-Leibler divergence instead, we refer to [chapter 4](#) where this extension is proposed and developed. In this chapter, the regularity of the model will not be questioned and always be assumed sufficient for our mathematical derivations to stand. In particular the Fisher information matrix (denoted by  $\mathcal{I}$ ) exists as well as the Jeffreys prior (whose density is denoted  $J$ ).

Note that in the expression of the  $D_f$ -mutual information above, the involved distributions do not always exist when  $\Pi$  is improper. In this case, the quantity is not always well-defined. That is why in the context of reference priors this quantity is considered taking restrictions of  $\Pi$  on compact subsets of  $\Theta$ . Reference priors are then defined as asymptotic maximizer of the restricted mutual information as  $k \rightarrow \infty$ .

In this work, we aim at “globally” maximizing the  $D_f$ -mutual information, i.e. considering its expression written as in eq. (6.1) for priors defined on the whole space  $\Theta$ . Moreover, our method limits itself to the maximization of the mutual information as expressed in eq. (6.1), i.e. for a fixed value of  $k$ . All in all, we focus the solving of an optimization problem that takes the form:

$$\text{find } \Pi^* \in \arg \max_{\Pi \in \mathcal{P}_N} \mathcal{I}_{D_f}^k(\Pi), \quad (6.2)$$

where  $\mathcal{P}_N$  is a set of normalized proper continuous priors:  $\mathcal{P}_N = \{\Pi \in \mathcal{M}_C^\nu, \Pi(\Theta) = 1\}$ . In this work, we will always treat cases where  $\Theta \subset \mathbb{R}^d$  and  $\nu$  is the Lebesgue measure. Actually, restricting the research of reference priors to the optimization problem written above is not very limiting. Indeed, the set  $\mathcal{P}_N$  contains priors that are “very close” to any improper priors: considering the topology on  $\mathcal{M}/\simeq$  induced by the Q-vague convergence (see Bioche and Druilhet, [2016](#)),  $\mathcal{P}_N$  is dense in  $\mathcal{M}_C^\nu$ . Also, we expect that if  $k$  is large enough, the solution of the optimization problem should get closer to the reference prior. That intuition has already been developed by Berger, Bernardo, and Sun ([2009](#)) or Le, [2014](#) who proved it in the one dimensional case.

Furthermore, as mentioned in the introduction, this work aims also at approximating “properly constrained reference priors”, i.e. priors that are maximizer of the mutual information under some constraints. The idea is to leverage the development conducted in [chapter 5](#) to make reference priors being proper. We process by addressing the resolution of the following optimization problem:

$$\begin{aligned} \text{find } \tilde{\Pi}^* \in & \arg \max_{\substack{\Pi \in \mathcal{P}_N \\ \text{s.t. } C(\Pi) < \infty}} \mathcal{I}_{D_f}^k(\Pi), \end{aligned} \quad (6.3)$$

where  $C(\Pi)$  defines a constraint of the form  $\int_\Theta a(\theta)d\Pi(\theta)$ ,  $a$  being a positive function. We remind that, when the mutual information in the above optimization problem is defined from a  $\delta$ -divergence, and when  $a$  verifies

$$\int_\Theta J(\theta)a(\theta)^{1/\delta}d\theta < \infty \quad \text{and} \quad \int_\Theta J(\theta)a(\theta)^{1+1/\delta}d\theta < \infty, \quad (6.4)$$

the developments done in [chapter 5](#) state that the constrained solution  $\tilde{\Pi}^*$  admits a density  $\tilde{\pi}^*$  that asymptotically takes the form:

$$\tilde{\pi}^*(\theta) \propto J(\theta)a(\theta)^{1/\delta}, \quad (6.5)$$

which is proper.

### 6.2.2 Implicit expression of the prior using neural networks

In this work, the goal is to solve the optimization problems expressed in eqs. (6.2) and (6.3) by variational inference. The idea is to restrict the study to a parametric set of priors  $\{\Pi_\lambda, \lambda \in \Lambda\}$ ,  $\Lambda \subset \mathbb{R}^L$ , reducing the optimization problem to a finite dimensional one: find  $\arg \max_{\lambda \in \Lambda} \mathcal{I}_{D_f}^k(\Pi_\lambda)$ . The construction of this parametric set comes by defining our priors implicitly, from a simpler random variable  $\varepsilon$  and a non-linear function  $g$ :

$$\theta \sim \Pi_\lambda \iff \theta = g(\lambda, \varepsilon) \quad \text{and} \quad \varepsilon \sim \mathbb{P}_\varepsilon. \quad (6.6)$$

The function  $g$  is a measurable function parameterized by some  $\lambda \in \Lambda \subset \mathbb{R}^L$ . Typically,  $g$  is a neural network with  $\lambda$  corresponding to its weights and biases, and  $g$  is assumed to be differentiable w.r.t.  $\lambda$ . The variable  $\varepsilon$  can be seen as a latent variable. It has an easy-to-sample distribution  $\mathbb{P}_\varepsilon$  with a simple density function. In practice, we use the centered multivariate Gaussian  $\mathcal{N}(0, I_{p \times p})$ . If one chooses a dense and heavy parameterized function  $g$ , then the family  $\mathcal{P}_\Lambda = \{\Pi_\lambda, \lambda \in \Lambda\}$  will be expected to be vast, so that many priors could be approached using this method. However, this implicit construction comes with one inconvenient: except in very simple cases, the density of  $\pi_\lambda$  is not known and cannot be evaluated. Only samples of  $\theta \sim \Pi_\lambda$  can be obtained.

We note that in Nalisnick and Smyth, 2017, this implicit sampling method is compared to several other algorithms used to learn reference priors in one-dimensional cases. Among these methods, we can mention an algorithm proposed in Berger, Bernardo, and Sun, 2009 which does not sample from the reference prior but only evaluates it for specific points, or an MCMC-based approach in Lafferty and Wasserman, 2001, which is inspired from the previous one but can sample from the reference prior. According to this comparison, implicit sampling is, in the worst case, competitive with the other methods, but achieves state-of-the-art results in the best case. Hence, computing the variational approximation of the reference prior, which we will refer to as the VA-RP, seems to be a promising technique.

### 6.2.3 Objective functions and their gradient

The VA-RP is formally formulated as the solution of an optimization problem of this such:

$$\text{find } \Pi_{\lambda^*} \text{ where } \lambda^* = \arg \max_{\lambda \in \Lambda} \mathcal{O}_{D_f}^k(\Pi_\lambda), \quad (6.7)$$

where  $\Pi_\lambda$  refers to the implicit parametrization described in section 6.2.2. The function  $\mathcal{O}_{D_f}^k$  is called the objective function, it is maximized using stochastic gradient optimization. It is intuitive to fix  $\mathcal{O}_{D_f}^k$  to equal  $\mathcal{I}_{D_f}^k$ , in order to maximize the mutual information of interest. Yet, in this section, we suggest alternative objective functions that can be considered to compute the VA-RP. Our method is adaptable to any objective function  $\mathcal{O}_{D_f}^k$  admitting a gradient w.r.t.  $\lambda$  that takes the form of the following definition.

**Definition 6.1.** We recall that  $\mathcal{P}_\Lambda$  denotes the set of priors defined by the implicit parametrization described in section 6.2.2. A mapping  $\Pi_\lambda \in \mathcal{P}_\Lambda \mapsto \mathcal{O}_{D_f}^k(\Pi_\lambda)$  is said to be an acceptable objective function if it admits a gradient w.r.t.  $\lambda$  that takes the form

$$\nabla_\lambda \mathcal{O}_{D_f}^k(\Pi_\lambda) = \mathbb{E}_{\varepsilon \sim \mathbb{P}_\varepsilon} \left[ \sum_{j=1}^d \partial_{\theta_j} \tilde{\mathcal{O}}_{D_f}^k(g(\lambda, \varepsilon)) \nabla_\lambda g_j(\lambda, \varepsilon) \right] \quad (6.8)$$

where  $\tilde{\mathcal{O}}_{D_f}^k : \Theta \rightarrow \mathbb{R}$  is a function that is independent of  $\lambda$ .

Such objective functions allow for flexible implementation, as it permits the separation of sampling and differentiation operations:

- The gradient of  $\tilde{\mathcal{O}}_{D_f}^k$  mostly relies on random sampling and depends only on the likelihood  $\ell_k$  and the function  $f$ .
- The gradient of  $g$  is computed independently. In practice, it is possible to leverage usual automatic differentiation techniques for the neural network. An example is PyTorch’s automatic differentiation feature “autograd”.

This separation is advantageous as automatic differentiation tools —such as autograd— are well-suited to differentiating complex networks but struggle with functions incorporating randomness. This way, the optimization problem can be addressed using stochastic gradient optimization, approximating at each step the gradient in eq. (6.8) via Monte Carlo estimates. In our experiments, the implementation of the algorithm is done using Adam optimizer.

Several acceptable objective functions exist, such as the  $D_f$ -mutual information  $\mathcal{I}_{D_f}^k$ , as stated in the proposition below.

**Proposition 6.1.** *The  $D_f$ -mutual information  $\mathcal{I}_{D_f}^k$  is an acceptable objective function. Its gradient equals*

$$\begin{aligned}\nabla_{\lambda} \mathcal{I}_{D_f}^k(\Pi_{\lambda}) &= \mathbb{E}_{\varepsilon \sim \mathbb{P}_{\varepsilon}} \left[ \sum_{j=1}^d \partial_{\theta_j} \tilde{I}(g(\lambda, \varepsilon)) \nabla_{\lambda} g_j(\lambda, \varepsilon) \right] \\ &\quad + \mathbb{E}_{\theta \sim \Pi_{\lambda}} \left[ \mathbb{E}_{\mathbf{Y}_k \sim \mathbb{P}_{\mathbf{Y}_k | \theta}} \left[ \frac{1}{\ell_k(\mathbf{Y}_k | \theta)} \nabla_{\lambda} p_{\lambda, \mathbf{Y}_k}(\mathbf{Y}_k) f' \left( \frac{p_{\lambda, \mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k | \theta)} \right) \right] \right],\end{aligned}\quad (6.9)$$

where:

$$\partial_{\theta_j} \tilde{I}(\theta) = \mathbb{E}_{\mathbf{Y}_k \sim \mathbb{P}_{\mathbf{Y}_k | \theta}} \left[ \partial_{\theta_j} \log \ell_k(\mathbf{Y}_k | \theta) F \left( \frac{p_{\lambda, \mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k | \theta)} \right) \right], \quad (6.10)$$

with  $F(x) = f(x) - x f'(x)$  and  $p_{\lambda, \mathbf{Y}_k}$  denoting the marginal density when the prior is  $\Pi_{\lambda}$ .

*Proof.* We remind that  $\mathcal{I}_{D_f}^k(\Pi_{\lambda}) = \mathbb{E}_{\theta \sim \Pi_{\lambda}} \mathbb{E}_{\mathbf{Y}_k \sim \mathbb{P}_{\mathbf{Y}_k | \theta}} f \left( \frac{p_{\lambda, \mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k | \theta)} \right)$ . We compute its gradient w.r.t.  $\lambda$  as follows:

$$\nabla_{\lambda} \mathcal{I}_{D_f}^k(\Pi_{\lambda}) = \nabla_{\lambda} \left[ \mathbb{E}_{\theta \sim \Pi_{\lambda}} \tilde{I}(\theta) \right] + \mathbb{E}_{\theta \sim \Pi_{\lambda}} \mathbb{E}_{\mathbf{Y}_k \sim \mathbb{P}_{\mathbf{Y}_k | \theta}} \left[ \frac{1}{\ell_k(\mathbf{Y}_k | \theta)} \nabla_{\lambda} p_{\lambda, \mathbf{Y}_k}(\mathbf{Y}_k) f' \left( \frac{p_{\lambda, \mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k | \theta)} \right) \right], \quad (6.11)$$

with  $\tilde{I}(\theta) = \mathbb{E}_{\mathbf{Y}_k \sim \mathbb{P}_{\mathbf{Y}_k | \theta}} [f(p_{\lambda, \mathbf{Y}_k}(\mathbf{Y}_k) / \ell_k(\mathbf{Y}_k | \theta))]$  so that for any  $j$  its derivative w.r.t.  $\theta_j$  is

$$\partial_{\theta_j} \tilde{I}(\theta) = \int_{\mathcal{Y}^k} \frac{-p_{\lambda, \mathbf{Y}_k}(\mathbf{y})}{\ell_k(\mathbf{y} | \theta)} f' \left( \frac{p_{\lambda, \mathbf{Y}_k}(\mathbf{y})}{\ell_k(\mathbf{y} | \theta)} \right) \partial_{\theta_j} \ell_k(\mathbf{y} | \theta) d\mu^{\otimes k}(\mathbf{y}) + \int_{\mathcal{Y}^k} f \left( \frac{p_{\lambda, \mathbf{Y}_k}(\mathbf{y})}{\ell_k(\mathbf{y} | \theta)} \right) \partial_{\theta_j} \ell_k(\mathbf{y} | \theta) d\mu^{\otimes k}(\mathbf{y}); \quad (6.12)$$

that being equal to the expression in eq. (6.10). Using the trick  $\nabla_{\lambda} \mathbb{E}_{\theta \sim \Pi_{\lambda}} \tilde{I}(\theta) = \mathbb{E}_{\varepsilon \sim \mathbb{P}_{\varepsilon}} \nabla_{\lambda} \tilde{I}(g(\lambda, \varepsilon))$ , we recovered the expression of the gradient stated in the proposition.

Now we have to show that the second term of the gradient can be written in the same way as the first one to ensure the gradient is acceptable. We call  $\mathcal{G}$  this second term:

$$\mathcal{G} = \mathbb{E}_{\varepsilon} \mathbb{E}_{\mathbf{Y}_k \sim \mathbb{P}_{\mathbf{Y}_k | g(\lambda, \varepsilon)}} \frac{1}{\ell_k(\mathbf{Y}_k | g(\lambda, \varepsilon))} \nabla_{\lambda} p_{\lambda, \mathbf{Y}_k}(\mathbf{Y}_k) f' \left( \frac{p_{\lambda, \mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k | g(\lambda, \varepsilon))} \right), \quad (6.13)$$

with  $\mathbb{E}_{\varepsilon}$  being a shortcut for  $\mathbb{E}_{\varepsilon \sim \mathbb{P}_{\varepsilon}}$ . Remarking that  $\nabla_{\lambda} p_{\lambda, \mathbf{Y}_k}(\mathbf{Y}_k) = \mathbb{E}_{\varepsilon_2} \sum_{j=1}^d \partial_{\theta_j} \ell_k(\mathbf{Y}_k | g(\lambda, \varepsilon_2)) \nabla_{\lambda} g_j(\lambda, \varepsilon_2)$ , we develop  $\mathcal{G}$  as:

$$\begin{aligned}\mathcal{G} &= \mathbb{E}_{\varepsilon_1} \mathbb{E}_{\mathbf{Y}_k \sim \mathbb{P}_{\mathbf{Y}_k | g(\lambda, \varepsilon_1)}} \mathbb{E}_{\varepsilon_2} \sum_{j=1}^d \frac{1}{\ell_k(\mathbf{Y}_k | g(\lambda, \varepsilon_1))} f' \left( \frac{p_{\lambda, \mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k | g(\lambda, \varepsilon_1))} \right) \partial_{\theta_j} \ell_k(\mathbf{Y}_k | g(\lambda, \varepsilon_2)) \nabla_{\lambda} g_j(\lambda, \varepsilon_2) \\ &= \mathbb{E}_{\varepsilon_2} \sum_{j=1}^d \nabla_{\lambda} g_j(\lambda, \varepsilon_2) \mathbb{E}_{\varepsilon_1} \mathbb{E}_{\mathbf{Y}_k \sim \mathbb{P}_{\cdot | g(\lambda, \varepsilon_1)}} \frac{1}{\ell_k(\mathbf{Y}_k | g(\lambda, \varepsilon_1))} f' \left( \frac{p_{\lambda, \mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k | g(\lambda, \varepsilon_1))} \right) \partial_{\theta_j} \ell_k(\mathbf{Y}_k | g(\lambda, \varepsilon_2)).\end{aligned}\quad (6.14)$$

The above can be written as  $\mathcal{G} = \mathbb{E}_{\varepsilon_2} \sum_{j=1}^d \partial_{\theta_j} \tilde{\mathcal{G}}(g(\lambda, \varepsilon_2)) \nabla_{\lambda} g_j(\lambda, \varepsilon_2)$ , with  $\tilde{\mathcal{G}}$  being defined as

$$\tilde{\mathcal{G}} : \theta \mapsto \tilde{\mathcal{G}}(\theta) = \mathbb{E}_{\varepsilon_1} \mathbb{E}_{\mathbf{Y}_k \sim \mathbb{P}_{\mathbf{Y}_k | g(\lambda, \varepsilon_1)}} \left[ \frac{1}{\ell_k(\mathbf{Y}_k | g(\lambda, \varepsilon_1))} f' \left( \frac{p_{\lambda, \mathbf{Y}_k}(\mathbf{Y}_k)}{\ell_k(\mathbf{Y}_k | g(\lambda, \varepsilon_1))} \right) \ell_k(\mathbf{Y}_k | \theta) \right]. \quad (6.15)$$

Eventually, we obtain that  $\mathcal{I}_{D_f}^k$  is acceptable with  $\tilde{\mathcal{I}}_{D_f}^k = \tilde{I} + \tilde{\mathcal{G}}$ .  $\square$

In their work, Nalisnick and Smyth (2017) propose an alternative objective function to optimize, that we denote  $\mathcal{B}_{D_f}^k$  and call the lower-bound objective function. This function corresponds to a lower bound

of the mutual information. It is derived from an upper bound on the marginal distribution and relies on maximizing the likelihood. Their approach is only presented for  $f = -\log$ , we generalize the lower bound for any decreasing function  $f$ :

$$\mathcal{B}_{D_f}^k(\Pi_\lambda) = \mathbb{E}_{\theta \sim \Pi_\lambda} \mathbb{E}_{\mathbf{Y}_k \sim \mathbb{P}_{\mathbf{Y}_k|\theta}} \left[ f \left( \frac{\ell_k(\mathbf{Y}_k|\hat{\theta}^{\text{MLE}})}{\ell_k(\mathbf{Y}_k|\theta)} \right) \right] \quad (6.16)$$

where  $\hat{\theta}^{\text{MLE}}$  refers to the maximum likelihood estimator (MLE), i.e.  $\ell_k(\mathbf{Y}_k|\hat{\theta}^{\text{MLE}}) = \sup_{\theta \in \Theta} \ell_k(\mathbf{Y}_k|\theta)$ . It only depends on the likelihood and not on  $\lambda$ , making the gradient computation simpler:

$$\nabla_\lambda \mathcal{B}_{D_f}^k(\Pi_\lambda) = \mathbb{E}_{\varepsilon \sim \mathbb{P}_\varepsilon} \left[ \sum_{j=1}^d \partial_{\theta_j} \tilde{\mathcal{B}}_{D_f}^k(g(\lambda, \varepsilon)) \nabla_\lambda g_j(\lambda, \varepsilon) \right], \quad (6.17)$$

where:

$$\partial_{\theta_j} \tilde{\mathcal{B}}_{D_f}^k(\theta) = \mathbb{E}_{\mathbf{Y}_k \sim \mathbb{P}_{\mathbf{Y}_k|\theta}} \left[ \partial_{\theta_j} \log \ell_k(\mathbf{Y}_k|\theta) F \left( \frac{\ell_k(\mathbf{Y}_k|\hat{\theta}^{\text{MLE}})}{\ell_k(\mathbf{Y}_k|\theta)} \right) \right]. \quad (6.18)$$

This gradient is such that the following results stands.

**Proposition 6.2.** *The lower-bound objective function  $\mathcal{B}_{D_f}^k$  is an acceptable objective function. Its gradient is expressed in eq. (6.17).*

We emphasize that, given  $p_{\mathbf{Y}_k}^\lambda(\mathbf{Y}_k) \leq \ell_k(\mathbf{Y}_k|\hat{\theta}^{\text{MLE}})$  for all  $\lambda$ , we have  $\mathcal{B}_{D_f}^k(\Pi_\lambda) \leq \mathcal{I}_{D_f}^k(\Pi_\lambda)$  while  $f$  is decreasing. We recall that, in this work, we aim at using  $\delta$ -divergences (with  $\delta \in (0, 1)$ ), i.e.  $f = f_\delta$  with  $f_\delta(x) = \frac{x^\delta - \delta x - (1-\delta)}{\delta(\delta-1)}$ . Unfortunately,  $f_\delta$  is not a decreasing function, yet we can circumvent this by replacing it with  $\hat{f}_\delta$  defined hereafter, because  $D_{f_\delta} = D_{\hat{f}_\delta}$ :

$$\hat{f}_\delta(x) = \frac{x^\delta - 1}{\delta(\delta-1)} = f_\delta(x) + \frac{1}{\delta-1}(x-1). \quad (6.19)$$

The use of this function results in a more stable computation overall. Moreover, one argument for the use of  $\delta$ -divergences rather than the KL-divergence, is that we have a universal and explicit upper bound on the mutual information:

$$\mathcal{I}_{D_{\hat{f}_\delta}}^k(\Pi) = \mathcal{I}_{D_{\hat{f}_\delta}}^k(\Pi) \leq \hat{f}_\delta(0) = \frac{1}{\delta(1-\delta)}. \quad (6.20)$$

This bound can be an indicator on how well the mutual information is optimized, although there is no guarantee that it can be attained in general.

## Approximation of the gradients

The gradients of the objective functions proposed in this section can be approximated via successive Monte-Carlo, sampling from  $\mathbb{P}_\varepsilon$  and from the conditional distributions  $\mathbb{P}_{\mathbf{Y}_k|\theta}$ . Concerning the marginal density  $p_{\lambda, \mathbf{Y}_k}$  and the maximum likelihood  $\ell_k(\mathbf{Y}_k|\hat{\theta}^{\text{MLE}})$ , they can be approximated using:

$$\begin{aligned} p_{\lambda, \mathbf{Y}_k}(\mathbf{Y}_k) &= \mathbb{E}_{\varepsilon \sim \mathbb{P}_\varepsilon} \ell_k(\mathbf{Y}_k|g(\lambda, \varepsilon)) \approx \frac{1}{M} \sum_{m=1}^M \ell_k(\mathbf{Y}_k|g(\lambda, \varepsilon_m)) \quad \text{where } \varepsilon_1, \dots, \varepsilon_M \sim \mathbb{P}_\varepsilon; \\ \nabla_\lambda p_{\lambda, \mathbf{Y}_k}(\mathbf{Y}_k) &\approx \sum_{m=1}^M \sum_{j=1}^d \partial_{\theta_j} \ell_k(\mathbf{Y}_k|g(\lambda, \varepsilon_m)) \nabla_\lambda g_j(\lambda, \varepsilon_m) \quad \text{where } \varepsilon_1, \dots, \varepsilon_M \sim \mathbb{P}_\varepsilon; \\ \ell_k(\mathbf{Y}_k|\hat{\theta}^{\text{MLE}}) &\approx \max_{m \in \{1, \dots, M\}} \ell_k(\mathbf{Y}_k|g(\lambda, \varepsilon_m)) \quad \text{where } \varepsilon_1, \dots, \varepsilon_M \sim \mathbb{P}_\varepsilon. \end{aligned} \quad (6.21)$$

### 6.2.4 Adaptation for the constrained VA-RP

In this section, we suggest an adaptation of the variational optimization problem to incorporate some linear constraints on the prior. As mentioned in section 6.2.1 and developed in chapter 5, there exist specific constraints that would make the theoretical solution proper. This is also a way to incorporate expert knowledge to some extent. We consider  $R$  constraints of the form:

$$\forall r \in \{1, \dots, R\}, \mathcal{C}_r(\Pi_\lambda) = \mathbb{E}_{\theta \sim \Pi_\lambda} [a_r(\theta)] - b_r, \quad (6.22)$$

with  $a_r : \Theta \mapsto \mathbb{R}^+$  integrable and linearly independent functions, and  $b_r \in \mathbb{R}$ . We then adapt the optimization problem in eq. (6.7) to propose the following constrained optimization problem:

$$\begin{aligned} & \text{find } \Pi_{\lambda^*}^C \quad \text{where } \lambda^* \in \arg \max_{\lambda \in \Lambda} \mathcal{O}_{D_f}^k(\Pi_\lambda) \\ & \text{subject to } \forall r \in \{1, \dots, R\}, \mathcal{C}_r(\Pi_\lambda) = 0. \end{aligned} \quad (6.23)$$

The prior  $\Pi_{\lambda^*}^C$  is called the constrained VA-RP. When  $D_f$  corresponds to a  $\delta$ -divergence, the maximizer of the mutual information is expected to approximate the explicit solution whose density  $\pi^C$  satisfies

$$\pi^C(\theta) \propto J(\theta) \left( 1 + \sum_{r=1}^R \eta_r a_r(\theta) \right)^{1/\delta}, \quad (6.24)$$

where  $\eta_1, \dots, \eta_R \in \mathbb{R}$  are constants determined by the constraints.

According to the results stated in chapter 5, when the constraint is defined from only a function  $a$  that satisfies eq. (6.4), the reference prior  $\Pi^C$  under the constraint  $\mathcal{C}(\Pi^C) < \infty$  has a density of the form  $\pi^C \propto J \cdot a^{1/\delta}$  and is proper. We propose the following method to approximate the VA-RP under such constraint:

- Denote by  $J$  the Jeffreys prior. Compute the VA-RP  $\Pi_{\hat{\lambda}} \approx J$ , in the same manner as for the unconstrained case.
- Estimate the constants  $\mathcal{K}$  and  $c$  using Monte Carlo samples from the VA-RP, as:

$$\begin{aligned} \mathcal{K}_{\hat{\lambda}} &= \int_{\Theta} (\theta) a(\theta)^{1/\delta} d\Pi_{\hat{\lambda}}(\theta) \approx \int_{\Theta} J(\theta) a(\theta)^{1/\delta} d\theta = \mathcal{K}, \\ c_{\hat{\lambda}} &= \int_{\Theta} a(\theta)^{1+(1/\delta)} d\Pi_{\hat{\lambda}}(\theta) \approx \int_{\Theta} J(\theta) a(\theta)^{1+(1/\delta)} d\theta = c. \end{aligned} \quad (6.25)$$

- Since we have the equality for the constrained reference prior  $\Pi^C$ :

$$\mathbb{E}_{\theta \sim \Pi^C} [a(\theta)] = \frac{1}{\mathcal{K}} \int_{\Theta} J(\theta) a(\theta)^{1+(1/\delta)} d\theta = \frac{c}{\mathcal{K}}, \quad (6.26)$$

we compute the constrained VA-RP using the constraint :  $\mathbb{E}_{\theta \sim \Pi_{\lambda^*}} [a(\theta)] = c_{\hat{\lambda}} / \mathcal{K}_{\hat{\lambda}}$  to approximate  $\Pi_{\lambda^*}^C \approx \Pi^C$ .

The idea is to solve the constrained optimization problem as an unconstrained problem but with a Lagrangian as the objective function. We take the work of Nocedal and Wright (2006) as support.

We denote  $\eta$  the Lagrange multiplier. Instead of using the usual Lagrangian function, Nocedal and Wright (2006) suggest adding two terms in the optimization problem: (i) a term denoted  $\tilde{\eta}$ , which a vector with positive components that serves as penalization coefficients, and (ii) a vector  $\eta'$ , which can be thought of an initial estimate of  $\eta$ . The objective is to find a saddle point  $(\lambda^*, \eta^*)$  which is a solution of the updated optimization problem:

$$\max_{\lambda} \left( \min_{\eta} \mathcal{O}_{D_f}^k(\Pi_\lambda) + \sum_{r=1}^R \eta_r \mathcal{C}_r(\Pi_\lambda) + \sum_{r=1}^R \frac{1}{2\tilde{\eta}_r} (\eta_r - \eta'_r)^2 \right). \quad (6.27)$$

One can see that the third term serves as a penalization for large deviations from  $\eta'$ . The minimization on  $\eta$  is feasible because it is a convex quadratic, and we get  $\eta = \eta' - \tilde{\eta} \cdot \mathcal{C}(\Pi_\lambda)$ . Replacing  $\eta$  by its expression leads to the resolution of the problem:

$$\max_{\lambda} \mathcal{O}_{D_f}^k(\Pi_\lambda) + \sum_{r=1}^R \eta'_r \mathcal{C}_r(\Pi_\lambda) - \sum_{r=1}^R \frac{\tilde{\eta}_r}{2} \mathcal{C}_r(\Pi_\lambda)^2. \quad (6.28)$$

This motivates the definition of the augmented Lagrangian:

$$\mathcal{L}_A(\lambda, \eta, \tilde{\eta}) = \mathcal{O}_{D_f}^k(\Pi_\lambda) + \sum_{r=1}^R \eta_r \mathcal{C}_r(\Pi_\lambda) - \sum_{r=1}^R \frac{\tilde{\eta}_r}{2} \mathcal{C}_r(\Pi_\lambda)^2. \quad (6.29)$$

Its gradient has a form that is acceptable (see definition 6.1) while  $\mathcal{O}_{D_f}^k$  is acceptable as well:

$$\begin{aligned} \nabla_\lambda \mathcal{L}_A(\lambda, \eta, \tilde{\eta}) &= \nabla_\lambda \mathcal{O}_{D_f}^k(\Pi_\lambda) + \mathbb{E}_{\varepsilon \sim \mathbb{P}_\varepsilon} \left[ \sum_{j=1}^d \left( \sum_{r=1}^R \partial_{\theta_j} a_r(g(\lambda, \varepsilon)) (\eta_r - \tilde{\eta}_r \mathcal{C}_r(\Pi_\lambda)) \right) \nabla_\lambda g_j(\lambda, \varepsilon) \right] \\ &= \mathbb{E}_{\varepsilon \sim \mathbb{P}_\varepsilon} \left[ \sum_{j=1}^d \left( \partial_{\theta_j} \tilde{\mathcal{O}}_{D_f}^k(g(\lambda, \varepsilon)) + \sum_{r=1}^R \partial_{\theta_j} a_r(g(\lambda, \varepsilon)) (\eta_r - \tilde{\eta}_r \mathcal{C}_r(\Pi_\lambda)) \right) \nabla_\lambda g_j(\lambda, \varepsilon) \right]. \end{aligned} \quad (6.30)$$

In practice, the augmented Lagrangian algorithm takes the sequential form:

$$\begin{cases} \lambda^{t+1} = \arg \max_{\lambda \in \Lambda} \mathcal{L}_A(\lambda, \eta^t, \tilde{\eta}) \\ \forall r \in \{1, \dots, R\}, \eta_r^{t+1} = \eta_r^t - \tilde{\eta}_r \cdot \mathcal{C}_r(\pi_{\lambda^{t+1}}). \end{cases} \quad (6.31)$$

In our implementation,  $\eta$  is updated every 100 epochs. The penalty parameter  $\tilde{\eta}$  can be interpreted as the learning rate of  $\eta$ , we use an adaptive scheme inspired by Basir and Senocak, 2023 where we check if the largest constraint value  $\|\mathcal{C}(\Pi_\lambda)\|_\infty$  is higher than a specified threshold TH or not. If  $\|\mathcal{C}(\Pi_\lambda)\|_\infty > \text{TH}$ , we multiply  $\tilde{\eta}$  by  $v$ , otherwise we divide by  $v$ . We also impose a maximum value  $\tilde{\eta}_{\max}$ .

## 6.2.5 Posterior sampling from the implicit definition

Although our main object of study is the prior distribution, one needs to find the posterior distribution given an observed dataset  $\mathbf{y} \in \mathcal{Y}^k$  in order to do the inference on  $\theta$ . If the prior  $\Pi_\lambda$  admits a density  $\pi_\lambda$ , the posterior density is of the form :

$$p_\lambda(\theta | \mathbf{y}) = \frac{\pi_\lambda(\theta) \ell_k(\mathbf{y} | \theta)}{p_{\lambda, \mathbf{Y}_k}(\mathbf{y})}. \quad (6.32)$$

In practice, we cannot evaluate from this density, as the prior density is not known in general. This complicates the approximation of the posterior distribution and the computation of *a posteriori* quantities. To circumvent this problem, we suggest considering the posterior distribution of the latent variable  $\varepsilon$ , which admits the density:

$$q_\lambda(\varepsilon | \mathbf{y}) = \frac{q_\varepsilon(\varepsilon) \ell_k(\mathbf{y} | g(\lambda, \varepsilon))}{p_{\lambda, \mathbf{Y}_k}(\mathbf{y})}, \quad (6.33)$$

where  $q_\varepsilon$  is the probability density function of  $\varepsilon$ . The latter is known, so that the posterior density is known up to a constant. Therefore, it is possible to sample  $\theta$  *a posteriori* by sampling  $\varepsilon$  w.r.t. its posterior distribution  $\mathbb{P}_{\varepsilon | \mathbf{y}}$ :

$$\theta \sim \mathbb{P}_{T | \mathbf{y}} \iff \theta = g(\lambda, \varepsilon) \quad \text{with} \quad \varepsilon \sim \mathbb{P}_{\varepsilon | \mathbf{y}}. \quad (6.34)$$

This statement can be simply verified by deriving, for any bounded measurable function  $\varphi$  on  $\Theta$ :

$$\mathbb{E}_{\varepsilon \sim \mathbb{P}_{\varepsilon | \mathbf{y}}} \varphi(g(\lambda, \varepsilon)) = \mathbb{E}_{\varepsilon \sim \mathbb{P}_\varepsilon} \left[ \varphi(g(\lambda, \varepsilon)) \frac{\ell_k(\mathbf{y} | g(\lambda, \varepsilon))}{p_{\lambda, \mathbf{Y}_k}(\mathbf{y})} \right] = \mathbb{E}_{\theta \sim \Pi_\lambda} \left[ \varphi(\theta) \frac{\ell_k(\mathbf{y} | \theta)}{p_{\lambda, \mathbf{Y}_k}(\mathbf{y})} \right] = \mathbb{E}_{\theta \sim \mathbb{P}_{T | \mathbf{y}}} \varphi(\theta) \quad (6.35)$$

Samples of  $\varepsilon$  distributed w.r.t. its posterior distribution can be obtained using different methods from the expression of the density in eq. (6.33). For instance, variational inference techniques could be implemented. In this work, we choose MCMC sampling, namely an adaptive Metropolis-Hastings sampler.

## 6.3 Numerical experiments on different models

In this section, we apply our algorithm to different statistical models, and we assess the correct approximation of the Jeffreys prior by the obtained VA-RP.

As for the output of the neural networks, the activation function just before the output is different for each statistical model, the same can be said for the learning rate. In some cases, we apply an affine transformation on the variable  $\theta$  to avoid divisions by zero during training. In every test case, we consider simple networks for an easier fine-tuning of the hyperparameters and also because the precise computation of the loss function is an important bottleneck. The neural networks weights are initialized randomly w.r.t. a Gaussian distribution, and their biases are initially set at 0.

### 6.3.1 Multinomial model

#### Description of the model

The multinomial model is a case study that is simple in the sense that the target distributions —the Jeffreys prior and posterior— have explicit expressions and are part of a usual parametric family of proper distributions. This study allows testing and validating our algorithm on a multidimensional problem.

The multinomial distribution can be interpreted as the generalization of the binomial distribution for higher dimensions. In this case,  $\mathcal{Y} = \{y \in \{0, \dots, n\}^d, \sum_{i=1}^d y^i = n\}$  for a fixed  $n \in \mathbb{N}^*$ , and  $\Theta = (0, 1)^d$ . The data follow, conditionally to  $\theta$ , a multinomial distribution characterized by the likelihood:

$$\ell(y|\theta) = \frac{n!}{y^1! \dots y^d!} \prod_{i=1}^d \theta_i^{y^i}; \quad (6.36)$$

whose log-gradient w.r.t.  $\theta$  equals

$$\partial_{\theta_j} \log \ell(y|\theta) = \frac{y^j}{\theta_j}; \quad (6.37)$$

for all  $\theta \in \Theta$ ,  $y = (y^1, \dots, y^d) \in \mathcal{Y}$ . In this model, an explicit expression of the maximum likelihood estimator can be derived as

$$\ell_k(\mathbf{y}|\hat{\theta}^{\text{MLE}}) = \ell_k\left(\mathbf{y} \mid \frac{1}{nk} \sum_{i=1}^k y_i\right) \quad \text{for all } \mathbf{y} = (y_1, \dots, y_k) \in \mathcal{Y}^k; \quad (6.38)$$

where  $\frac{1}{nk} \sum_{i=1}^k y_i$  is the  $d$ -dimensional vector whose  $j$ -th coordinate is  $\frac{1}{nk} \sum_{i=1}^k y_i^j$ .

The reference prior (the prior maximizing the mutual information) is the Jeffreys prior which is a Dirichlet distribution for this model:  $\mathbf{J} = \text{Dir}_d\left(\frac{1}{2}, \dots, \frac{1}{2}\right)$ , which is proper. The Jeffreys posterior is a conjugate Dirichlet distribution:

$$\mathbf{J}_{\text{post}}(T|\mathbf{Y}_k) = \text{Dir}_d(\zeta) \quad \text{with} \quad \zeta = (\zeta_j)_{j=1}^d, \quad \zeta_j = \frac{1}{2} + \sum_{i=1}^k Y_i^j. \quad (6.39)$$

We recall that the probability density function  $p_{\text{Dir}_d(\zeta)}$  of a Dirichlet distribution is the following:

$$p_{\text{Dir}_d(\zeta)}(t) = \frac{\Gamma(\sum_{j=1}^d \zeta_j)}{\prod_{j=1}^d \Gamma(\zeta_j)} \prod_{j=1}^d t_j^{\zeta_j - 1} \quad \text{for all } t \in (0, 1)^d. \quad (6.40)$$

In chapter 3, we mentioned that although the Jeffreys prior is the prior that maximizes the mutual information, Berger and Bernardo (1992b) and Berger, Bernardo, and Sun (2015) argue that other priors for the multinomial model are more suited in terms of non-informativeness as the dimension of  $\theta$  increases. According to them, an appropriate reference prior should result from a sequential construction, referred to as the  $r$ -group reference prior. This prior is derived through hierarchical maximization of the mutual information where the parameter coordinates are grouped into  $r$  groups. Their process is explained in chapter 3 (section 3.3.2). According to the cited authors, while the Jeffreys prior corresponds to the  $r = 1$ -group reference prior using this construction, the  $d$ -group one should be favored for this model. Nevertheless, our approach

consists in approaching the maximizer of the mutual information, hence, the Jeffreys prior remains our target in this regard.

Regarding the neural network that defines the VA-RP in this study, we opt for a simple neural network with one linear layer and a Softmax activation function assuring that all components are positive and sum to 1. Explicitly, we have that:

$$\theta = \text{Softmax}(W^\top \varepsilon + b), \quad (6.41)$$

with  $W \in \mathbb{R}^{p \times d}$  the weight matrix and  $b \in \mathbb{R}^d$  the bias vector. The density function of  $\theta$  does not have a closed expression. The following results are obtained with  $d = 4$  and  $\delta = 0.5$  for the divergence and the lower bound is used as the objective function.

## Results

We notice (fig. 6.1) that the  $D_{f_\delta}$ -mutual information lies between 0 and  $1/\delta(1 - \delta) = 4$  during the training, which is coherent with the theory, the confidence interval is rather large because it is computed with few samples, but the mean value has an increasing trend. In order to obtain more reliable values for the mutual information, one can use more samples in the Monte Carlo estimates at the cost of heavier computations.

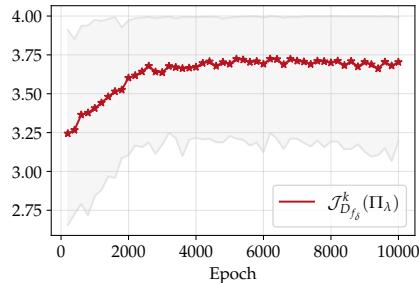


Figure 6.1: Monte Carlo estimation of the  $D_{f_\delta}$ -mutual information with  $\delta = 0.5$  and  $k = 200$  of  $\Pi_\lambda$  through the training of the neural network. The red curve is the mean value and the gray zone is the 95% confidence interval of the Monte-Carlo estimate. The learning rate used in the optimization is 0.0025.

Figures 6.2 and 6.3 propose qualitative comparisons between the obtained VA-RP at the end of the training and the Jeffreys prior. The first comparison takes the form of histograms of samples distributed w.r.t. each of the priors. The second differs from the fact that the samples are drawn w.r.t. the posterior distributions for a given observed dataset of size 10. We notice that although the shape of the fitted prior resembles the one of the Jeffreys prior, one can notice that it tends to put more weight towards the extremities of the interval (fig. 6.2). The posterior distribution however is quite similar to the target Jeffreys posterior on every component (fig. 6.3).

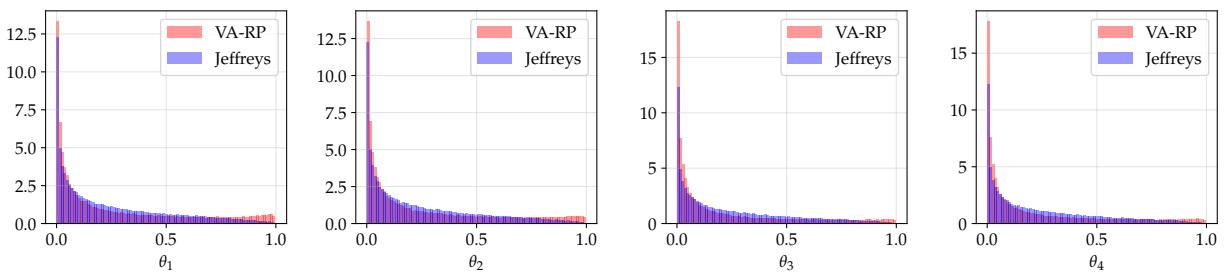


Figure 6.2: Histograms of the VA-RP and the Jeffreys prior for each dimension of  $\theta$ , each one is obtained from  $10^5$  samples.

Since the multinomial model is simple and computationally practical, we would like to quantify the effect of the divergence with different  $\delta$  values on the output of the algorithm. In order to do so, we utilize the

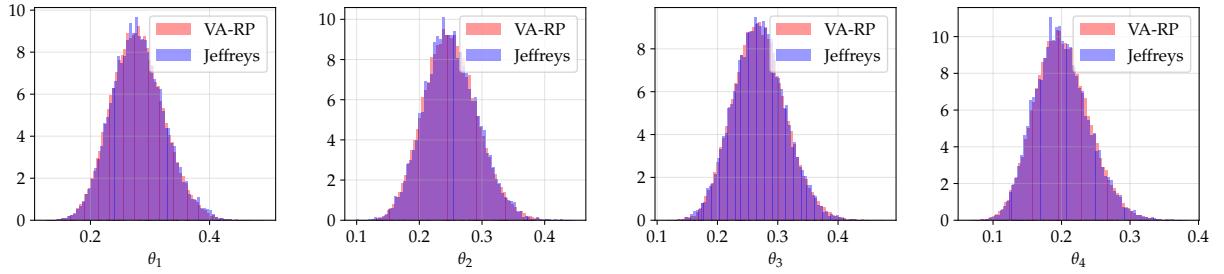


Figure 6.3: Histograms of the posteriors obtained from the VA-RP and the Jeffreys prior for each dimension of  $\theta$ , each one is obtained from  $10^5$  samples. The posteriors are derived given a random dataset of size 10, and distributed using  $\theta_{\text{true}} = (\frac{1}{4}, \dots, \frac{1}{4})$ .

maximum mean discrepancy (MMD) defined as :

$$\text{MMD}^2(\mathcal{H}; P||Q) = \mathbb{E}_{X, X' \sim P \otimes P}[K(X, X')] + \mathbb{E}_{Y, Y' \sim Q \otimes Q}[K(Y, Y')] - 2\mathbb{E}_{X, Y \sim P \otimes Q}[K(X, Y)]; \quad (6.42)$$

for any distributions  $P, Q$ , and where  $\mathcal{H}$  denotes a reproducible kernel Hilbert space whose  $K$  is a reproducing kernel. The MMD is used for instance in the context of two-sample tests (Gretton et al., 2012), whose purpose is to compare distributions. We use in our computations the Gaussian or RBF kernel :

$$K(\theta, \theta') = \exp(-0.5 \cdot \|\theta - \theta'\|_2^2) \quad (6.43)$$

for which the MMD is a metric, this means that the following implication:

$$\text{MMD}(\mathcal{H}; P||Q) = 0 \implies P = Q \quad (6.44)$$

is verified with the other axioms. In practice, we consider an unbiased estimator of the  $\text{MMD}^2$  given by:

$$\widehat{\text{MMD}}^2(P||Q) = \frac{1}{m(m-1)} \sum_{i \neq j} K(x_i, x_j) + \frac{1}{m'(m'-1)} \sum_{i \neq j} K(y_i, y_j) - \frac{2}{mm'} \sum_{i,j} K(x_i, y_j), \quad (6.45)$$

where  $(x_1, \dots, x_m)$  and  $(y_1, \dots, y_{m'})$  are sampled from  $P$  and  $Q$  respectively. In our case,  $P$  is the VA-RP and  $Q$  is the target Jeffreys prior. Since the MMD can be time-consuming or memory inefficient to compute in practice for very large samples, we consider only the last  $2 \cdot 10^4$  entries of our priors and posterior samples. According to Table 6.1, the difference between  $\delta$  values in terms of the MMD criterion is essentially inconsequential. One remark is that the mutual information tends to be more unstable as  $\delta$  gets closer to 1. The explanation is that when  $\delta$  tends to 1, we have the approximation :

$$\hat{f}_\delta(x) \approx \frac{x-1}{\delta(\delta-1)} + \frac{x \log(x)}{\delta}, \quad (6.46)$$

which diverges for all  $x$  because of the first term. Hence, we advise the user to avoid  $\delta$  values that are too close to 1. In the following, we use  $\delta = 0.5$  for the divergence.

The last quantitative study conducted for this model validates the result developed in chapter 4 about the asymptotic expansion of the  $D_{f_\delta}$ -mutual information. In the case  $d = 1$ , we show in fig. 6.4 that  $I_{D_{f_\delta}}^k(\mathbf{J})$  and  $\overline{\mathcal{I}}_{D_{f_\delta}}^k(\Pi_\lambda) := \mathcal{I}_{D_{f_\delta}}^k(\Pi_\lambda) + \frac{1}{\delta(1-\delta)}$  are close, where  $I_{D_{f_\delta}}^k(\mathbf{J})$  denotes the theoretical asymptotic rate w.r.t.  $k$  of  $\mathcal{I}_{D_{f_\delta}}^k(\mathbf{J})$ . According to the result of chapter 4, the latter admits the following asymptotic expansion:

$$\mathcal{I}_{D_{f_\delta}}^k(\mathbf{J}) \underset{k \rightarrow \infty}{\sim} \frac{-1}{\delta(1-\delta)} + I_{D_{f_\delta}}^k(\mathbf{J}), \quad \text{with} \quad I_{D_{f_\delta}}^k(\mathbf{J}) = \frac{-1}{\delta(1-\delta)^{3/2}} \left( \frac{\pi k}{2} \right)^{-\delta/2}. \quad (6.47)$$

This comparison supports two claims: (i) the  $D_{f_\delta}$  mutual information derived for a fixed  $k$  resembles to its asymptotic form w.r.t.  $k$ , which is the form that is maximized by the reference prior by definition; (ii) as the VA-RP maximizes the former, it is close to the maximal argument of the latter.

$\delta$	Prior	Posterior
0.1	$7.07 \times 10^{-2}$	$2.09 \times 10^{-3}$
0.25	$7.42 \times 10^{-2}$	$3.39 \times 10^{-3}$
0.5	$5.26 \times 10^{-2}$	$1.96 \times 10^{-3}$
0.75	$7.80 \times 10^{-2}$	$1.50 \times 10^{-3}$
0.9	$6.15 \times 10^{-2}$	$4.84 \times 10^{-4}$

Table 6.1: MMD values for different  $\delta$ -divergences at prior and posterior levels. As a reference on the prior level, when computing the criterion between two independent Dirichlet  $\text{Dir}_d(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2})$  distributions (ie the Jeffreys prior) on  $2 \cdot 10^4$  samples, we obtain an order of magnitude of  $10^{-3}$ .

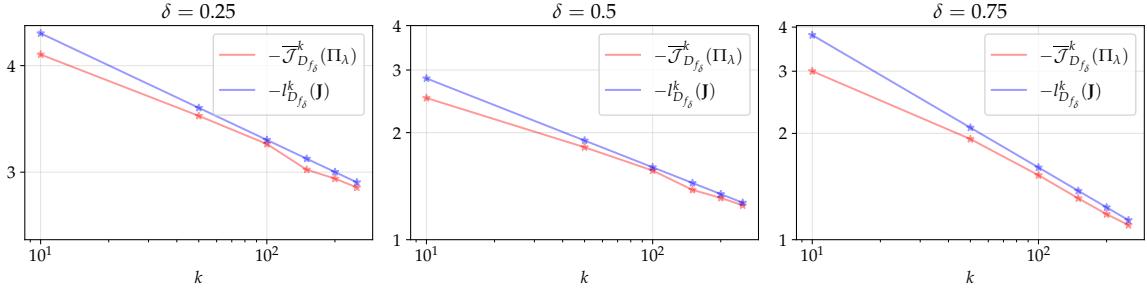


Figure 6.4: Values of the quantities  $-\bar{\mathcal{J}}_{D_{f_\delta}}^k(\Pi_\lambda)$  and  $-l_{D_{f_\delta}}^k(J)$  when  $d = 1$  as a function of  $k$ , in log scale, for different values of  $\delta$ :  $\delta = 0.25$  (left),  $\delta = 0.5$  (middle), and  $\delta = 0.75$  (right).

### 6.3.2 Normal model

#### Description of the model and results without constraints

The second case study that is developed in this chapter is the Gaussian model with known mean. Explicitly,  $\mathcal{Y} = \mathbb{R}$ ,  $\Theta = (0, \infty)$ , and the model is characterized by the likelihood

$$\ell(y|\theta) = \frac{1}{\sqrt{2\pi\theta}} \exp\left(-\frac{1}{2\theta}(y-\mu)^2\right), \quad (6.48)$$

with  $\mu \in \mathbb{R}$  being fixed. The gradient of the log-likelihood equals

$$\partial_{\theta_j} \log \ell_k(\mathbf{y}|\theta) = -\frac{k}{2\theta} + \frac{1}{2\theta^2} \sum_{i=1}^k (y_i - \mu)^2 \quad \text{for all } \mathbf{y} = (y_1, \dots, y_k) \in \mathcal{Y}^k. \quad (6.49)$$

There exists an explicit expression of the MLE as well:  $\ell_k(\mathbf{Y}_k|\hat{\theta}^{\text{MLE}}) = \ell_k(\mathbf{Y}_k|\frac{1}{k} \sum_{i=1}^k Y_i)$ . In this model, the Jeffreys prior admits a simple analytical density  $J(\theta) \propto 1/\theta$ , which is improper. It is conjugate with the likelihood: the distribution of the Jeffreys posterior is an inverse-gamma distribution:

$$J_{\text{post}}(T|\mathbf{y}) = \Gamma^{-1}\left(\frac{k}{2}, \frac{1}{2} \sum_{i=1}^k (y_i - \mu)^2\right) \quad \text{for all } \mathbf{y} \in \mathcal{Y}^k, \quad (6.50)$$

which is proper.

It is possible to implement the training of the VA-RP to approximate the improper Jeffreys prior. However, it becomes complicated to sample from the obtained prior, and one cannot directly compare it with the target. Yet, we can circumvent this problem by comparing the posterior distributions. In fig. 6.5, we present the result obtained from training a neural network with one layer and a Softplus activation function. The dimension of  $\varepsilon$  is set at  $p = 10$ . We observe that the evolution of the  $D_{f_\delta}$ -mutual information (fig. 6.5-left) remains stable, and the posterior distribution obtained from the VA-RP is close to the one obtained from Jeffreys (fig. 6.5-right).

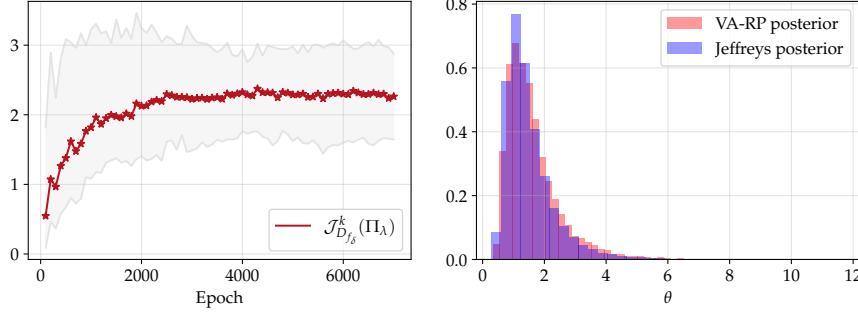


Figure 6.5: Left: Monte-Carlo estimation (mean in red, 95% confidence interval of the estimate in gray) of the  $D_{f_\delta}$ -mutual information with  $\delta = 0.5$  and  $k = 200$  of  $\Pi_\lambda$  through the training of the neural network. Right: histograms of the posteriors obtained from the VA-RP and from the Jeffreys prior. The posteriors are derived given a random dataset of size 10 and distributed using  $\theta_{\text{true}} = 1$ .

### Constraint incorporation

We also propose the introduction of constraints in this model, in order to ensure the VA-RP approximates a proper prior. We define  $a(\theta) = \frac{1}{\theta^{e_1} + \theta^{e_2}}$  with  $e_1 < 0 < e_2$ . The constant  $\mathcal{K}$  and  $c$  defined in section 6.2.4 are finite:

$$\forall \xi \geq 1, \int_0^\infty \frac{1}{\theta} \left( \frac{1}{\theta^{e_1} + \theta^{e_2}} \right)^\xi d\theta \leq \frac{1}{\xi} \left( \frac{1}{e_2} - \frac{1}{e_1} \right). \quad (6.51)$$

Thus, the constrained reference prior that is the target of the constrained VA-RP is the prior whose density  $\pi^C$  verifies  $\pi^C(\theta) \propto J(\theta)a(\theta)^{1/\delta}$  and is proper.

We implement the above taking  $\delta = 1/2$ ,  $e_1 = -1$ ,  $e_2 = 1$ , so that  $\mathcal{K} = 1/2$  and  $c = \pi/16$ , making the explicit constraint being  $\mathcal{C}(\Pi) = \mathbb{E}_{\theta \sim \Pi}[a(\theta)] = \pi/8$ . In fig. 6.6-left, we compare the prior distributions, which are both proper. The figure supports the robustness of the algorithm to approximate the reference prior even with added constraints. The posterior distributions are compared as well in fig. 6.6-right, which still reinforce the correctness of the approximation.

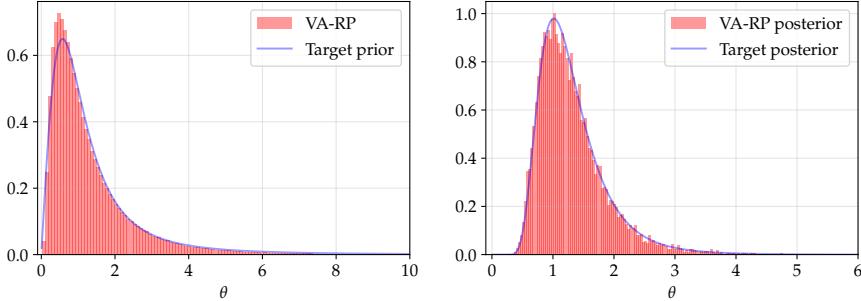


Figure 6.6: Comparisons between the VA-RP and the Jeffreys prior. Left: histogram of the VA-RP (red) and density of the target constrained reference prior  $\pi^C$  (blue). Right: Histogram of the posterior obtained from the VA-RP (red) and density of the posterior obtained from the target constrained reference prior (blue). The posteriors are derived given a random dataset of size 10 and distributed using  $\theta_{\text{true}} = 1$ .

## 6.4 Conclusion

In this work, we developed an algorithm to perform variational approximation of reference priors using the generalized definition of mutual information based on  $f$ -divergences that was developed in chapter 4. Additionally, because the reference priors of interest are often improper and often yield improper posteriors,

we adapted the variational definition of the problem to incorporate constraints that ensure the posteriors are proper, following the development conducted in [chapter 5](#).

Numerical experiments have been carried out on two test cases of different complexities in order to validate our approach. The results demonstrate the usefulness of our approach in estimating both prior and posterior distributions across the considered problems. This study thus provides a flexible solution for approximating the theoretical solutions that were proposed in this [part I](#) of the manuscript. We argue that the implementation can be adapted to a wide range of statistical models, including the ones where the target prior is not known explicitly. In [part II](#), an example of such statistical model will be studied.

To go further, future work could address the implementation of the method using deeper neural networks, allowing the VA-RP to cover a wider set of priors. However, we have observed that more complex neural network architectures tend to cause a significant rise of the computational cost, complicating the training process. The same applies when augmenting of the dimension of the parameter  $\theta$ .

We conclude that, beyond these challenges, this works proposes an approach that becomes especially relevant in cases where the target reference prior —or constrained reference prior— has a theoretical expression that is too complicated or too costly to derive.

## II

### SEISMIC FRAGILITY CURVE ESTIMATION



# 7

## Seismic fragility curves in probabilistic risk assessment studies

**Abstract** Seismic fragility curves are key quantities of interest in the seismic probabilistic risk assessment framework because they efficiently describe the fragility of a mechanical system of interest under a seismic excitation. They are studied since the 1980s on various types of data and characterization of the seismic motion and of the structure's response. In this chapter, we propose a review of the numerous methods that exist to estimate the curves. We particularly describe the role played by the characteristics of the available data to select a modeling of the fragility curve. This chapter is also the occasion to present some mechanical equipments that will be studied in the following chapters of this manuscript.

<b>7.1</b>	<b>Introduction</b>	<b>109</b>
<b>7.2</b>	<b>Data: from seismic signals to equipments failures</b>	<b>110</b>
<b>7.3</b>	<b>Modeling of the fragility curve</b>	<b>114</b>
7.3.1	Non-parametric modeling	114
7.3.2	Parametric modeling: probit-lognormal model	115
<b>7.4</b>	<b>Example of case studies</b>	<b>117</b>
7.4.1	An elasto-plastic oscillator	117
7.4.2	A piping system from a pressurized water reactor	118
7.4.3	Stacked structure for storage of packages	119
<b>7.5</b>	<b>Conclusion: which case study and which data for a Bayesian estimation of fragility curves?</b>	<b>120</b>

### 7.1 Introduction

The Seismic Probabilistic Risk Assessment (SPRA) defines a framework and a methodology for the study of seismic structural reliability. It has been introduced since 1968 (Cornell, 1968) to incorporate the seismic risk evaluation in the probabilistic risk assessment studies. The SPRA has been mostly developed and carried out in the 1980s on nuclear facilities (see e.g. Kennedy, Cornell, et al., 1980; Kennedy and Ravindra, 1984). It includes: the determination of the seismic hazard, the analysis of the seismic fragility of structural components, the evaluation of the risks combination and their consequences in the system, as described in the report of the Electric Power Research Institute (EPRI, 2013), which provides guidelines for its implementation. It is since widely used in the nuclear industry (e.g. Ellingwood, 1990; Y.-J. Park, Hofmayer, and Chokshi, 1998; Kennedy, 1999) and its consideration is established among the safety standards adopted by the International Atomic Energy Agency (IAEA, 2020).

Within the SPRA, seismic fragility curves represent a key asset that play a prominent role in the analysis of structural components' fragility step (see EPRI, 2011). They express in probabilistic terms the fragility of structures under seismic excitation. We can mention that they are also a tool of interest of performance-based earthquake engineering (PBEE; Ghobarah, 2001; Noh, Lallement, and Kiremidjian, 2014), which aims at relating performance objectives to level of damage to the structure. In any case, they are defined as a function of the seismic hazard, which—driven by the magnitude (M), the source-site distance (R), and other earthquake parameters—is reduced to a scalar value derived from the seismic signal: the intensity measure (IM), under the so-called “sufficiency assumption” (Cornell, 2004; Luco and Cornell, 2007). In practice, a fragility curve, denoted  $P_f$ , therefore expresses the probability of failure of a mechanical structure as a function of an IM value of interest such as peak ground acceleration (PGA) or pseudo-spectral acceleration (PSA), among others:

$$P_f(a) = \mathbb{P}(\text{"failure"} | \text{IM} = a). \quad (7.1)$$

Within the SPRA framework, the fragility curve expression is expected to be combined with the seismic hazard frequency to compute the annual damage frequency of the component  $C_f$ :

$$C_f = \int_0^\infty P_f(a) |dH(a)|. \quad (7.2)$$

Here,  $|dH(a)| = |\frac{d}{da}\mathbb{P}(\text{IM} > a)|da$  is defined from the mean annual frequency of exceedance of a ground motion of level  $a$ . To derive the damage frequency  $C_f$ , seismologists determine  $|dH|$  using not only records of seismic signals but also historical data, geological information, and geotechnical analysis. It should be noted that the sufficiency assumption was introduced to reduce estimation costs since it assumes that the fragility curve of a given structure is identical regardless of the seismic scenario. Actually, this assumption embeds some uncertainty in the problem since, as shown in Radu and Grigoriu, 2018; 2021, different seismic scenarios can lead to identical distributions of some IMs, although the underlying seismic signals have significantly different frequency contents. The uncertainty rooted in the fragility curve by this assertion could be classified as an *aleatoric* (irreducible) kind of uncertainty (we refer to chapter 1 for the definition of the uncertainty quantification framework).

In this second part of the manuscript, we focus on the estimation of seismic fragility curves. Thus, we do not address the determination of the seismic hazard  $H$ , which is the responsibility of seismologists, nor the derivation of the damage frequency of the component  $C_f$ . Nevertheless, estimating the fragility curve itself remains a daunting task. It is commonly done statistically, using different kind of methods depending on the source of data available, their characteristics, and their quantity. This chapter proposes a review of these methods. In the next section, we define the different kind of data involved when estimating seismic fragility curves. In particular, we present a stochastic seismic signal generator, and we present how the failure of mechanical equipment is generally defined. In section 7.3, different models and methods are presented, they are grouped in two categories: the non-parametric ones and the parametric ones. Explicit examples of case studies are then given in section 7.4, from a theoretical one to an experimental one for which really few data are available. Conclusive thoughts are proposed in section 7.5.

## 7.2 Data: from seismic signals to equipments failures

In the literature, different sources of data are exploited to estimate seismic fragility curves. We can cite, for instance: (i) expert assessments supported by test data (e.g. Kennedy, Cornell, et al., 1980; Kennedy and Ravindra, 1984; Zentner, Gündel, and Bonfils, 2017), (ii) experimental data (e.g. Y.-J. Park, Hofmayer, and Chokshi, 1998), (iii) empirical data from past earthquakes (e.g. Shinozuka, Feng, et al., 2000; Straub and Der Kiureghian, 2008; Lallement, Kiremidjian, and Burton, 2015; Buratti et al., 2017; Laguerre, Salehi, and Desroches, 2024), and (iv) analytical results obtained from various numerical models using artificial or natural seismic excitations (e.g. Ellingwood, 2001; S.-H. Kim and Shinozuka, 2004; Zentner, 2010; Koutsourelakis, 2010; Mai, Konakli, and Sudret, 2017; Trevlopoulos, Feau, and Zentner, 2019; F. Wang and Feau, 2020; Mandal, Siddhartha Ghosh, and Pujari, 2016; Z. Wang, Pedroni, et al., 2018; Z. Wang, Zentner, and Zio, 2018; Zhao, Yu, and Mo, 2020; Katayama et al., 2021; Gauchy, Feau, and Garnier, 2021; Khansefid et al., 2023; Lee, Kwag, and Ju, 2023). In every of these studies, each sample in the dataset regroups:

1. Information about a seismic ground motion. The considered ground motions are sometimes natural and sometimes artificial. The information can take several forms, as stated in the introduction it is generally used to derive one or several scalars called intensity measures (IMs).
2. Information about the response of the structure or the equipment of interest to the seismic excitation. To define the fragility curve, this information must permit to characterize the failure of the studied system.

Since available records of real seismic excitations for a given site are often scarce, it is common to construct a dataset of artificial earthquake accelerograms using a seismic signal generator. Various techniques exist for this purpose. According to Rezaeian and Der Kiureghian (2008), the methods can be categorized among the “source-based” ones, which model the occurrence of earthquake rupture at some sources and the propagation of seismic waves to the studied site, and the “site-based” ones, which model the seismic signals for the site of interest from the consideration of its characteristics and historical recorded earthquakes. A review of the first kind is proposed in Zerva, 1988, and a review of the second can be found in Shinotuka and Deodatis, 1988. As more recent examples for the latter, we can also cite Trevlopoulos, Feau, and Zentner, 2019; and Zentner and Poirion, 2012. In the following, we present a stochastic seismic signal generator that is proposed by Rezaeian and Der Kiureghian (2010). It lies among the site-based models, and has been implemented by Sainct et al. (2020), who calibrated it using 97 real accelerograms selected in the European Strong Motion Database for a magnitude  $M$  such that  $5.5 \leq M \leq 6.5$ , and a source-to-site distance  $R < 20$  km (Ambraseys et al., 2000). An example of one of these real signals is plotted in fig. 7.1.

As said in the introduction, this thesis does not seek to question thoroughly the seismic hazard, and we aim at providing methods that can be applied to any modeling of the seismic signals. Nevertheless, in the following chapters, our methods are applied and validated on different case studies that are presented later on. These case studies take the form of mechanical equipments that have been submitted to artificial seismic signals generated using the generator implemented by Sainct et al. (2020), and presented below.

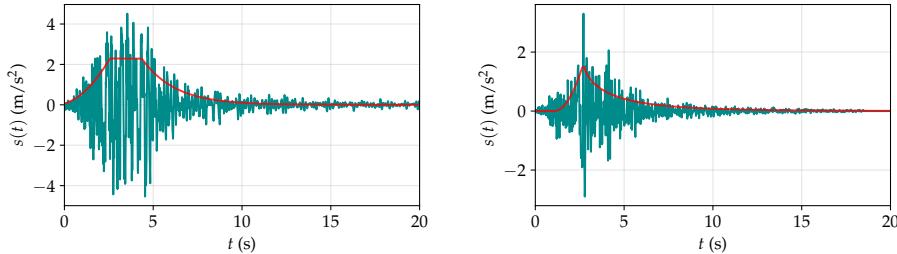


Figure 7.1: Example of two accelerogram records of real seismic signals (cyan). For each signal, its envelope  $q(t, \rho)$  is plotted in red.

### Seismic signals generator and IMs

The generator models a seismic excitation as a temporal signal that is the realization  $s : t \in [0, T] \mapsto s(t) \in \mathbb{R}$  of a random process defined on a probability space  $(\Omega, \Sigma, \mathbb{P})$ . The process presented here corresponds to a filtered stochastic white noise with time dependent parameters:

$$s(t) = s(t; w, \rho, \lambda) = q(t, \rho) \left[ \frac{1}{\sigma_f(t)} \int_{-\infty}^t h_f(t - \tau, \lambda) w(\tau) d\tau \right], \quad (7.3)$$

where  $w$  is a realization of a Gaussian white noise process  $W$ . In other terms  $W : [0, T] \times \Omega \rightarrow \mathbb{R}$  is such that for all  $t_1 < t_2, t_3 < t_4$ ,  $\int_{t_1}^{t_2} W(\tau) d\tau \sim \mathcal{N}(0, (t_2 - t_1)\sigma^2)$ , and  $\mathbb{E} \int_{t_1}^{t_2} W(\tau) d\tau \int_{t_3}^{t_4} W(\tau) d\tau = (\min(t_2, t_4) - \max(t_1, t_3))^+ \sigma^2$ .

The integral in eq. (7.3) corresponds to the filtering of  $w$ ,  $h_f(t, \lambda)$  being the impulse response function

(IRF) of the linear filter and  $\sigma_f^2$  being its variance:  $\sigma_f^2(t) = \int_{-\infty}^t h^2(t-\tau, \lambda(\tau)) d\tau$ . The IRF is defined by

$$h_f(t-\tau, \lambda) = \frac{\omega_f(\tau)}{\sqrt{1-\zeta_f^2}} \exp[-\zeta_f \omega_f(\tau)(t-\tau)] \sin\left(\omega_f(\tau) \sqrt{1-\zeta_f^2}(t-\tau)\right) \mathbb{1}_{t \geq \tau}, \quad (7.4)$$

with  $\omega_f(\tau) := \omega_0 + \frac{\tau}{T}(\omega_n - \omega_0)$ . The IRF  $h_f$  corresponds to the response of a linear oscillator with damping ratio  $\zeta_f$  and time-dependent natural frequency  $\omega_f(\tau)$ . It is parameterized by  $\lambda = (\omega_0, \omega_n, \zeta_f) \in (0, \infty)^2 \times [0, 1]$ .

The function  $q(t, \rho)$  is a non-negative modulating function that represents the envelope of the accelerogram, it is defined by

$$q(t, \rho) = \begin{cases} \rho_1 t^2 / T_1^2 & \text{if } 0 \leq t < T_1 \\ \rho_1 & \text{if } T_1 \leq t < T_2 \\ \rho_1 \exp[-\rho_2(t-T_2)\rho_3] & \text{if } T_2 \leq t \end{cases} \quad (7.5)$$

where  $\rho = (\rho_1, \rho_2, \rho_3, T_1, T_2) \in (0, \infty)^5$ .

A real accelerogram record is associated to parameters  $(\lambda, \rho)$  by conducting signal processing techniques. We refer to Rezaeian and Der Kiureghian, 2008 for more details. In fig. 7.1, the envelopes  $q(\cdot, \rho)$  associated with two real seismic signals are plotted. As previously announced, we consider  $N_r = 97$  real acceleration records from the European Strong Motion Database for a magnitude  $M$  such that  $5.5 \leq M \leq 6.5$  and a source-to-site distance  $R < 20$  km. They have been associated to parameters  $\bar{\phi}_i := (\bar{\lambda}_i, \bar{\rho}_i)$ ,  $i = 1, \dots, N_r$ .

To generate artificial seismic signals, we propose sampling realizations of the process  $s(\cdot; W, \phi)$ , where  $\phi \in \mathbf{D} := (0, \infty)^7 \times [0, 1]$  is stochastic and follows a distribution that corresponds to the empirical distribution of the  $(\bar{\phi}_i)_{i=1}^{N_r}$ . More precisely, the distribution of  $\phi$  is given by the Gaussian kernel density estimation of the one of the  $(\bar{\phi}_i)_{i=1}^{N_r}$ , i.e., its density  $p_\phi$  is given by

$$p_\phi(x) \propto \sum_{i=1}^{N_r} \exp\left(-\frac{1}{2}(x - \bar{\phi}_i)^\top \Sigma^{-1}(x - \bar{\phi}_i)\right) \mathbb{1}_{x \in \mathbf{D}}, \quad (7.6)$$

where  $\Sigma$  is derived from the  $(\bar{\phi}_i)_{i=1}^{N_r}$  (see Kristan, Leonardis, and Skočaj, 2011).

To sample from the generator, a tuple of parameters  $\tilde{\phi}$  is sampled given the distribution of  $\phi$ . Then, the filtered Gaussian noise is approximated by a truncation of the integral in eq. (7.3):

$$\int_{-\infty}^t h_f(t-\tau, \lambda) W(\tau) d\tau \approx \sum_{j=0}^M h_f(t-\tau(j), \lambda) \int_{\tau(j)}^{\tau(j+1)} W(\tau) d\tau, \quad (7.7)$$

where  $\tau(0), \dots, \tau(M)$  is a subdivision of a subset of  $(-\infty, t)$ . The above approximation can be sampled using  $\int_{\tau(j)}^{\tau(j+1)} W(\tau) d\tau \sim \mathcal{N}(0, (\tau(j+1) - \tau(j))\sigma^2)$ .

From a seismic signal  $s$  that is a realization of the process described above, different intensity measure indicators can be derived. The choice of the appropriate IM to estimate seismic fragility curves remains a complex question. According to Giovenale, Cornell, and Esteva (2004), the appropriateness of an IM must be defined in terms of efficiency, sufficiency, and hazard compatibility. However, the most efficient or sufficient IM is not the same for two different case studies (see Mackie and Stojadinović, 2001; Hariri-Ardebili and Saouma, 2016). Moreover, while we do not thoroughly research the best IM in this thesis, we will see in the following chapters that the best choice is not necessarily simply the one that is the most correlated with the structure's response. Below we propose a non-exhaustive list of common IMs, a more complete one can be found in Luco and Cornell, 2007:

- the peak ground acceleration (PGA) is defined as  $\text{PGA} = \max_{t \in [0, T]} |s(t)|$ ;
- the peak ground velocity (PGV) is defined as  $\text{PGV} = \max_{t \in [0, T]} \left| \int_0^t s(\tau) d\tau \right|$ ;
- the peak ground displacement (PGD) is defined as  $\text{PGD} = \max_{t \in [0, T]} \left| \int_0^t \int_0^\tau s(u) du d\tau \right|$ ;

- the pseudo spectral acceleration (PSA) at frequency  $f_L$  and damping ratio  $\xi$  is defined as  $\text{PSA} = (2\pi f_L)^2 \max_{t \in [0, T]} |x(t)|$ , where  $x$  is the solution of the linear equation

$$x''(t) + 2\xi 2\pi f_L x'(t) + (2\pi f_L)^2 x(t) = -s(t). \quad (7.8)$$

This IM is component-dependent since, in practice, the damping ratio  $\xi$  and the frequency  $f_L$  are evaluated from the mechanical characteristic of the studied system. Generally, the frequency  $f_L$  considered for deriving the PSA corresponds to the first mode of the structure's response under excitation. Equation (7.8) corresponds to the response equation of the linear single degree of freedom system associated to the structure.

For conducting the numerical experiments of this thesis,  $10^5$  seismic signals have been computed using this generator. In fig. 7.2, we draw the distribution of the PGA and the PSA (at frequency 5 Hz and damping ratio 1%) given by this generator. It is compared with the kernel density approximation of the PGA and PSA values of the 97 real seismic signals that have been used for fitting the generator. These comparisons show that the synthetic signals have realistic features. It is important to note, that the distribution of the IM produced by this generator may differ in practice from the distribution used to derive the damage frequency evoked in eq. (7.2). The latter is typically based on a more comprehensive characterization of the IM, incorporating not only recorded signals but also additional sources such as historical data, geological information, and geotechnical analyses.

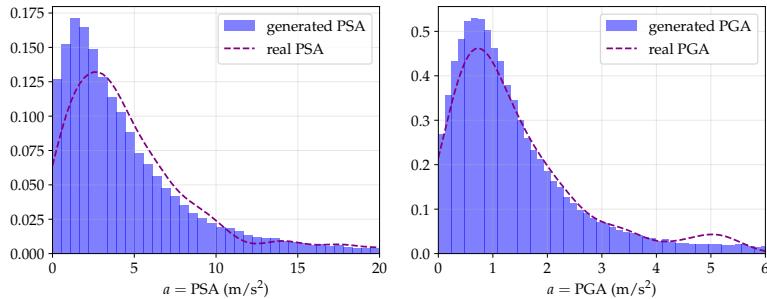


Figure 7.2: Histograms of IMs derived from  $10^5$  generated synthetic signals : the IM is the PSA (left) and the PGA (right). They are compared with the densities of IMs coming from real accelerograms estimated by Gaussian kernel estimation (dashed lines).

### Engineering demand parameter and failure

As evoked in the preamble of this section, numerous data sources exist for identifying the fragility (and the failure) of mechanical structures and components. In most cases a particular engineering demand parameter (EDP) of the system is focused on. The EDP is a scalar quantity that is observed (numerically or practically) during the seismic excitation, it can be the maximal displacement of a specific part of mechanical equipment for instance. More explicit examples will be given when specific case studies will be presented later on in this chapter. In those cases, the failure is defined when the EDP exceeds a threshold limit.

All in all, available datasets often take the form of tuples  $((S_i, D_i))_{i=1}^k$ , where  $S_i$  is the  $i$ -th seismic signal submitted to the system and  $D_i$  is the EDP that was observed during the experiment. In the case studies that are treated in this manuscript a specific IM is chosen, and a reduced dataset of the form  $((a_i, z_i))_{i=1}^k$  is studied, where  $a_i$  is the IM of the  $i$ -th seismic signal, and  $z_i = \mathbb{1}_{D_i > C}$ , where  $C$  is the threshold that defines the failure of the system. In other words,  $z_i = 1$  if the system has failed when submitted to the  $i$ -th seismic signal, and  $z_i = 0$  otherwise. The dataset permits to statistically estimate the fragility curve of the system of interest. In this thesis, we restrict the study to cases where the observations can be considered independent. Naturally, this restriction excludes certain case studies, such as experimental campaigns carried out on certain structures subjected to a series of seismic solicitations of increasing level, such as reinforced concrete structures, because these systems are damaged with each seismic solicitation. The advantage of considering binary data (failure or non-failure) is that it allows to deal with cases for which (i) the failure is defined by multiple criteria (e.g. stacked structures) and (ii) an EDP is not directly observed (e.g. electrical devices).

## 7.3 Modeling of the fragility curve

Different models and methods have been suggested in the literature to estimate seismic fragility curves. In the genesis of the SPRA, seismic fragility curves were estimated based on expert judgments, principally due to the scarcity of available data to conduct an accurate statistical estimation (see Kennedy, Cornell, et al., 1980). Of course, the accuracy of such an estimation is not guaranteed either, and its reliability depends on that of the experts.

Nowadays, estimates of seismic fragility curves are mostly derived using statistical techniques, leveraging (i) dataset that are sometimes less scarce and more precise, and also (ii) statistical techniques that are more efficient. The expert judgment still complement those in some works in the literature.

Non-parametric and parametric methods are both used in the literature to estimate seismic fragility curves. In the former, the probabilistic relation between the failure and the IM is sought. When an EDP is available, authors seek to estimate first the probabilistic relation between the IM and the EDP, to deduce the resulting fragility curve. In the latter, the fragility curve is assumed to belong to a parameterized set, reducing the problem to the estimation of these parameters. While the first method is more general in the sense that it covers a wider range of possible estimates, it is also often less efficient than the second. Indeed, reducing the problem to a finite-dimensional one allows providing satisfying estimates with fewer data. However, their reliability depends on the trustworthiness of the chosen parametric modeling.

In the following, we review briefly the state-of-the art on the estimation of seismic fragility curves. We start by a review of non-parametric methods in section 7.3.1. In that section we also describe a non-parametric estimation that is based on Monte-Carlo estimates, and that can serve as a reference when a large dataset is available. In section 7.3.2 we discuss the parametric modelings of the fragility curves, and we present the most common one: the probit-lognormal model.

### 7.3.1 Non-parametric modeling

Most of non-parametric models for seismic fragility curves estimation leverage—if possible—the knowledge of an EDP that describes more precisely the response of the structure under seismic excitation than just the binary outcome—failure or non-failure. In this case it is possible to estimate the conditional distribution of the EDP to the IM (or IMs) via a surrogate modeling of the system  $\text{EDP} = \mathcal{M}(\text{IM})$ .

Among common surrogates we can mention the Gaussian processes (GPs). As examples of their use for seismic fragility curves estimation, we can cite Gidaris, Taflanidis, and Mavroeidis, 2015, in which GPs are used to estimate the EDP given multiple parameters characterizing the ground motion; and Gauchy, Feau, and Garnier, 2024, in which the GPs-based estimation of the EDP is coupled with a sensitivity analysis of the fragility curve to the mechanical parameters of the studied system.

We also mention polynomial chaos expansion (PCE) as a surrogate numerously used for seismic fragility curves estimation. For instance, PCEs are combined in Mai, Spiridonakos, et al., 2016 with non-linear autoregressive with exogenous input models to estimate the temporal response of the mechanical system as a function of the seismic signal. In X. Zhu, Broccardo, and Sudret, 2023, a stochastic PCE is conducted, consisting in the addition of a latent variable and an additive noise to the deterministic PCE expression.

Outside surrogate modeling, what one would call machine-learning-based techniques are also implemented to estimate seismic fragility curves. For examples, we quote linear regression or generalized linear regression (Lallemand, Kiremidjian, and Burton, 2015), classification-based methods (see the review suggested in Kiani, Camp, and Pezeshk, 2019) such as logistic regression (as in Bernier and Padgett, 2019) or support vector machine (as in Saint et al., 2020). In the latter work, support vector machine is used to classify EDPs whether they led to failures or non-failures as a function of combinations of IMs. We also mention artificial networks based methods, such as used in Mitropoulou and Papadrakakis, 2011; Z. Wang, Pedroni, et al., 2018.

### A reference fragility curve constructed with a large dataset

All the methods cited above provide estimations whose efficiency will depend on the size of the available dataset and the correctness of the assumptions they involve on the probability relation between the system's response and the IMs. It is possible to minimize such assumptions considerably, trying to estimate directly the probabilities  $\mathbb{P}(\text{"failure"}|\text{IM} = a)$ , for all value of  $a$ , via Monte-Carlo estimates for instance. Such a methodology must give the best estimation of the fragility curve in terms of robustness, in the sense that it does not rely on any assumption. Of course, one does not have infinitely many data for all value of  $a$  to provide an exact estimation of the whole curve, yet it is possible to approximate the curve locally in different sub-areas of the domain  $\mathcal{A}$  in which  $a$  lives. More explicitly consider a dataset  $((a_i, z_i))_{i=1}^N$  of IMs and binary outcomes ( $z_i = 1$  if the  $i$ -th seismic signal led to a failure, and 0 otherwise), and choose  $N_c$  clusters of the  $(a_i)_{i=1}^N$  that we denote  $(K_j)_{j=1}^{N_c}$ :  $\bigcup_{j=1}^{N_c} K_j = \{a_i, i = 1, \dots, N\}$ . Then it is possible to approximate the fragility curve evaluated at the centroids  $(c_j)_{j=1}^{N_c}$  of the clusters:

$$P_f^{\text{MC}}(c_j) = \frac{1}{n_j} \sum_{i, a_i \in K_j} z_i, \quad (7.9)$$

where  $n_j$  is the sample size of cluster  $K_j$ . This non-parametric estimation is implemented by Trevlopoulos, Feau, and Zentner (2019), who suggest defining the clusters  $(K_j)_{j=1}^{N_c}$  using K-means since the IM values in available datasets are generally not uniformly distributed.

When the number of available data is small and when they are poorly diverse in terms of values of the IM, this estimation method becomes limited for estimating seismic fragility curves. However, in the other case, we consider that it provides a robust result. For this reason, while this thesis addresses the estimation of fragility curves when few data are available, we will evaluate our methods on case studies for which a large validation dataset exists. The latter is used to derive what we call a reference fragility curve that is identified to  $P_f^{\text{MC}}$ , and which will be compared with the estimates provided by our methods.

### 7.3.2 Parametric modeling: probit-lognormal model

#### The probit-lognormal model

Although based on stronger assumptions on the structure's response than non-parametric methods, parametric fragility curves were historically considered, especially in cases with small dataset limited to binary outcome (i.e. no EDP is available). In the SPRA and PBEE frameworks, the so-called probit-lognormal model was chosen, and it remains prevalent to this day (see e.g. Shinozuka, Feng, et al., 2000; Straub and Der Kiureghian, 2008; Zentner, 2010; F. Wang and Feau, 2020; Mandal, Siddhartha Ghosh, and Pujari, 2016; Zhao, Yu, and Mo, 2020; Ellingwood, 2001; S.-H. Kim and Shinozuka, 2004; Mai, Konakli, and Sudret, 2017; Trevlopoulos, Feau, and Zentner, 2019; Katayama et al., 2021; Lee, Kwag, and Ju, 2023). This model consists in introducing a parameter  $\theta = (\alpha, \beta)$  and defining the fragility curve as follows

$$P_f(a) = \mathbb{P}(\text{"failure"}|\text{IM} = a) = \Phi\left(\frac{\log a - \log \alpha}{\beta}\right), \quad (7.10)$$

where  $\Phi$  is the cumulative distribution function of a standard normal variable. An example of such curves is given in fig. 7.3. We mention that, sometimes, the parametrization of the model is slightly different. As recalled by Zentner, Gündel, and Bonfils, 2017, the distinguishing of epistemic and aleatoric uncertainty in  $\beta$  is commonly suggested, leading to rewriting eq. (7.10) with  $\beta = \sqrt{\beta_U^2 + \beta_R^2}$ . The literature suggests that the uncertainty embedded in  $\alpha$  is only epistemic.

Different strategies exist to estimate the two parameters, namely the median  $\alpha$  and the log standard deviation  $\beta$ . Historically, maximum likelihood estimation (MLE) techniques are recommended. When the data are independent, the bootstrap technique can be used to obtain confidence intervals relating to the size of the sample considered (e.g. Shinozuka, Feng, et al., 2000; Zentner, 2010; F. Wang and Feau, 2020). To describe this technique, we define first the associated statistical model: the observations are modeled as

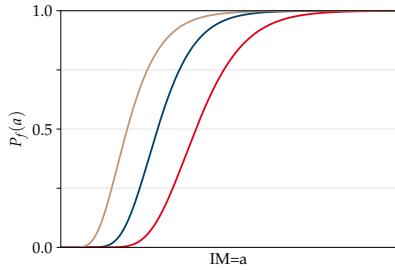


Figure 7.3: Examples of probit-lognormal fragility curves.

realizations of the random variable  $(A, Z) \in \mathcal{A} \times \{0, 1\}$ . It is supposed that observations are independent conditionally to  $\theta$  and that  $(A, Z)$  are distributed conditionally to  $\theta \in \Theta = (0, \infty)^2$  as:

$$A|\theta \sim A \sim H, \quad \text{and} \quad Z|A, \theta \sim \mathcal{B}\left(\Phi\left(\beta^{-1} \log \frac{A}{\alpha}\right)\right), \quad (7.11)$$

where  $H$  is the distribution of  $A$  and  $\mathcal{B}(p)$  denotes a Bernoulli distribution of mean  $p$ . Thus, the likelihood  $\ell_k$  given observations  $(\mathbf{a}^k, \mathbf{z}^k)$  with  $\mathbf{a}^k = (a_i)_{i=1}^k$  and  $\mathbf{z}^k = (z_i)_{i=1}^k$  is expressed as

$$\ell_k(\mathbf{z}^k, \mathbf{a}^k | \theta) = \prod_{i=1}^k \ell(z_i, a_i | \theta) = \prod_{i=1}^k \Phi\left(\frac{\log a_i - \log \alpha}{\beta}\right)^{z_i} \left(1 - \Phi\left(\frac{\log a_i - \log \alpha}{\beta}\right)\right)^{1-z_i} h(a_i), \quad (7.12)$$

where  $h$  is the p.d.f. of  $A$ . Since the  $(h(a_i))_i$  are constants of  $\theta$  and are not known in general, it is common to consider the following alternative form of the likelihood:

$$\ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) = \prod_{i=1}^k \ell(z_i | a_i, \theta) = \prod_{i=1}^k \Phi\left(\frac{\log a_i - \log \alpha}{\beta}\right)^{z_i} \left(1 - \Phi\left(\frac{\log a_i - \log \alpha}{\beta}\right)\right)^{1-z_i}. \quad (7.13)$$

The maximum likelihood estimator is  $\theta^{\text{MLE}}(\mathbf{z}^k, \mathbf{a}^k) = \arg \max_{\theta \in \Theta} \ell_k(\mathbf{z}^k, \mathbf{a}^k | \theta)$ . The bootstrap technique consists in deriving a stochastic estimator  $\hat{\theta}_k^{\text{MLE}}(\mathbf{z}^k, \mathbf{a}^k)$  whose distribution is defined from expressing it as:  $\hat{\theta}^{\text{MLE}}(\mathbf{z}^k, \mathbf{a}^k) = \theta^{\text{MLE}}((z_{U_i}, a_{U_i})_{i=1}^k)$ , where  $U_1, \dots, U_k$  are i.i.d. random variables distributed w.r.t. a uniform distribution in  $\{1, \dots, k\}$ .

### About the implementation of the Bayesian framework to estimate $\theta = (\alpha, \beta)$

As a strategy that allows to estimate the parameters defining the fragility curve, the Bayesian framework has recently become increasingly popular in seismic fragility analysis (see e.g. Gardoni, Der Kiureghian, and Mosalam, 2002; Z. Wang, Zentner, and Zio, 2018; Katayama et al., 2021; Koutsourelakis, 2010; Damblin et al., 2014; Tadinada and A. Gupta, 2017; Kwag and A. Gupta, 2018; Jeon et al., 2019; Tabandeh, Asem, and Gardoni, 2020). It is praised for its capacity to solve the irregularity issues encountered when estimating fragility curves with classical methods. For instance, the MLE-bootstrap method is known to lead to unrealistic estimates such as unit-step functions when few data are available. However, a challenge remains when implementing the Bayesian framework, and lies in the selection of the prior. As a matter of fact, as announced since [chapter 1](#), selecting a prior is a critical step in Bayesian analysis, and when exploring the literature, a wide range of different consideration can be found regarding its construction.

For example, in Tadinada and A. Gupta, 2017 and Kwag and A. Gupta, 2018, the median prior values come from equivalent linearized mechanical models. In Z. Wang, Zentner, and Zio, 2018, both aleatory and epistemic uncertainties are taken into account in the parametric model originally introduced in Kennedy, Cornell, et al., 1980: an artificial neural network is trained and used to characterize (i) the aleatory uncertainty and (ii) the prior median value of  $\alpha$ , while the associated epistemic uncertainty is taken from the existing literature. The log-normal prior distribution of  $\alpha$  is then updated with empirical data. In Katayama et al.,

2021, the results of incremental dynamic analysis are used to obtain a prior value of  $\alpha$ , whereas the prior value of  $\beta$  is determined through a parametric study. This results in satisfactory convergence, whatever its target value, before application to practical problems. In Straub and Der Kiureghian, 2008, the authors mainly focus on the implications for fragility analyses of statistical dependencies within the data. The prior is defined as the product of a normal distribution for  $\ln(\alpha)$ , and the improper distribution  $1/\beta$  for  $\beta$ . The definition of the normal distribution is based on engineering assessments. This prior was preferred to  $1/\alpha$  on the grounds that it led to unrealistically large posterior values of  $\alpha$ . A sensitivity analysis is further performed to examine the impact of the choice of the prior distribution on the final results. Finally, we note that the Bayesian framework is also relevant for fitting numerical models (e.g., mathematical expressions based on engineering assessments or physics-based models) to experimental data in order to estimate fragility curves (Gardoni, Der Kiureghian, and Mosalam, 2002; Tabandeh, Asem, and Gardoni, 2020) or meta-models such as logistic regressions (Koutsourelakis, 2010; Jeon et al., 2019).

## 7.4 Example of case studies

### 7.4.1 An elasto-plastic oscillator

The first case study that we present is a single-degree-of-freedom elasto-plastic oscillator with kinematic hardening. This simple mechanical system illustrates the essential features which can be found in the non-linear responses of some real-world structures under seismic excitation and has, for this reason, already been used in several studies (Trevlopoulos, Feau, and Zentner, 2019; Saint et al., 2020; Gauchy, Feau, and Garnier, 2021). In addition, it provides reference results as a reasonable numerical cost. It is depicted in fig. 7.4, and its equation of motion when submitted to an excitation  $s$  is the following:

$$x''(t) + 2\xi 2\pi f_L x'(t) + f^{\text{nl}}(t) = -s(t), \quad (7.14)$$

with  $x''$  and  $x'$  respectively the relative acceleration and velocity of the mass,  $\xi$  the damping ratio,  $f_L$  the circular frequency, and  $f^{\text{nl}}$  the nonlinear resisting force.

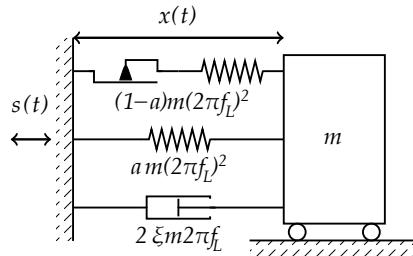


Figure 7.4: Elasto-plastic oscillator of mass  $m$  with kinematic hardening, with parameters  $f_L = 5$  Hz and  $\xi = 2\%$ . The yield limit is  $Y = 5.10^{-3}$  m, and the post-yield stiffness is 20% of the elastic stiffness, i.e.,  $a = 0.2$ .

The relevant EDP is the absolute maximum value of the mass' displacement, i.e.,  $\text{EDP} = \max_{t \in [0, T]} |x(t)|$ , where  $T$  is the duration of the seismic excitation. In fig. 7.5, we plot that EDP that has been derived for the  $10^5$  seismic signals generated and presented in section 7.2.

Since a very large number of data are available for this case study, it is possible to derive a reference fragility curve using the non-parametric method that is described in section 7.3.1. Such reference fragility curves are plotted in fig. 7.6, for different thresholds  $C$  that define the failure. In general, the failure criterion that corresponds to the 90%-level quantile of the maximum displacements calculated with the  $10^5$  artificial signals is chosen, i.e.,  $C = 8.0 \cdot 10^{-3}$  m. The reference curves are compared with the probit-lognormal fragility curves (described in section 7.3.2) where  $\theta$  is estimated by MLE using the full dataset as well. This comparison is presented for both the PSA and the PGA as IM. It demonstrates that the probit-lognormal model allows a good approximation of the reference curve for this case study.

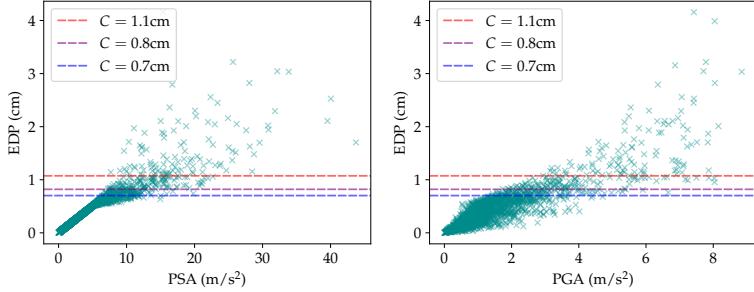


Figure 7.5: Results of the  $10^5$  simulations conducted on the elasto-plastic oscillator. Each cross is an element of the dataset (IM, EDP) where the IM is the PSA (left) and the PGA (right). Different critical rotation thresholds  $C$  are plotted in dashed lines. They yield different proportions of failures in the dataset: respectively 95% (red), 90% (purple) and 85% (blue).

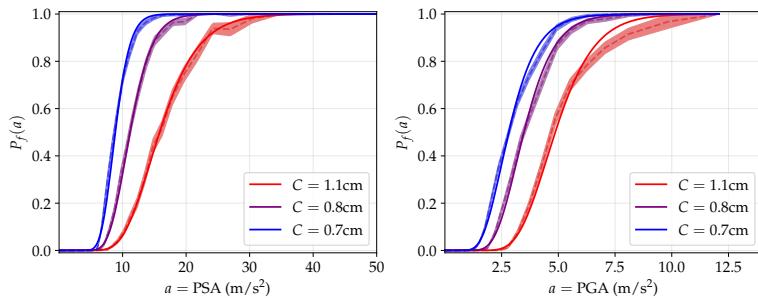


Figure 7.6: Reference non-parametric fragility curves of the elasto-plastic oscillator obtained via Monte Carlo estimates (dashed lines) surrounded by their 95% confidence intervals, for different critical displacement threshold  $C$  with (left) the PSA and (right) the PGA as IM. The thresholds yield different proportions of failures in the dataset: respectively 95% (red), 90% (purple) and 85% (blue). For each value of  $C$  are plotted (same color, solid line) the corresponding probit-lognormal MLE.

#### 7.4.2 A piping system from a pressurized water reactor

The case study presented in this section is a piping system that was tested on the Azalee shaking table at the EMSI laboratory of CEA/Saclay, as shown in fig. 7.7-left. Figure 7.7-right depicts the finite element model (FEM), based on beam elements and implemented through the proprietary FE code CAST3M (CEA, 2019). The validation of the FEM was carried out thanks to an experimental campaign described in Touboul, Sollogoub, and Blay, 1999.

The mock-up comprises a carbon steel TU42C pipe with an outer diameter of 114.3 mm, a thickness of 8.56 mm, and a 0.47 elbow characteristic parameter. This pipe, filled with water without pressure, includes three elbows, with a valve-mimicking mass of 120 kg, constituting over 30% of the mock-up's total mass. One end of the mock-up is clamped, while the other is guided to restrict displacements in the X and Y directions. Additionally, a rod is positioned atop the specimen to limit mass displacements in the Z direction (refer to fig. 7.7-right). During testing, excitation was applied exclusively in the X direction.

In order to conduct comparative performance studies, numerous simulations have been performed. They were carried out from a subset of  $8 \cdot 10^4$  of the  $10^5$  artificial seismic signals. Nevertheless, as in practice the piping system is located in a building, the artificial signals were filtered using a fictitious 2% damped linear single-mode building at 5 Hz, which corresponds to the first eigenfrequency of the 1% damped piping system. In such a situation, for some seismic signals, the behavior of the piping system is nonlinear. Regarding the nonlinear constitutive law of the material, a bilinear law exhibiting kinematic hardening was used to reproduce the overall nonlinear behavior of the mock-up with satisfactory agreement compared to the results of the seismic tests (Touboul, Sollogoub, and Blay, 1999). Following the recommendation in Touboul, Blay, et al., 2006, we consider that the engineering demand parameter (EDP) is the out-of-plane rotation of the

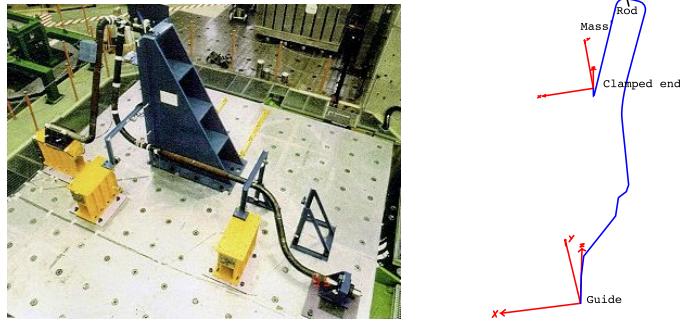


Figure 7.7: (left) Overview of the piping system—which is part of a French pressurized water reactor—on the Azalee shaking table and (right) associated finite element model.

elbow near the clamped end of the mock-up. As a result we have a dataset of  $8 \cdot 10^4$  independent tuples of the form (IM, EDP) for different IMs. In fig. 7.8 are plotted the elements of the datasets (PSA, EDP) and (PGA, EDP), along with different critical thresholds.

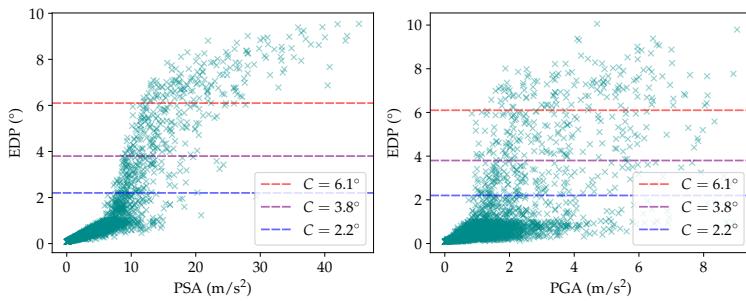


Figure 7.8: Results of the  $8 \cdot 10^4$  numerical simulations conducted on the piping system. Each cross is an element of the dataset (IM, EDP) where the IM is the PSA (left) and the PGA (right). Different critical rotation thresholds  $C$  are plotted in dashed lines. They yield different proportions of failures in the dataset: respectively 95% (red), 90% (purple) and 85% (blue).

The complete set of  $8 \cdot 10^4$  simulations provides a satisfactory dataset to derive a reference of the seismic fragility curve for this case study, still using the non-parametric method depicted in section 7.3.1. In fig. 7.9, we compare this non-parametric reference with the parametric fragility curve using the probit-lognormal model (section 7.3.2) where  $\alpha$  and  $\beta$  are estimated by MLE using the  $8 \cdot 10^4$  data items as well. This comparison is presented considering different critical rotation thresholds  $C$  of the equipment for both the PSA and the PGA. First, it should be noted that, with the PGA as IM, it is not possible to completely describe the fragility curve. For the maximum PGA values observed, the failure probabilities stagnate between 0.5 and 0.8 depending on the failure criterion considered. Therefore, the PGA is not the most suitable IM of the two. For the type of structure considered here, this point is well documented in the literature. Then, the comparisons demonstrate, in our setting, the adequacy of a probit-lognormal modeling of the fragility curves, even with high failure thresholds. Its bias with non-parametric fragility curve exists but remains limited. When the number of observations is small, this model bias is negligible in front of the uncertainty on the estimates.

### 7.4.3 Stacked structure for storage of packages

The case study considered hereafter concerns a freestanding stacked structure composed of three pallets intended for the storage of packages. As shown in fig. 7.10, it is a square-based structure 3 meters high with a slenderness of 2.4.

The mass of one package is equal to 265 kg, whereas the mass of one pallet is equal to 60 kg. The total mass of the structure is equal to 3,360 kg. The pallets are made of 3 mm-thick hollow-section aluminum square tubes. These tubes are welded and form the base and uprights of the pallets. The base of a pallet

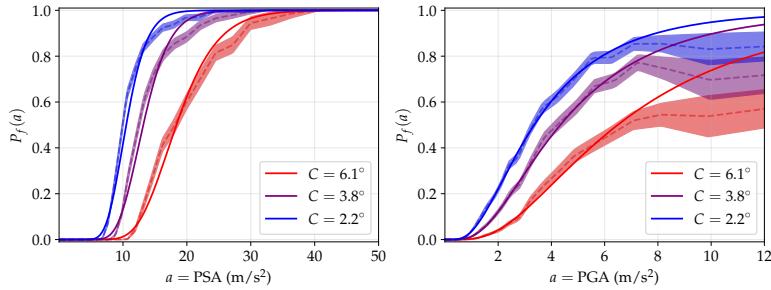


Figure 7.9: Reference non-parametric fragility curves of the piping system obtained via Monte Carlo estimates (dashed lines) surrounded by their 95% confidence intervals, for different critical rotation threshold  $C$  with (left) the PSA and (right) the PGA as IM. The thresholds yield different proportions of failures in the dataset: respectively 95% (red), 90% (purple) and 85% (blue). For each value of  $C$  are plotted (same color, solid line) the corresponding probit-lognormal MLE.

supports four freestanding packages, whereas the uprights support the upper freestanding pallet(s). The stacked structure was studied in Beylat, 2020 by means, among others, of a complete experimental campaign, which was carried out on the 1D shaking table “Vesuve” of CEA/Saclay. In particular, the stack was subjected to 21 of the 97 real signals that were used to fit the generator depicted in section 7.2. For this structure, the EDP corresponds to the maximal displacement of the top of the stack. An example of a test result is shown in fig. 7.10, which depicts the horizontal displacements over time of the top of the stack. The initial position is indicated in red, while the different positions in time are indicated in blue. Due to uplift, sliding, and rotation motions, the top of the stack exceeds, for 2 of the 21 tests, the admissibility criterion, which is materialized in black. Only a little number of experimental results are available for this case study, so that it is not possible to derive a reference fragility curve here.

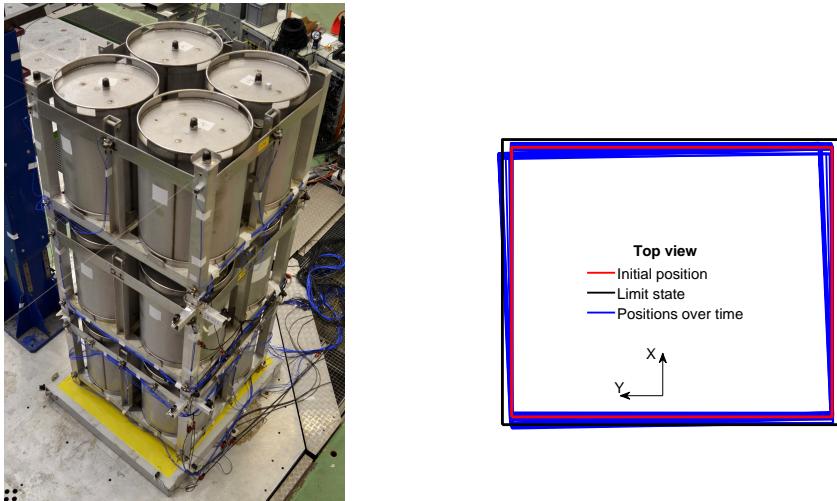


Figure 7.10: (left) Overview of the stacked structure placed on the Vesuve 1D shaking table and (right) example of test result: horizontal displacements of the top of the stack when subjected to seismic excitation in the X direction.

## 7.5 Conclusion: which case study and which data for a Bayesian estimation of fragility curves?

In the literature, many works address the estimation of seismic fragility curves, introducing many approaches to handle that task. Nevertheless, the appropriate method to provide robust estimates heavily depends on

### *7.5. Conclusion: which case study and which data for a Bayesian estimation of fragility curves?*

the characteristics of the studied equipments, on the type of data that are available, and on their number. In this chapter, we have described standard forms of datasets considered in the literature, and we have depicted how they are used to providing approximations of fragility curves using different approaches.

In the worst case, the dataset is made of tuples composed by (i) a value of the IM, and (ii) a binary outcome about the system's response: failure or non-failure. In those cases, the non-parametric models provide limited estimations, and a parametric model, such as the prominent probit-lognormal one has to be favored. If the dataset is also scarce, then most classical frequentist methods become unsatisfying as well, letting the Bayesian framework to arise as a cornerstone.

However, a major issue connects all the studies suggesting a Bayesian estimation of seismic fragility curves that we reviewed: the prior selection is not thoroughly questioned. Moreover, the choices done regarding its construction are hard to justify, and one could question the reliability of the provided estimates.

All in all, this chapter and that last thought introduce the essential interplay that exists between the [part I](#) and [part II](#) of this manuscript. Since our work is sought to be applied on case studies taken from the nuclear industry, we aim at providing a reliable and auditable methodology when we estimate seismic fragility curves. The expected auditability also concerns the choice of the prior, which hence must be constructed in a way that ensures its objectivity.



# 8

## Bayesian estimation of seismic fragility curves using a reference prior

This chapter is mainly based on the published works: Antoine Van Biesbroeck, Clément Gauchy, Cyril Feau, and Josselin Garnier (2024). “Reference prior for Bayesian estimation of seismic fragility curves”. In: *Probabilistic Engineering Mechanics* 76, p. 103622. doi: [10.1016/j.probengmech.2024.103622](https://doi.org/10.1016/j.probengmech.2024.103622)

**Abstract** For many structures of interest, estimating seismic fragility curves is a daunting task because of the limited amount of data available, and the limited amount of information they contain (such as binary data, i.e. only describing the structure as being in a failure or non-failure state). In such a context, a probit-lognormal modeling of the fragility curve is prominent, and among various methods that exist to estimate the parameters of the model, the Bayesian inference became popular in the literature. However, the trustworthiness of the latter relies on the prior selection, jeopardizing the approach if the selection is motivated by subjective considerations. This chapter proposes a comprehensive objective Bayesian estimation of seismic fragility curves, implementing the Jeffreys prior for the probit-lognormal model. This prior is completely studied along with the decay rates of the likelihood to ensure the validity of our methodology. It is also implemented on the three case studies detailed in [chapter 7](#). Our theoretical and practical results show that the Jeffreys prior leads to more robust and efficient fragility curves estimates in comparison with other methods of the literature.

<b>8.1</b>	<b>Introduction</b>	<b>124</b>
<b>8.2</b>	<b>Probit-lognormal model and Bayesian framework</b>	<b>125</b>
<b>8.3</b>	<b>Jeffreys prior construction for the probit-lognormal model</b>	<b>126</b>
8.3.1	Derivation of the Jeffreys prior	126
8.3.2	Practical implementation	127
8.3.3	Thorough study of the prior's decay rates	128
<b>8.4</b>	<b>Competing approaches and estimation tools</b>	<b>133</b>
8.4.1	Bayesian estimates of the seismic fragility curve	133
8.4.2	Competing prior taken from the literature	133
8.4.3	Maximum likelihood estimation with bootstrapping	134
<b>8.5</b>	<b>Limits of the estimates given by the three approaches: the curse of degeneracy</b>	<b>134</b>
<b>8.6</b>	<b>Performance evaluation metrics</b>	<b>137</b>
<b>8.7</b>	<b>Numerical applications</b>	<b>138</b>
8.7.1	Case study 1: the elasto-plastic oscillator	138
8.7.2	Case study 2: the piping system from a pressurized reactor	140

8.7.3	Case study 3: the stacked structure for storage of packages . . . . .	141
<b>8.8</b>	<b>A review of the properties of the SK posterior . . . . .</b>	<b>142</b>
8.8.1	Statistical model and likelihood . . . . .	143
8.8.2	Likelihood asymptotics . . . . .	144
8.8.3	Posterior asymptotics . . . . .	145
<b>8.9</b>	<b>Conclusion . . . . .</b>	<b>145</b>

## 8.1 Introduction

Seismic fragility curves are key quantities of interest of the seismic probabilistic risk assessment (SPRA) framework. They are defined as the probability of failure of a mechanical system conditionally to a scalar value that is derived from a seismic signal, coined intensity measure (IM). A detailed review of those curves and of the methods that exist to estimate them is proposed in [chapter 7](#).

Although various data sources can be exploited to evaluate such curves, most of them suffer from their scarcity, complicating the uncertainty quantification in the estimates. Moreover, we focus in this work on cases where the dataset contains information of the mechanical system's response under seismic excitation that is limited to binary outcome, i.e. failure or non-failure. In this sense, this work will mainly address equipment problems for which only binary results of seismic qualification tests (e.g., tests of electrical relays, etc.) or empirical data such as presented in Straub and Der Kiureghian, [2008](#) are available. However, the methodology developed here could perfectly be applied to simulation-based approaches as well.

In those cases, the Bayesian framework became heavily popular in the literature, being praised for its capability to regularize the estimates. Which means that the framework is known to avoid the generation of unrealistic fragility curves such as unit-step functions, which are common with classical frequentist methods. In these settings where only binary outcomes of the system's response are available, Bayesian analysis is used to infer a parameterized fragility curve from the data. The parametric model generally takes the form of a probit-lognormal fragility curve, which is prominent in the field of earthquake engineering. However, the estimates given by the posterior distribution in the Bayesian paradigm are significantly influenced by the prior, compounding their reliability since there is no consensus on some rule for the choice of the latter. Actually, there are plethora of different considerations (which are sometimes questionable because of the subjectivity they introduce) to define the prior in the literature. As an example, Straub and Der Kiureghian ([2008](#)) consider independent prior distributions for the parameters defining the fragility curve, namely the median  $\alpha$  and the log standard deviation  $\beta$ . The prior is defined as the product of a normal distribution for  $\ln(\alpha)$ , and the improper distribution  $1/\beta$  for  $\beta$ . The definition of the normal distribution is based on engineering assessments, assuming that, for the relevant component and considering the peak ground acceleration (PGA) as IM for example, the median lies between 0.02 g and 3 g with a probability of 90%. This prior was preferred to  $1/\alpha$  on the grounds that it led to unrealistically large posterior values of  $\alpha$ .

In this study, the goal is to choose the prior while eliminating, insofar as it is possible, any subjectivity which would unavoidably lead to open questions regarding the impact of the prior on the final results. To achieve this goal, the reference prior theory permits constructing priors that can be qualified as objective, because they are sought to maximize the influence of the real observations over that of the prior on the posterior distribution. We remind that a complete review of the theory is proposed in [chapter 3](#). This allows us to focus on the well-known Jeffreys prior, the asymptotic optimum of the mutual information w.r.t. the size of the data set (see [chapter 3](#) or [chapter 4](#)), and which will be explicitly derived and studied for the first time in this context.

Of course, from a subjectivity perspective, the choice of a parametric model for the fragility curve is debatable. However, numerical experiments based on the seismic responses of mechanical systems suggest that the choice of an appropriate IM makes it possible to reduce the potential biases between reference fragility curves (that can be obtained by massive Monte-Carlo methods) and their log-normal estimations (Gauchy, Feau, and Garnier, [2021](#)). This observation is reinforced by recent studies on the impact of IMs on fragility curves (Sainct et al., [2020](#); Ciano, Gioffrè, and Grigoriu, [2020](#); Ciano, Gioffrè, and Grigoriu, [2022](#)). In this paper, we will ensure the relevance of the estimations by comparing them to the results of massive

Monte-Carlo methods on academic examples. The results are illustrated with the PGA, yet we recall that our method can be implemented with any other IM of interest, without additional complexity. We also indicate that the influence of the IM —PGA vs. PSA— on the convergence of the estimates is investigated in [appendix A](#).

This chapter is organized as follows. In the next section, we recall the definition of the probit-lognormal model, and we express the implementation of reference prior theory framework in this context. Then, the Jeffreys prior is derived and completely studied in [section 8.3](#). The conduction of Bayesian inference for estimating fragility curves using the Jeffreys prior is the proposed approach of this study, it is compared with other approaches taken from the literature that are depicted in [section 8.4](#). In [section 8.5](#), we discuss the limitation of the different methods implemented, that results from the form of the likelihood of the probit-lognormal model, whose decay rates are derived. The comparison of the approaches is done by computing the performance evaluation metrics proposed in [section 8.6](#), and by evaluating them on the three case studies depicted in [chapter 7](#) in [section 8.7](#). Eventually, [section 8.8](#) precedes the conclusion. It proposes a deeper study of the competing prior considered in this chapter.

## 8.2 Probit-lognormal model and Bayesian framework

We remind that the probit-lognormal model of the fragility curve has been described in [chapter 7](#). It defines the fragility curve of the mechanical system of interest as

$$P_f(a) = \mathbb{P}(\text{"failure"} | \text{IM} = a) = \Phi\left(\frac{\log a - \log \alpha}{\beta}\right), \quad (8.1)$$

where  $\Phi$  is the c.d.f. of a standard Gaussian, and  $\alpha, \beta$  are parameters that we seek to estimate. To be precise,  $\alpha \in (0, \infty)$  is the median and  $\beta \in (0, \infty)$  is the log standard deviation of the curve. We denote  $\theta = (\alpha, \beta) \in \Theta = (0, \infty)^2$ .

In statistical terms, we consider that the failure of the equipment is the realization of a random variable  $Z$ , which takes values in  $\mathcal{Z} = \{0, 1\}$  (1 for failure, 0 for non-failure). We also denote by  $A$  the random variable of the IM. It takes value in a set  $\mathcal{A} \subset (0, \infty)$  and is supposed to follow a distribution  $H$ . Conditionally to  $\theta$ , the tuple  $Y = (Z, A)$  follows a distribution defined by  $A \sim H$  and  $Z|A, \theta \sim \mathcal{B}(P_f(A))$ , where  $\mathcal{B}(p)$  denotes the Bernoulli distribution of parameter  $p$ , and  $P_f$  is defined in eq. [\(8.1\)](#).

We recall that given realizations  $(\mathbf{z}^k, \mathbf{a}^k)$ , where  $\mathbf{z}^k = (z_i)_{i=1}^k$ ,  $\mathbf{a}^k = (a_i)_{i=1}^k$ , of the r.v.  $Y$ , this model admits the following likelihood:

$$\ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) = \prod_{i=1}^k \ell(z_i | a_i, \theta) = \prod_{i=1}^k \Phi\left(\frac{\log a_i - \log \alpha}{\beta}\right)^{z_i} \left(1 - \Phi\left(\frac{\log a_i - \log \alpha}{\beta}\right)\right)^{1-z_i}. \quad (8.2)$$

To conduct a Bayesian estimation of the fragility curve, we have to consider a prior distribution on  $\theta$ . From a prior whose density is  $\pi$ , and given observations  $(\mathbf{z}^k, \mathbf{a}^k)$  the posterior density  $p(\theta | \mathbf{z}^k, \mathbf{a}^k)$  is derived using the Bayes' theorem:

$$p(\theta | \mathbf{z}^k, \mathbf{a}^k) = \frac{\ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) \pi(\theta)}{\int_{\Theta} \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta') \pi(\theta') d\theta'}. \quad (8.3)$$

The construction of that prior is done using the reference prior theory. We recall that this theory is comprehensively introduced in [chapter 3](#). We remind that it consists in choosing the prior that maximizes the mutual information  $\mathcal{I}^k$  as  $k \rightarrow \infty$ . The mutual information for a prior whose density is  $\pi$  is defined here as

$$\mathcal{I}^k(\pi) = \int_{\Theta} \int_{\mathcal{A}^k} \sum_{\mathbf{z}^k \in \{0, 1\}^k} f\left(\frac{\int_{\Theta} \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta') \pi(\theta') d\theta'}{\ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta)}\right) \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) \mathbf{h}(\mathbf{a}^k) \pi(\theta) d\mathbf{a}^k d\theta, \quad (8.4)$$

where  $f = -\log$  classically, or follows the suggestion we made in the [chapter 4](#); and  $\mathbf{h}(\mathbf{a}^k) := \prod_{i=1}^k h(a_i)$ ,  $h$  being the density of  $A$ .

The solution of this asymptotic optimization problem is the Jeffreys prior (see chapters 3 and 4). The Jeffreys prior is already praised in Bayesian inference for its property to be invariant by re-parametrization of the model. Its density  $J$  is defined up to a constant as follows

$$J(\theta) \propto \sqrt{|\det \mathcal{I}(\theta)|} \quad \text{with} \quad \mathcal{I}(\theta) = - \sum_{z \in \{0,1\}} \int_{\mathcal{A}} \nabla_{\theta}^2 \log \ell(z|a, \theta) h(a) da. \quad (8.5)$$

## 8.3 Jeffreys prior construction for the probit-lognormal model

Based on the elements discussed in the previous section, the Jeffreys prior seems to be the best objective candidate for this problem. In this section, we will therefore study the Jeffreys prior in order to estimate probit-lognormal seismic fragility curves with binary data. To our knowledge, the application of the reference prior theory to this field of study is completely new. The explicit calculation of this prior is carried out in section 8.3.1. It is followed in section 8.3.2 by an explanation about the practical implementation suggested. In section 8.3.3, we propose a thorough study of the prior's decay rates and its proper characteristic.

### 8.3.1 Derivation of the Jeffreys prior

The first step consists in computing the Fisher information matrix  $\mathcal{I}(\theta)$  in our model, defined in eq. (8.5). Here,  $\theta = (\theta_1, \theta_2) = (\alpha, \beta) \in \Theta$  and

$$\mathcal{I}(\theta)_{i,j} = - \sum_{z \in \{0,1\}} \int_{\mathcal{A}} \ell(z|a, \theta) \partial_{\theta_i \theta_j}^2 \log \ell(z|a, \theta) h(a) da \quad (8.6)$$

for  $i, j \in \{1, 2\}$ , with

$$\log \ell(z|a, \theta) = z \log \Phi \left( \frac{\log a - \log \alpha}{\beta} \right) + (1-z) \log \left( 1 - \Phi \left( \frac{\log a - \log \alpha}{\beta} \right) \right). \quad (8.7)$$

Denoting  $\gamma = \gamma(a) = \beta^{-1} \log(a/\alpha)$ , the first-order partial derivatives of  $\log \ell(z|a, \theta)$  with respect to  $\theta$  are:

$$\partial_{\alpha} \log \ell(z|a, \theta) = - \frac{1}{\alpha \beta} z \frac{\Phi'(\gamma)}{\Phi(\gamma)} + \frac{1}{\alpha \beta} (1-z) \frac{\Phi'(\gamma)}{1 - \Phi(\gamma)}, \quad (8.8)$$

$$\partial_{\beta} \log \ell(z|a, \theta) = - \frac{\log \frac{a}{\alpha}}{\beta^2} z \frac{\Phi'(\gamma)}{\Phi(\gamma)} + \frac{\log \frac{a}{\alpha}}{\beta^2} (1-z) \frac{\Phi'(\gamma)}{1 - \Phi(\gamma)}, \quad (8.9)$$

and the second-order partial derivatives are:

$$\begin{aligned} \partial_{\alpha \beta}^2 \log \ell(z|a, \theta) &= - \frac{1}{\beta} \partial_{\alpha} \ell(z|a, \theta) + \frac{\log \frac{a}{\alpha}}{\alpha \beta^3} z \frac{\Phi''(\gamma) \Phi(\gamma) - \Phi'(\gamma)^2}{\Phi(\gamma)^2} \\ &\quad - \frac{\log \frac{a}{\alpha}}{\alpha \beta^3} (1-z) \frac{\Phi''(\gamma) (1 - \Phi(\gamma)) + \Phi'(\gamma)^2}{(1 - \Phi(\gamma))^2}, \end{aligned} \quad (8.10)$$

$$\begin{aligned} \partial_{\alpha \alpha}^2 \log \ell(z|a, \theta) &= - \frac{1}{\alpha} \partial_{\beta} \log \ell(z|a, \theta) + \frac{1}{\alpha^2 \beta^2} z \frac{\Phi''(\gamma) \Phi(\gamma) - \Phi'(\gamma)^2}{\Phi(\gamma)^2} \\ &\quad - \frac{1}{\alpha^2 \beta^2} (1-z) \frac{\Phi''(\gamma) (1 - \Phi(\gamma)) + \Phi'(\gamma)^2}{(1 - \Phi(\gamma))^2}, \end{aligned} \quad (8.11)$$

and

$$\begin{aligned} \partial_{\beta \beta}^2 \log \ell(z|a, \theta) &= - \frac{2}{\beta} \partial_{\beta} \log \ell(z|a, \theta) + \frac{\log^2 \frac{a}{\alpha}}{\beta^4} z \frac{\Phi''(\gamma) \Phi(\gamma) - \Phi'(\gamma)^2}{\Phi(\gamma)^2} \\ &\quad - \frac{\log^2 \frac{a}{\alpha}}{\beta^4} (1-z) \frac{\Phi''(\gamma) (1 - \Phi(\gamma)) + \Phi'(\gamma)^2}{(1 - \Phi(\gamma))^2}. \end{aligned} \quad (8.12)$$

The expressions in eqs. (8.10) to (8.12) of the second-order partial derivatives of  $\ell(z|a, \theta)$  need to be integrated over  $\mathcal{Z}$  and  $\mathcal{A}$ . Summing over the discrete variable  $z$  first replaces  $z$  by  $\Phi(\gamma)$  in the equations. Finally, if we denote  $A_{01}, A_{02}, A_{11}, A_{12}, A_{21}, A_{22}$  by

$$\begin{aligned} A_{0u} &= \int_{\mathcal{A}} \frac{\Phi'(\gamma(a))^2}{\Phi((-1)^{u+1}\gamma(a))} h(a) da, \\ A_{1u} &= \int_{\mathcal{A}} \log \frac{a}{\alpha} \frac{\Phi'(\gamma(a))^2}{\Phi((-1)^{u+1}\gamma(a))} h(a) da, \\ A_{2u} &= \int_{\mathcal{A}} \log^2 \frac{a}{\alpha} \frac{\Phi'(\gamma(a))^2}{\Phi((-1)^{u+1}\gamma(a))} h(a) da, \end{aligned} \quad (8.13)$$

for  $u \in \{1, 2\}$ , then the Fisher information matrix  $\mathcal{I}(\theta)$  takes on the following form:

$$\mathcal{I}(\theta) = \begin{pmatrix} \frac{1}{\alpha^2 \beta^2} (A_{01} + A_{02}) & \frac{1}{\alpha \beta^3} (A_{11} + A_{12}) \\ \frac{1}{\alpha \beta^3} (A_{11} + A_{12}) & \frac{1}{\beta^4} (A_{21} + A_{22}) \end{pmatrix}. \quad (8.14)$$

The integrals in eq. (8.13) are computed using Simpson's rule on a regular grid. The distribution  $h(a)$  is approximated by Gaussian kernel density estimation based on seismic signals that we present in the next subsection. Finally, the Jeffreys prior density is obtained by taking the square root of the determinant of the matrix defined in eq. (8.14).

### 8.3.2 Practical implementation

The derivation conducted in the previous subsection showed that knowing the probability distribution of the IM is required in order to calculate the Fisher information matrix. In this thesis we have derived the Jeffreys prior using both the PGA and the PSA as IM. Still, it is important to emphasize that this choice is purely illustrative and bears no consequence for the proposed methodology. The distributions of both IMs are approximated given their values for  $10^5$  artificial signals generated using the generator presented in chapter 7 (section 7.2). The empirical distribution of the PGA and of the PSA of these signals is recalled in fig. 8.1, and is compared with a log-normal distribution with same mean and same log deviation. The density  $h$  of their distribution can be identified as that log-normal density. In fig. 8.1, we also show that the Jeffreys prior resembles to that log-normal approximation w.r.t.  $\alpha$ .

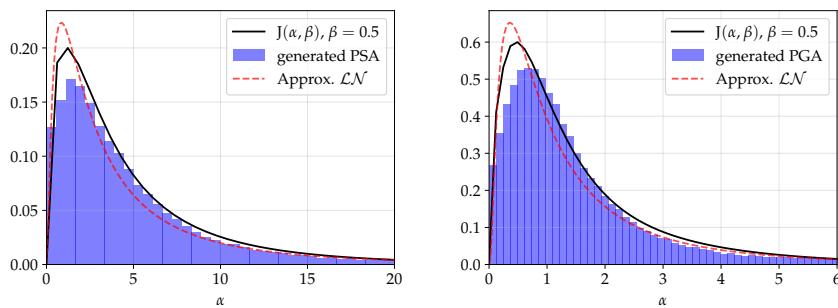


Figure 8.1: Comparison of a sectional view of the Jeffreys' prior w.r.t.  $\alpha$  (black) with the histograms of the generated signals (blue) and the log-normal approximations (red) for both IMs (PSA in the left figure, PGA in the right figure).

In practice, the use of Markov Chain Monte-Carlo (MCMC) methods (see section 8.4) to sample the *a posteriori* distribution means that the prior must be evaluated (up to a multiplicative constant) many times during the calculations. Considering the computational complexity stemming from the integrals that need to be calculated, it was decided to perform the evaluations of the prior on an experimental design based on a fine-mesh regular linear grid of  $\Theta$  (here  $(0, +\infty)^2$ ) and to build an interpolated approximation of the Jeffreys prior matching this design. This strategy is more suitable for our numerical applications and very

tractable because  $\Theta$  is a two-dimensional domain. Figure 8.2 shows plots of the Jeffreys prior for the two IMs considered. To be precise,  $2000 \times 2000$  prior values were computed for  $(\alpha, \beta) \in [10^{-5}, 50] \times [10^{-5}, 2]$  in the case of the PSA, and  $(\alpha, \beta) \in [10^{-5}, 10] \times [10^{-3}, 2]$  in the case of the PGA. These values are then processed in order to obtain a linear interpolation.

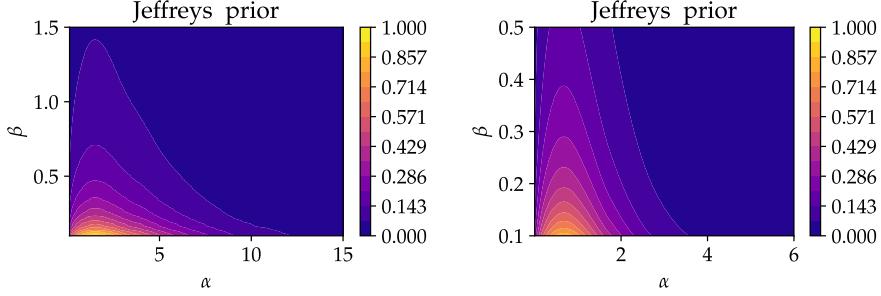


Figure 8.2: The Jeffreys priors calculated from PSA (left) and PGA (right) on subdomains of  $\Theta = (0, +\infty)^2$ .

The computational complexity of the Jeffreys prior is not in itself a major drawback. Since it depends exclusively on the distribution of the IM, the initial cost of the Jeffreys prior's complex calculation would quickly be recovered in installations such as nuclear power plant, where it is necessary to determine the fragility curves of many Structures and Components (SCs). Compared to methodologies that aim to define a prior based on mechanical calculations which are, by definition, specific to SCs, the generic character of the Jeffreys prior is a clear advantage. This will be explored in the applications section of this paper (section 8.7). Moreover, the Jeffreys prior is completely defined and does not depend on any additional subjective choices.

### 8.3.3 Thorough study of the prior's decay rates

Since the Jeffreys prior is known to be often improper, we propose in this section to determine its asymptotic decay rates. In section 8.5, the decay rates of the likelihood will be computed as well, leading us to elucidate whether the posteriors yielded by Jeffreys prior are proper or not.

We focus on the asymptotic behavior of Jeffreys in four bounds of the space  $\Theta$ , the study is compiled in the following propositions. The statements rely on the assumption below about the distribution of the IM. The comparison in fig. 8.1 illustrates the correctness of this assumption.

**Assumption 8.1.** The IM is distributed according to a log-normal distribution, i.e., there exists  $\mu \in \mathbb{R}$  and  $\sigma \in (0, +\infty)$  such that

$$h(a) = \frac{1}{\sqrt{2\pi\sigma^2}a} \exp\left(-\frac{(\log a - \mu)^2}{2\sigma^2}\right). \quad (8.15)$$

**Proposition 8.1.** Fixing  $\alpha > 0$ , there exists a  $D'(\alpha) > 0$  such that

$$J(\theta) \underset{\beta \rightarrow 0}{\sim} \frac{D'(\alpha)}{\beta}. \quad (8.16)$$

**Proposition 8.2.** There exists a constant  $E' > 0$  such that for any  $\alpha > 0$

$$J(\theta) \underset{\beta \rightarrow \infty}{\sim} \frac{E'}{\alpha\beta^3}. \quad (8.17)$$

**Proposition 8.3.** Fixing  $\beta > 0$ , there exists a  $G''(\beta) > 0$  such that

$$J(\theta) \underset{|\log \alpha| \rightarrow \infty}{\sim} G''(\beta) \frac{|\log \alpha|}{\alpha} \exp \left( -\frac{(\log \alpha - \mu)^2}{2\beta^2 + 2\sigma^2} \right). \quad (8.18)$$

To prove these statements, we rely on the two following lemmas, which provide upper bounds for the function  $\gamma \mapsto [\Phi(\gamma)(1 - \Phi(\gamma))]^{-1}$ .

**Lemma 8.1.** For any  $\gamma \in \mathbb{R}$ ,  $[\Phi(\gamma)(1 - \Phi(\gamma))]^{-1} \leq 4 \exp(2\gamma^2/\pi)$ .

**Lemma 8.2.** For any  $\gamma \in \mathbb{R}$ ,

$$\Phi(\gamma)(1 - \Phi(\gamma)) \geq \frac{\sqrt{2/\pi} \exp(-\gamma^2/2)}{(|\gamma| + \sqrt{\gamma^2 + 4})}. \quad (8.19)$$

*Proof of lemma 8.1.* From the following inequality about the erf function (Chu, 1955):

$$\forall \gamma > 0, \sqrt{1 - e^{-\frac{\gamma^2}{2}}} \leq \text{erf}(\gamma/\sqrt{2}) \leq \sqrt{1 - e^{-2\frac{\gamma^2}{\pi}}}, \quad (8.20)$$

we can deduce that, for any  $\gamma > 0$ ,

$$\begin{aligned} e^{-\frac{2\gamma^2}{\pi}} &\leq 1 - \text{erf}(\gamma/\sqrt{2})^2 \leq e^{-\frac{\gamma^2}{2}}, \\ \frac{1}{4}e^{-\frac{2\gamma^2}{\pi}} &\leq \frac{1}{4} \left(1 - \text{erf}(\gamma/\sqrt{2})\right) \left(1 + \text{erf}(\gamma/\sqrt{2})\right) \leq \frac{1}{4}e^{-\frac{\gamma^2}{2}}, \end{aligned} \quad (8.21)$$

the middle term being equal to  $\Phi(\gamma)(1 - \Phi(\gamma))$ . This implies that:

$$[\Phi(\gamma)(1 - \Phi(\gamma))]^{-1} \leq 4e^{\frac{2\gamma^2}{\pi}}, \quad (8.22)$$

hence the result for  $\gamma > 0$ .

While it is clear that the inequality still stands for  $\gamma = 0$ , notice that from  $\Phi(-\gamma) = 1 - \Phi(\gamma) \forall \gamma \in \mathbb{R}$  it follows that  $\gamma \mapsto \Phi(\gamma)(1 - \Phi(\gamma))$  is an even function. Thus, the inequality still stands for any  $\gamma < 0$ ; this concludes the proof of the lemma.  $\square$

*Proof of lemma 8.2.* Komatsu's inequality (Ito and McKean, 1974, p. 17):

$$\forall \gamma > 0, \frac{2}{\sqrt{\gamma^2 + 4} + \gamma} \leq e^{\frac{\gamma^2}{2}} \int_{\gamma}^{\infty} e^{-\frac{t^2}{2}} dt \leq \frac{2}{\sqrt{\gamma^2 + 2} + \gamma} \quad (8.23)$$

implies

$$\forall \gamma > 0, \frac{2e^{-\frac{\gamma^2}{2}}}{\sqrt{\gamma^2 + 4} + \gamma} \leq \sqrt{2\pi}(1 - \Phi(\gamma)) \leq \frac{2e^{-\frac{\gamma^2}{2}}}{\sqrt{\gamma^2 + 2} + \gamma}. \quad (8.24)$$

Since  $0 < \Phi < 1$  it follows for  $\gamma > 0$  that:

$$\Phi(\gamma)(1 - \Phi(\gamma)) \geq \frac{2e^{-\frac{\gamma^2}{2}}}{\sqrt{\gamma^2 + 4} + \gamma} \left(1 - \frac{2e^{-\frac{\gamma^2}{2}}}{\sqrt{\gamma^2 + 2} + \gamma}\right) \geq \frac{\sqrt{2/\pi} e^{-\frac{\gamma^2}{2}}}{\sqrt{\gamma^2 + 4} + \gamma}. \quad (8.25)$$

Finally, since  $\Phi(-\gamma) = 1 - \Phi(\gamma)$ ,  $\gamma \mapsto \Phi(\gamma)(1 - \Phi(\gamma))$  is an even function and we obtain for any  $\gamma \in \mathbb{R}$

$$\Phi(\gamma)(1 - \Phi(\gamma)) \geq \frac{\sqrt{2/\pi} e^{-\frac{\gamma^2}{2}}}{\sqrt{\gamma^2 + 4} + |\gamma|}. \quad (8.26)$$

□

### Proofs of the propositions

*Proof of proposition 8.1.* Let  $\alpha > 0$ . For  $0 \leq k \leq 2$ , let us consider  $A_{k1,k2} = A_{k1} + A_{k2}$  with  $A_{kj}$  defined in eq. (8.13):

$$A_{k1,k2} = \int_0^\infty \log^k \frac{a}{\alpha} \frac{\Phi'(\gamma(a))^2}{\Phi(\gamma(a))(1 - \Phi(\gamma(a)))} h(a) da. \quad (8.27)$$

A substitution gives

$$A_{k1,k2} = \beta^{k+1} \int_{-\infty}^\infty F_{A_{k1,k2}}(\gamma) d\gamma, \quad \text{with} \quad F_{A_{k1,k2}}(\gamma) = \frac{\gamma^k}{2\sqrt{\pi^3 \sigma^2}} \frac{e^{-\gamma^2} e^{-\frac{(\beta\gamma - \mu + \log \alpha)^2}{2\sigma^2}}}{\Phi(\gamma)(1 - \Phi(\gamma))}. \quad (8.28)$$

Using lemma 8.1 an upper bound can be derived for  $F_{A_{k1,k2}}$ : for any  $\gamma \in \mathbb{R}$ ,  $\beta > 0$ ,

$$|F_{A_{k1,k2}}(\gamma)| \leq \tilde{C}(\alpha) |\gamma|^k e^{-\frac{1}{3}\gamma^2}, \quad (8.29)$$

which defines an integrable function on  $\mathbb{R}$ ,  $\tilde{C}$  being a constant independent of  $\gamma$  and  $\beta$ . Hence the limit

$$\lim_{\beta \rightarrow 0} \int_{-\infty}^\infty F_{A_{k1,k2}}(\gamma) d\gamma = \int_{-\infty}^\infty \frac{\gamma^k}{2\sqrt{\pi^3 \sigma^2}} \frac{e^{-\gamma^2} e^{-\frac{(\mu - \log \alpha)^2}{2\sigma^2}}}{\Phi(\gamma)(1 - \Phi(\gamma))} d\gamma. \quad (8.30)$$

The last integral is null when  $k = 1$  since the integrand is odd in this case. When  $k$  is even, the integrand is positive valued almost everywhere, which implies that the integral is positive. From this, we can establish that  $A_{k1,k2} \underset{\beta \rightarrow 0}{\sim} D_k(\alpha) \beta^{k+1}$  for some  $D_k(\alpha) > 0$  if  $k = 0, 2$ , and  $A_{11,12} \underset{\beta \rightarrow 0}{=} o(\beta^2)$ .

Looking back at the Fisher information matrix, we can state that

$$\det \mathcal{I}(\theta) \underset{\beta \rightarrow 0}{=} \frac{D_0(\alpha) D_2(\alpha)}{\alpha^4 \beta^2} + o\left(\frac{1}{\beta^2}\right). \quad (8.31)$$

Finally, we obtain:

$$J(\theta) \underset{\beta \rightarrow 0}{\sim} \frac{D'(\alpha)}{\beta}, \quad (8.32)$$

where  $D'(\alpha) > 0$  is a constant independent of  $\beta$ . □

*Proof of proposition 8.2.*

The asymptotic expansion of the erf function in 0 is:

$$\operatorname{erf}(\gamma) \underset{\gamma \rightarrow 0}{=} \frac{2}{\sqrt{\pi}} \gamma + O(\gamma^2), \quad (8.33)$$

which allows us to state the behavior of  $\Phi(\gamma)$  when  $\gamma \rightarrow 0$ :

$$\Phi(\gamma) \underset{\gamma \rightarrow 0}{=} \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \gamma + O(\gamma^2). \quad (8.34)$$

Let us now fix  $\alpha > 0$  and consider  $A_{k1,k2} = A_{k1} + A_{k2}$  written as:

$$A_{k1,k2} = \int_{-\infty}^{\infty} \tilde{F}_{A_{k1,k2}}(x) dx, \quad \text{with} \quad \tilde{F}_{A_{k1,k2}}(x) = \frac{x^k}{2\sqrt{\pi^3\sigma^2}} \frac{e^{-\frac{x^2}{\beta^2}} e^{-\frac{(x-\mu+\log\alpha)^2}{2\sigma^2}}}{\Phi(\beta^{-1}x)(1-\Phi(\beta^{-1}x))}. \quad (8.35)$$

Let us note the convergence of  $\tilde{F}_{A_{k1,k2}}(x)$  towards an integrable function when  $\beta \rightarrow \infty$ . Moreover, lemma 8.1 allows us to bound  $\tilde{F}_{A_{k1,k2}}$ :

$$|\tilde{F}_{A_{k1,k2}}(x)| \leq \frac{2|x|^k}{\sqrt{\pi^3\sigma^2}} e^{-\frac{(x-\mu+\log\alpha)^2}{2\sigma^2}} e^{2\frac{x^2}{\pi\beta^2}} \leq \frac{2|x|^k}{\sqrt{\pi^3\sigma^2}} e^{-\frac{(x^2-2(\mu-\log\alpha))^2}{4\sigma^2}} e^{\frac{(\mu+\log\alpha)^2}{2\sigma^2}}, \quad (8.36)$$

for any  $x \in \mathbb{R}$  and  $\beta > 2\sigma/\sqrt{\pi}$ . This dominating function is integrable on  $\mathbb{R}$ . Thus, when  $\beta \rightarrow \infty$ ,  $A_{k1,k2}$  admits a limit expressed by:

$$\lim_{\beta \rightarrow \infty} A_{k1,k2} = E_k(\alpha) = \int_{-\infty}^{\infty} \frac{2x^k}{\sqrt{\pi^3\sigma^2}} e^{-\frac{(x-\mu+\log\alpha)^2}{2\sigma^2}} dx = \frac{2\sqrt{2}}{\pi} \mathbb{E}[X^k], \quad (8.37)$$

with  $X \sim \mathcal{N}(\mu - \log\alpha, \sigma^2)$ . Recalling the expression of the Jeffreys prior:

$$J(\theta)^2 = \left| \frac{1}{\alpha^2\beta^6} A_{01,02} A_{21,22} - \frac{1}{\alpha^2\beta^6} A_{11,12}^2 \right|, \quad (8.38)$$

we can deduce that it is equivalent to  $(E_0(\alpha)E_2(\alpha) - E_1^2(\alpha))/\alpha^2\beta^6$  when  $\beta \rightarrow \infty$ . Finally,

$$J(\theta) \underset{\beta \rightarrow \infty}{\sim} \frac{E'}{\alpha\beta^3}, \quad (8.39)$$

with  $E' = \sqrt{E_0(\alpha)E_2(\alpha) - E_1^2(\alpha)} = 2\sigma/\pi$ .

□

*Proof of proposition 8.3.*

As a preliminary result, let us recall that  $\text{erf}(x) \underset{x \rightarrow \infty}{=} 1 - \frac{e^{-x^2}}{x\sqrt{\pi}} + o\left(\frac{e^{-x^2}}{x}\right)$ , to establish

$$\Phi(\gamma)(1 - \Phi(\gamma)) \underset{|\gamma| \rightarrow \infty}{\sim} \frac{e^{-\frac{\gamma^2}{2}}}{|\gamma|\sqrt{2\pi}}. \quad (8.40)$$

Let us consider  $A_{k1,k2} = A_{k1} + A_{k2}$ :

$$A_{k1,k2} = C' \int_0^{\infty} \left(\log \frac{a}{\alpha}\right)^k \frac{e^{-\frac{1}{\beta^2} \log^2 \frac{a}{\alpha}} e^{-\frac{(\log a - \mu)^2}{2\sigma^2}}}{\Phi(\beta^{-1} \log \frac{a}{\alpha})(1 - \Phi(\beta^{-1} \log \frac{a}{\alpha}))} \frac{da}{a}, \quad (8.41)$$

denoting  $C' = \sqrt{4\pi^3\sigma^2}^{-1}$ . By substituting

$$\nu = \log a - \frac{\sigma^2}{\sigma^2 + \beta^2} \log \alpha - \frac{\beta^2}{\sigma^2 + \beta^2} \mu = \log a - r \log \alpha - s\mu, \quad (8.42)$$

we obtain

$$A_{k1,k2} = C' \int_{-\infty}^{\infty} \hat{F}_{A_{k1,k1}}(\nu) d\nu, \quad (8.43)$$

with

$$\hat{F}_{A_{k1,k1}}(\nu) = (\nu + (r-1) \log \alpha + s\mu)^k \frac{e^{-\frac{(\nu+(r+1)\log\alpha+s\mu)^2}{\beta^2}} e^{-\frac{(\nu+r\log\alpha+(s-1)\mu)^2}{2\sigma^2}}}{[\Phi(1-\Phi)](h_\beta(\nu))}, \quad (8.44)$$

and where  $h_\beta(\nu) = \beta^{-1}(\nu + (r-1) \log \alpha + s\mu)$ . Using eq. (8.40), we obtain

$$[\Phi(1-\Phi)](h_\beta(\nu)) \underset{|\log \alpha| \rightarrow \infty}{\sim} \frac{\beta e^{-\frac{(\nu+(r-1)\log\alpha+s\mu)^2}{2\beta^2}}}{|\nu + (r-1) \log \alpha + s\mu| \sqrt{2\pi}}. \quad (8.45)$$

Then for a clear understanding of the asymptotic behavior of  $\hat{F}_{A_{k1,k2}}$ , let us compute

$$\begin{aligned} & -\frac{(\nu + (r-1) \log \alpha + s\mu)^2}{2\beta^2} - \frac{(\nu + r \log \alpha + (s-1)\mu)^2}{2\sigma^2} \\ &= -\left(\frac{1}{2\beta^2} + \frac{1}{2\sigma^2}\right)\nu^2 - \left(\frac{((r-1) \log \alpha + s\mu)^2}{2\beta^2} + \frac{(r \log \alpha + (s-1)\mu)^2}{2\sigma^2}\right) \\ &= -\left(\frac{1}{2\beta^2} + \frac{1}{2\sigma^2}\right)\nu^2 - \frac{(\log \alpha - \mu)^2}{2(\beta^2 + \sigma^2)}. \end{aligned} \quad (8.46)$$

We expand  $\hat{F}_{k1,k2}(\nu) = \sum_{j=1}^k C_k^j (r-1)^j \log^j \alpha (\nu + s\mu)^{k-j} g(\nu) = \sum_{j=0}^k \hat{F}_{k1,k2}^{(j)}(\nu)$ , with  $g(\nu)$  defined as

$$g(\nu) = \frac{e^{-\frac{(\nu+(r+1)\log\alpha+s\mu)^2}{\beta^2}} e^{-\frac{(\nu+r\log\alpha+(s-1)\mu)^2}{2\sigma^2}}}{[\Phi(1-\Phi)](\beta^{-1}(\nu + (r-1) \log \alpha + s\mu))}. \quad (8.47)$$

Therefore, by combining eqs. (8.45) and (8.46), we obtain that  $\hat{F}_{k1,k2}^{(j)}$  satisfies

$$\hat{F}_{k1,k2}^{(j)}(\nu) e^{\frac{(\log \alpha - \mu)^2}{2\beta^2 + 2\sigma^2}} (\log \alpha)^{-j} |\log \alpha|^{-1} \underset{|\log \alpha| \rightarrow \infty}{\longrightarrow} (\sqrt{2\pi}\beta)^{-1} C_k^j (-1)^j (1-r)^{j+1} (\nu + s\mu)^{k-j} e^{-\left(\frac{1}{2\beta^2} + \frac{1}{2\sigma^2}\right)\nu^2}. \quad (8.48)$$

Using lemma 8.2, we can also define an integrable upper bound for the above function as:

$$|\hat{F}_{k1,k2}^{(j)}(\nu)| e^{\frac{(\log \alpha - \mu)^2}{2\beta^2 + 2\sigma^2}} |\log \alpha|^{-j-1} \leq \frac{\sqrt{2/\pi} e^{-\left(\frac{1}{2\beta^2} + \frac{1}{2\sigma^2}\right)\nu^2}}{|h_\beta(\nu)| + \sqrt{h_\beta(\nu)^2 + 4}} \leq \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{1}{2\beta^2} + \frac{1}{2\sigma^2}\right)\nu^2}. \quad (8.49)$$

Thus, we can switch the limits and the integration, and the following results are obtained:

$$C'' \beta A_{01,02} e^{\frac{(\log \alpha - \mu)^2}{2\beta^2 + 2\sigma^2}} \underset{\log \alpha \rightarrow \infty}{=} (1-r) G \log \alpha - s\mu G + o(1) \quad (8.50)$$

$$C'' \beta A_{11,12} e^{\frac{(\log \alpha - \mu)^2}{2\beta^2 + 2\sigma^2}} \underset{\log \alpha \rightarrow \infty}{=} -(1-r)^2 G \log^2 \alpha - s^2 \mu^2 G + 2(1-r)s\mu G \log \alpha - G' + o(1), \quad (8.51)$$

$$\begin{aligned} C'' \beta A_{22,22} e^{\frac{(\log \alpha - \mu)^2}{2\beta^2 + 2\sigma^2}} \underset{\log \alpha \rightarrow \infty}{=} & (1-r)^3 G \log^3 \alpha - s^3 \mu^3 G - 3(1-r)^2 s\mu G \log^2 \alpha \\ & - 3(1-r)s^2 \mu^2 G \log \alpha + 3(1-r)G' \log \alpha + o(1), \end{aligned} \quad (8.52)$$

with  $C'' = (C'\sqrt{2\pi})^{-1}$ ,  $G = \sigma\beta\sqrt{2\pi(\beta^2 + \sigma^2)^{-1}}$  and  $G' = G^2/2\pi$ . This way

$$\begin{aligned} (C'')^2\alpha^2\beta^8J(\theta)^2e^{\frac{2(\log\alpha-\mu)^2}{2\beta^2+2\sigma^2}} &= 3(r-1)^2s^2\mu^2G^2\log^2\alpha + 3(r-1)^2s^2\mu^2G^2\log^2\alpha \\ &\quad + 3(r-1)^2GG'\log^2\alpha - 4(r-1)^2s^2\mu^2G^2\log^2\alpha \\ &\quad - 2(r-1)^2s^2\mu^2G^2\log^2\alpha - 2(r-1)^2GG'\log^2\alpha + o(\log^2\alpha). \end{aligned} \quad (8.53)$$

Note that the above equality is still valid when  $\log\alpha \rightarrow -\infty$ . Finally,

$$J(\theta) \underset{|\log\alpha| \rightarrow \infty}{\sim} G''(\beta) \frac{|\log\alpha|}{\alpha} \exp\left(-\frac{(\log\alpha-\mu)^2}{2\beta^2+2\sigma^2}\right), \quad (8.54)$$

with

$$G''(\beta) = C''^{-1}(r-1)^2GG'\beta^{-4} = \frac{2\sigma^3\beta^3}{\sqrt{\pi}(\sigma^2 + \beta^2)^{7/2}}. \quad (8.55)$$

□

This study states that the Jeffreys prior is improper, since it is not integrable w.r.t.  $\beta$ .

## 8.4 Competing approaches and estimation tools

### 8.4.1 Bayesian estimates of the seismic fragility curve

The most relevant method in order to benefit from the Bayesian theory and the reference prior construction presented in section 8.2 consists in deriving the posterior defined in eq. (8.3). When the latter is proper, it then becomes possible to generate, according to that distribution, samples of  $\theta$  conditioned on the observed data. These *a posteriori* generations of  $\theta$  can be obtained using MCMC methods. In this study, we have implemented an adaptive Metropolis-Hastings (M-H) algorithm with a Gaussian transition kernel and a covariance adaptation process (following the suggestions of Haario, Saksman, and Tamminen, 2001). Such an algorithm allows sampling from a probability density known up to a multiplicative constant. In this context, the *a posteriori* samples of  $\theta$  can be used to define credibility intervals for the probit-lognormal estimations of the fragility curves.

### 8.4.2 Competing prior taken from the literature

For comparison purposes, we selected the prior suggested by Straub and Der Kiureghian (2008) —called the SK prior— whose density is defined as the product of a normal distribution for  $\ln(\alpha)$  and the improper distribution  $1/\beta$  for  $\beta$ , namely:

$$\pi_{SK}(\theta) \propto \frac{1}{\alpha\beta} \exp\left(-\frac{(\log\alpha-\mu)^2}{2\sigma^2}\right). \quad (8.56)$$

In Straub and Der Kiureghian, 2008, the parameters  $\mu$  and  $\sigma$  of the log-normal distribution are chosen to generate a non-informative prior. For a fair comparison with the approach proposed in this chapter, we decided to pick  $\mu$  and  $\sigma$  as equal to the mean and the standard deviation of the logarithm of the IM. This choice is consistent with the fact that the Jeffreys prior is similar to a log-normal distribution with these parameters (see fig. 8.1). The prior  $\pi_{SK}(\theta)$  is plotted in fig. 8.3 for both the PSA and the PGA as IM.

An analysis of the posterior obtained from the SK prior is given in section 8.5. It shows that the posterior is always improper, which jeopardizes the validity of any *a posteriori* estimation using MCMC methods. This could however be mitigated by truncating w.r.t.  $\beta$  (yet this involves a subjective choice), which we will do in the numerical application conducted in this work. This issue persists in the authors' original framework, which is slightly different from ours. This is confirmed in section 8.8.

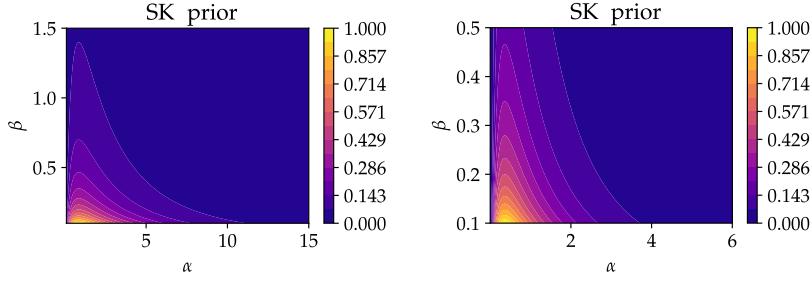


Figure 8.3: Prior suggested by Straub and Der Kiureghian (2008), expressed in eq. (8.56) and scaled to a log-normal estimation of the IM distribution, for both the PSA (left) and the PGA (right) as IM.

**Asymptotic comparison of the Jeffreys and SK priors** By comparing the Jeffreys and SK prior asymptotics, it can be observed that:

- Regarding the asymptotics w.r.t.  $\beta$ , while the divergence rates are the same when  $\beta \rightarrow 0$ , the Jeffreys prior performs better when  $\beta \rightarrow \infty$ :

$$J(\theta) \underset{\beta \rightarrow \infty}{\propto} \beta^{-2} \pi_{SK}(\theta). \quad (8.57)$$

Consequently, the SK posterior will result in higher probabilities for high values of  $\beta$  compared to the Jeffreys prior.

- Regarding the asymptotics w.r.t.  $\alpha$ , both are asymptotically close to a log-normal distribution, with a slight “disadvantage” for the Jeffreys prior, for which the asymptotic variance is derived by adding  $\beta^2$  to the variance of the SK prior. This means that while for small values of  $\beta$  (smaller than  $\sigma$ ), both priors remain comparable w.r.t.  $\alpha$ , the Jeffreys prior gives higher probabilities to  $\alpha$  outliers when  $\beta$  also has a high value. However, as seen above, the probability for large values of  $\beta$  is quite low for the Jeffreys prior compared to the SK prior.

### 8.4.3 Maximum likelihood estimation with bootstrapping

The most common non-Bayesian estimator of the seismic fragility curves under the probit-lognormal modeling is the maximum likelihood estimator (MLE). This estimator has been introduced in chapter 7 (section 7.3). It is expressed as follows given a set of  $k$  observations  $(\mathbf{z}^k, \mathbf{a}^k)$ :

$$\theta^{\text{MLE}}(\mathbf{z}^k, \mathbf{a}^k) = \arg \max_{\theta \in \Theta} \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta). \quad (8.58)$$

It is common to compute multiple MLE in order to obtain confidence intervals. Doing so, the bootstrap technique defines the stochastic estimator of  $\theta$ :

$$\hat{\theta}_k^{\text{BMLE}}(\mathbf{z}^k, \mathbf{a}^k) = \theta^{\text{MLE}}((z_{U_i}, a_{U_i})_{i=1}^k), \quad (8.59)$$

where  $U_1, \dots, U_k$  are i.i.d. random variables distributed w.r.t. a uniform distribution in  $\{1, \dots, k\}$ . This estimator is used to compute confidence intervals by estimating empirically its quantiles from many samples. This is a very common approach for fragility curves (e.g., Shinozuka, Feng, et al., 2000; Gehl, Douglas, and Seyed, 2015; Baker, 2015). The convergence of the MLE and the relevance of this method are detailed in Vaart, 1992. However, the relevance of the bootstrap method is often limited by the irregularity of its results for small values of  $k$  (see e.g., Zentner, Gündel, and Bonfils, 2017).

## 8.5 Limits of the estimates given by the three approaches: the curse of degeneracy

In this section, we study the likelihood decay rates in order to (i) explain irregular phenomena commonly encountered with classical frequentist methods, and (ii) verify the proper characteristics of the posteriors in

the Bayesian framework. Our study of the likelihood's asymptotic behavior is summarized in the following proposition.

**Proposition 8.4.** *Let us consider  $k > 1$  and a data sample  $(\mathbf{a}^k, \mathbf{z}^k)$ . Let us introduce the vectors  $\mathbf{N} = (z_i \mathbb{1}_{a_i < \alpha} + (1 - z_i) \mathbb{1}_{a_i > \alpha})_{i=1}^k$ ,  $\log^2 \frac{\mathbf{a}^k}{\alpha} = (\log^2 \frac{a_i}{\alpha})_{i=1}^k$ .*

- Fixing  $\alpha > 0$ , then

$$\ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) \xrightarrow[\beta \rightarrow \infty]{} 2^{-k} \quad \text{and} \quad \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) \xrightarrow[\beta \rightarrow 0]{} O\left(\beta^{|\mathbf{N}|} e^{-\frac{\mathbf{N}^T \log^2 \frac{\mathbf{a}^k}{\alpha}}{2\beta^2}}\right), \quad (8.60)$$

where  $|\mathbf{N}| = \sum_{i=1}^k N_i$ .

- Fixing  $\beta > 0$ , then

$$\ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) \xrightarrow[\alpha \rightarrow 0]{} O\left(|\log \alpha|^{|\mathbf{z}^k|-k} e^{-\frac{1}{2\beta^2} \sum_{i=1}^k (1-z_i)(\log a_i - \log \alpha)^2}\right) \quad (8.61)$$

and

$$\ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) \xrightarrow[\alpha \rightarrow \infty]{} O\left(\log(\alpha)^{-|\mathbf{z}^k|} e^{-\frac{1}{2\beta^2} \sum_{i=1}^k z_i(\log a_i - \log \alpha)^2}\right), \quad (8.62)$$

where  $|\mathbf{z}^k| = \sum_{i=1}^k z_i$  is the number of failures in the observed sample.

*Proof of proposition 8.4.* As a reminder, the likelihood is expressed as:

$$\begin{aligned} \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) &= \prod_{i=1}^k \Phi\left(\frac{\log a_i - \log \alpha}{\beta}\right)^{z_i} \left(1 - \Phi\left(\frac{\log a_i - \log \alpha}{\beta}\right)\right)^{1-z_i} \\ &= \exp\left[\sum_{i=1}^k (z_i \log \Phi(\gamma_i) + (1-z_i) \log(1 - \Phi(\gamma_i)))\right], \end{aligned} \quad (8.63)$$

denoting  $\gamma_i = \beta^{-1} \log \frac{a_i}{\alpha}$ .

To treat the case where  $\beta \rightarrow \infty$  we can observe that while  $\alpha$  is fixed, the quantities  $\Phi(\gamma_i)$  all converge to 1/2. The product of those limits gives the limit  $\ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) \xrightarrow[\beta \rightarrow \infty]{} 2^{-k}$ .

For the other cases, it should be reminded that  $\Phi(x) = \frac{1}{2}(1 + \text{erf}(x/\sqrt{2}))$  and  $\text{erf}(x) \xrightarrow[x \rightarrow \infty]{} 1 - \frac{e^{-x^2}}{x\sqrt{\pi}} + o\left(\frac{e^{-x^2}}{x}\right)$ , which leads to

$$\Phi(x) \xrightarrow[x \rightarrow \infty]{} 1 - \frac{e^{-\frac{x^2}{2}}}{x\sqrt{2\pi}} + o\left(\frac{e^{-\frac{x^2}{2}}}{x}\right). \quad (8.64)$$

Let us consider an  $i \in \{1, \dots, k\}$  and compute

$$\begin{aligned}
 & z_i \log \Phi(\gamma_i) + (1 - z_i) \log(1 - \Phi(\gamma_i)) \\
 & \underset{\gamma_i \rightarrow \infty}{=} z_i \log \left( 1 - \frac{e^{-\frac{\gamma_i^2}{2}}}{\gamma_i \sqrt{2\pi}} + o\left(\frac{e^{-\frac{\gamma_i^2}{2}}}{\gamma_i}\right) \right) + (1 - z_i) \log \left( \frac{e^{-\frac{\gamma_i^2}{2}}}{\gamma_i \sqrt{2\pi}} + o\left(\frac{e^{-\frac{\gamma_i^2}{2}}}{\gamma_i}\right) \right) \\
 & \underset{\gamma_i \rightarrow \infty}{=} -z_i \frac{e^{-\frac{\gamma_i^2}{2}}}{\gamma_i \sqrt{2\pi}} + (1 - z_i) \log \left( \frac{e^{-\frac{\gamma_i^2}{2}}}{\gamma_i \sqrt{2\pi}} \right) + o(1) \\
 & \underset{\gamma_i \rightarrow \infty}{=} -(1 - z_i) \frac{\gamma_i^2}{2} - (1 - z_i) \log(\gamma_i \sqrt{2\pi}) + o(1).
 \end{aligned} \tag{8.65}$$

Using the relation  $\Phi(-x) = 1 - \Phi(x)$ , it follows that, in the case  $a_i < \alpha$ :

$$z_i \log \Phi(\gamma_i) + (1 - z_i) \log(1 - \Phi(\gamma_i)) \underset{\gamma_i \rightarrow -\infty}{=} -z_i \frac{\gamma_i^2}{2} - z_i \log(-\gamma_i \sqrt{2\pi}) + o(1). \tag{8.66}$$

Going back to the likelihood asymptotics, let us first suppose  $\alpha > 0$  and  $\beta \rightarrow 0$ . Thus, denoting the vectors  $\mathbf{N} = (z_i \mathbb{1}_{a_i < \alpha} + (1 - z_i) \mathbb{1}_{a_i > \alpha})_{i=1}^k$  and  $\log^2 \frac{\mathbf{a}^k}{\alpha} = (\log^2 \frac{a_i}{\alpha})_{i=1}^k$ , we obtain

$$\ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) \underset{\beta \rightarrow 0}{=} \frac{C(\alpha)}{\sqrt{2\pi}^{|\mathbf{N}|}} \beta^{|\mathbf{N}|} e^{-\frac{\mathbf{N}^T \log^2 \frac{\mathbf{a}^k}{\alpha}}{2\beta^2} + o(1)} \underset{\beta \rightarrow 0}{=} O\left(\beta^{|\mathbf{N}|} e^{-\frac{\mathbf{N}^T \log^2 \frac{\mathbf{a}^k}{\alpha}}{2\beta^2}}\right), \tag{8.67}$$

where  $|\mathbf{N}| = \sum_{i=1}^k N_i$  and  $C(\alpha) = \prod_{i=1}^k |\log \frac{a_i}{\alpha}|^{N_i}$ .

Let us then fix  $\beta > 0$  to get

$$\begin{aligned}
 \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) & \underset{\alpha \rightarrow \infty}{=} \frac{\beta^{|\mathbf{N}|}}{\sqrt{2\pi}^{|\mathbf{N}|}} \left( \prod_{i=1}^k \log \left( \frac{\alpha}{a_i} \right)^{-z_i} \right) \exp \left( -\frac{1}{2\beta^2} \sum_{i=1}^k z_i (\log a_i - \log \alpha)^2 + o(1) \right) \\
 & \underset{\alpha \rightarrow \infty}{=} O\left(\log(\alpha)^{-|\mathbf{z}^k|} e^{-\frac{1}{2\beta^2} \sum_{i=1}^k z_i (\log a_i - \log \alpha)^2}\right),
 \end{aligned} \tag{8.68}$$

where  $|\mathbf{z}^k| = \sum_{i=1}^k z_i$  is the number of failures in the observed sample. Similarly, we obtain

$$\begin{aligned}
 \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) & \underset{\alpha \rightarrow 0}{=} \frac{\beta^{|\mathbf{N}|}}{\sqrt{2\pi}^{|\mathbf{N}|}} \left( \prod_{i=1}^k \log \left( \frac{a_i}{\alpha} \right)^{-(1-z_i)} \right) \exp \left( -\frac{1}{2\beta^2} \sum_{i=1}^k (1 - z_i) (\log a_i - \log \alpha)^2 + o(1) \right) \\
 & \underset{\alpha \rightarrow 0}{=} O\left(|\log \alpha|^{|\mathbf{z}^k|-k} e^{-\frac{1}{2\beta^2} \sum_{i=1}^k (1-z_i) (\log a_i - \log \alpha)^2}\right).
 \end{aligned} \tag{8.69}$$

□

This statement expresses that, under general circumstances, the likelihood converges rapidly to 0 in every direction but the one where  $\beta \rightarrow \infty$ . The latter is a concern, because it requires that the decay rate of the prior in that direction is proper to yield proper posteriors. Moreover, under some circumstances, the vector  $\mathbf{N}$  is null. It is the case when we fall in one of the three scenarios described in the following definition. In that case, we say that the likelihood is degenerate, and it does not converge towards 0 when  $\beta \rightarrow 0$ .

**Definition 8.1** (Likelihood degeneracy). If the observed samples  $(\mathbf{z}^k, \mathbf{a}^k)$  belong to one of the following three types:

- type 1 : no failure is observed:  $z_i = 0$  for any  $i$ ;

- type 2 : only failures are observed:  $z_i = 1$  for any  $i$ ;
- type 3 : the failures and non-failures are partitioned into two disjoint subsets when classified according to their IM values: there exists  $a \in \mathcal{A}$  such that for any  $i, j$ ,  $a_i < a < a_j \iff z_i \neq z_j$  (see the illustration in fig. 8.4);

then the likelihood is degenerate.

An example of degenerate likelihood is presented in fig. 8.4 with an illustration of this phenomenon. In the following, we deduce from this study the decay rates of the posterior densities yielded by the two priors considered in this work (the SK prior and the Jeffreys prior). In particular, we show that the Jeffreys prior may produce improper posteriors in degenerate scenarios. For this reason, this chapter addresses the estimation of seismic fragility curves from dataset issuing non-degenerate likelihoods.

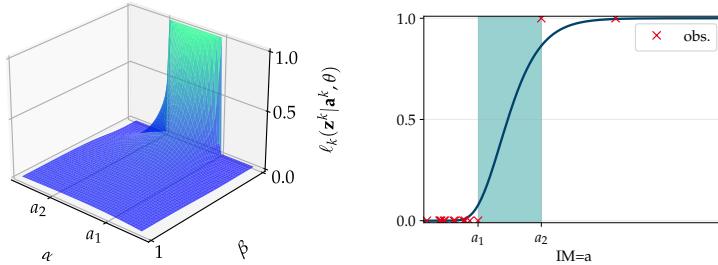


Figure 8.4: Example of a type 3 data sample  $(\mathbf{z}^k, \mathbf{a}^k)$  which gives a degenerate likelihood. Graphs of (left) the likelihood given the degenerate data sample as a function of the tuple  $(\alpha, \beta)$  and (right) the fragility curve (blue curve) according to which the points (red crosses) are sampled : each red cross is a tuple  $(z, a)$ , where its X-axis equals  $a$  and its Y-axis equals  $z$ . In both figures,  $a_1$  is the maximal observed IM value among “non-failures”, and  $a_2$  is minimal one among “failures”. The interval  $(a_1, a_2)$  (in cyan in the right figure) separates the failures and the non-failure observed.

### Decay rates of the posterior densities and properness

The Jeffreys and SK priors are not proper with respect to  $\beta$  (see section 8.3 for the Jeffreys prior and eq. (8.56) for the SK prior). For the Jeffreys prior, the divergence and convergence rates with respect to  $\beta$  only make the resulting posterior proper when the prior is coupled with non-degenerate likelihoods. However, one can see that this is not the case for the SK posterior, which is not integrable w.r.t.  $\beta$  because of a convergence rate that is too low at  $+\infty$ . This prevents the validation of the MCMC estimates for this posterior, unless a truncation of the distribution is considered. Note that this prior was originally designed within a Bayesian framework that slightly differs from ours. In section 8.8, we confirm that the posterior is not proper even when derived in the exact framework of the author’s paper.

## 8.6 Performance evaluation metrics

In order to assess and compare the performances of the proposed approaches, we propose different quantitative metrics. Beforehand, we suppose we have access to a reference fragility curve  $P_f^{\text{ref}}$ . In two of the case studies that we treat in the following, a large validation dataset is available and can be used to derive such reference using non-parametric Monte-Carlo estimates following the suggestion of Trevlopoulos, Feau, and Zentner (2019). This method was described in chapter 7 (section 7.3) and implemented in chapter 7 onto two case studies for which a large validation dataset was available.

Given observations  $(\mathbf{z}^k, \mathbf{a}^k)$ , we denote by  $a \mapsto P_f^{|\mathbf{z}^k, \mathbf{a}^k}(a)$  the random process defined as the fragility curve conditional to the  $(\mathbf{z}^k, \mathbf{a}^k)$ . Its distribution inherits from the posterior distribution of  $\theta$  in the case of the Bayesian method, or it inherits from the distribution of the bootstrap estimator  $\hat{\theta}^{\text{BMLE}}(\mathbf{z}^k, \mathbf{a}^k)$  in the case of

the MLE. For each value  $a$ , the  $r$ -quantile of the random variable  $P_f^{|\mathbf{z}^k, \mathbf{a}^k}(a)$  is denoted by  $q_r^{|\mathbf{z}^k, \mathbf{a}^k}(a)$ . We then define:

- The quadratic error:  $\mathcal{E}^{|\mathbf{z}^k, \mathbf{a}^k} = \mathbb{E}_{\theta|\mathbf{z}^k, \mathbf{a}^k} \left[ \|P_f^{|\mathbf{z}^k, \mathbf{a}^k} - P_f^{\text{ref}}\|_{L^2}^2 \right]$ .
- The  $1 - r$  square credibility width:  $\mathcal{W}^{|\mathbf{z}^k, \mathbf{a}^k} = \|q_{1-r/2}^{|\mathbf{z}^k, \mathbf{a}^k} - q_{r/2}^{|\mathbf{z}^k, \mathbf{a}^k}\|_{L^2}^2$ .

The  $L^2$  norm is taken on the domain  $[a_{\min}, a_{\max}]$  that covers the available seismic signals:

$$\|P\|_{L^2}^2 = \int_{a_{\min}}^{a_{\max}} P(a)^2 da. \quad (8.70)$$

The  $L^2$  norms are approximated using Simpson's interpolation, and the expectation are estimated by Monte-Carlo sums, using samples following either the posterior distributions or the distribution of the bootstrap estimator.

## 8.7 Numerical applications

In this section, we will examine three case studies. Two of them leverage the many simulation datasets available that have been previously computed for validation purposes. They will be used in the derivation of a reference fragility curve (as explained in section 8.6), and allow us to validate the corresponding probit-lognormal models. The three case studies were presented in chapter 7. The first one is a nonlinear oscillator for which  $10^5$  nonlinear simulations have been implemented for validation purposes. The second case study deals with a piping system which is part of the secondary cooling system of a French Pressurized Water Reactor. Due to the high computational cost, only  $10^4$  simulations have been performed for this case. In both cases, estimations are performed using different testing data sets of a size  $k$  chosen as negligible compared to the size of the validation dataset. These testing datasets are taken from the set of available nonlinear dynamical simulation results. A third case study is eventually presented in order to showcase how our method could be applied to practical experiments.

### 8.7.1 Case study 1: the elasto-plastic oscillator

The first case study is a single-degree-of-freedom elasto-plastic oscillator with kinematic hardening. This mechanical system was presented in chapter 7 (section 7.4.1). For this case study,  $10^5$  nonlinear dynamical simulations have been conducted to constitute a validation dataset. In fig. 8.5, we recall (i) a scheme of the oscillator, (ii) the validation dataset in the plan (PGA,EDP), (iii) reference fragility curves derived using non-parametric estimation from the validation dataset (see section 8.6). In this study, the failure is defined when the EDP exceeds a threshold  $C = 0.8$  cm.

Examples of fragility curve estimations are shown in fig. 8.6. They are obtained from 5000 samples of  $\theta$  generated with the implemented statistical methods (see section 8.4), which are based on two observation sets of sizes  $k = 20$  and  $k = 30$ . Although the nature of the two intervals compared is different —credibility interval for the Bayesian framework and confidence interval for the MLE—, these results clearly illustrate the advantage of the Bayesian framework over the MLE for small samples. With the MLE, irregularities characterized by null estimates of  $\beta$  appear, resulting in “vertical” confidence intervals. In section 8.5, we established that the likelihood is easily maximized for  $\beta = 0$  when samples are partitioned into two disjunct subsets when classified according to IM values: one subset for which there is no failure and one for which there is failure (i.e. when the likelihood is degenerate because of a type 3 data sample). Moreover, when few failures are observed in the initial sample, the bootstrap technique can lead to the consideration of many sub-samples that maximize the likelihood at  $\beta = 0$ . This is better evidenced by an examination of the raw values of  $\theta$  generated in fig. 8.7. The degenerate  $\beta$  values resulting from the MLE appear clearly but, although it should theoretically also be affected, the Bayesian framework shows no evidence of a similar phenomenon for this type of samples.

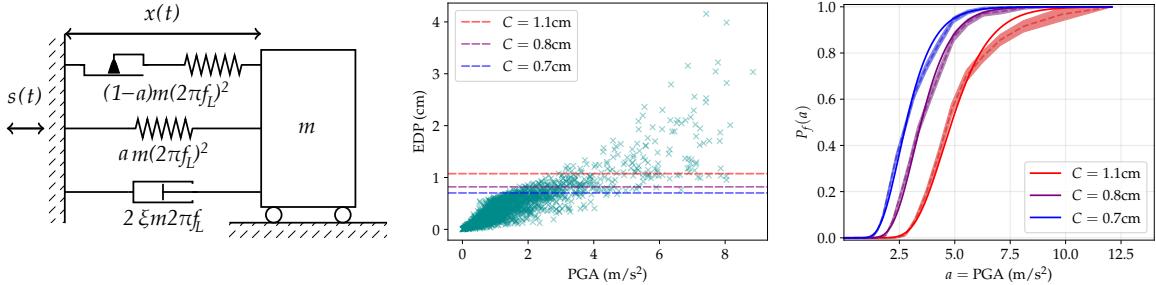


Figure 8.5: Illustration of the first case study. Left: scheme of an elasto-plastic oscillator of mass  $m$ . Middle: results of  $10^5$  non-linear dynamical simulations. Right: reference non-parametric fragility curves (dashed lines) surrounded by their 95%-confidence intervals, compared with the probit-lognormal fragility curves derived by MLE from the validation dataset of size  $10^5$ . Different thresholds  $C$  are considered, each yields different proportions of failures in the validation dataset: resp. 95% (red), 90% (purple) and 85% (blue).

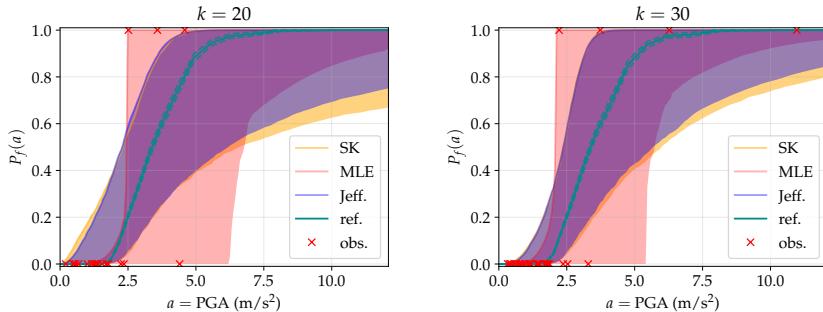


Figure 8.6: 95% credibility (for Bayesian estimation) or confidence (for the MLE) intervals of fragility curve estimations for the elasto-plastic oscillator, using the methods described in section 8.4. The cyan curve is the reference  $P_f^{\text{ref}}$  suggested in section 8.6, surrounded by its 95%-confidence interval (dashed lines). The red crosses represent the observed data.

Since the SK prior is calibrated to look like the Jeffreys prior, fig. 8.6 shows a strong similarity between the Bayesian estimations of the fragility curves obtained with these two priors. However, fig. 8.7 shows that many outliers are obtained with the SK prior. These values explain why the credibility intervals of the fragility curves are larger with the SK prior when  $k = 20$ . This observation is supported theoretically by the calculation provided in section 8.3. There is actually a better convergence of the Jeffreys prior toward 0 when  $\beta \rightarrow \infty$ . This superior asymptotic behavior obviously results in posteriors that happen to assign a lower probability to outlier points (a phenomenon particularly noticeable when the data sample is small) as well as to the weight of the likelihood within the posterior.

For a better understanding of this phenomenon, we calculated the quantitative metrics defined in section 8.6. For any  $k$  ranging from 15 to 100, we considered  $m = 200$  different draws of observations  $(\mathbf{z}_{(j)}^k, \mathbf{a}_{(j)}^k)_{j=1}^m$  in order to derive the average values and 95%-confidence intervals of the metrics  $\mathcal{E}^{|\mathbf{z}_{(j)}^k, \mathbf{a}_{(j)}^k|}$ ,  $\mathcal{W}^{|\mathbf{z}_{(j)}^k, \mathbf{a}_{(j)}^k|}$ ,  $j \in \{1, \dots, m\}$ , for each of the three methods. They are plotted in fig. 8.8. Firstly, these diagrams demonstrate the benefits of the Bayesian framework compared to the MLE approach for small observation sets. Secondly, the compared performance of the Jeffreys and SK posteriors is highlighted by the confidence interval endpoints of the quadratic error and the credibility interval. Specifically, the latter highlights the effect of the superior asymptotic behavior of the Jeffreys prior along the width of the credibility interval. It shows variations similar to but smaller than the SK prior, thus highlighting its capacity to generate fewer outliers for the pair  $(\alpha, \beta)$ , as expected.

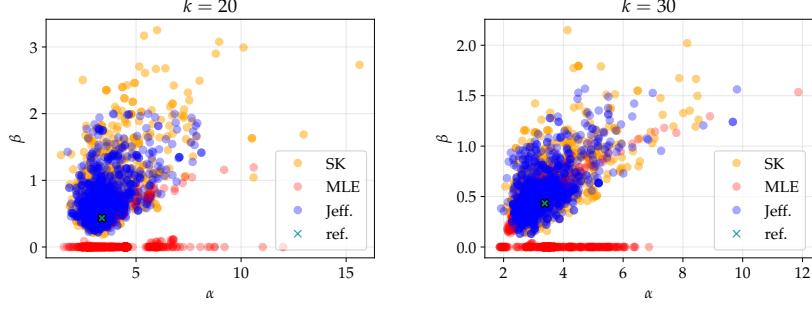


Figure 8.7: Scatter plots of the generated  $\theta$  for the estimation of the fragility curves presented in fig. 8.6 for the elasto-plastic oscillator. For all three statistical methods, we plotted 5000  $\theta = (\alpha, \beta)$  estimated with two observed datasets of size  $k = 20$  (left figure) and  $k = 30$  (right figure). The cyan cross represents  $\theta^{\text{MLE}}$  derived from the  $10^5$  samples composing the validation dataset.

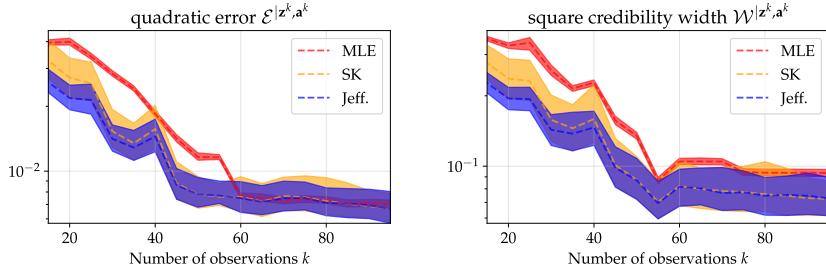


Figure 8.8: Average values of the performance evaluation metrics  $\mathcal{E}|\mathbf{z}^k, \mathbf{a}^k$  (left) and  $\mathcal{W}|\mathbf{z}^k, \mathbf{a}^k$  (right) for the elasto-plastic oscillator computed by replications of the experiment with different observation sets and for different sample sizes, ranging from  $k = 15$  to 100. In each figure, the shaded areas show the 95%-confidence intervals of the metrics.

### 8.7.2 Case study 2: the piping system from a pressurized reactor

This second case study deals with a piping system which is part of the secondary cooling system of a French pressurized water reactor. This piping system was presented in details in chapter 7 (section 7.4.2). Nonlinear dynamical simulations of this system's response have been carried out for  $8 \cdot 10^4$  of the seismic signals we dispose. The resulting validation dataset in the plan (PGA,EDP) is depicted in fig. 8.9, along with a picture of the system's mock-up and reference fragility curves derived using non-parametric estimation from the validation set (as explicated in section 8.6). In this study, the failure is defined when the EDP exceeds a threshold  $C = 3.8^\circ$ .

Estimations similar to the first case study's were performed here, and highlighted the same trends. As expected, for sets of 5000 values of  $\theta$ —generated with each statistical method considered in this work—and for two sample sizes  $k = 20$  and  $k = 30$  of nonlinear dynamical simulations, fig. 8.10 shows the superiority of the Bayesian framework over the coupled MLE and bootstrap approach. Just like in the case study with the oscillator, irregularities appear with the MLE-based approach: the confidence intervals are similarly “quasi-vertical”, reflecting the fact that many estimations of  $\beta$  are equal to 0. Moreover, the credibility intervals are wider with the SK prior than with the Jeffreys prior, which here too can be interpreted as an increased number of outliers. These observations are clearly supported by the results presented in fig. 8.11.

For a more complete overview of their relative performances, the evaluation metrics described in section 8.6 have been computed in the same way as for the first case study:  $m = 200$  draws of data samples  $(\mathbf{z}_{(j)}^k, \mathbf{a}_{(j)}^k)_{j=1}^m$  have been randomly chosen to compute, for any value of  $k$  ranging from 15 to 100, the average values and 95%-confidence intervals of the metrics  $\mathcal{E}|\mathbf{z}_{(j)}^k, \mathbf{a}_{(j)}^k$ ,  $\mathcal{W}|\mathbf{z}_{(j)}^k, \mathbf{a}_{(j)}^k$ , for each of the three methods. They are presented in fig. 8.12. These results confirm the superior performance of the Jeffreys prior compared to the other two methods.

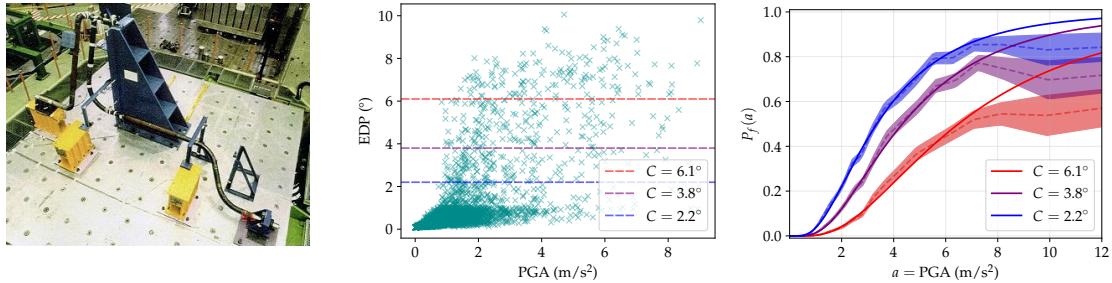


Figure 8.9: Illustration of the second case study. Left: picture of the mock-up on the Azalee shaking table of CEA. Middle: results of  $8 \cdot 10^4$  non-linear dynamical simulations. Right: reference non-parametric fragility curves (dashed lines) surrounded by their 95%-confidence intervals, compared with the probit-lognormal fragility curves derived by MLE from the validation dataset of size  $8 \cdot 10^4$ . Different thresholds  $C$  are considered, each yields different proportions of failures in the validation dataset: resp. 95% (red), 90% (purple) and 85% (blue).

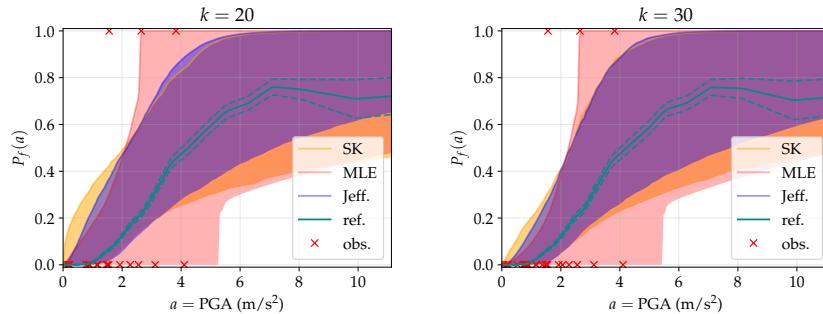


Figure 8.10: 95% credibility (for Bayesian estimation) or confidence (for the MLE) intervals of fragility curve estimations for the piping system, using the methods described in section 8.4. The cyan curve is the reference  $P_f^{\text{ref}}$  suggested in section 8.6, surrounded by its 95%-confidence interval (dashed lines). The red crosses represent the observed data.

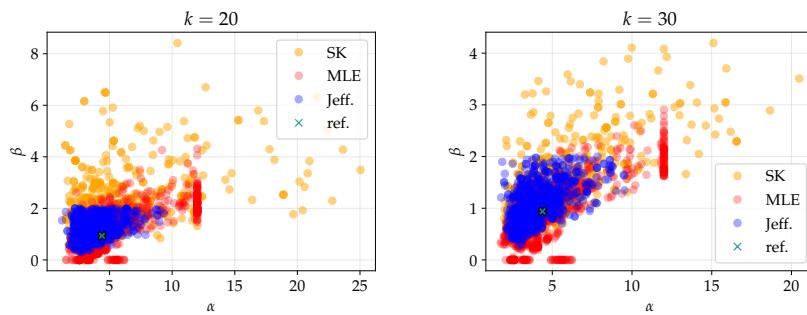


Figure 8.11: Scatter plots of the generated  $\theta$  for the estimation of the fragility curves presented in fig. 8.6 for the piping system. For all three statistical methods, we plotted 5000  $\theta = (\alpha, \beta)$  estimated with two observed datasets of size  $k = 20$  (left figure) and  $k = 30$  (right figure). The cyan cross represents  $\theta^{\text{MLE}}$  derived from the  $10^5$  samples composing the validation dataset.

### 8.7.3 Case study 3: the stacked structure for storage of packages

The case study considered hereafter concerns a freestanding stacked structure composed of three pallets intended for the storage of packages. A complete description of the system was proposed in chapter 7. In fig. 8.13-left. For this case study, only  $k = 21$  experimental results have been performed on a shaking table. An example of a result is shown in fig. 8.13-right, which depicts the horizontal displacements over time of the top of the stack. The initial position is indicated in red, while the different positions in time are indicated

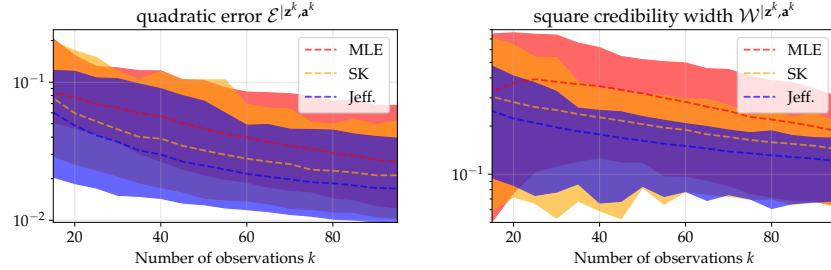


Figure 8.12: Average values of the performance evaluation metrics  $\mathcal{E}|z^k, a^k$  (left) and  $\mathcal{W}|z^k, a^k$  (right) for the piping system computed by replications of the experiment with different observation sets and for different sample sizes, ranging from  $k = 15$  to 100. In each figure, the shaded areas show the 95%-confidence intervals of the metrics.

in blue. Due to uplift, sliding, and rotation motions, the top of the stack exceeds, for 2 of the 21 tests, the admissibility criterion, which is materialized in black. Due to the signals used for these tests, the Jeffreys prior is suitable for estimating a fragility curve of the stack, as is the SK prior calibrated to approximate the Jeffreys prior.

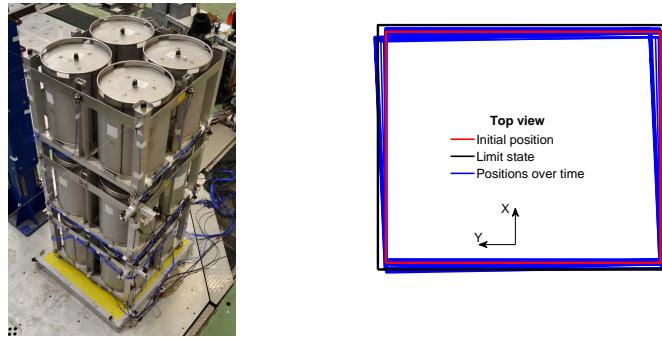


Figure 8.13: (left) Overview of the stacked structure placed on the Vesuve 1D shaking table and (right) example of test result: horizontal displacements of the top of the stack when subjected to seismic excitation in the X direction.

The small number of data points does not allow comparison with a reference. Therefore, only a qualitative study can be carried out. Figure 8.14 shows the results from 5000 estimations of  $\theta$  performed with the three statistical methods considered in this study. These results do not call into question the previous ones because they, once again, highlight the same phenomena, namely: (i) some irregularities of the MLE-based approach characterized by "quasi-vertical" confidence intervals; (ii) wider credibility intervals with the SK prior than with the Jeffreys prior; and (iii) the generation of many outliers of  $\theta$  with the SK posterior compared to the Jeffreys posterior, which explains the slightly larger credibility intervals on the log-normal estimates of the fragility curves.

## 8.8 A review of the properties of the SK posterior

In this chapter, we have compared our approach with the one that results from an adaptation of the prior suggested by Straub and Der Kiureghian (2008). We proved in section 8.5 that within our framework, this prior results in an improper posterior. This puts the validity of the MCMC estimations into question, and could explain the lower performance of the SK prior compared to the Jeffreys prior. In their paper, the authors use the Bayesian methodology the same way we do, yet the consideration of uncertainties over the observed earthquake intensity measures and the equipment capacities leads to a slightly different likelihood. In order to verify that the drawbacks of their prior highlighted in this chapter are not due to our statistical choices,

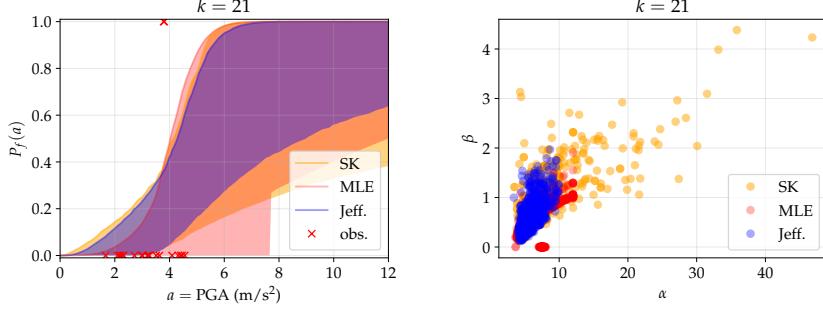


Figure 8.14: Left: 95% credibility (for Bayesian estimation) or confidence (for the MLE) intervals of fragility curve estimation for the stacked structure. The red crosses represent the observed data. Right: scatter plot of the 5000 generated  $\theta$  used for the computation of the credibility or confidence zones depicted in the left figure.

we dedicated this section to the study of the asymptotic expansions of the posterior in the exact framework presented in Straub and Der Kiureghian, 2008. We shall first introduce the exact model of their paper for the estimation of seismic fragility curves in section 8.8.1, using notations consistent with our study. We will then derive the likelihood and its asymptotics in section 8.8.2. Finally, we will express the convergence rates of the posterior in section 8.8.3, which will allow us to conclude that the SK posterior is indeed improper.

### 8.8.1 Statistical model and likelihood

Let us consider the observations of earthquakes labeled  $l = 1, \dots, L$  at equipment labeled  $i = 1, \dots, I_j$  located in substations labeled  $j = 1, \dots, J$ . The observed items are  $(\mathbf{z}_{jl}, \hat{a}_{jl})_{j,l}$ ,  $\mathbf{z}_{jl} = (z_{ijl})_i$  being the failure occurrences of the  $I_j$  pieces of equipment at substation  $j$  during earthquake  $l$  ( $z_{ijl} \in \{0, 1\}$ ), and  $\hat{a}_{jl}$  being the observed IM at substation  $j$  during earthquake  $l$  ( $\hat{a}_{jl} \in (0, +\infty)$ ). They are assumed to follow the latent model presented below.

At substation  $j$  the  $l$ -th earthquake results in an IM value  $a_{jl}$  that is observed with an uncertainty multiplicative noise:  $\log \hat{a}_{jl} = \log a_{jl} + \varepsilon_{jl}$  where  $\varepsilon_{jl} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ . The noise variance  $\sigma_\varepsilon^2$  is supposed to be known. The uncertain intrinsic capacity of equipment  $i$  at substation  $j$  is  $r_{ij} \sim \mathcal{N}(\mu_r, \sigma_r^2)$  and  $y_{jl} \sim \mathcal{N}(0, \sigma_y^2)$  is the uncertain factor common to all equipment capacities at substation  $j$  during earthquake  $l$ . The random variables  $r_{ij}$ ,  $y_{jl}$  and  $\varepsilon_{jl}$  are supposed to be independent.

A failure of equipment  $i$  at substation  $j$  during earthquake  $l$  is considered when the performance of the structural component  $g_{ijl}$  satisfies  $g_{ijl} > 0$ . This performance can be expressed as

$$g_{ijl} = \log \hat{a}_{jl} + \varepsilon_{jl} - y_{jl} - r_{ij} = x_{jl} - r_{ij}$$

with  $x_{jl} = \log \hat{a}_{jl} + \varepsilon_{jl} - y_{jl}$ .

This establishes the following conditional relation between the observed data:

$$p(z_{ijl} | \hat{a}_{jl}, \Sigma) = \int_{\mathbb{R}} p(z_{ijl} | x_{jl}, \hat{a}_{jl}, \Sigma) \frac{\exp\left(-\frac{(x_{jl} - \log \hat{a}_{jl})^2}{2(\sigma_\varepsilon^2 + \sigma_y^2)}\right)}{\sqrt{2\pi(\sigma_\varepsilon^2 + \sigma_y^2)}} dx_{jl}, \quad (8.71)$$

denoting  $\Sigma = (\sigma_r, \sigma_y, \mu_r)$ , and with

$$p(z_{ijl} | x_{jl}, \hat{a}_{jl}, \Sigma) = \Phi\left(\frac{x_{jl} - \mu_r}{\sigma_r}\right)^{z_{ijl}} \left(1 - \Phi\left(\frac{x_{jl} - \mu_r}{\sigma_r}\right)\right)^{1-z_{ijl}}, \quad (8.72)$$

when substation  $j$  is only affected by one earthquake. The method proposed in Straub and Der Kiureghian, 2008 actually considers cases in which a substation may be impacted by two successive earthquakes and takes into account the fact that its response to the second would be correlated to its response to the first one. This

would lead to a different likelihood. However, it is mentioned that this possibility only concerns a small number of data points. We can therefore limit our calculations to the simplest case and assume  $l = L = 1$ . The subscript  $l$  will therefore be dropped in what follows.

Finally, the likelihood for this model can be expressed as:

$$\ell_J(\mathbf{z}|\hat{\mathbf{a}}, \Sigma) = \prod_{j=1}^J \int_{\mathbb{R}} \prod_{i=1}^{I_j} p(z_{ij}|x_j, \log \hat{a}_j, \Sigma) \frac{\exp\left(-\frac{(x_j - \log \hat{a}_j)^2}{2(\sigma_\varepsilon^2 + \sigma_y^2)}\right)}{\sqrt{2\pi(\sigma_\varepsilon^2 + \sigma_y^2)}} dx_j, \quad (8.73)$$

denoting  $\mathbf{z} = (\mathbf{z}_j)_{j=1}^J$ ,  $\hat{\mathbf{a}} = (\hat{a}_j)_{j=1}^J$ , and with the integrated conditional distribution defined in eq. (8.72).

In the Bayesian framework introduced in Straub and Der Kiureghian, 2008, the model parameter is  $\Sigma$ . Let us denote  $\alpha = \exp \mu_r$ ,  $\beta = \sqrt{\sigma_r^2 + \sigma_y^2}$  and  $\rho = \sigma_y^2 / \beta^2$ . Denoting  $\theta = (\alpha, \beta, \rho)$ , the knowledge of  $\Sigma$  then becomes equivalent to the one of  $\theta$  and the likelihood of eq. (8.73) can be expressed conditionally to  $\theta$  instead of  $\Sigma$ :

$$\ell_J(\mathbf{z}|\hat{\mathbf{a}}, \theta) = \prod_{j=1}^J \int_{\mathbb{R}} \prod_{i=1}^{I_j} \Psi^{z_{ij}} \left( \frac{x - \log \alpha}{\beta \sqrt{1 - \rho}} \right) \frac{\exp\left(-\frac{(x - \log \alpha)^2}{2(\sigma_\varepsilon^2 + \rho \beta^2)}\right)}{\sqrt{2\pi(\sigma_\varepsilon^2 + \rho \beta^2)}} dx, \quad (8.74)$$

where the notation  $\Psi^{z_{ij}}(\gamma)$  is used to denote  $\Phi(\gamma)^{z_{ij}}(1 - \Phi(\gamma))^{1-z_{ij}}$ .

Straub and Der Kiureghian (2008) propose the following improper prior distribution for the parameter  $\theta$ :

$$\pi_{SK}(\theta) \propto \frac{1}{\beta \alpha} \exp\left(-\frac{(\log \alpha - \mu)^2}{2\sigma^2}\right) \mathbb{1}_{0 \leq \rho \leq 1}. \quad (8.75)$$

*A posteriori* estimations of  $\theta$  are consequently generated from MCMC methods

$$p(\theta|\mathbf{z}, \hat{\mathbf{a}}) \propto \ell_J(\mathbf{z}|\hat{\mathbf{a}}, \theta) \pi_{SK}(\theta). \quad (8.76)$$

### 8.8.2 Likelihood asymptotics

In this subsection, we will study the asymptotics of the likelihood defined in eq. (8.74) when  $\beta \rightarrow \infty$ . Let us first consider the substitution  $u = (x - \log \hat{a}_j) / \sqrt{\sigma_\varepsilon^2 + \rho \beta^2}$  to express the likelihood as

$$\ell_J(\mathbf{z}|\hat{\mathbf{a}}, \theta) = \prod_{j=1}^J \int_{\mathbb{R}} f_j^\beta(u) du, \quad \text{with} \quad f_j^\beta(u) = \prod_{i=1}^{I_j} \Phi(h_j^\beta(u))^{z_{ij}} (1 - \Phi(h_j^\beta(u)))^{1-z_{ij}} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}, \quad (8.77)$$

with

$$h_j^\beta(u) = \frac{(u + \log \hat{a}_j) \sqrt{\sigma_\varepsilon^2 + \rho \beta^2} - \log \alpha}{\beta \sqrt{1 - \rho}}. \quad (8.78)$$

This way, remembering that  $0 \leq \Phi(1 - \Phi) \leq 1$ , an upper bound  $u \mapsto e^{-u^2/2} / \sqrt{2\pi}$  can be found for  $f_j^\beta$  for any  $\beta, u$ . It converges when  $\beta \rightarrow +\infty$  as follows:

$$\lim_{\beta \rightarrow \infty} f_j^\beta(u) = \prod_{i=1}^{I_j} \Psi^{z_{ij}} \left( \frac{(u + \log \hat{a}_j) \sqrt{\rho}}{\sqrt{1 - \rho}} \right) \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}. \quad (8.79)$$

This gives the following limit for the likelihood:

$$\lim_{\beta \rightarrow \infty} \ell_J(\mathbf{z}|\hat{\mathbf{a}}, \theta) = \prod_{j=1}^J \int_{\mathbb{R}} \prod_{i=1}^{I_j} \Psi^{z_{ij}} \left( \frac{(u + \log \hat{a}_j) \sqrt{\rho}}{\sqrt{1 - \rho}} \right) \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} du, \quad (8.80)$$

which is a positive quantity.

### 8.8.3 Posterior asymptotics

By combining eqs. (8.75), (8.76) and (8.80) we obtain the posterior asymptotics

$$p(\theta|\mathbf{z}, \hat{\mathbf{a}}) \underset{\beta \rightarrow \infty}{\sim} \frac{C}{\beta}, \quad (8.81)$$

with  $C$  being a positive constant. This makes the posterior improper w.r.t.  $\beta$ , with the same convergence rate as the one derived in our framework.

## 8.9 Conclusion

Assessing the seismic fragility of SCs is a daunting task when data is limited. The performance of the Bayesian framework in this kind of situation is well-known. Nevertheless, choosing a prior remains difficult because its impact on the *a posteriori* distribution cannot be neglected, and therefore neither can its impact on the estimation of any relevant element linked to the fragility curves.

Elaborating on the reference prior theory in order to define an objective prior, we derived, for the first time in this field of study, the Jeffreys prior for the log-normal model, with binary data that indicates the state of the structure (i.e. failure or non-failure). This prior is completely defined and thoroughly studied, and does not depend on any additional subjective choice.

This work is also an opportunity to develop a better theoretical understanding of the conditions that result in irregular fragility curves such as unit-step functions in practice. This issue is quite inevitable when data is limited, since such curves are a result of the very composition of the sample. Although this issue affects every approach, our results prove the robustness of the proposed approach over the traditional ones in terms of regularization (i.e. the absence of irregular estimation of the fragility curve) and stability (i.e. the absence of outliers when sampling the *a posteriori* distribution of the parameters). Nevertheless, we recall that the study conducted in the present chapter does not permit generating reliable estimate when the observed data yield what we call a degenerate likelihood. In the latter case, we have proven theoretically that all the methods carried out in this chapter—including ours—are jeopardized. That statement paves the way for further works that could (i) reinforce the prior construction to suppress its posteriors' weaknesses under degenerate likelihoods; and (ii) study methods for reducing the probability of the existence of degenerate likelihoods. These perspectives motivated the works that are presented in the following chapters of this manuscript.

Although the numerical implementation of the Jeffreys prior is complex—more so than a prior defined as the product of two classical distributions such as log-normal distributions, for instance—it is not a major issue. As a matter of fact, since it depends solely on the distribution of the IM, the “cost” of the initial calculation would quickly be recovered on the scale of an industrial installation containing several SCs whose fragility curves must be estimated. For example, compared to methodologies that aim to define a prior based on mechanical calculations for a given SC, the advantage of the Jeffreys prior lies in its generic nature. The fact that it can be applied to all SCs subjected to the same seismic scenario largely compensates for the implementation of mechanical studies dedicated to each relevant SC. Additionally, the methodology can be implemented with any relevant IM without creating additional complexity.



# 9

## Appropriate constraint incorporation in the probit-lognormal reference prior

This chapter is mainly based on the published work: Nils Baillie, Antoine Van Biesbroeck, Cyril Feau, and Clément Gauchy (2025). "Bayesian estimation of seismic fragility curves based on variational reference priors using neural networks". In: *Proceedings of the 6th UNCECOMP Conference*. URL: <https://2025.uncecomp.org/proceedings/pdf/21225.pdf>

**Abstract** Estimating seismic fragility curves under conditions of limited data and binary structural responses remains a challenge that is commonly addressed using Bayesian inference to update a probit-lognormal modeling of fragility curves. Since the prior selection remains a critical step in this context, we aim at constructing in this chapter a prior that (i) minimizes the subjectivity it embeds, and (ii) is robust in terms of generation of irregular estimates (such as unit-step functions). To do so, we introduce a constrained reference prior that is designed to regularize the posterior distribution while conserving an objective characteristic. Two implementation strategies are explored: a numerical approximation of the constrained prior density and a variational approximation using a neural network that implicitly parameterizes the prior. We compare both approaches through synthetic examples and a real case study. Our results highlight the capacity of the constrained prior to provide accurate and efficient estimates.

<b>9.1</b>	<b>Introduction</b>	<b>148</b>
<b>9.2</b>	<b>Reminder of the probit-lognormal model and of the Bayesian framework</b>	<b>149</b>
<b>9.3</b>	<b>Constrained reference prior for probit-lognormal fragility curves</b>	<b>150</b>
9.3.1	Constrained reference prior framework	150
9.3.2	Application to the probit-lognormal reference prior	151
9.3.3	Practical derivation with integral interpolations	151
<b>9.4</b>	<b>Variational approximation of the constrained reference prior</b>	<b>152</b>
9.4.1	Definition of the VA-RP	152
9.4.2	Optimization under constraints	152
9.4.3	Neural network architecture	153
<b>9.5</b>	<b>Comparison between the VA-RP and the interpolated reference prior</b>	<b>154</b>
<b>9.6</b>	<b>Application of the method on the piping system case study</b>	<b>157</b>
9.6.1	Benchmarking metrics	157
9.6.2	Results	157
<b>9.7</b>	<b>Impacts of the neural network's architecture on the VA-RP</b>	<b>159</b>
9.7.1	Decay rates of the VA-RP density	159
9.7.2	Variance of $\log \alpha$ a priori	160
<b>9.8</b>	<b>Conclusion</b>	<b>161</b>

## 9.1 Introduction

In the seismic probabilistic risk assessment (SPRA) framework, the estimation of seismic fragility curves is a critical step. We recall that those curves are defined in [chapter 7](#), where we review the methods that are used in the literature to estimate them. As a matter of fact, the information provided in the available databases to estimate the fragility curves is sometimes poor. In our study, we are particularly interested in cases where (i) the data are scarce, and (ii) the available knowledge about the mechanical system's response to the seismic excitation is limited to a binary outcome (i.e., failure or non-failure). Under those settings, the parameterization of the fragility curve using the probit-lognormal model is a widely recognized approach in earthquake engineering.

The Bayesian framework has been widely implemented for estimating probit-lognormal seismic fragility curves of mechanical structures and components. As a matter of fact, this framework has been proven efficient over classical frequentist methods in avoiding the generation of irregular estimates —such as unit-step functions— that may arise when limited numbers of data are available.

The common weaknesses of most of these approaches result from their prior construction processes. Indeed, on the one hand, informative priors may be efficient but need to be avoided in reliability assessment studies because of the subjectivity they embed. On the other hand, the work conducted in [chapter 8](#) demonstrates that (i) insufficiently informative priors do not guarantee to produce proper posteriors given the likelihood decay rates of the probit-lognormal model, and that (ii) the reference prior is, most of the time, and in particular for the probit-lognormal model, defined by considering that the data are not what we call “degenerate”. That degeneracy phenomenon represents a major curse when estimating seismic fragility curves, and the work done in [chapter 8](#) shows that it does not only affect the Bayesian estimates. Frequentist methods such as MLE when coupled with bootstrap techniques are also very affected. Moreover, in [appendix A](#), we illustrate on an example that this degeneracy is likely to occur (i) under small samples, and (ii) when the IM that is considered is more correlated with the structure's response. The latter statement paradoxically jeopardizes the estimation of the fragility curve given an input that embeds more information on the output.

In this chapter, we propose a truly regularizing objective prior based Bayesian methodology, in the sense that the prior we propose is sought to avoid the generation of unrealistic fragility curves —even in degenerate scenarios— while conserving an objective characteristic (as much as possible). When looking for objective priors, we recall that the reference prior theory is prominent, since it suggests constructing priors that maximize the mutual information, which amounts maximizing the information brought by the data to the posterior (we refer to [chapter 3](#) for a complete review of the reference prior theory). In this study, we mainly rely on the theoretical work developed in [chapter 5](#), which suggests incorporating constraints to the reference prior. The constraints take the form of linear constraints that slightly regularize the reference prior's decay rate to ensure it yields proper posteriors.

To implement the constrained reference prior, we compare two methodologies. The first one consists in approximating numerically the prior density by interpolating the integrals involved in its theoretical non-explicit formulation. This approach corresponds to the one conducted in [chapter 8](#) with the unconstrained reference prior. It has the disadvantage of being computationally expensive. The second methodology consists in using a variational approximation of the reference prior (VA-RP). This method consists in applying the algorithm constructed in [chapter 6](#), which parameterizes the reference prior as the output of a neural network. The weights of the neural networks are trained to maximize the mutual information, while ensuring the incorporated constraint on the prior is satisfied. In this study, the neural network architecture is specifically tailored to the context of seismic fragility curves estimation. In both methods, we assess the performance of the posterior distributions, which are estimated using Markov Chain Monte-Carlo (MCMC) techniques.

After this introduction, we briefly recall the probit-lognormal modeling of fragility curves. Then, we derive the theoretical expression of a constrained reference prior in section [9.3](#). This expression is non-explicit, yet we explain how it can be used to numerically approximate the prior's density. In section [9.4](#), we propose a second approach that is based on variational inference to approximate the constrained reference prior. The two approaches are then compared in section [9.5](#), before being applied on a real case study in

section 9.6. Before to conclude, we propose some results in section 9.7 to provide more insight about the impact of the neural network's architecture on the variational approximation of the reference prior.

## 9.2 Reminder of the probit-lognormal model and of the Bayesian framework

We remind that the probit-lognormal model was defined in chapter 7 (section 7.3). We recall that it consists in a statistical model where the data take the form of realizations of the random vector  $Y = (Z, A)$ , where  $A \in \mathcal{A} \subset (0, \infty)$  refers to the IM of the seismic signal, and  $Z \in \{0, 1\}$  refers to the outcome ( $Z = 1$  in case of a failure,  $Z = 0$  otherwise). The model consists in considering the following parameterized distribution of  $Y$  conditionally to  $\theta = (\alpha, \beta) \in \Theta$ :

$$A \sim A|\theta \sim H, \quad \text{and} \quad Z|A, \theta \sim \mathcal{B}(P_f(A)), \quad (9.1)$$

where  $H$  is the distribution of the IM, its density is denoted  $h$ ;  $\mathcal{B}(p)$  denotes the Bernoulli distribution of parameter  $p$ ; and  $P_f$  refers to the probit-lognormal fragility curve:

$$P_f(a) = \mathbb{P}(Z = 1 | \text{IM} = a, \theta) = \Phi\left(\frac{\log a - \log \alpha}{\beta}\right), \quad (9.2)$$

with  $\Phi$  being the c.d.f. of a standard Gaussian distribution. This model is parameterized by the two parameters  $\alpha \in (0, \infty)$  and  $\beta \in (0, \infty)$ .

Given  $k$  realizations  $(\mathbf{z}^k, \mathbf{a}^k)$  of  $Y$  ( $\mathbf{z}^k = (z_i)_{i=1}^k$  and  $\mathbf{a}^k = (a_i)_{i=1}^k$ ), the likelihood of this model is

$$\ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) = \prod_{i=1}^k \ell(z_i | a_i, \theta) = \prod_{i=1}^k \Phi\left(\frac{\log a_i - \log \alpha}{\beta}\right)^{z_i} \left(1 - \Phi\left(\frac{\log a_i - \log \alpha}{\beta}\right)\right)^{1-z_i}. \quad (9.3)$$

We recall that incorporating the Bayesian framework in this model amounts to choose a prior on  $\theta$  whose density is denoted by  $\pi$ . It allows to define a posterior whose density  $p(\cdot | \mathbf{z}^k, \mathbf{a}^k)$  is given by the Bayes' theorem:

$$p(\theta | \mathbf{z}^k, \mathbf{a}^k) = \frac{\ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) \pi(\theta)}{\int_{\Theta} \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta') \pi(\theta') d\theta'}. \quad (9.4)$$

Note that if the integral in the above equation is finite, then the posterior has to be proper. If it is not, it is still possible to define the posterior up to a multiplicative constant (see chapter 3), but it does not permit to generate *a posteriori* samples of  $\theta$  for inference anymore.

### Degeneracy of the likelihood

We recall below the asymptotic decay rates of the model's likelihood, that are derived in chapter 8:

$$\forall \alpha > 0, \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) \xrightarrow[\beta \rightarrow \infty]{} 2^{-k} \quad \text{and} \quad \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) \xrightarrow[\beta \rightarrow 0]{} O\left(\beta^{|\mathbf{N}|} e^{-\frac{\mathbf{N}^T \log^2 \frac{\mathbf{a}^k}{\alpha}}{2\beta^2}}\right), \quad (9.5)$$

$$\forall \beta > 0, \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) \xrightarrow[\alpha \rightarrow 0]{} O\left(|\log \alpha|^{|\mathbf{z}^k|-k} e^{-\frac{1}{2\beta^2} \sum_{i=1}^k (1-z_i)(\log a_i - \log \alpha)^2}\right), \quad (9.6)$$

$$\ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) \xrightarrow[\alpha \rightarrow \infty]{} O\left(\log(\alpha)^{-|\mathbf{z}^k|} e^{-\frac{1}{2\beta^2} \sum_{i=1}^k z_i(\log a_i - \log \alpha)^2}\right), \quad (9.7)$$

where  $\mathbf{N} = (z_i \mathbb{1}_{a_i < \alpha} + (1 - z_i) \mathbb{1}_{a_i > \alpha})_{i=1}^k$ ,  $\log^2 \frac{\mathbf{a}^k}{\alpha} = (\log^2 \frac{a_i}{\alpha})_{i=1}^k$ ,  $|\mathbf{N}| = \sum_{i=1}^k N_i$  and  $|\mathbf{z}^k| = \sum_{i=1}^k z_i$ .

These formulas indicate that the likelihood asymptotic behavior may drastically change given different data samples  $(\mathbf{z}^k, \mathbf{a}^k)$ . This remark led to the definition of degeneracy of the likelihood, recalled below.

**Definition** (Likelihood degeneracy; reminder of definition 8.1). If the observed samples  $(\mathbf{z}^k, \mathbf{a}^k)$  belong to one of the following three types:

- type 1 : no failure is observed:  $z_i = 0$  for any  $i$ ;
- type 2 : only failures are observed:  $z_i = 1$  for any  $i$ ;
- type 3 : the failures and non-failures are partitioned into two disjoint subsets when classified according to their IM values: there exists  $a \in \mathcal{A}$  such that for any  $i, j$ ,  $a_i < a < a_j \iff z_i \neq z_j$ ;

then the likelihood is degenerate.

When the likelihood is degenerate, the vector  $\mathbf{N}$  is null and the likelihood does not shrink at  $\beta \rightarrow 0$ . As a consequence, the posterior will be proper only if the prior is proper w.r.t.  $\beta$ . Moreover, in case of degeneracy of type 1 or 2, the likelihood does not shrink at either  $\alpha \rightarrow \infty$  or  $\alpha \rightarrow 0$ . In such a scenario, it becomes also essential to consider a prior that is proper w.r.t.  $\alpha$ .

In practice, cases of degeneracy of type 3 are the most critical ones since they are more likely to appear, especially when (i) the data are scarce, and (ii) the IM is more correlated with the system's response (see appendix A).

## 9.3 Constrained reference prior for probit-lognormal fragility curves

### 9.3.1 Constrained reference prior framework

To construct the prior, we take as a support the reference prior theory, whose principle is to maximize w.r.t. the prior the mutual information  $\mathcal{J}^k$ . The mutual information is expected to measure the amount of information brought by the data over the *a priori* information in the workflow. This methodology ensures the prior can be qualified as objective. We recall that this theory is completely reviewed in chapter 3. The mutual information can be written as a function of the prior  $\Pi$ :

$$\mathcal{J}^k(\Pi) = \int_{\Theta} \int_{\mathcal{A}^k} \sum_{\mathbf{z}^k \in \{0,1\}^k} f \left( \frac{\int_{\Theta} \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta') d\Pi(\theta')}{\ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta)} \right) \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) dH^{\otimes k}(\mathbf{a}^k) d\Pi(\theta), \quad (9.8)$$

where  $f = -\log$  in the original theory, yet in this study we consider the mutual information written with  $f = f_{\delta} : x \mapsto \frac{x^{\delta} - 1}{\delta(\delta-1)}$ ,  $\delta \in (0, 1)$ . This corresponds to the generalized mutual information based on  $\delta$ -divergence, introduced in chapter 4.

To be precise, the mutual information expression written in eq. (9.8) is not directly maximized w.r.t.  $\Pi$ , it is maximized asymptotically as  $k \rightarrow \infty$  w.r.t.  $\Pi$ . We refer to chapter 3 or chapter 4 for a more detailed definition of the reference prior. This asymptotic maximization is noted in this chapter as follows:

$$\Pi^* \in \arg \max_{\pi \in \mathcal{P}; k \rightarrow \infty} \mathcal{J}^k(\Pi), \quad (9.9)$$

where  $\mathcal{P}$  is the set of all possible priors, namely the set of continuous  $\sigma$ -finite measures on  $\Theta$ .

It has been proven (see chapters 3 and 4) that the optimization problem (9.9) admits the Jeffreys prior as a solution. That prior is defined from its density  $J$ , expressed up to a constant as:

$$J(\theta) \propto \sqrt{|\det \mathcal{I}(\theta)|} \quad \text{with} \quad \mathcal{I}(\theta) = - \sum_{z \in \{0,1\}} \int_{\mathcal{A}} \nabla_{\theta}^2 \log \ell(z | a, \theta) \ell(z | a, \theta) dH(a). \quad (9.10)$$

As expressed in the introduction, we aim at applying the work done in chapter 5, in which it is suggested to add a constraint in the optimization problem (9.9):

$$\Pi^* \in \arg \max_{\Pi \in \mathcal{P}; k \rightarrow \infty} \mathcal{J}^k(\Pi), \quad \text{subject to } \mathcal{C}(\Pi) < \infty, \quad (9.11)$$

where  $\mathcal{C}(\Pi) = \int a_c(\theta) d\Pi(\theta)$  for some function  $a_c(\cdot)$ . That function is sought to regularize the Jeffreys prior decay rates. If it verifies

$$\int_{\Theta} J(\theta) a_c(\theta)^{1+1/\delta} d\theta < \infty, \quad \text{and} \quad \int_{\Theta} J(\theta) a_c(\theta)^{1/\delta} d\theta < \infty \quad (9.12)$$

then the prior whose density is proportional to  $\theta \mapsto J(\theta) \cdot a_c^{1/\delta}(\theta)$  is the solution of problem (9.11), and is a proper prior.

### 9.3.2 Application to the probit-lognormal reference prior

In chapter 8, the Jeffreys prior decay rates have been comprehensively studied. We recall that

$$\forall \alpha > 0, J(\theta) \underset{\beta \rightarrow 0}{=} \frac{D'(\alpha)}{\beta} \quad \text{and} \quad J(\theta) \underset{\beta \rightarrow \infty}{=} \frac{E'}{\alpha \beta^3}, \quad (9.13)$$

$$\forall \beta > 0, J(\theta) \underset{\log \alpha \rightarrow \pm \infty}{=} G''(\beta) \frac{|\log \alpha|}{\alpha} \exp\left(-\frac{(\log \alpha - \mu)^2}{2\beta^2 + 2\sigma^2}\right), \quad (9.14)$$

where  $D'(\alpha) > 0$  depends only on  $\alpha$ ,  $E' > 0$  is a constant, and  $G''(\beta) > 0$  depends only on  $\beta$ .

The Jeffreys prior has an improper tail at  $\beta \rightarrow 0$  which does not shrink when it is multiplied by a degenerate likelihood. While we do not exactly know the shape of the prior on the whole domain, we assume that we can associate it to (i) an improper marginal distribution w.r.t.  $\beta$  whose density follows asymptotically the rates in eq. (9.13), and (ii) a proper distribution of  $\alpha$  conditionally to  $\beta$ .

It is thus possible to construct a constraint that will adjust the decay rate at  $\beta \rightarrow 0$  of the Jeffreys prior in order to make it integrable. We chose  $a_c(\theta) = \theta^\kappa$ , that satisfies eq. (9.12) for  $\kappa \in (0, \frac{2\delta}{1+\delta})$ . The solution of the optimization problem (9.11) is then the prior whose density  $\pi^*$  verifies  $\pi^*(\theta) \propto J(\theta)\beta^{\kappa/\delta}$ . In this work, we denote by  $\Pi_\gamma^*$  this prior and by  $\pi_\gamma^*$  its density, where  $\gamma := \kappa/\delta \in (0, \frac{2}{1+\delta}) \subset (0, 2)$ :

$$\pi_\gamma^*(\theta) \propto J(\theta)\beta^\gamma. \quad (9.15)$$

### 9.3.3 Practical derivation with integral interpolations

The prior  $\Pi_\gamma^*$  can be implemented by approximating numerically its density  $\pi_\gamma^*$ , using the theoretical expression of the Jeffreys prior. We recall that, in chapter 8, we proved that the Fisher information matrix can be expressed as

$$\mathcal{I}(\theta) = \begin{pmatrix} \frac{1}{\alpha^2 \beta^2} (A_{01} + A_{02}) & \frac{1}{\alpha \beta^3} (A_{11} + A_{12}) \\ \frac{1}{\alpha \beta^3} (A_{11} + A_{12}) & \frac{1}{\beta^4} (A_{21} + A_{22}) \end{pmatrix}, \quad (9.16)$$

by defining  $A_{01}, A_{02}, A_{11}, A_{12}, A_{21}, A_{22}$  as

$$\begin{aligned} A_{0u} &= \int_{\mathcal{A}} \frac{\Phi'(\beta^{-1} \log \frac{a}{\alpha})^2}{\Phi((-1)^{u+1} \beta^{-1} \log \frac{a}{\alpha})} h(a) da, \\ A_{1u} &= \int_{\mathcal{A}} \log \frac{a}{\alpha} \frac{\Phi'(\beta^{-1} \log \frac{a}{\alpha})^2}{\Phi((-1)^{u+1} \beta^{-1} \log \frac{a}{\alpha})} h(a) da, \\ A_{2u} &= \int_{\mathcal{A}} \log^2 \frac{a}{\alpha} \frac{\Phi'(\beta^{-1} \log \frac{a}{\alpha})^2}{\Phi((-1)^{u+1} \beta^{-1} \log \frac{a}{\alpha})} h(a) da, \end{aligned} \quad (9.17)$$

for  $u \in \{1, 2\}$ , and where  $h$  denotes the p.d.f. of the IM.

Using the above, it is possible to approximate  $J(\theta)$  by conducting Simpson's interpolations of the involved integrals in the Fisher information matrix. The interpolations are performed from a large dataset of IM values. We refer to chapter 7 where the seismic signal generator used in this work to provide a large dataset of seismic signals is presented. The empirical distribution of the PGA of the seismic signals in that dataset is recalled in section 9.3.3. The interpolations are done a regular grid of  $\mathcal{A} = [a_{\min}, a_{\max}]$ , where  $a_{\min}$  (resp.

$a_{\max}$ ) is the minimal (reps. the maximal) value of the IM in the dataset. The density  $h$  is approximated by Gaussian kernel density estimation.

Due to the computational cost of the numerical interpolations involved when approximating the Jeffreys prior, we conducted the same methodology as in [chapter 8](#). This methodology consists in evaluating it on an experimental design based on a fine-mesh grid  $\Theta$ . These evaluations allow building an interpolated approximation of the Jeffreys prior density matching the design. The density  $\pi_\gamma^*$  is then obtained by multiplying the interpolated approximated Jeffreys by  $\beta^\gamma$ .

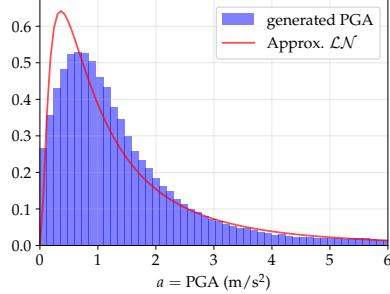


Figure 9.1: Comparison of the histograms of the PGA values of generated signals (blue) with a lognormal density (red) that has same median (i.e. 0) and same log deviation (i.e. 1).

## 9.4 Variational approximation of the constrained reference prior

### 9.4.1 Definition of the VA-RP

The second approximation of the constrained reference prior that we conduct results form a variational inference methodology. We recall that variational inference encompasses a range of techniques designed to approximate probability distributions by solving an optimization problem (such as maximizing the evidence lower bound, as in Kingma and Welling, [2014](#)). Since our goal is to find the prior that maximizes the mutual information defined in eq. [\(9.8\)](#), variational inference naturally lends itself to the approximation of reference priors.

The approach that we conduct is based on the variational approximation of the reference prior (VA-RP) proposed in [chapter 6](#). It consists in limiting the family of priors to a parametric class  $\{\Pi_\lambda, \lambda \in \Lambda\}$ , with  $\Lambda \subset \mathbb{R}^L$ , thereby reducing the problem to a finite-dimensional setting. Consequently, the optimization tasks in equations eqs. [\(9.9\)](#) and [\(9.11\)](#) translate into maximizing  $\mathcal{I}^k(\Pi_\lambda)$  over  $\lambda \in \Lambda$ . We define the priors  $\Pi_\lambda$  implicitly through the transformation:

$$\theta \sim \Pi_\lambda \iff \theta = g(\lambda, \varepsilon), \quad \text{where } \varepsilon \sim \mathbb{P}_\varepsilon. \quad (9.18)$$

Here,  $g$  is a parameterized measurable function —typically a neural network— where  $\lambda$  represents its weights and biases, it has to be differentiable with respect to  $\lambda$ . The variable  $\varepsilon$  serves as a latent variable, drawn from a simple, easy-to-sample distribution  $\mathbb{P}_\varepsilon$ . In practice, we choose a centered multivariate Gaussian  $\mathcal{N}(0, I_{p \times p})$ .

### 9.4.2 Optimization under constraints

The VA-RP is sought to approximate the solution of the optimization problems [\(9.9\)](#) or [\(9.11\)](#).

To do so, the problem is considered for a fixed value of  $k$ . Also, instead of maximizing the mutual information, the method is maximizes the following objective function, which is a lower bound of the mutual information:

$$\lambda^* \in \arg \max_{\lambda \in \Lambda} \mathcal{B}^k(\Pi_\lambda), \quad \mathcal{B}^k = \mathbb{E}_{\theta \sim \Pi_\lambda} \sum_{\mathbf{z}^k \in \{0,1\}^k} \int_{\mathcal{A}^k} f_\delta \left( \frac{\ell_k(\mathbf{z}^k | \mathbf{a}^k, \hat{\theta}^{\text{MLE}})}{\ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta)} \right) \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) h(a) da, \quad (9.19)$$

where  $\hat{\theta}^{\text{MLE}}$  refers to the maximum likelihood estimator:  $\ell_k(\mathbf{z}^k | \mathbf{a}^k, \hat{\theta}^{\text{MLE}}) = \max_{\theta \in \Theta} \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta)$ .

The optimization is conducted by stochastic gradient ascent, provided that the gradient of the objective function  $\mathcal{B}^k$  w.r.t.  $\lambda$  is given by

$$\mathbb{E}_{\varepsilon \sim \mathbb{P}_\varepsilon} \left[ \sum_{j=1}^2 \partial_{\theta_j} \tilde{\mathcal{B}}^k(g(\lambda, \varepsilon)) \nabla_\lambda g_j(\lambda, \varepsilon) \right], \quad (9.20)$$

where the gradients of the neural network  $\nabla_\lambda g(\lambda, \varepsilon)$  are computed in the algorithm via automatic differentiation, and the terms  $\partial_{\theta_j} \tilde{\mathcal{B}}^k(g(\lambda, \varepsilon))$  depend on the likelihood of the probit-lognormal model:

$$\partial_{\theta_j} \tilde{\mathcal{B}}^k(\theta) = \sum_{\mathbf{z}^k \in \{0,1\}^k} \int_{\mathcal{A}^k} \partial_{\theta_j} \log \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) F_\delta \left( \frac{\ell_k(\mathbf{z}^k | \mathbf{a}^k, \hat{\theta}^{\text{MLE}})}{\ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta)} \right) \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) h(a) da, \quad (9.21)$$

where  $F_\delta(x) = f_\delta(x) - xf'_\delta(x)$ . The derivatives of the log-likelihood here equal the following:

$$\begin{aligned} \partial_\alpha \log \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) &= -\frac{1}{\alpha\beta} z \frac{\Phi'(\beta^{-1} \log \frac{a}{\alpha})}{\Phi(\beta^{-1} \log \frac{a}{\alpha})} + \frac{1}{\alpha\beta} (1-z) \frac{\Phi'(\beta^{-1} \log \frac{a}{\alpha})}{1 - \Phi(\beta^{-1} \log \frac{a}{\alpha})}, \\ \partial_\beta \log \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) &= -\frac{\log \frac{a}{\alpha}}{\beta^2} z \frac{\Phi'(\beta^{-1} \log \frac{a}{\alpha})}{\Phi(\beta^{-1} \log \frac{a}{\alpha})} + \frac{\log \frac{a}{\alpha}}{\beta^2} (1-z) \frac{\Phi'(\beta^{-1} \log \frac{a}{\alpha})}{1 - \Phi(\beta^{-1} \log \frac{a}{\alpha})}. \end{aligned} \quad (9.22)$$

Still following the method of [chapter 6](#), the gradient ascent algorithm is adapted to approximate explicitly the solution of the optimization problem [\(9.11\)](#). This adaptation consists in actually solving

$$\lambda^* \in \arg \max_{\lambda \in \Lambda} \mathcal{B}^k(\Pi_\lambda) \quad \text{subject to } \mathcal{C}(\Pi_\lambda) = 0, \quad (9.23)$$

where  $\mathcal{C}(\Pi_\lambda) = \int_{\Theta} a(\theta) d\Pi_\lambda(\theta) - c/\mathcal{K}$ , with

$$\mathcal{K} = \int_{\Theta} J(\theta) a(\theta)^{1/\delta} d\theta \quad \text{and} \quad c = \int_{\Theta} J(\theta) a(\theta)^{1+1/\delta} d\theta. \quad (9.24)$$

The above constants are derived using the numerical approximation of the Jeffreys prior density proposed in section [9.3.3](#). For more details on the conduction of the optimization process, we refer to [chapter 6](#).

### 9.4.3 Neural network architecture

A single layer neural network is implemented for this problem. It takes the following form

$$\varepsilon \mapsto \begin{pmatrix} \mathbf{t}_1(\varepsilon) \\ \mathbf{t}_2(\varepsilon) \end{pmatrix} \mapsto \begin{pmatrix} \tilde{\varepsilon}_1 = \mathbf{w}_1^\top \mathbf{t}_1(\varepsilon) + b_1 \\ \tilde{\varepsilon}_2 = \mathbf{w}_2^\top \mathbf{t}_2(\varepsilon) + b_2 \end{pmatrix} \mapsto \begin{pmatrix} \alpha = u_1(\tilde{\varepsilon}_1) \\ \beta = u_2(\tilde{\varepsilon}_2) \end{pmatrix}. \quad (9.25)$$

The coordinates of the vectors  $\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{R}^p$  and  $\mathbf{b} = (b_1, b_2) \in \mathbb{R}^2$  constitute the weights of the neural network, they are initialized independently and randomly w.r.t. a Gaussian distribution.  $\mathbf{t}_1, \mathbf{t}_2, u_1, u_2$  are activation functions. They are specifically chosen in order to ensure the VA-RP's decay rates match with the ones of the target prior:  $u_1 = u_2 = \exp$  and for  $\mathbf{x} \in \mathbb{R}^p$ ,  $\mathbf{t}_1(\mathbf{x}) = (t_1(x_i))_i$  and  $\mathbf{t}_2(\mathbf{x}) = (t_2(x_i))_i$  with

$$t_1(x) = x; \quad t_2(x) = \log(1 - \Phi(x)), \quad (9.26)$$

In section [9.7](#), we prove that under those settings the VA-RP yields marginal distributions  $p_\alpha, p_\beta$  w.r.t.  $\alpha, \beta$ , that take the following form:

$$\begin{aligned} p_\alpha(\alpha) &= \frac{1}{\alpha \sqrt{2\pi} \|\mathbf{w}_1\|_2} \exp \left( -\frac{(\log \alpha - b_1)^2}{2 \|\mathbf{w}_1\|_2^2} \right) \\ p_\beta(\beta) &= \sum_{i, w_{2i} > 0} K_i \beta^{-\frac{1}{w_{2i}} - 1} \mathbb{1}_{\beta > e^{b_2}} + \sum_{i, w_{2i} < 0} K_i \beta^{-\frac{1}{w_{2i}} - 1} \mathbb{1}_{\beta < e^{b_2}}, \end{aligned} \quad (9.27)$$

where  $(w_{2i})_{i=1}^p$  denote the coordinates of  $\mathbf{w}_2$ , and the  $K_i$  are constants that depend on  $\mathbf{w}_2$  and  $b_2$ . The marginal distribution  $p_\alpha$  is a log-normal distribution, which is similar, up to the  $\log \alpha$  term, to the Jeffreys prior decay rates w.r.t.  $\alpha$  (eq. (9.14)). Concerning  $p_\beta$ , calling  $w_{2j}^{-1} = \min\{w_{2i}^{-1}, w_{2i} > 0\}$  and  $w_{2l}^{-1} = \max\{w_{2i}^{-1}, w_{2i} < 0\}$ , it admits the following decay rates:

$$p_\beta(\beta) \underset{\beta \rightarrow 0}{\sim} K_j \beta^{-\frac{1}{w_{2l}} - 1}; \quad p_\beta(\beta) \underset{\beta \rightarrow \infty}{\sim} K_l \beta^{-\frac{1}{w_{2j}} - 1}. \quad (9.28)$$

They are consistent with the ones of the Jeffreys prior.

## 9.5 Comparison between the VA-RP and the interpolated reference prior

In this section, we compare the approximation of the target prior using the VA-RP approach with the approximation by numerical interpolation of the theoretical expression of the Jeffreys prior density (denoted AJ). We consider that the AJ is a more accurate approximation of the target than the VA-RP. For this reason, this section mostly serves the assessment of the VA-RP methodology in the context of the probit-lognormal model.

We implemented the probit-lognormal modeling assuming that the p.d.f.  $h$  of the IM is defined by

$$h(a) = \frac{1}{\sqrt{2\pi\sigma_a^2 a}} \exp\left(-\frac{(\log a - \mu_a)^2}{2\sigma_a^2}\right) \quad (9.29)$$

with  $\mu_a = 0$  and  $\sigma_a = 1$ . These parameters, which define the distribution of  $a$ , are the only elements that have a theoretical impact on the reference prior. Their value are derived to fit the empirical distribution of the PGA values of the seismic signals that we consider. That empirical distribution is compared with the p.d.f. of eq. (9.29) in section 9.3.3.

**Training of the unconstrained VA-RP** First, we present the training process of the unconstrained VA-RP. It is expected to approach the unconstrained reference prior, i.e. the Jeffreys prior. As for the different parameters for the training, we take  $k = 50$  and  $p = 50$ . The training consists in the maximization of  $\mathcal{B}^k$  with  $\delta = 0.5$ . The optimization has been done through a number of 9000 epochs, with a learning rate of 0.005.

Figure 9.2 shows the evolutions of the weights of the neural network that define the marginal distributions of  $p_\alpha$  and  $p_\beta$  (see eq. (9.27)), as a function of the number of epochs. Thus, the bias  $b_1$  is constant and equal to 0, which is consistent with the asymptotic distribution of  $\alpha$ .  $1/w_{2l}$  and  $1/w_{2j}$  both tend to 0. As a result, the marginal distribution of  $\beta$  tends to  $1/\beta$ . This is an improper distribution which does not exactly match the expected improper Jeffreys prior (see eq. (9.13)): it has a heavier tail as  $\beta \rightarrow \infty$  than the original Jeffreys. However, we will see at the end of this section that the information contained in the distribution is not concentrated in its tails. In reality, low values of  $\beta$  have huge weights (see the paragraph "Posterior evaluation"). Finally,  $\|\mathbf{w}_1\|_2$  tends to  $+\infty$ . This means that the variance of  $\log \alpha$  tends to  $+\infty$  as it is expected (see section 9.7), even though this does not lead to the target distribution. This last result is nevertheless not surprising. It is the consequence of the chosen architecture of the VA-RP:  $p_\alpha$  is a lognormal distribution (eq. (9.28)), which is not exactly the case for the Jeffreys prior (eq. (9.14)). What fundamentally distinguishes these two distributions is that (i) in the first case, that of Jeffreys, the variance of  $\log \alpha$  is finite conditionally to  $\beta$ , and is infinite otherwise because of the heavy tail of the distribution w.r.t.  $\beta$  while (ii) in the second case, the variance becomes infinite to correspond, via the training process, to the target variance of the Jeffreys prior. In doing so, the marginal distribution of  $\alpha$  tends towards the improper distribution  $1/\alpha$ . Although this result is open to criticism, it appears to be an "acceptable" compromise insofar as the model's likelihood tends to substantially attenuate the prior's tails with respect to  $\alpha$ , thereby limiting the influence of the prior's decay rates on the posterior. An alternative approach could involve modifying the neural network architecture to better capture the correlation between  $\alpha$  and  $\beta$ . However, such adjustments would inevitably increase the computational complexity of the training process.

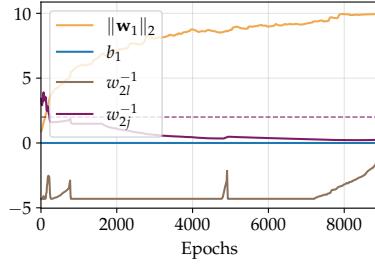


Figure 9.2: Weights of interest of the neural network during the unconstrained optimization. The brown and violet curves represent respectively  $1/w_{2l}$  and  $1/w_{2j}$  that are defined in section 9.4.3. In violet dashed line is plotted the target value of  $1/w_{2j}$ .

**Training of the constrained VA-RP** As for the constrained case, we remind that we introduce a constraint of the form  $\mathcal{C}(\Pi_\lambda) = 0$  with  $\mathcal{C}(\Pi_\lambda) = \int_{\Theta} \beta^\kappa d\Pi_\lambda - \mathcal{K}/c$ . We fix  $\kappa = 0.15$ , so that  $\gamma = \kappa/\delta = 0.3$ . In this case the VA-RP is expected to approximate the constrained reference prior, i.e. the prior whose density  $\pi^*$  is proportional to  $\theta \mapsto J(\theta)\beta^\gamma$ . Its asymptotic rates are identical to those of the Jeffreys prior for  $\alpha$ . If we fix  $\alpha$ , we have then:

$$\begin{cases} \pi^*(\theta) \propto \beta^{\gamma-1} & \text{as } \beta \rightarrow 0 \\ \pi^*(\theta) \propto \beta^{\gamma-3} & \text{as } \beta \rightarrow +\infty. \end{cases} \quad (9.30)$$

In this case, the development done in section 9.7 suggests that this prior verifies  $\mathbb{E}_{\theta \sim \pi^*} |\log \alpha|^2 = \infty$ . The training process of the constrained VA-RP is done during 12000 epochs with a learning rate of 0.001.

Figure 9.3.(a) suggests that the constraint seems reasonably verified during the whole optimization as the constraint gap is close to zero and appears to be decreasing. Figure 9.3(b) shows that the inverse weights  $1/w_{2l}$  and  $1/w_{2j}$  appear to steadily approach their theoretical values, respectively  $-\gamma = -0.3$  and  $2 - \gamma = 1.7$ . As in the unconstrained case, we observe that  $\|w_1\|_2$  tends to  $+\infty$ . This is consistent with the theoretical results of section 9.7.2 which suggest, recall, that for all  $\gamma \geq 0$ ,  $\mathbb{E}[|\log \alpha|^2] = \infty$ . So, we verify that adding constraints therefore allows, as expected, to “control” the distribution of  $\beta$ . Concerning  $\alpha$ , we find the same trend as in the unconstrained case, which is theoretically expected from the point of view of the variance of the marginal distribution.

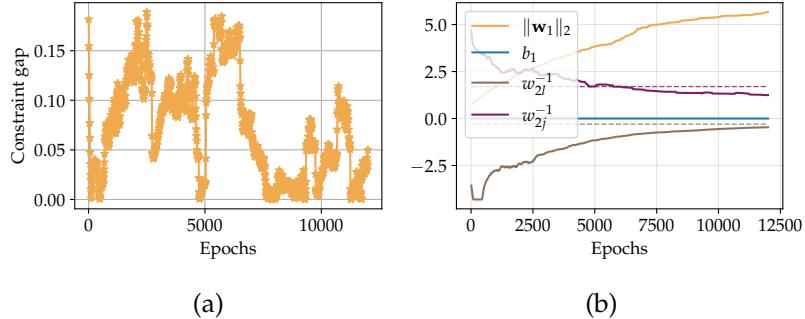


Figure 9.3: (a) Evolution of the constraint value gap during training. It corresponds to the difference between the target and current values for the constraint (in absolute value). (b) Weights of interest of the neural network during the constrained optimization. The yellow and blue curves represent respectively  $\|w_1\|_2$  and  $b_1$ , where  $\|w_1\|_2^2$  is the variance of the log of the first component, and  $b_1$  its mean. The brown and violet curves represent respectively  $1/w_{2l}$  and  $1/w_{2j}$  that are defined in section 9.4.3. In same color dashed line are plotted their target value.

**Posterior evaluation** We evaluate the posterior that the VA-RPs derived in this section yield. For this purpose, we take as dataset  $k = 50$  samples  $(\mathbf{z}^k, \mathbf{a}^k)$ ,  $\mathbf{z}^k = (z_i^k)_{i=1}^k$ ,  $\mathbf{a}^k = (a_i^k)_{i=1}^k$ , that are generated given a

probit-lognormal model with  $\theta_{\text{true}}$  close to  $(3.37, 0.43)$ .

*A posteriori* samples are generated for both approaches by using an adaptive Metropolis-Hastings (M-H). In the case of the AJ approach, the M-H is implemented to generate samples using the approximated posterior density. In the case of the VA-RP, the M-H is implemented to generate samples using the posterior density  $q(\cdot | \mathbf{z}^k, \mathbf{a}^k)$  density of  $\varepsilon$ :

$$q(\varepsilon | \mathbf{z}^k, \mathbf{a}^k) \propto q(\varepsilon) \ell_k(\mathbf{z}^k | \mathbf{a}^k, g(\lambda, \varepsilon)), \quad (9.31)$$

where  $q(\cdot)$  denotes the p.d.f. of  $\varepsilon$ . *A posteriori* samples of  $\theta$  are obtained by applying the neural network  $g(\lambda, \cdot)$  to the *a posteriori* samples of  $\varepsilon$ . For both methods, a number of 5000 samples of  $\theta$  are generated this way. In this example, we precise that the likelihood is not degenerate.

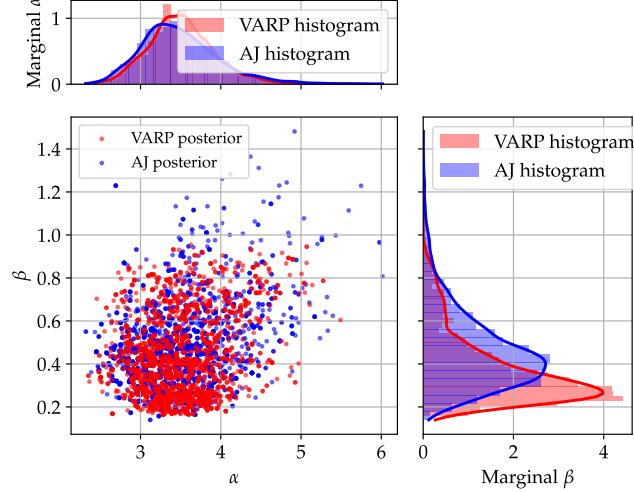


Figure 9.4: Scatter histogram of the unconstrained VA-RP posterior and the approximated Jeffreys (AJ) posterior distributions obtained from 5000 samples. Kernel density estimation is used on the marginal distributions in order to approximate their density functions with Gaussian kernels.

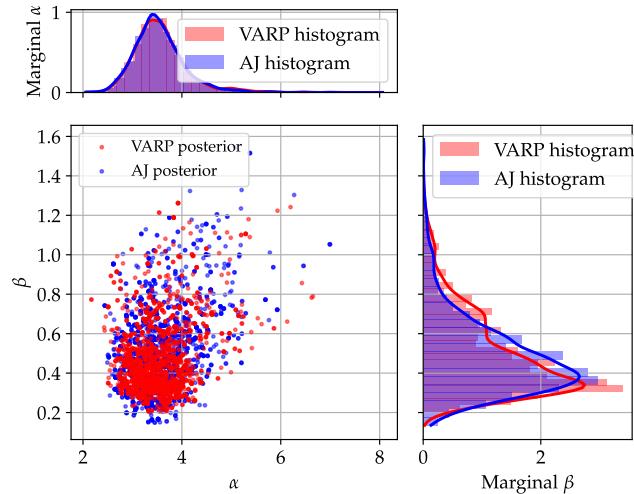


Figure 9.5: Scatter histogram of the constrained VA-RP posterior and the approximated Jeffreys (AJ) posterior distributions obtained from 5000 samples. Kernel density estimation is used on the marginal distributions in order to approximate their density functions with Gaussian kernels.

Figures 9.4 and 9.5 show that the posterior distribution obtained from the VA-RP looks more like Jeffreys prior in the constrained case than in the unconstrained one. More specifically, w.r.t.  $\alpha$ , the VA-RP is similar to

the Jeffreys prior in both cases. That suggests that the limitation of our neural network architecture discussed earlier has little impact on the posterior distribution. Regarding  $\beta$ , in the unconstrained case, the VA-RP ascribes predominant weight to smaller values of  $\beta$  compared to the Jeffreys prior, with a mode that is closer to 0. In terms of fragility curve estimation that could represent an issue, as it may lead to estimates with thinner credibility intervals around a fragility curve.

## 9.6 Application of the method on the piping system case study

In this section, we propose to apply the VA-RP that we have constructed and optimized above to an industrial problem. The industrial problem, consists in the piping system that was introduced in [chapter 7](#) (section [7.4](#)). This application serves as a proof of concept, we intend to show that the constrained VA-RP allows estimating efficiently and accurately seismic fragility curves in practice.

We recall that a validation dataset made of  $8 \cdot 10^4$  samples is available for this case study and permit constructing a reference fragility curves  $P_f^{\text{ref}}$  following the non-parametric method described in [chapter 7](#) (section [7.3](#)). In this study, we consider the PGA as IM.

### 9.6.1 Benchmarking metrics

In this chapter, we evaluate our estimates by comparing them to the reference fragility curve  $P_f^{\text{ref}}$  that is derived using a validation dataset. We recall that the reference was compared in [chapter 7](#) to a probit-lognormal curve  $P_f^{\text{MLE}}$  whose parameters are estimated by maximum likelihood estimation using the same validation dataset. That comparison demonstrates the adequacy of such modeling of the fragility curve. However, it also emphasizes an existing bias that the model has compared with the non-parametric estimate. In the following, we refer to this difference between  $P_f^{\text{MLE}}$  and  $P_f^{\text{ref}}$  as the square model bias, denoted by  $\mathcal{MB}$ :

$$\mathcal{MB} = \|P_f^{\text{ref}} - P_f^{\text{MLE}}\|_{L^2}^2; \quad (9.32)$$

where the norm  $\|\cdot\|_{L^2}$  is defined by  $\|P\|_{L^2}^2 = \int_{a_{\min}}^{a_{\max}} P(a)^2 da$  where the domain  $[a_{\min}, a_{\max}]$  covers the available seismic signals. The norm is derived numerically from Simpson's interpolation.

Given an observed sample of size  $k$ :  $(\mathbf{z}^k, \mathbf{a}^k)$ ,  $\mathbf{z}^k = (z_i)_{i=1}^k$ ,  $\mathbf{a}^k = (a_i)_{i=1}^k$ , we denote by  $P_f^{|\mathbf{z}^k, \mathbf{a}^k} : a \mapsto \Phi(\beta^{-1} \log(a/\alpha))$  the process whose distribution inherits from the posterior distribution of  $\theta$  conditionally to  $(\mathbf{z}^k, \mathbf{a}^k)$ . For any value of  $a$ , we note  $q_r^{|\mathbf{z}^k, \mathbf{a}^k}(a)$  the  $r$ -quantile of  $P_f^{|\mathbf{z}^k, \mathbf{a}^k}(a)$ , and  $m^{|\mathbf{z}^k, \mathbf{a}^k}(a)$  its median. We propose to evaluate our estimates using the following metrics:

- The square bias to the median  $\mathcal{B}^{|\mathbf{z}^k, \mathbf{a}^k} = \|m^{|\mathbf{z}^k, \mathbf{a}^k} - P_f^{\text{ref}}\|_{L^2}^2$ .
- The square credibility width  $\mathcal{W}^{|\mathbf{z}^k, \mathbf{a}^k} = \|q_{1-r/2}^{|\mathbf{z}^k, \mathbf{a}^k} - q_{r/2}^{|\mathbf{z}^k, \mathbf{a}^k}\|_{L^2}^2$ , with  $r = 0.05$ .
- The upper quantile error  $\mathcal{Q}^{|\mathbf{z}^k, \mathbf{a}^k} = \|q_{1-r/2}^{|\mathbf{z}^k, \mathbf{a}^k} - P_f^{\text{ref}}\|_{L^2}^2$ .

The quantiles and the medians are approximated using generated samples.

### 9.6.2 Results

In fig. [9.6](#) we provide examples of *a posteriori* fragility curves credibility intervals yielded by the two priors, with and without incorporating a constraint on the decay rates. These examples arise from given samples of  $k = 80$  observations. They illustrate that the VA-RP is capable of providing fragility curve estimates that are equivalent of the ones provided by the AJ, which is an approximation of the exact expression of the reference prior.

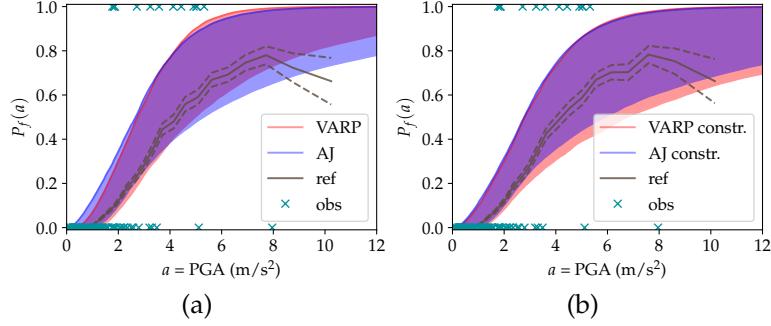


Figure 9.6: Example of *a posteriori* 95% credibility intervals yielded by the posterior distribution resulting from a sample of  $k = 80$  data and from the VA-RP (in red) or the AJ (in blue) : (a) unconstrained priors and (b) constrained priors. The reference curve  $P_f^{\text{ref}}$  (solid line) is surrounded by its 95% confidence interval (dashed line) in each figure. The cyan crosses represent the observed sample.

To deepen the comparison, we implemented the methods multiple times for values of  $k$  varying in  $[10, 250]$ . More precisely, for each  $k \in \{10, 20, \dots, 250\}$ , 10 random data samples of size  $k$  have been drawn. Each of these samples were then used to compute the metrics  $\mathcal{B}|z^k, a^k$ ,  $\mathcal{W}|z^k, a^k$  and  $\mathcal{Q}|z^k, a^k$ .

The average values of these metrics along with their 95% confidence intervals are presented in fig. 9.7 (for the unconstrained case) and fig. 9.8 (for the constrained case) as functions of  $k$ .

Figures 9.7 and 9.8 confirm previous results on a larger scale, namely that, with the current architecture, the constrained VA-RP outperforms the unconstrained VA-RP. Regarding bias, for example, it produces less erratic results. Compared to the AJ approximation, the VA-RP approximations exhibit slightly larger biases. The gap decreases as the number of data increases as expected, but only in the constrained case. Regarding the sizes of the credibility intervals (the values of  $\mathcal{W}|z^k, a^k$ ), they are generally smaller in the unconstrained case. This result is consistent with the observation made in fig. 9.4 concerning the marginal distribution of  $\beta$ . This shows smaller values than with the AJ approximation of the Jeffreys prior. It is important to emphasize that it is primarily the lower bounds of the credibility intervals that are “penalized” with the constrained VA-RP approximation. The upper bound remains consistent with that of the AJ, as reflected in the values of  $\mathcal{Q}|z^k, a^k$ . This result illustrates the relevance of using the constrained VA-RP approach for safety study. Indeed, in such a context, focusing on the upper quantile is consistent with a conservative approach.

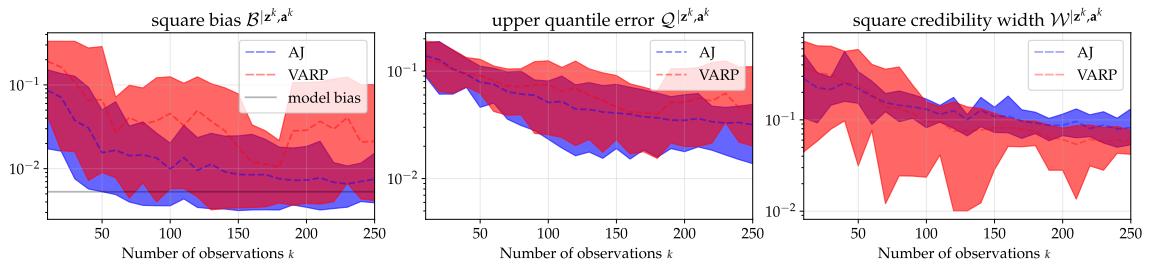


Figure 9.7: Average errors (dashed line)  $\mathcal{B}|z^k, a^k$ ,  $\mathcal{Q}|z^k, a^k$ ,  $\mathcal{W}|z^k, a^k$  and their 95% confidence intervals, derived from numerical replications of the unconstrained methods for each value of  $k$  in  $\{10, 20, \dots, 250\}$ . The values derived using the VA-RP (red) are compared with the ones using a numerical derivation to approximate the Jeffreys prior (blue). In the left figure, the square model bias  $\mathcal{MB}$  is plotted (solid line).

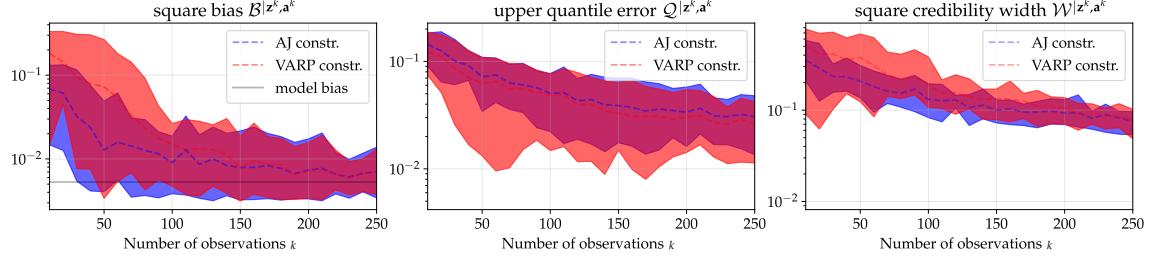


Figure 9.8: Same as fig. 9.7, but when a constraint is incorporated in the priors.

## 9.7 Impacts of the neural network's architecture on the VA-RP

### 9.7.1 Decay rates of the VA-RP density

In this section we elucidate the form taken by the marginal distribution of the VA-RP w.r.t.  $\beta$ . The VA-RP is assumed to be expressed as presented in section 9.4.3 (eq. (9.25)): the random variable  $\beta$  is defined as

$$\beta = \exp \left[ \sum_{i, w_{2i} \neq 0} w_{2i} \log(1 - \Phi(\varepsilon_i)) + b_i \right], \quad (9.33)$$

where the  $(\varepsilon_i)_i$  are independent standard Gaussian variables. We recall that the weights  $(w_{2i})_i$  are initialized randomly w.r.t. a Gaussian distribution. Therefore, they have almost surely distinct absolute values. We rely on the two following lemmas.

**Lemma 9.1.** *Let  $X$  be a random variable following a standard Gaussian distribution, and  $\mu > 0$ . Then*

- the r.v.  $Z = -\mu^{-1} \log(1 - \Phi(X))$  admits the density  $f_Z$  defined by  $f_Z(z) = \mu e^{-\mu z} \mathbb{1}_{z>0}$ , it is an exponential distribution with parameter  $\mu$ ;
- the r.v.  $\tilde{Z} = \mu^{-1} \log(1 - \Phi(X))$  admits the density  $f_{\tilde{Z}}$  defined by  $f_{\tilde{Z}}(z) = \mu e^{\mu z} \mathbb{1}_{z<0}$ , we call it an opposite exponential distribution with parameter  $-\mu$ .

**Lemma 9.2.** *Let  $Z_1, \dots, Z_n$  be independent exponential or opposite exponential distribution with parameters  $(\mu_i)_{i=1}^n$ , distinct in absolute value. Their sum  $\bar{Z}$  admits a density  $f_{\bar{Z}}$  that takes the form*

$$f_{\bar{Z}}(z) = \sum_{i, \mu_i > 0} K_i e^{-\mu_i z} \mathbb{1}_{z>0} + \sum_{i, \mu_i < 0} K_i e^{-\mu_i z} \mathbb{1}_{z<0}, \quad (9.34)$$

for some constants  $K_1, \dots, K_n$ .

The expression of  $\beta$  as written in eq. (9.33) expresses  $\log(\beta - b_2)$  as a sum of independent exponential or opposite exponential distributions with parameters  $(|w_{2i}|^{-1})_i$ , which are distinct. Thus, the r.v.  $\log(\beta - b_2)$  admits a density  $\tilde{p}$  that is defined as in eq. (9.34):

$$\tilde{p}(z) = \sum_{i, w_{2i} > 0} K_i e^{-\frac{z}{w_{2i}}} \mathbb{1}_{z>0} + \sum_{i, w_{2i} < 0} K_i e^{-\frac{z}{w_{2i}}} \mathbb{1}_{z<0}. \quad (9.35)$$

Given that  $p_\beta(\beta e^{b_2}) e^{b_2} \propto \tilde{p}(\log \beta) / \beta$  there exists some constants  $(\tilde{K}_i)_{i=1}^n$  such that

$$p_\beta(\beta) = \sum_{i, w_{2i} > 0} \tilde{K}_i \beta^{-\frac{1}{w_{2i}} - 1} \mathbb{1}_{\beta > e^{b_2}} + \sum_{i, w_{2i} < 0} \tilde{K}_i \beta^{-\frac{1}{w_{2i}} - 1} \mathbb{1}_{\beta < e^{b_2}}. \quad (9.36)$$

*Proof of lemma 9.1.* First,  $\Phi$  being the c.d.f. of a standard Gaussian distribution,  $\Phi(X)$  follows an uniform distribution in  $(0, 1)$ . The density  $f_Z(z) = \frac{e^{-\frac{z}{\mu}}}{\mu}$  involved in the first statement of the lemma is the p.d.f. of an exponential distribution with parameter  $\mu^{-1}$ . The c.d.f. of that distribution is defined by  $F_Z(z) = -\mu \log(1 - z)$ . Thus,  $-\mu \log(1 - \Phi(X))$  is an exponential distribution with parameter  $\mu^{-1}$ .

To conclude, notice that the second statement of the lemma simply results by elucidating the p.d.f. of  $\tilde{Z}$  from the one of  $Z$  given that  $\tilde{Z} = -Z$ .  $\square$

*Proof of lemma 9.2.* We prove this result by induction over  $n$ . When  $n = 1$ , the form in eq. (9.34) is consistent with the p.d.f. of an exponential or opposite exponential distribution.

We suppose this statement true for  $n - 1$  such r.v., we denote by  $\tilde{f}$  the p.d.f. of  $\sum_{i=1}^{n-1} Z_i$ , and by  $f_{Z_n}$  the one of  $Z_n$ . As  $\sum_{i=1}^{n-1} Z_i$  and  $Z_n$  are independent,  $f_{\bar{Z}}$  equals the convolution between  $\tilde{f}$  and  $f_{Z_n}$ :  $f_{\bar{Z}} = \tilde{f} * f_{Z_n}$ .

Before going further, let us derive the following integrals, for any  $0 < \nu_1 < \nu_2$ :

$$\begin{aligned} \int_{\mathbb{R}} e^{-\nu_1(y-x)} e^{-\nu_2 x} \mathbb{1}_{y-x>0} \mathbb{1}_{x>0} dx &= \frac{e^{-\nu_2 y} - e^{-\nu_1 y}}{\nu_2 - \nu_1} \mathbb{1}_{y>0}, \\ \int_{\mathbb{R}} e^{-\nu_1(y-x)} e^{\nu_2 x} \mathbb{1}_{y-x>0} \mathbb{1}_{x<0} dx &= \frac{e^{\nu_2 y}}{\nu_1 + \nu_2} \mathbb{1}_{y<0} + \frac{e^{-\nu_1 y}}{\nu_1 + \nu_2} \mathbb{1}_{y>0}, \\ \int_{\mathbb{R}} e^{\nu_1(y-x)} e^{\nu_2 x} \mathbb{1}_{y-x<0} \mathbb{1}_{x<0} dx &= \frac{e^{\nu_1 y} - e^{\nu_2 y}}{\nu_2 - \nu_1} \mathbb{1}_{y<0}. \end{aligned} \quad (9.37)$$

Thus, if  $f_{Z_n}(x) = \mu_n e^{-\mu_n x} \mathbb{1}_{x>0}$  for  $\mu_n > 0$  we get

$$\begin{aligned} f_{\bar{Z}}(z) &= \sum_{i < n, \mu_i > 0} \frac{K_i \mu_n}{\mu_i - \mu_n} e^{-\mu_i z} \mathbb{1}_{z>0} + \sum_{i < n, \mu_i < 0} \frac{K_i \mu_n}{\mu_i + \mu_n} e^{-\mu_i z} \mathbb{1}_{z<0} \\ &\quad - e^{-\mu_n z} \left[ \sum_{i, \mu_i > 0} \frac{K_i \mu_n}{\mu_i - \mu_n} + \sum_{i, \mu_i < 0} \frac{K_i \mu_n}{\mu_i + \mu_n} \right] \mathbb{1}_{z>0}; \end{aligned} \quad (9.38)$$

and if  $f_{Z_n}(x) = \mu_n e^{\mu_n x} \mathbb{1}_{x<0}$  for  $\mu_n > 0$  then

$$\begin{aligned} f_{\bar{Z}}(z) &= \sum_{i, \mu_i > 0} \frac{K_i \mu_n}{\mu_i + \mu_n} e^{-\mu_i z} \mathbb{1}_{z>0} + \sum_{i, \mu_i < 0} \frac{K_i \mu_n}{\rho - \mu_n} e^{-\mu_i z} \mathbb{1}_{z<0} \\ &\quad + e^{\mu_n z} \left[ \sum_{i, \mu_i > 0} \frac{K_i \mu_n}{\mu_i + \mu_n} + \sum_{i, \mu_i < 0} \frac{K_i \mu_n}{\mu_i - \mu_n} \right] \mathbb{1}_{z<0}. \end{aligned} \quad (9.39)$$

In any case, it fits the expected form.  $\square$

### 9.7.2 Variance of $\log \alpha$ a priori

In this section, we assume that the constrained reference prior has a density that takes the following form:

$$\pi^\gamma(\theta) = p(\alpha|\beta)p^\gamma(\beta), \quad p(\alpha|\beta) \propto \frac{|\log \alpha|}{\alpha} \exp\left(-\frac{(\log \alpha - \mu)^2}{2\beta^2 + 2\sigma^2}\right), \quad p^\gamma(\beta) \propto \frac{1}{\beta^{1-\gamma} + \beta^{3-\gamma}}, \quad (9.40)$$

where  $\mu \in \mathbb{R}$ ,  $\sigma > 0$ , and  $\gamma \in (0, 2)$  that is null if the reference prior is not constrained. The prior corresponds to (i) a marginal distribution w.r.t.  $\beta$  that has the same decay rates as the theoretical target, and (ii) a distribution of  $\alpha$  conditionally to  $\beta$  whose decay rates corresponds to the ones in eq. (9.14). We recall that  $\gamma \in (0, 2)$ .

We fix  $\beta > 0$ , let us consider a random variable  $X \sim \mathcal{N}(\mu, \Sigma^2)$ , where  $\Sigma = \Sigma(\beta) = \sqrt{\sigma^2 + \beta^2}$ . We can

write

$$\mathbb{E}|X| = \int_{-\infty}^{\infty} \frac{|x|}{\sqrt{2\pi\Sigma^2}} e^{-\frac{(x-\mu)^2}{2\Sigma^2}} dx = \frac{1}{\sqrt{2\pi\Sigma^2}} \int_0^{\infty} \frac{|\log y|}{y} e^{-\frac{(\log y - \mu)^2}{2\Sigma^2}} dy \quad (9.41)$$

$$\text{and } \mathbb{E}|X|^3 = \int_{-\infty}^{\infty} \frac{|x|^3}{\sqrt{2\pi\Sigma^2}} e^{-\frac{(x-\mu)^2}{2\Sigma^2}} dx = \frac{1}{\sqrt{2\pi\Sigma^2}} \int_0^{\infty} \frac{|\log y|^3}{y} e^{-\frac{(\log y - \mu)^2}{2\Sigma^2}} dy. \quad (9.42)$$

Equation (9.41) elucidates the conditional distribution of  $\alpha|\beta$  as the one whose density is

$$p_a^\gamma(\alpha|\beta) = (\mathbb{E}|X|\sqrt{2\pi\Sigma^2})^{-1} \frac{|\log \alpha|}{\alpha} \exp\left(-\frac{(\log \alpha - \mu)^2}{2\beta^2 + 2\sigma^2}\right). \quad (9.43)$$

This way,

$$\mathbb{E}[|\log \alpha|^2 | \beta] = \frac{\mathbb{E}|X|^3}{\mathbb{E}|X|}. \quad (9.44)$$

To derive the right-hand term of the above equation, we use the following (see Winkelbauer, 2014):

$$\begin{aligned} \mathbb{E}|X|^3 &= 2^{3/2}\Sigma^3 \frac{\Gamma(2)}{\Gamma(-\frac{3}{2})} \sum_{k=0}^{\infty} \frac{\Gamma(-\frac{3}{2}+k)}{\Gamma(\frac{1}{2}+k)} \frac{\left(-\frac{\mu}{2\Sigma^2}\right)^k}{k!}, \\ \text{and } \mathbb{E}|X| &= 2^{1/2}\Sigma \frac{\Gamma(1)}{\Gamma(-\frac{1}{2})} \sum_{k=0}^{\infty} \frac{\Gamma(-\frac{1}{2}+k)}{\Gamma(\frac{1}{2}+k)} \frac{\left(-\frac{\mu}{2\Sigma^2}\right)^k}{k!}. \end{aligned} \quad (9.45)$$

Using the identity  $z\Gamma(z) = \Gamma(z+1)$ , we derive

$$\frac{\Gamma(-\frac{3}{2}+k)}{\Gamma(\frac{1}{2}+k)} = \frac{\Gamma(-\frac{1}{2}+k)}{(-\frac{3}{2}+k)\Gamma(\frac{1}{2}+k)} = \frac{1}{(-\frac{3}{2}+k)(-\frac{1}{2}+k)}. \quad (9.46)$$

Going back to eq. (9.44), we deduce:

$$\mathbb{E}|\log \alpha|^2 = \mathbb{E}[\mathbb{E}[|\log \alpha|^2 | \beta]] = K' \int_0^{\infty} \frac{\beta^2 + \sigma^2}{\beta^{1-\gamma} + \beta^{3-\gamma}} \frac{\sum_{k=0}^{\infty} \frac{(-\mu(2\sigma^2 + \beta^2)^{-1})^k}{(2k-3)(2k-1)k!}}{\sum_{k=0}^{\infty} \frac{(-\mu(2\sigma^2 + \beta^2)^{-1})^k}{-(2k-1)k!}} d\beta, \quad (9.47)$$

with  $K' > 0$  being a normalization constant. To conclude, we perform an asymptotic analysis of the integrated function in the above equation as  $\beta \rightarrow \infty$ . The power series

$$\sum_{k=0}^{\infty} \frac{z^k}{(2k-3)(2k-1)k!} \quad \text{and} \quad \sum_{k=0}^{\infty} \frac{z^k}{-(2k-1)k!} \quad (9.48)$$

have infinite radius of convergence, so that when  $z \rightarrow 0$ , they converge respectively towards  $\frac{1}{3}$  and 1. Also, while  $z$  is a negative real number, the power series are positive.

Therefore, the term in the integral of eq. (9.47) is positive and asymptotically equivalent to  $\beta^{-1+\gamma}$  as  $\beta \rightarrow \infty$ . We deduce that for any  $\gamma \geq 0$ ,

$$\mathbb{E}[|\log \alpha|^2] = \infty. \quad (9.49)$$

## 9.8 Conclusion

In this work, we introduced a novel prior for conducting a Bayesian estimation of seismic fragility curves. This prior results from the application of the development of the reference prior theory that we conducted in the part I of this manuscript.

First, we defined an appropriate constraint that regularizes the Jeffreys prior decay rates, in order to make it proper. In this way, the prior is ensured to yield proper posteriors, even in cases where the likelihood is degenerate.

Second, the prior was implemented using two approaches: one that is close to its theoretical expression, but that is computationally expensive to sample from; one that result from a variational approximation. The latter requires a complex training of a neural network, which, once accomplished, gives a parameterized form of the prior that can be sampled from more efficiently. This implementation should still be regarded as a proof of concept. Indeed, the form of the VA-RP remains dependent of the chosen architecture. Also, while we proved that the chosen architecture provided a satisfying approximation of the target prior in our modeling, the first approach (i.e., the numerical derivation of the prior density based on integral interpolations) seemed to remain more reliable.

Third, the two approaches were compared and applied on a real case study. We believe that our results demonstrate that (i) the incorporated constraint provides an accurate and non-altered estimation of the fragility curves, and (ii) the VA-RP has been capable of accurately capturing the asymptotic behavior of the target prior, so that it provides competing estimates to the latter.

Looking forward, complexing the VA-RP architecture should allow to include a wider range of priors among the reachable priors by the method, including more accurate approximations of the target. However, such a prospect would increase the training cost of the neural network. Furthermore, The impact that the chosen constraint has on the estimates requires further investigation. Such an investigation will be conducted in the following chapter, where the constrained reference prior that is suggested in this chapter will be implemented.

# 10

## Design of experiments with constrained reference priors for robust inference

This chapter is mainly based on the submitted work: Antoine Van Biesbroeck, Clément Gauchy, Cyril Feau, and Josselin Garnier (2025b). “Robust a posteriori estimation of probit-lognormal seismic fragility curves via sequential design of experiments and constrained reference prior”. arXiv:2503.07343. DOI: [10.48550/arXiv.2503.07343](https://doi.org/10.48550/arXiv.2503.07343)

**Abstract** To estimate seismic fragility curves using the Bayesian framework, the observed data constitute the most objective source of information to consider. In this chapter, we conduct a Bayesian estimation of seismic fragility curves that aims to optimize that information. We consider datasets that have the disadvantages of (i) being scarce, and (ii) reducing structural responses to binary outcomes. However, we propose an experimental design that is sought to select seismic signals that are expected to maximize the impact of data on the posterior distribution. Moreover, our approach is supported by the reference prior theory, to ensure the prior itself maximizes the influence of data on the posterior. The reference prior is slightly constrained in this work in order to tackle occurrences of so-called degenerate likelihoods. The results demonstrate the ability of the method to efficiently and robustly estimate fragility curves, and to avoid degeneracy even with limited data. Additionally, we demonstrate that the estimates quickly reach the model bias induced by the probit-lognormal modeling.

<b>10.1</b> Introduction . . . . .	164
<b>10.2</b> Probit-lognormal modeling and constrained reference prior . . . . .	165
<b>10.3</b> Sequential design of experiments . . . . .	166
10.3.1 Methodology description . . . . .	166
10.3.2 Approximation of the index . . . . .	167
10.3.3 Stopping criterion . . . . .	168
<b>10.4</b> Estimates and benchmarking metrics . . . . .	169
<b>10.5</b> Application of the method on the piping system . . . . .	170
10.5.1 Short presentation of the case study . . . . .	170
10.5.2 Qualitative study . . . . .	170
10.5.3 Quantitative study . . . . .	171
10.5.4 Synthesis . . . . .	175
<b>10.6</b> Behavior of the method on toy case studies . . . . .	175
10.6.1 Description of the case studies . . . . .	175
10.6.2 Results . . . . .	176
10.6.3 Simple suggestion of a domain where IMs should be theoretically selected . . . . .	177

10.6.4 Additional mathematical derivations . . . . .	178
10.6.5 Equivalence between toy case studies . . . . .	180
<b>10.7 Conclusion . . . . .</b>	<b>180</b>

## 10.1 Introduction

In this chapter, we conduct a Bayesian estimation of seismic fragility curves. Seismic fragility curves, which are defined as the probability of failure of a mechanical equipment conditional to a given intensity measure (IM) of the seismic scenario, are reviewed in [chapter 7](#). We continue to focus on case studies where (i) the information about the structure’s response is limited to a binary outcome (i.e., failure or non-failure) and (ii) the available data are scarce.

In the Bayesian workflow, two sources of information are combined to issue the posterior distribution, which provides estimates of the quantities of interest. Specifically, *a priori* information coexists with information that comes from the observations, and that is incorporated through the statistical model.

This thought is fundamental when it resorts to estimating seismic fragility curves. Indeed, the developments done in [chapter 8](#) and [chapter 9](#) have shown that (i) the *a priori* information must be carefully defined, since any subjectivity embedded into its design may significantly impact the estimates, and (ii) the information that comes from the data can take “degenerate” forms, jeopardizing efficiency. In these chapters, we proved that appropriate prior designs can yield accurate estimates of seismic fragility curves without being affected by degenerate scenarios. These appropriate priors correspond to constrained reference priors, where the constraints are sought to regularize the decay rates of the objective Jeffreys prior in order to ensure it yields proper posteriors. Nevertheless, the influence of data on the posterior remains important. In practice, two datasets of the same size may differ considerably in the information they convey, with one contributing more effectively to the posterior than the other. Enhancing this contribution will inevitably guide the result “aside” from the prior, which will improve objectivity as well.

In this chapter, we propose a strategy of design of experiments (DoE), to get the most out of the probit-lognormal model, namely the most accurate and robust estimation possible with the minimum of data. Given a large database of synthetic seismic signals —generated to match the seismic scenario of interest— the proposed strategy intends to sequentially select the synthetic signals with which to perform the tests, in order to optimally estimate —by minimizing their number— the probit-lognormal estimations of fragility curves. Different strategies exist to conduct an experimental design. Generally, they are based on the definition of a criterion, such as the stepwise uncertainty reduction (SUR) one, introduced by Villemonteix, Vazquez, and Walter ([2009](#)). As for examples of studies in reliability analysis that conduct a criterion-based experimental design, we can cite Bect, Bachoc, and Ginsbourger, [2019](#); Azzimonti et al., [2021](#); Agrell and Dahl, [2021](#); Lartaud, Humbert, and Garnier, [2025](#). Our work proposes a strategy inherited from the reference prior theory, i.e., based on the information theory. We recall that the reference prior theory is introduced in [chapter 3](#). This strategy aims to maximize the impact that the selected data has on the posterior distribution of the fragility curve. This paradigm is not only used to construct an objective prior, but also to define a criterion from which is based our experimental design.

It is also important to recognize that the model itself provides information to the Bayesian workflow and influences the posterior. For instance, we draw attention to the work conducted in [appendix B](#), where a simple model that assumes a linear correlation between the logarithm of the structural response and that of the IM is implemented with a similar experimental design. The results presented in that appendix illustrate the possible limitations of models that lead to estimation biases whose highlight the limitations of the model, entailed by its irreducible bias. In this chapter, we acknowledge that the probit-lognormal model is itself inherently biased. However, we aim to demonstrate that this model can accurately estimate seismic fragility curves in scenarios with limited data if used as part of an approach built on a full understanding of the model’s limitations.

Finally, note that the strategy proposed in this chapter echoes the conclusions of the recent reference X. Zhu, Broccardo, and Sudret, [2023](#). In this reference, it is indeed recommended to use adaptive strategies to deal with problems involving high failure thresholds, even with the use of more sophisticated models than

the probit-lognormal model. We show here that this conclusion obviously applies to the probit-lognormal model. Our strategy therefore allows us to extend its domain of validity, by exploiting it as best we can whatever the failure threshold of interest.

The rest of the chapter is organized as follows. In the next section we start by briefly recalling the probit-lognormal modeling of fragility curves, and we present the constrained priors that we propose. The experimental design is explained and detailed in section 10.3, and benchmarking metrics that serve to evaluate the performances of our approach are defined in section 10.4. The method is then implemented on a case study in section 10.5, where a thorough analysis of our results is conducted. In order to provide more insights about the behavior of the DoE, we suggest another implementation of the method on toy case studies in section 10.6. Finally, a conclusion terminates this chapter in section 10.7.

## 10.2 Probit-lognormal modeling and constrained reference prior

We recall briefly the probit-lognormal model that was introduced in chapter 7. We suppose to observe realizations of the random vector  $(Z, A)$ , where  $A \in \mathcal{A} \subset (0, \infty)$  represents the IM value, and  $Z \in \{0, 1\}$  represents the binary outcome ( $Z = 1$  if the equipment fails and  $Z = 0$  otherwise). The probit-lognormal model consist in assuming the following parameterized distribution of  $(Z, A)$  conditionally to the parameter  $\theta = (\alpha, \beta) = \Theta = (0, \infty)^2$ :

$$A \sim A | \theta \sim H, \quad \text{and} \quad Z | A, \theta \sim \mathcal{B}(P_f(A)) \quad \text{with} \quad P_f(a) = \Phi\left(\frac{\log a - \log \alpha}{\beta}\right), \quad (10.1)$$

where  $\mathcal{B}$  refers to a Bernoulli distribution, and  $H$  is the distribution of  $A$ , its density is  $h$ . Given observations  $\mathbf{z}^k = (z_i)_{i=1}^k$ ,  $\mathbf{a}^k = (a_i)_{i=1}^k$ , the likelihood of this model is the following:

$$\ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) = \prod_{i=1}^k \ell(z_i | a_i, \theta) = \prod_{i=1}^k \Phi\left(\frac{\log a_i - \log \alpha}{\beta}\right)^{z_i} \left(1 - \Phi\left(\frac{\log a_i - \log \alpha}{\beta}\right)\right)^{1-z_i}. \quad (10.2)$$

The decay rates of this likelihood vary significantly as a function of the observations. We remind below the definition of a degenerate likelihood that we introduced in chapter 8, where the decay rates of the likelihood were comprehensively studied.

**Definition** (Likelihood degeneracy; reminder of definition 8.1). If the observed samples  $(\mathbf{z}^k, \mathbf{a}^k)$  belong to one of the following three types:

- type 1 : no failure is observed:  $z_i = 0$  for any  $i$ ;
- type 2 : only failures are observed:  $z_i = 1$  for any  $i$ ;
- type 3 : the failures and non-failures are partitioned into two disjoint subsets when classified according to their IM values: there exists  $a \in \mathcal{A}$  such that for any  $i, j$ ,  $a_i < a < a_j \iff z_i \neq z_j$ ;

then the likelihood is degenerate.

To conduct a Bayesian estimation of probit-lognormal fragility curves, one should select a prior carefully to ensure it issues a proper posterior even in degenerate scenarios. We define below adequate requirements for a valid *a posteriori* estimation, i.e., an estimation that results from a proper posterior.

**Definition 10.1.** The *a posteriori* estimation is valid if the posterior is proper. For seismic fragility curves it is the case if:

- the likelihood is not degenerate and the prior verifies

$$\forall \beta > 0, \pi(\theta) \underset{\log \alpha \rightarrow \pm \infty}{=} O(1),$$

$$\forall \alpha > 0, \pi(\theta) \underset{\beta \rightarrow 0}{=} O(1), \text{ and } \pi(\theta) \text{ is integrable in the neighborhood of } \beta \rightarrow \infty.$$

- the likelihood is degenerate, the prior is proper w.r.t  $\beta$  and verifies  $\forall \beta > 0, \pi(\theta) \underset{\log \alpha \rightarrow \pm \infty}{=} O(1)$ .

In this chapter, we select a prior that is supported by the reference prior theory, in order to minimize its subjectivity. We remind that the theory is reviewed and developed in the part I of this manuscript. In chapter 8 a reference prior that takes the form of a Jeffreys prior was derived and studied in the context of probit-lognormal fragility curves. This prior being inefficient in degenerate scenario, one constraint was proposed in chapter 9 to circumvent this problem. The prior approached in that work was given by the density:

$$\pi_\gamma(\theta) \propto J(\theta)\beta^\gamma, \quad \text{where } \gamma \in (0, 2). \quad (10.3)$$

The same approach is conducted here, where we propose a novel prior that is more computationally efficient. Indeed, since the prior will be evaluated multiple time to provide *a posteriori* estimates using Markov chain Monte Carlo (MCMC) techniques, we aim to suggest a prior that is not expensive to evaluate. That prior, which has a density that we denote  $\pi_\gamma^*$ , is sought to resemble to the properly constrained reference prior expressed in eq. (10.3):

$$\pi_\gamma^*(\theta) \propto \frac{1}{\alpha(\beta^{1-\gamma} + \beta^{3-\gamma})} \exp\left(-\frac{(\log \alpha - \mu_A)^2}{2\sigma_A^2 + 2\beta^2}\right), \quad (10.4)$$

where  $\mu_A$  and  $\sigma_A$  respectively denote the mean and the standard deviation of the r.v.  $\log A$ . The form of the prior density  $\pi_\gamma^*$  results partly from the consideration of Jeffreys prior's decay rates. With respect to  $\beta$ , the decay rates are the same as the ones of the constrained reference prior density in eq. (10.3) (see the study of Jeffreys prior's decays in chapter 8, section 8.3.3). With respect to  $\alpha$ , the prior density  $\pi_\gamma^*$  admits slightly different decay rates (a term equivalent to  $|\log \alpha|$  vanished). That choice results from two considerations: (i) especially outside degenerate likelihoods of types 1 or 2, this change has a negligible impact on the posterior, and (ii) in chapter 8 (section 8.3.2) we elucidated that the reference prior w.r.t.  $\alpha$  is close to the distribution of the IM, which is associated to a log-normal distribution. We precise that both priors in eqs. (10.3) and (10.4) have been tested when we conducted our numerical results to ensure they give non-distinguishable results.

While  $0 < \gamma < 2$ , the prior satisfies both criteria required for a valid *a posteriori* estimation (definition 10.1). The case  $\gamma = 0$  corresponds to the critical case where  $\pi_\gamma^*$  is similar to the Jeffreys prior density and thus provides a valid *a posteriori* estimation if and only if the likelihood is not degenerate. The limit case  $\gamma = 2$  never provides a valid *a posteriori* estimation. We precise that the methodology presented in the next section requires that the prior tackles degenerate-likelihood cases. The tuning of this hyper-parameter must be thought as the research for a balance between objectivity ( $\gamma$  closer to 0) and suitability for inference. The influence of  $\gamma$  is studied in the practical application of our method in section 10.5.

## 10.3 Sequential design of experiments

### 10.3.1 Methodology description

The priors defined in eq. (10.4) handle the degeneracy of the likelihood to provide valid *a posteriori* estimation of the fragility curve. However, it remains important to reduce the occurrence of this phenomenon because:

- even if it partially disappears thanks to the slightly informed priors we suggest, the possibility of obtaining estimates that tend towards unrealistic fragility curves always exists with very small data sizes;

- (ii) by nature, a likelihood becomes degenerate consequently to a lack of information within the observed data samples. Extinguishing degeneracy should thus lead to a better understanding of the structure's response and its fragility curve.

In this chapter, we propose to tackle degeneracy with an appropriate experimental design strategy. Different sequential methods exist. Suppose that we have observed the sample  $(\mathbf{z}^k, \mathbf{a}^k)$ , we need to create a criterion to choose the next input  $a_{k+1}$ . Given our analysis (statement (ii) above) from an information theory viewpoint, the strategy to build this criterion is to maximize the impact that the selected observation would have on the posterior distribution. This strategy echoes the one that supports the reference prior definition as a maximal argument of the mutual information, so that we suggest a similar criterion to select  $a_{k+1}$ :

$$a_{k+1} = \arg \max_{a \in \mathcal{A}} \mathcal{I}_{k+1}(a); \quad \mathcal{I}_{k+1}(a_{k+1}) = \mathbb{E}_{z_{k+1} | \mathbf{a}^{k+1}, \mathbf{z}^k} [D_\delta(p(\theta | \mathbf{z}^k, \mathbf{a}^k) || p(\theta | \mathbf{z}^{k+1}, \mathbf{a}^{k+1}))]. \quad (10.5)$$

The index  $\mathcal{I}_{k+1}$  can be seen as a sensitivity index measuring the sensitivity of the posterior w.r.t. the data (Da Veiga, 2015). Its sequential maximization amounts to maximize the impact of the data to the posterior. Our index can be seen as derived from the popular framework of Stepwise Uncertainty Reduction techniques (Villemonteix, Vazquez, and Walter, 2009). Typically, those methods are formulated with an *a posteriori* variance within the expectation in eq. (10.5) instead of the dissimilarity measure suggested here. Algorithm 1 proposes a practical pseudo-code of our methodology. It requires to derive an approximation of the index  $\mathcal{I}_{k+1}$ , which we elucidate in the following section.

---

#### Algorithm 1 Planning of experiments

---

**Notations:** Seismic signal :  $\mathcal{S}$

Intensity measure of the seismic signal:  $\text{IM}(\mathcal{S})$

Mechanical response to the seismic signal (failure or success):  $R(\mathcal{S})$

**Initialization:**  $k_0 > 0$  (in practice  $k_0 = 2$ ), initial seismic signals  $\mathcal{S}_1, \dots, \mathcal{S}_{k_0}$

Define the initial data:  $\mathbf{z}^{k_0} = (R(\mathcal{S}_1), \dots, R(\mathcal{S}_{k_0}))$ ,  $\mathbf{a}^{k_0} = (\text{IM}(\mathcal{S}_1), \dots, \text{IM}(\mathcal{S}_{k_0}))$

**for**  $k = k_0 \dots k_{\max} - 1$  **do**

    Approximate  $\mathcal{I}_{k+1}$  via Monte Carlo sampling

    Compute  $a_{k+1} = \arg \max_a \mathcal{I}_{k+1}(a)$

    Choose a seismic signal  $\mathcal{S}_{k+1}$  such that  $\text{IM}(\mathcal{S}_{k+1}) = a_{k+1}$

    Perform the experiment and define  $z_{k+1} = R(\mathcal{S}_{k+1})$

**end for**

---

### 10.3.2 Approximation of the index

Let us adopt the short notation

$$\Psi_a^z(\theta) = \Phi \left( \frac{\log a - \log \alpha}{\beta} \right)^z \left( 1 - \Phi \left( \frac{\log a - \log \alpha}{\beta} \right) \right)^{1-z}. \quad (10.6)$$

The index to maximize has the following form:

$$\mathcal{I}_{k+1}(a_{k+1}) = (\delta(\delta - 1))^{-1} \sum_{z \in \{0,1\}} \mathbb{P}(z_{k+1} = z | \mathbf{a}^{k+1}, \mathbf{z}^k) \Delta_{k+1}(\mathbf{a}^{k+1}, \mathbf{z}^{k+1}) \quad (10.7)$$

with :  $\mathbb{P}(z_{k+1} = z | \mathbf{a}^{k+1}, \mathbf{z}^{k+1}) = \int_{\Theta} \Psi_{a_{k+1}}^z(\theta) p(\theta | \mathbf{z}^k, \mathbf{a}^k) d\theta$ , and,

$$\begin{aligned} \Delta_{k+1}(\mathbf{a}^{k+1}, \mathbf{z}^{k+1}) &= \int_{\Theta} \left( \frac{p(\theta | \mathbf{z}^k, \mathbf{a}^k)}{p(\theta | \mathbf{z}^{k+1}, \mathbf{a}^{k+1})} \right)^\delta p(\theta | \mathbf{z}^{k+1}, \mathbf{a}^{k+1}) d\theta \\ &= \int_{\Theta} \left( \frac{\int_{\Theta} p(\vartheta | \mathbf{z}^k, \mathbf{a}^k) \Psi_{a_{k+1}}^{z_{k+1}}(\vartheta) d\vartheta}{\Psi_{a_{k+1}}^{z_{k+1}}(\theta)} \right)^\delta \frac{\Psi_{a_{k+1}}^{z_{k+1}}(\theta) p(\theta | \mathbf{z}^k, \mathbf{a}^k)}{\int_{\Theta} p(\vartheta | \mathbf{z}^k, \mathbf{a}^k) \Psi_{a_{k+1}}^{z_{k+1}}(\vartheta) d\vartheta} d\theta \\ &= \left( \int_{\Theta} \Psi_{a_{k+1}}^{z_{k+1}}(\theta) p(\theta | \mathbf{z}^k, \mathbf{a}^k) d\theta \right)^{\delta-1} \int_{\Theta} \Psi_{a_{k+1}}^{z_{k+1}}(\theta)^{1-\delta} p(\theta | \mathbf{z}^k, \mathbf{a}^k) d\theta. \end{aligned} \quad (10.8)$$

The index can thus be approximated via Monte-Carlo from a sample of  $\theta$  distributed according to a preceding posterior distribution  $p(\theta|\mathbf{z}^q, \mathbf{a}^q)$ , with  $q \leq k$ ;  $q$  can equal 0 in which case the sample is distributed w.r.t. the prior. For this purpose, one can rely on the following formulas:

$$\begin{aligned}\int_{\Theta} \Psi_{a_{k+1}}^{z_{k+1}}(\theta) p(\theta|\mathbf{z}^k, \mathbf{a}^k) d\theta &= \int_{\Theta} \Psi_{a_{k+1}}^{z_{k+1}}(\theta) \prod_{j=q+1}^k \Psi_{a_j}^{z_j}(\theta) p(\theta|\mathbf{z}^q, \mathbf{a}^q) \frac{1}{L_q^k} d\theta, \\ \int_{\Theta} \Psi_{a_{k+1}}^{z_{k+1}}(\theta)^{1-\delta} p(\theta|\mathbf{z}^k, \mathbf{a}^k) d\theta &= \int_{\Theta} \Psi_{a_{k+1}}^{z_{k+1}}(\theta)^{1-\delta} \prod_{j=q+1}^k \Psi_{a_j}^{z_j}(\theta) p(\theta|\mathbf{z}^q, \mathbf{a}^q) \frac{1}{L_q^k} d\theta.\end{aligned}\quad (10.9)$$

where  $L_q^k = \int_{\Theta} \prod_{j=q+1}^k \Psi_{a_j}^{z_j}(\theta) p(\theta|\mathbf{z}^q, \mathbf{a}^q) d\theta$  does not depend on  $a_{k+1}$ . This way, if  $\theta_1, \dots, \theta_M$  is an i.i.d. sample distributed according to the posterior distribution  $p(\theta|\mathbf{z}^q, \mathbf{a}^q)$ , we can define for  $\zeta = 1$  and  $\zeta = 1 - \delta$ :

$$\begin{aligned}Q_{\zeta}^0 &= \frac{1}{M} \sum_{i=1}^M \Psi_{a_{k+1}}^0(\theta_i)^{\zeta} \prod_{j=q+1}^k \Psi_{a_j}^{z_j}(\theta_i), \\ Q_{\zeta}^1 &= \frac{1}{M} \sum_{i=1}^M \Psi_{a_{k+1}}^1(\theta_i)^{\zeta} \prod_{j=q+1}^k \Psi_{a_j}^{z_j}(\theta_i),\end{aligned}\quad (10.10)$$

to approximate easily  $\mathcal{I}_{k+1}(a_{k+1})$  up to the constant  $(L_q^k)^{\delta+1}$ :

$$\begin{aligned}\mathcal{I}_{k+1}(a_{k+1}) &\simeq (L_q^k)^{-\delta-1} (\delta(\delta-1))^{-1} \left( (Q_{\zeta}^0)^{\delta} Q_{1-\delta}^0 + (Q_{\zeta}^1)^{\delta} Q_{1-\delta}^1 \right) \\ &\propto (\delta(\delta-1))^{-1} \left( (Q_{\zeta}^0)^{\delta} Q_{1-\delta}^0 + (Q_{\zeta}^1)^{\delta} Q_{1-\delta}^1 \right).\end{aligned}\quad (10.11)$$

Eventually,  $a_{k+1}$  is chosen as the maximal argument of the right-hand side term in the above equation.

### 10.3.3 Stopping criterion

Algorithm 1 consists into a loop where  $k$  iterates from  $k_0$  to  $k_{\max} - 1$ . The upper value  $k_{\max}$  represents the total number of experiments at the end of the campaign. In practical studies, that value is limited by the cost of experiments. As a matter of fact, additional experiments are expected to provide a quantity of information that enhances the estimates. Practitioners could judge to cease the campaign if the cost of an experiment overtakes its benefits.

Such quantity of information provided by data samples is derived through our method in the index  $\mathcal{I}$ , that is why we suggest to study its variation to elucidate a stopping criterion:

$$\mathcal{VI}_k = \frac{|\mathcal{I}_{k+1}(a_{k+1}) - \mathcal{I}_k(a_k)|}{|\mathcal{I}_k(a_k)|}. \quad (10.12)$$

When the index  $\mathcal{VI}_k$  falls below a certain threshold value, the method has ceased to leverage enough from the observations. This index can be appreciated alongside the variation of the estimated fragility curve itself:

$$\mathcal{VP}_k = \frac{\|m^{|\mathbf{z}^k, \mathbf{a}^k} - m^{|\mathbf{z}^{k+1}, \mathbf{a}^{k+1}}\|_{L^2}}{\|m^{|\mathbf{z}^k, \mathbf{a}^k}\|_{L^2}}, \quad (10.13)$$

where  $m^{|\mathbf{z}^k, \mathbf{a}^k}$  is the median of the fragility curve estimate given the observations  $(\mathbf{z}^k, \mathbf{a}^k)$ ; details concerning the practical definition of the norm  $\|\cdot\|_{L^2}$  are given in section 10.4. When the index  $\mathcal{VP}_k$  falls, the estimated fragility curve given by the method has stopped to distinctly evolve.

The index  $\mathcal{VP}_k$  constitutes a criterion which is more perceptible for practitioners. In the next section where we apply our method to a practical case study, we verify that  $\mathcal{VI}_k$  and  $\mathcal{VP}_k$  give pragmatic and consistent information.

## 10.4 Estimates and benchmarking metrics

Estimates of seismic fragility curves given observation  $(\mathbf{z}^k, \mathbf{a}^k)$  are obtained by sampling i.i.d values of  $\theta$  from the posterior distribution  $p(\theta|\mathbf{z}^k, \mathbf{a}^k)$ . That sampling can be done via MCMC methods. In our work we use an adaptive Metropolis–Hastings algorithm (Haario, Saksman, and Tamminen, 2001) that necessitates iterative evaluations of the posterior up to a multiplicative constant. For a given value of  $\gamma$ , we can evaluate the posterior from the density  $\pi_\gamma^*$  suggested in eq. (10.4). We recall that in chapter 9, a method is suggested to evaluate the density  $\pi_\gamma$  expressed in eq. (10.3). We precise that we have verified that the results obtained from this method to compute  $\pi_\gamma$  or from the density  $\pi_\gamma^*$  are indiscernible. Their comparison is not the point of this chapter and is not discussed in the following.

To evaluate our estimates and the performances of our methodology, we propose to define metrics on the *a posteriori* fragility curves that is expressed conditionally to  $\theta$ . The latter can be defined as the random process  $a \mapsto P_f^{|\mathbf{z}^k, \mathbf{a}^k}(a)$  where  $P_f^{|\mathbf{z}^k, \mathbf{a}^k}(a) = \Phi(\beta^{-1} \log a / \alpha)$ . It has a distribution that naturally inherits from the posterior distribution of  $\theta$ . For each value of  $a$ , we note  $q_r^{|\mathbf{z}^k, \mathbf{a}^k}(a)$  the  $r$ -quantile of  $P_f(a)$ , and  $m^{|\mathbf{z}^k, \mathbf{a}^k}(a)$  its median. These are defined for any  $a \in \mathcal{A} = [0, a_{\max}]$ , where  $a_{\max}$  corresponds to the highest value of the IM that exists in the dataset of generated seismic signals that we use (see chapter 7), i.e.,  $a_{\max} = 12 \text{ m/s}^2$  for the PGA, and  $a_{\max} = 60 \text{ m/s}^2$  for the PSA. We suppose that a validation dataset allows to define a reference fragility curve  $P_f^{\text{ref}}$  (see chapter 7, section 7.3 for more details about the construction of  $P_f^{\text{ref}}$  given a validation dataset). We define:

- The square bias to the median:  $\mathcal{B}^{|\mathbf{z}^k, \mathbf{a}^k} = \|m^{|\mathbf{z}^k, \mathbf{a}^k} - P_f^{\text{ref}}\|_{L^2}^2$ .
- The quadratic error:  $\mathcal{E}^{|\mathbf{z}^k, \mathbf{a}^k} = \mathbb{E}_{\theta|\mathbf{z}^k, \mathbf{a}^k} \left[ \|P_f^{|\mathbf{z}^k, \mathbf{a}^k} - P_f^{\text{ref}}\|_{L^2}^2 \right]$ .
- The  $1 - r$  square credibility width:  $\mathcal{W}^{|\mathbf{z}^k, \mathbf{a}^k} = \|q_{1-r/2}^{|\mathbf{z}^k, \mathbf{a}^k} - q_{r/2}^{|\mathbf{z}^k, \mathbf{a}^k}\|_{L^2}^2$ .
- The square model bias :  $\mathcal{MB} = \|P_f^{\text{ref}} - P_f^{\text{MLE}}\|_{L^2}^2$ .  $P_f^{\text{MLE}}$  denotes a probit-lognormal curve whose parameter  $\theta$  is obtained by maximum likelihood estimation using the whole validation dataset.

The norm  $\|\cdot\|_{L^2}$  is defined by

$$\|P\|_{L^2}^2 = \frac{1}{\tilde{a}_{\max} - \tilde{a}_{\min}} \int_{\tilde{a}_{\min}}^{\tilde{a}_{\max}} P(a)^2 da. \quad (10.14)$$

For a consistent quantification of the above errors no matter the selected IM,  $\tilde{a}_{\min}$  and  $\tilde{a}_{\max}$  are selected to match a quantile of the reference fragility curve: we fix  $0 < q_1 < q_2 < 1$  and choose  $\tilde{a}_{\min}$  and  $\tilde{a}_{\max}$  such that  $q_1 = P_f^{\text{ref}}(\tilde{a}_{\min})$  and  $q_2 = P_f^{\text{ref}}(\tilde{a}_{\max})$ . In our work, the quantities defined in this section are derived considering the PGA and the PSA as IMs. In the context of the piping system that is studied in the next section, the domain of the reference fragility curve available is more limited in the case of the PGA (see fig. 10.2): in this case the reference probability of failure lives between around  $10^{-3}$  and 0.9. Thus, we fixed the values of  $q_1$  and  $q_2$  to match these in any case. The implementation of these metrics is done through Monte-Carlo derivation and numerical approximations of the integrals from Simpsons' interpolation on a sub-division of  $[\tilde{a}_{\min}, \tilde{a}_{\max}]$ .

The model bias  $\mathcal{MB}$  corresponds to the deviations between  $P_f^{\text{ref}}$  and  $P_f^{\text{MLE}}$  which can be visualized in fig. 10.2. By construction, it is influenced by the distribution of the IM since  $P_f^{\text{MLE}}$  is estimated by considering all the available samples. We cannot therefore speculate whether this is the absolute minimum bias and its value is only presented for information purposes. We will return to this point in the section devoted to the interpretation of the results.

## 10.5 Application of the method on the piping system

### 10.5.1 Short presentation of the case study

The case study that is considered in this work is the piping system that was presented in chapter 7. This case study has been studied in chapter 8 and chapter 9. As a difference with the two preceding chapters, we propose here to study the equipment's fragility curves given two different IMs: the PGA and the PSA at 5 Hz for a damping ratio of 1% (in accordance with the equipment's characteristics). We recall that we have access for this case study to a validation dataset of  $8 \cdot 10^4$  samples. In fig. 10.1, we recall the validation datasets (IM, EDP) for both of the IMs. In fig. 10.2, we remind the reference fragility curves  $P_f^{\text{ref}}$  for this case study given different thresholds, and we compare them with the probit-lognormal fragility curves  $P_f^{\text{MLE}}$  that were derived using the whole validation dataset. It should be noted that, with the PGA as IM, it is not possible to completely describe the fragility curve. For the maximum PGA values observed, the failure probabilities stagnate between 0.5 and 0.8 depending on the failure criterion considered. Therefore, the PGA is not the most suitable IM of the two. In this study, the considered failure criterion was  $C = 3.8^\circ$ .

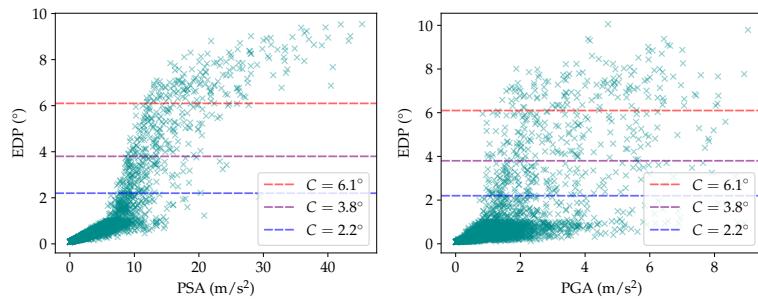


Figure 10.1: Results of the  $8 \cdot 10^4$  numerical simulations. Each cross is an element of the dataset (IM, EDP) where the IM is the PSA (left) and the PGA (right). Different critical rotation thresholds  $C$  are plotted in dashed lines. They yield different proportions of failures in the dataset: respectively 95% (red), 90% (purple) and 85% (blue).

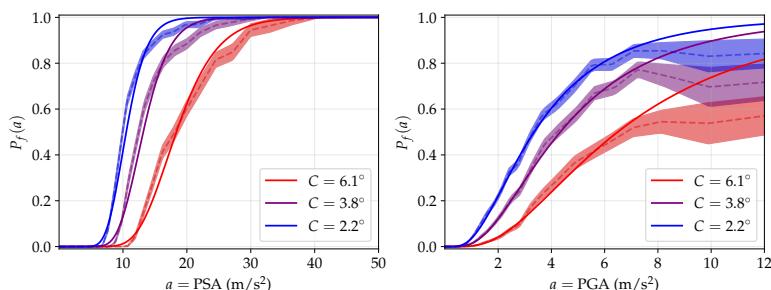


Figure 10.2: Reference non-parametric fragility curves obtained via Monte Carlo estimates (dashed lines) surrounded by their 95% confidence intervals, for different critical rotation threshold  $C$  with (left) the PSA and (right) the PGA as IM. The thresholds yield different proportions of failures in the dataset: respectively 95% (red), 90% (purple) and 85% (blue). For each value of  $C$  are plotted (same color, solid line) the corresponding probit-lognormal MLE.

### 10.5.2 Qualitative study

In fig. 10.3 we present examples of *a posteriori* fragility curve estimations. They take the form of 95%-credibility intervals. These qualitative graphs allow us to appreciate the performance of the estimation when the IMs are selected using our DoE method compared to the standard approach. In this example, the DoE was

implemented by setting  $\gamma = 0.5$ . The same comments stay valid for any of the two IMs we have taken into account in this study —the PSA and the PGA—, namely:

- (i) When the number of observed data is really small (20 samples), the standard method tends to provide “vertical” fragility curves estimates, which result in “vertical” credibility intervals. By verticality, understand here that the tangent of the estimated curve at the median has an infinite slope (i.e.,  $\beta = 0$ ). This phenomenon is a consequence of the fact that the likelihood is degenerate in this case. The credibility intervals are thus tighter but strongly incorrect. When the DoE is implemented with the same number of samples, the phenomenon fades and the credibility intervals follow accurately the reference curve.
- (ii) When a more decent number of data are observed (80 samples), the credibility intervals given from the DoE follow quite closely the reference curve. Within the observed seismic signals, we then notice a better balance between the ones with large IMs and the ones with small IMs compared to the standard method (see the “red crosses” in the figures). The consequence is a thinner credibility interval given by the DoE method.

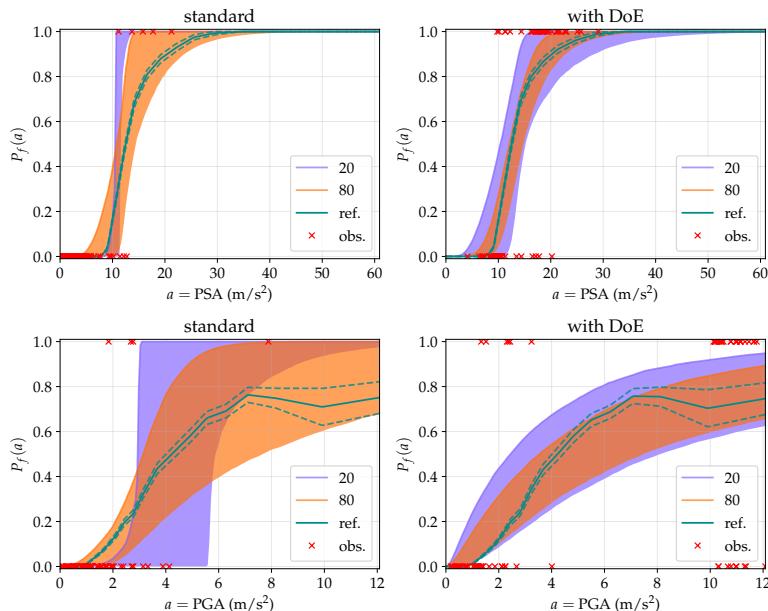


Figure 10.3: Examples of fragility curves estimation. Top figures: the IM is the PSA; bottom figures: the IM is the PGA. Left figures: the seismic signals are chosen w.r.t. their standard distribution. Right figures: they are chosen w.r.t. our DoE strategy. On each figure: credibility intervals from 20 (purple) or 80 (orange) observations, reference fragility curves (green solid line) from Monte Carlo simulations and their 95%-confidence intervals (green dashed lines). The red crosses represent the 80 observations.

### 10.5.3 Quantitative study

Figures 10.4 to 10.8 present quantitative results that go beyond a single example. Within those, 12 methods are compared: the standard method and the DoE methods for any  $\gamma \in \{0, 0.1, 0.3, \dots, 1.9\}$ . The case  $\gamma = 0$  is particular: in this case the prior  $\pi_\gamma^*$  does not always satisfy the criteria for a valid *a posteriori* estimation (definition 10.1) even though that is essential for the DoE to be implemented. Thus,  $\gamma = 0$  corresponds to an alteration of the DoE method: first, data are collected given the DoE method carried out with  $\gamma = 0.1$ ; second, the prior  $\pi_{0,1}^*$  is replaced with  $\pi_0^*$  to compute the estimates. This particular treatment is carried out in order to quantify how that parametrized constraint affects the estimates.

For each of these methods, we have carried out 100 replications of the same numerical experiment, namely:  
 (i) generate an observed sample of  $k_{\max} = 250$  data items and (ii) derive for any  $k = 10, 20, \dots, k_{\max}$  the

posterior distribution  $p(\theta|z^k, \mathbf{a}^k)$ . These replications provide evaluations of the mean values of the metrics defined in section 10.4. The results are shown in fig. 10.5 for the PSA and in fig. 10.6 for the PGA.

**Overall comments on the influence of  $\gamma$**  For both IMs, figs. 10.5 and 10.6 show that the discrepancies between all the DoE-based results are negligible. We do not clearly distinguish a hierarchy of these with respect to the value of  $\gamma$ . We have verified that the selected IMs by the DoE methods are actually poorly influenced by the value of  $\gamma$  and that any change of performance impacted by the value of  $\gamma$  is hard to distinguish. We conclude that its influence on the posterior estimation remains small. Thus, our constrained reference prior is close enough to the unconstrained ( $\gamma = 0$ ) one and conserves its “objectivity” qualification.

**Degeneracy disappearance** Definition 8.1 lists the different types of situations for which likelihood degeneracy can occur. For our case study, fig. 10.4 presents the average number of occurrences of the three types of situations that lead to a degenerate likelihood, as a function of the number of observed samples. We observe that the DoE method clearly outperforms in reducing the degeneracy for small numbers of observations compared with the standard method (from  $k \simeq 20$  in the case of the PGA, and at most  $k \simeq 50$  in the case of the PSA). Echoing this observation, we notice that this probability is, in general, higher with the PSA as IM, especially when the seismic signals are selected according to the standard method.

These results must however be qualified because they arise directly from the way in which we implemented the two estimation methods —DoE and standard—, that is to say from a large database of artificial signals (i.e., of size  $8 \cdot 10^4$ ), generated beforehand. In section 10.6, it is indeed mathematically demonstrated that in a well-specified case, we can define a change of variable between two fragility curves which are defined by two different values of  $\theta$ , all other things being equal. In other words, from the point of view of the estimation of the parameter  $\theta$ , all cases are equivalent if there is sufficient data in the domain of interest, that is, the domain in which the fragility curve evolves “significantly” from 0 to 1. As this requirement is not met in practice, especially in our case study, the estimation performances are not identical when considering two distinct IMs such as the PGA or the PSA. Thus, in the standard case, whether for the PSA or PGA, as we have considered a high failure threshold, the main cause of degeneracy comes from the fact that the observed data are of type 1, that is to say that they do not reveal any failure. The distributions are concentrated at low IM values. Then, with the PSA, as the value of  $\beta$  of the fragility curve is smaller than with the PGA, while its distribution has a larger standard deviation, this favors degeneracy of type 3. To a lesser extent, this also affects the DoE method.

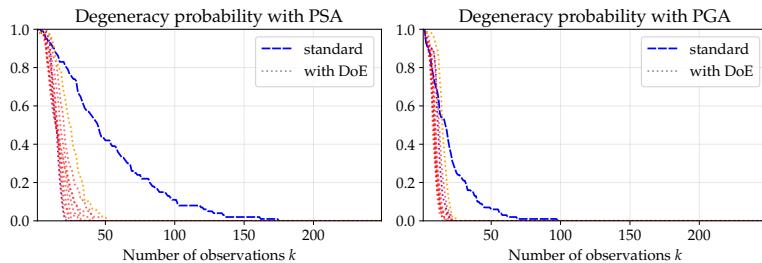


Figure 10.4: Probability that a sample of size  $k$  yields a degenerate likelihood, as a function of  $k$ . The considered IM is (left) the PSA and (right) the PGA. The degeneracy probability without DoE (blue dashed line) is compared with the ones with DoE (dotted lines), for different values of  $\gamma$ . Two extreme values of  $\gamma$  are highlighted:  $\gamma = 0.1$  (purple dotted line) and  $\gamma = 1.9$  (orange dotted line).

**Performances with few observations** Focusing first on the quadratic errors presented in figs. 10.5 and 10.6, we show that when the number of observations  $k$  is small, the outperformance of the DoE method over the standard one is clearly visible for any IM considered. Indeed, on the domain  $k < 50$ , the decrease of any of the errors is faster, reaching satisfying values at  $k = 50$ , which is consistent with the qualitative study of the previous section. Regarding the cases where  $k > 50$ , the behavior of the metrics differ between the case where the PSA is used and the one where it is the PGA. Globally, it is obvious that the PSA provides the

best results. The consideration of this IM rather than the PGA is more relevant for this case study. In the following paragraphs, we discuss more specifically the estimation biases for both IMs, since the quadratic error is a combination of the bias and the credibility width.

**Study of the bias when the IM is the PSA** As mentioned in section 10.5.1, the PSA is the best of the two IMs in our case study. As a result, the performance of the DoE method is beyond doubt (fig. 10.5). Up to values of  $k$  close to 50, the decrease in the estimation bias is rapid towards the so-called model bias value which is materialized in the figure (see section 10.4). This bias reflects the fact that the reference fragility curve,  $P_f^{\text{ref}}$ , does not correspond to an exact probit-lognormal curve as illustrated in fig. 10.2. Beyond a value of  $k$  close to 100, the bias stops decreasing significantly because it tends towards the model bias. From this threshold, additional observed data only make the credibility intervals thinner around the median. It is noted in this regard that the DoE method is able to provide estimates with a bias that is smaller than the so-called model bias. On average, the DoE-based estimations do not exactly match the model bias, although it is very close. This is not the case for the standard method, for which there is almost an order of magnitude difference, even with a sample size of 250. The difference between the bias obtained on average by the DoE method and the model bias simply reflects the fact that the distributions of the data used for the estimates are not the same in the two cases.  $P_f^{\text{MLE}}$  is estimated on the entire available database while the DoE methodology aims to select some of them by maximizing the criterion defined in eq. (10.5). Section 10.6 provides more insight into the DoE approach.

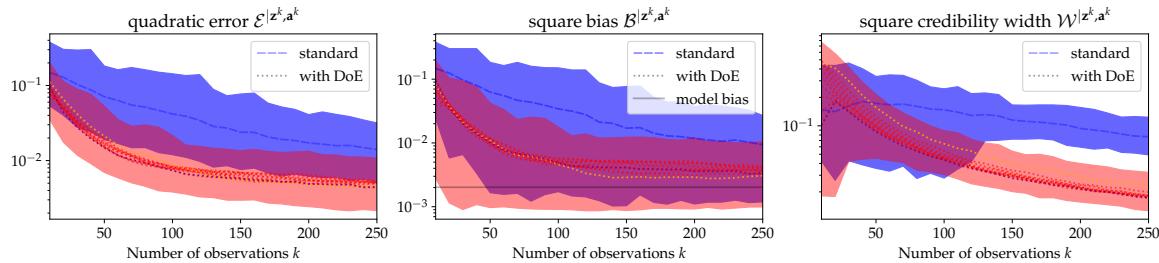


Figure 10.5: Average error metrics derived from numerous replications of the method, as a function of the size of the observed sample. In dashed blue lines: the average errors using the standard distribution of the IM surrounded by their 95%-confidence interval. In dotted lines: the average errors using our DoE strategy with different values of  $\gamma$ . They are surrounded (in red) by the maximal 95%-confidence intervals. Two extreme values are emphasized:  $\gamma = 0$  (orange) and  $\gamma = 1.9$  (purple). On the middle figure, the model square bias is plotted as a gray line. The IM is the PSA here.

**Study of the bias when the IM is the PGA** As mentioned in section 10.5.1, with the PGA as IM, it is not possible to fully describe the fragility curve. We therefore have no information on its evolution beyond the maximum value of the PGA observed, at about  $12 \text{ m/s}^2$ . Moreover, the discrepancies between  $P_f^{\text{ref}}$  and  $P_f^{\text{MLE}}$  are maximal for a PGA value of the order of  $8 \text{ m/s}^2$ , whereas before this value the fit is very good. For the strongest seismic signals at disposal, equipment failure is only observed in around 70% of cases.

Up to values of  $k$  close to 100, the DoE method outperforms the standard method. Beyond that, the behavior is different from that observed when the IM is the PSA and an evaluation of the performances of the DoE method is less obvious. On average, however, with the first 40 data points, the bias value obtained with the PGA is smaller than with the PSA.

In section 10.6 it is shown that with the DoE method, the distribution of the selected data tends to follow a bimodal distribution whose modes are equally distant from the logarithm of the median of the fragility curve (see also fig. 10.3 and the repartition of the “red crosses”). Since knowledge of the fragility curve over a restricted domain is likely to reduce the performance of the DoE method, a thorough study of the consequences of such a restriction is proposed in section 10.6, in a case where the model bias does not exist. This study shows that the performance of the DoE method is weakly affected by this limitation. We can therefore postulate that it is the presence of a significant bias in one of the two main domains in which the

DoE method selects data that is the cause of the deterioration in the method's performance for values of  $k > 100$ , compared to the standard approach. If we believe the results in fig. 10.3, with 80 data, we obtain however a completely reasonable estimate, taking into account the observed bias.

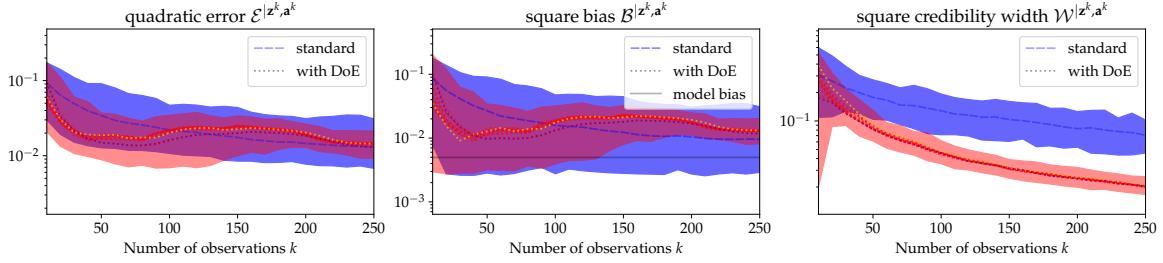


Figure 10.6: As in fig. 10.5 but here the IM is the PGA.

**Stopping criterion** To the extent that an irreducible estimation bias can be expected in practice, an indication that such a bias has been reached would constitute a suitable stopping criterion for the DoE method. Since such a criterion would imply knowing, in some way, the bias itself and this is not possible with small data sizes, we suggest to study (i) the index  $\mathcal{VI}_k$  that measures the variation of the quantity defined by eq. (10.5) that is used to select the seismic signals in the DoE method and (ii) the index  $\mathcal{VP}_k$  that measures the average evolution of the median of the fragility curve. These two quantities are defined in section 10.3.3. They reflect the information provided by adding a  $k$ -th data point.

Figure 10.7 shows the evolutions of the average values of  $\mathcal{VI}_k$  for respectively the PSA and the PGA as IM, whereas fig. 10.8 is devoted to the evolutions of  $\mathcal{VP}_k$ . These two figures, once again, show that learning with the DoE method is more effective than with the standard method.  $\mathcal{VI}_k$  and  $\mathcal{VP}_k$  are slightly influenced by the change in IM. Up to values of  $k$  close to  $k = 50$ , they even indicate that with PGA, the DoE method reaches more quickly a state in which the information brought by a new data point has less impact. This result is consistent with the one mentioned in the previous paragraph, namely that with the first 40 data points, the bias value obtained with the PGA is smaller than with the PSA, with the DoE method.

The practitioner can therefore use these indicators to define a stopping criterion depending on a threshold value. For instance, in our case, a threshold value set at  $\mathcal{VI}_k = 10^{-3}$  gives approximately  $k_{\max} = 60$  for the PSA and  $k_{\max} = 40$  for the PGA. Concerning  $\mathcal{VP}_k$ , the threshold value can be set at 5%. For both the PSA and the PGA, this ensures furthermore that the likelihood degeneracy probability is zero (see fig. 10.4). Remark that, even though the number of data required to reach a given stopping criterion is smaller with the PGA than with the PSA, this does not imply that the estimate is less accurate. Indeed, for small data sizes, the degeneracy phenomenon is more likely with the PSA than with the PGA (see fig. 10.4). Consequently, the widths of the credibility intervals remain large for large  $k$  with the PSA (see figs. 10.5 and 10.6). Our approach addresses this phenomenon by reducing the degeneracy even for a few observations. In practice, however, this number remains dependent on the IM of interest (see the paragraph "Degeneracy disappearance").

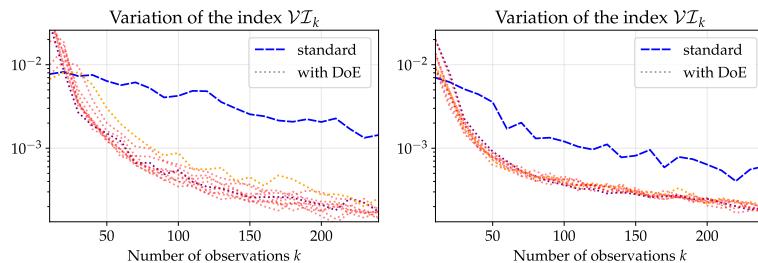


Figure 10.7: Average variations of the DoE index  $\mathcal{VI}_k$  as a function of the observed sample size  $k$ , using the PSA (left) and the PGA (right).

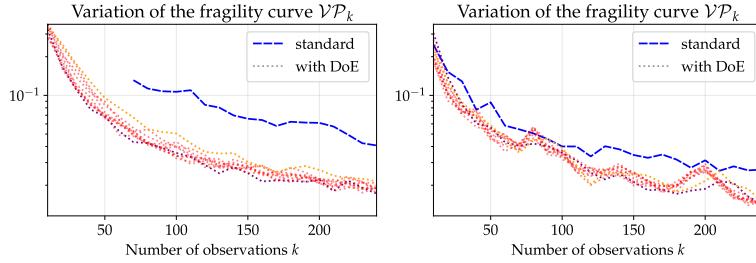


Figure 10.8: Average variations of the median *a posteriori* fragility curve  $\mathcal{VP}_k$  as a function of the observed sample size  $k$ , using the PSA (left) and the PGA (right).

### 10.5.4 Synthesis

The objective of the methodology proposed in this work is to make the most of the probit-lognormal model. Indeed, it appears to the practitioner as a model that is both pragmatic and relevant for estimating fragility curves with a reduced number of data, judging by the abundant use made of it in the literature. The limitation on the number of data is essential. It comes from the fact that (i) these are expensive in absolute terms, especially with experimental tests, and (ii) the probit-lognormal model being likely to be biased, there is no point in feeding it with a lot of data. To achieve this goal, we proposed a methodology for planning experiments in a Bayesian framework based on the reference prior theory.

Two current IMs —the PSA and the PGA— were considered to study the performance of the method which was applied to an equipment from the nuclear industry. In both cases, with a few data points —of the order of 80—, the performance of the proposed method was much better than that of the standard one. Performance is best if the method is used with an IM that has a good level of correlation with the structural response of interest.

In order to guide the user, we also proposed two indicators that allow learning to be stopped based on quantitative information. Both criteria appeared sensitive to the change of IM and therefore to the associated potential biases. Let us add, however, that since the degeneracy phenomenon does not make the estimation optimal from the point of view of the size of the credibility interval, it is appropriate, in practice, to continue learning until a non-degenerate sample is obtained (see definition 8.1).

## 10.6 Behavior of the method on toy case studies

### 10.6.1 Description of the case studies

In order to gain more insight on the method, we suggest in this section a brief analysis of its behavior on case studies that are free of any bias. Such case studies are purely theoretical, so that we qualify them as “toy case studies”. Here, the observed data  $(\mathbf{z}^k, \mathbf{a}^k)$  are perfectly generated according to the probit-lognormal statistical model presented in section 10.2: a reference parameter  $\theta^* = (\alpha^*, \beta^*)$  is chosen; then, from any input  $a$  (that simulates a theoretical IM), an output  $z \in \{0, 1\}$  is generated that is equal to 1 with probability  $\Phi(\beta^{*-1} \log \frac{a}{\alpha^*})$  and 0 otherwise.

The selected inputs belong to an interval  $\mathcal{A} = (0, a_{\max}]$ . They are picked by the design of experiments methodology that we present in this work with  $\gamma = 0.5$ . Note that when implementing algorithm 1 in this section, there is no need to generate seismic signals from a particular value of  $a$  here. Indeed, the “experiment” (understand here: the generation of  $z$ ) directly results from  $a$ . Regarding the  $k_0 = 2$  required initial values of  $a$ , we draw them uniformly in  $\mathcal{A}$ .

In those toy case studies, the value of  $\theta^*$  does not matter, in the sense that two such case studies with two different value of  $\theta^*$  would be equivalent up to a translation and a dilatation w.r.t.  $\log a$  (see section 10.6.5 for a proof of this statement).

In this section, we present three different toy case studies. They are all implemented using the same value of  $\theta^*$ . They differ from the different domains  $\mathcal{A}$  that are selected for them. The limits  $a_{\max}$  of these domains

are chosen to simulate an upper bound on the possible IM that can result from a generated seismic signal. They are selected to match a quantile of the reference fragility curve: if  $q \in [0, 1]$  one can derive  $a_q$  such that  $P_f^{\text{ref}}(a_q) := \Phi\left(\beta^{*-1} \log \frac{a_q}{\alpha^*}\right) = q$ , i.e.

$$a_q = \exp\left(\beta^* t_q + \log \alpha^*\right), \quad (10.15)$$

where  $t_q$  represents the  $q$ -quantile of a standard Gaussian distribution.

The three toy case studies implemented in this work result from defining  $a_{\max} = a_q$  with:

- $q \simeq 1$  (actually  $q = 1 - 10^{-3}$ ), to represent a case where no upper bound (or a very high one regarding the fragility curve) limits the generation of IMs.
- $q = 0.9$ .
- $q = 0.8$ .

The reference fragility curve of each of these toy case studies are plotted in fig. 10.9 (green solid lines) for each of the values of  $q$  aforementioned, and for  $\theta^* = (3, 0.3)$ .

### 10.6.2 Results

The design of experiments method has been replicated a hundred times for each case study. Each time, we have stopped the algorithm after it reached a number of 250 observed samples. The last 100 of these generated IMs have been kept for each replication. As we have carried out 100 replications, they constitute a sample of  $10^4$  values of IMs selected by the experimental design. The three empirical distributions of these samples resulting from the three values of  $q$  are compared in fig. 10.9.

When the method is not limited by any upper bound for the choice of IMs (i.e. case  $q = 1$ ), the logarithms of the selected ones are symmetrically distributed around  $\log \alpha^*$ . This comment was expected as in the statistical modeling,  $\log a$  is mathematically symmetric around  $\log \alpha^*$  as well.

The bimodal appearance of the distribution supports the “well-posedness” of the method. Indeed, in a case where the result is deterministic, the knowledge of the evaluation of the curve in two points is enough to recover its parameter  $\theta^*$ . Thus, it makes sense that the optimal distribution of the IM for estimating the whole fragility curve is to retrieve a proper estimation of its evaluation in two zones of the domain. A simple calculation is done in section 10.6.3 to suggest possible appropriate centers of the zones. These are called  $a_1^u$  and  $a_2^u$  and are appropriate in the sense that they are the points that most help distinguish a stochastic fragility curve from a deterministic one. In fig. 10.9, their values are emphasized (dashed pink line). They are clearly close from the actual modes of the empirical distribution of the IM in the case  $q = 1$ .

In the cases where the IM selection is limited ( $q < 1$ ), the method has been designed to select  $a_{\max}$  when the index tended to be maximized by a value which is higher. We notice that the method seems to target the same ideal domain as when  $q = 1$ : converging to a bimodal strategy with modes around optimal domains. The method is robust: for two given  $q_1, q_2$  and their  $a_{\max}^1, a_{\max}^2$  associated to two toy case studies, around the same proportion of IMs is selected below  $a$  for any  $a \leq \min(a_{\max}^1, a_{\max}^2)$  in the two cases. This results in numerously selected IMs equaling  $a_{\max}$  when it falls close to the second expected mode of the optimal distribution.

In fig. 10.10-left, we compare the average square bias resulting from the 100 replications of the three case studies. Their difference is difficult to distinguish, so the robustness of the method is confirmed: even when the limit among the available IMs sharply intersects the reference fragility curve, the distribution of the selected ones still provides an accurate estimation. However, a close look allows to notice the expected loss of performance induced by such limit: the smaller is  $q$ , the larger is the bias.

Regarding the other metrics exposed in fig. 10.10, namely the indices  $\mathcal{VI}_k$  and  $\mathcal{VP}_k$ , it is interesting to notice that their behaviors are comparable to that of the real case study. In particular,  $\mathcal{VI}_k$  reaches the threshold  $10^{-3}$  for a similar number of observations (around 50) in any of the case studies implemented in this work (real ones as well as toy ones). As a matter of fact, the index sequentially measures the quantity of information brought by the next seismic signal selected by the DoE. Thus, the stability of that quantity

between case studies supports the robustness of our approach in selecting IMs to inform the estimate and enhance the learning. However, the indices  $\mathcal{VP}_k$  are smaller for toy case studies. Without any model bias, the estimates reach more quickly the reference (see the average bias in fig. 10.10-left), so that they evolve less.

Finally, this section helps to clarify the results obtained with the two IMs that are the PSA and the PGA in section 10.5.3. The main difference between them is that with the PGA it is not possible to completely describe the fragility curve. With the PGA, the maximum value of its probit-lognormal estimate is approximately 0.8 but does not reach 1 as with the PSA. If we ignore the bias, this means that the case  $q = 1$  is similar to the situation where the PSA is used as IM, while the cases  $q = 0.8$  and  $q = 0.9$  are similar to the situation where the PGA is used. All things being equal, this study shows that the performance of the DoE method is weakly affected by a restricted learning domain. It is then likely that the presence of a significant bias in one of the two main domains in which the DoE method selects data is the cause of its performance degradation.

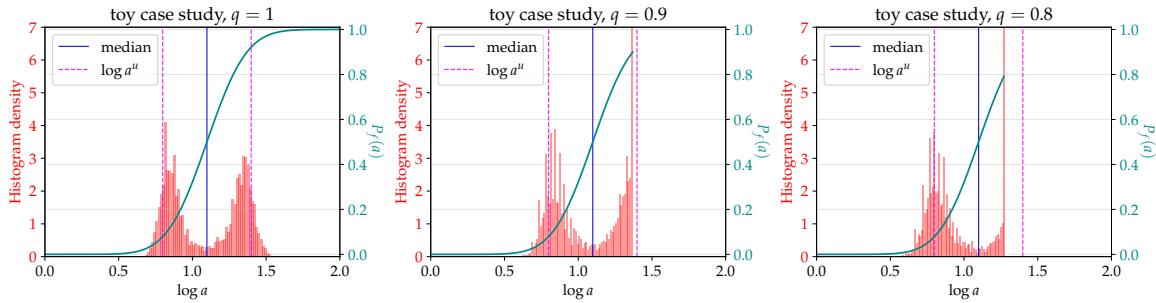


Figure 10.9: For each of the three toy case studies: the distribution of the 100 last selected IM by the method (red histograms) along with the associated reference fragility curve (green) of the toy case study, which is cut at a certain quantile  $q$  in each figure. The median of the reference (solid blue) and the values  $a_1^u, a_2^u$  (dashed pink) are emphasized.

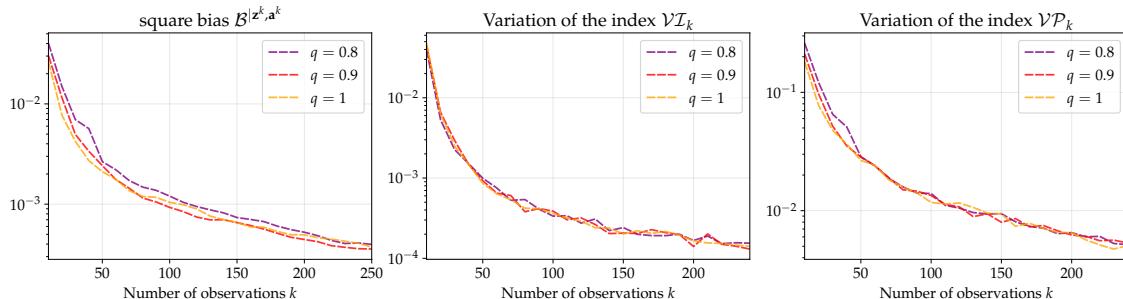


Figure 10.10: Left figure: average square bias of the three toy case studies as a function of the number of observations. Middle figure: index  $\mathcal{VI}_k$  for the three toy case studies as a function of the number of observations. Right figure: index  $\mathcal{VP}_k$  for the three toy case studies as a function of the number of observations.

### 10.6.3 Simple suggestion of a domain where IMs should be theoretically selected

In this section we suggest a criterion that gives insight on a domain where the fragility curve should be evaluated in order to optimize the estimation. For this purpose, we study the quantity  $\mathbb{E}[|P_f^{\text{ref}}(a) - P_f(a)|]$  as a function of  $a$ , where  $P_f^{\text{ref}}(a)$  represents a deterministic fragility curve and  $P_f(a)$  denotes a stochastic estimate. If  $a^*$  is a maximal argument of this quantity, it would be the one that helps the most distinguishing the two curves.

For this theoretical study, we assume to be in the settings of the toy case studies defined in the previous

subsection, i.e.  $P_f^{\text{ref}}(a) = \Phi(\beta^{*-1} \log \frac{a}{\alpha^*})$  for a couple  $(\alpha^*, \beta^*)$ . We derive

$$\begin{aligned} & \frac{d}{da} \mathbb{E}|P_f^{\text{ref}}(a) - P_f(a)| \\ &= \frac{1}{a\sqrt{2\pi}} \int_{\Theta} \left[ \frac{1}{\beta^*} \exp\left(-\frac{(\log a - \log \alpha^*)^2}{2\beta^{*2}}\right) - \frac{1}{\beta} \exp\left(-\frac{(\log a - \log \alpha)^2}{2\beta^2}\right) \right] p_{\alpha,\beta}(\alpha, \beta) d\alpha d\beta. \end{aligned} \quad (10.16)$$

To pursue, we can assume that  $(\alpha, \beta)$  follows an inverse-gamma-normal distribution. This class covers a wide range of different distributions and is conjugate with a normal distribution, so that the following computation is tractable. Additionally, the probit-lognormal modeling of the fragility curve is equivalent to assume that there exists a latent variable  $Y$  that is linearly correlated with the input:  $Y = \log A + \mathcal{N}(-\log \alpha^*, \beta^*)$ . Considering this latent modeling, we can prove that both the reference prior and the posterior distribution of  $\alpha, \beta$  belong to the class of inverse-gamma-normal ones (see [appendix B](#)). The inverse-gamma-lognormal distribution is defined by its density:

$$\begin{aligned} p_{\alpha,\beta}(\alpha, \beta) &= K(c, d, \tau, \zeta) \left(\frac{1}{\beta^2}\right)^{c+1/2} \exp\left(-\frac{d}{\beta^2}\right) \exp\left(-\tau \frac{(\log \alpha - \zeta)^2}{2\beta^2}\right), \\ \text{with } K(c, d, \tau, \zeta) &= \frac{d^c \sqrt{\tau}}{\Gamma(c) \sqrt{2\pi}}, \end{aligned} \quad (10.17)$$

where  $c > 0, d > 0, \tau > 0$  and  $\zeta \in \mathbb{R}$  are the parameters of the distribution.

Thus, the right-hand term in eq. (10.16) becomes:

$$\frac{1}{\sqrt{2\pi}} \frac{1}{a\beta^*} \exp\left(-\frac{(\log a - \log \alpha^*)^2}{2\beta^{*2}}\right) - \frac{d^c}{(d + \frac{\tau}{2\tau+2}(\zeta - \log a)^2)^{c+1/2}} \frac{\sqrt{\tau}}{\sqrt{\tau+1}} \frac{\Gamma(c + \frac{1}{2})}{\Gamma(c)} \frac{1}{a\sqrt{2\pi}}, \quad (10.18)$$

so that it equals 0 if and only if  $a$  verifies:

$$|\zeta - \log a| \frac{e^{1/2}}{\beta^*} \exp\left(-\frac{(\log a - \log \alpha^*)^2}{2\beta^{*2}}\right) = \frac{d^c |\zeta - \log a| e^{1/2}}{(d + \frac{\tau}{2\tau+2}(\zeta - \log a)^2)^{c+1/2}} \frac{\sqrt{\tau}}{\sqrt{\tau+1}} \frac{\Gamma(c + \frac{1}{2})}{\Gamma(c)}. \quad (10.19)$$

Given the derivations conducted in section [10.6.4](#), the right-hand term of the above equation tends to 1 when information is provided by observed data to the posterior. In this case and if  $\zeta = \log \alpha^*$  the equation leads to  $a = a_1^u$  or  $a = a_2^u$  where  $a_1^u, a_2^u$  are defined by

$$a_{1,2}^u = \alpha^* \exp(\pm \beta^*). \quad (10.20)$$

These values are built on a criterion that has a concrete and practical sense. However, they remain purely theoretical as they depend on the exact value of  $\theta^*$ .

#### 10.6.4 Additional mathematical derivations

The goal of this section is to study the domain of the right-hand term of eq. (10.19), especially in the “worst” cases (the distribution is very-informative or non-informative). We rely on the following lemmas:

**Lemma 10.1.** For any  $c > 0$ ,  $\frac{\Gamma(c + \frac{1}{2})}{\Gamma(c)} < \sqrt{c}$ . Also,  $\lim_{c \rightarrow \infty} \frac{\Gamma(c + \frac{1}{2})}{\Gamma(c)\sqrt{c}} = 1$ .

*Proof.* The first statement is a direct consequence of Gautschi’s inequality ([Gautschi, 1959](#)): for any  $c > 0$ ,  $s \in (0, 1)$ ,

$$c^{1-s} < \frac{\Gamma(c+1)}{\Gamma(c+s)} < (c+1)^{1-s}. \quad (10.21)$$

Fixing  $s = 1/2$  and using the identity  $\Gamma(c+1) = c\Gamma(c)$  leads to the result.

For the second statement, we can rely on Stirling's formula (Davis, 1959) that yields:

$$\Gamma(c+t) \underset{c \rightarrow \infty}{\sim} \Gamma(c)c^t \quad (10.22)$$

for any  $t \in \mathbb{C}$ . □

**Lemma 10.2.** Let  $f > 0$ , for any  $d, c > 0$  we define:

$$g(c, d) = \frac{\sqrt{c}}{(d+f)^{1/2}} \left( \frac{d}{d+f} \right)^c, \quad h(d) = \frac{1}{2} \log \left( 1 + \frac{f}{d} \right)^{-1}. \quad (10.23)$$

Thus, for any  $c, d > 0$ ,  $g(c, d) < (2ef)^{-1/2}$ , and  $\lim_{d \rightarrow \infty} g(h(d), d) = (2ef)^{-1/2}$ .

*Proof.* Let us differentiate  $g$  w.r.t.  $c$ :

$$\frac{\partial}{\partial c} g(c, d) = (d+f)^{-1/2} \left( \frac{d}{d+f} \right)^c \left[ \sqrt{c} \log \left( \frac{d}{d+f} \right) + \frac{1}{2\sqrt{c}} \right]. \quad (10.24)$$

The above quantity is decreasing w.r.t.  $c$  and equals 0 when  $c = h(d)$ . We deduce that for any  $d > 0$ ,

$$g(c, d) < g(h(d), d) = \frac{(2e)^{-1/2}}{(d+f)^{1/2}} \log \left( 1 + \frac{f}{d} \right)^{-1/2}. \quad (10.25)$$

Now, let us briefly study the function  $v(t) = (1+t^{-1}) \log(1+t)$  for  $t > 0$ . Firstly,  $\log(1+t) \underset{t \rightarrow 0}{\sim} t$  so that  $v(t) \underset{t \rightarrow 0}{\longrightarrow} 1$ . Secondly,  $v'(t) = \frac{1}{t} - \frac{1}{t^2} \log(1+t)$ , which is positive for any  $t > 0$ , and tends to  $\frac{1}{2}$  when  $t = 0$ . Therefore,  $v(t) > v(0) = 1$  for any  $t > 0$ .

Going back to eq. (10.25), we can write  $g(h(d), d) = (2ef)^{-1/2} v(f/d)^{-1/2}$  to obtain that

$$g(c, d) < (2ef)^{-1/2} \quad \text{and} \quad \lim_{d \rightarrow \infty} g(h(d), d) = (2ef)^{-1/2}. \quad (10.26)$$

□

The result of lemma 10.2 with  $f = \frac{1}{2} \frac{\tau}{\tau+1} (\log a - \zeta)^2$  lets us write the following inequality:

$$\frac{d^c \sqrt{c}}{(d + \frac{\tau}{2\tau+2} (\zeta - \log a)^2)^{c+1/2}} \frac{\sqrt{\tau}}{\sqrt{\tau+1}} < \frac{e^{-1/2}}{|\zeta - \log a|}, \quad (10.27)$$

this upper-bound being reached at the boundary of the domain. Combining this statement with lemma 10.1, we obtain:

$$\frac{d^c |\zeta - \log a| e^{1/2}}{(d + \frac{\tau}{2\tau+2} (\zeta - \log a)^2)^{c+1/2}} \frac{\sqrt{\tau}}{\sqrt{\tau+1}} \frac{\Gamma(c + \frac{1}{2})}{\Gamma(c)} < 1, \quad (10.28)$$

this upper-bound being reached at the boundary of the domain.

Additionally, the left-hand term of the above equation is non-negative and tends to 0 for some extreme values of the parameters. Therefore, there exist two extreme cases for the resolution of eq. (10.19): when its right-hand term equals 0 and when it equals 1. The former corresponds to the limit case where the solution is  $a = \pm\infty$  or  $a = \zeta$ , with  $a = \pm\infty$  minimizing the quantity of interest. The latter corresponds to the limit case where the solution verifies

$$\frac{|\zeta - \log a|^2}{\beta^{*2}} = \exp \left( \frac{|\log \alpha^* - \log a|^2}{\beta^{*2}} - 1 \right). \quad (10.29)$$

The solutions of the above equation when  $\zeta = \log \alpha^*$  are  $a = \alpha^* \exp(\pm \beta^*)$ . They correspond to the limit case when the distribution  $p_{\alpha, \beta}$  becomes informed and unbiased. If the model is appropriately specified, that

should be the fate of the posterior distribution: the posterior tends to match a normal distribution with mean  $\theta^*$  and variance  $\frac{1}{k}\mathcal{F}^{-1}(\theta^*)$  where  $\mathcal{F}$  refers to the Fisher information matrix (Vaart, 1992).

### 10.6.5 Equivalence between toy case studies

Let us consider two different toy case studies. They can differ by their reference parameters  $\theta_1^*$ ,  $\theta_2^*$  and their IMs  $a_1$ ,  $a_2$ , which are defined in the domains  $\mathcal{A}_1$ ,  $\mathcal{A}_2$ . We suppose that the bounds of  $\mathcal{A}_1$  and  $\mathcal{A}_2$  are defined to match the same quantiles of the reference fragility curves:  $\mathcal{A}_1 = (c_1^1, c_1^2)$ ,  $\mathcal{A}_2 = (c_2^1, c_2^2)$  with  $P_1^{\text{ref}}(c_1^1) = P_2^{\text{ref}}(c_1^1)$  and  $P_1^{\text{ref}}(c_1^2) = P_2^{\text{ref}}(c_2^2)$ , where  $P_i^{\text{ref}}$  (reps.  $P_2^{\text{ref}}$ ) denotes the reference fragility curves given by  $\theta_1^*$  (resp.  $\theta_2^*$ ).

Therefore, if we introduce  $\tilde{a}$  that defines the value of a new IM on the second toy case study, which verifies

$$\log \tilde{a} = \frac{\log c_1^2 - \log c_1^1}{\log c_2^2 - \log c_2^1} \log \frac{a_2}{c_2^1} + \log c_1^1, \quad (10.30)$$

we obtain that  $\tilde{a}$  lives in the domain  $\tilde{\mathcal{A}} = \mathcal{A}_1$ . Given this new IM, the reference fragility curve of the second case study can be re-defined as a function of  $\tilde{a}$ :

$$\tilde{P}_2^{\text{ref}}(\tilde{a}) = \Phi \left( \beta_2^{*-1} \left[ \frac{\log c_2^2 - \log c_2^1}{\log c_1^2 - \log c_1^1} (\log \tilde{a} - \log c_1^1) + \log c_1^1 - \log \alpha_2^* \right] \right) = \Phi \left( \tilde{\beta}_2^{*-1} \log \frac{\tilde{a}}{\tilde{\alpha}_2^*} \right), \quad (10.31)$$

for a certain  $\tilde{\theta}_2^* = (\tilde{\alpha}_2^*, \tilde{\beta}_2^*)$ . We have  $\tilde{P}_2^{\text{ref}}(c_1^1) = P_2^{\text{ref}}(c_1^1) = P_1^{\text{ref}}(c_1^1)$  and  $\tilde{P}_2^{\text{ref}}(c_2^2) = P_2^{\text{ref}}(c_2^2) = P_1^{\text{ref}}(c_1^2)$ , thus, as the parameter  $\theta$  defines uniquely a probit-lognormal fragility curve,  $\theta_1 = \tilde{\theta}_2$ .

As a conclusion, given a rescaling of the IM, the two case studies are equivalent.

## 10.7 Conclusion

When we seek to estimate seismic fragility curves with limited data, the information conveyed by any sources —the prior, the observations, the model— have a considerable influence on the quality of estimates.

In this work, we considered as in the previous chapters the probit-lognormal model, and had thus to acknowledge its irreducible bias compared with the actual reference fragility curves of the studied component.

Regarding *a priori* information, we relied on the reference prior theory in order to define the prior. This framework makes it possible to define a so-called objective prior in the sense that it favors the information brought by the data to the estimates. Because the most objective prior according to the theory does not tackle degenerate likelihoods and issues improper posteriors, we derived a slightly modified reference prior, more efficient and more robust. However, one has to note that this prior is slightly more informative than the original reference prior.

To optimize data allocation, we have proposed a design of experiments strategy that also inherits from the reference prior theory, i.e., that is based on information theory. Given a large database of synthetic seismic signals, the proposed strategy intends to sequentially select the synthetic signals with which to perform the calculations or the tests, in order to optimally estimate —by minimizing their number— the probit-lognormal estimations of fragility curves.

We proved that our method is robust and that *a priori* information incorporated through the modified prior is quickly negligible in front of the information brought by the data when they are selected by the experimental design. Compared to a standard approach that aims to select seismic signals in their initial distribution, considering the Jeffreys prior, we have shown the superiority of our method. It allows reaching more quickly —i.e., with a limited number of data— a small estimation bias with a small credibility interval, while reducing the phenomenon of degeneracy. The proposed methodology therefore makes the most of the log-normal model. Since estimation biases can be expected in practice, it is recommended to use the method with few data (i.e., less than 100). To this end, we propose two stopping criteria that reflect the information provided by any additional data.

# III

## CONCLUSION & PERSPECTIVES



# 11

## Conclusion & perspectives

In this part, we provide a general conclusion to this thesis. Following the structure below, we express the conclusive thoughts that terminate this manuscript.

<b>11.1 Summary</b> . . . . .	<b>183</b>
<b>11.2 On the contributions to the reference prior theory</b> . . . . .	<b>184</b>
<b>11.3 On the contributions to the estimation of seismic fragility curves</b> . . . . .	<b>185</b>
<b>11.4 General outlook</b> . . . . .	<b>185</b>
<b>11.5 Perspectives and connections with other fields</b> . . . . .	<b>186</b>
<b>11.6 Final words</b> . . . . .	<b>187</b>

### 11.1 Summary

In this final part, we present an *a posteriori* analysis of the research conducted throughout this thesis, that we developed in [part I](#) and [part II](#). The *a priori* takes the form here of the motivations and preliminary considerations that were outlined in [chapter 1](#), through a detailed examination of the problem and a discussion of the methodological avenues that, in our view, invited for further exploration. To structure this investigation, we recall that we identified six guiding questions that shaped the architecture of the thesis. These are recapitulated below:

**Question i** How can one define and support the objectivity of a prior?

**Question ii** What are the limitations of implementing such non-informative priors, and how can their use be reconciled with practical needs?

**Question iii** How can such priors be constructed and derived in practice?

**Question iv** In the context of SPRA, what forms do these objective priors take within a seismic fragility curve model?

**Question v** What are the implications of information scarcity in such models, and how can they be addressed?

**Question vi** How can the different sources of information (*a priori* information and data-based information) be best integrated across the full Bayesian workflow of the model under study?

These six questions naturally span two domains that may appear distinct at first glance: the reference prior theory and the study of seismic fragility curves. Accordingly, the initial formulation of the research

questions and the analysis of the core problem were complemented by an in-depth review of the state of the art in both fields.

This distinction helped to structure the work carried out in this thesis, as well as the organization of the manuscript itself, with each part aligned more closely with one of the two domains. Consequently, in the following sections (sections 11.2 and 11.3), we reflect separately on the contributions made to each area. However, particular emphasis is also placed on the original contribution outlined from the interaction between the two domains. This interaction represents an innovative aspect that ultimately address the main problem addressed by the thesis, we discuss this point in section 11.4. These works are various and given the wide range of issues they engage with, several open directions for future research emerge. These are discussed in section 11.5.

## 11.2 On the contributions to the reference prior theory

Objectivism in the Bayesian workflow refers to the quest for an elusive ideal. In general, the literature agrees that objectivity is best approached when data and the information they provide are prioritized over any subjective incorporation. It is difficult to assert definitively whether reference priors are truly objective priors; nevertheless, their construction is explicitly designed to minimize the influence of subjective thoughts.

Thus, one can view an objective prior as an ideal, and the framework of reference priors offers a clue to quantify how close a given prior comes to that ideal, by means of the mutual information defined between the prior and the posterior. In chapter 4, we extended the scope of this measure by proposing a generalization of its fundamental definition. Our results demonstrated that the solution to this generalized problem remains robust. In most cases, under minimal assumptions, the reference prior remains the Jeffreys prior. In other words, this result strengthens the objectivity of the Jeffreys prior since it is the solution of the reference prior problem for a wider class of dissimilarity measures that define the mutual information. This investigation provided our answer to **question i**.

Actually, to address the many known limitations of Jeffreys prior, it is necessary to question the amount of objectivity that is truly expected for the problem of interest. In practice, it is often found that the “most objective” priors are not always the most desirable. They may be too complex, impractical to implement, or insufficiently informative leading, for example, to improper posteriors. When these issues become limiting, objectivism must be revisited and adapted to address them. This is not about abandoning the objectivity of the approach. It is about attempting to provide a practical solution that “minimizes” the loss of objectivity. This is the conclusion we arrived at in chapter 5, which responds to **question ii**. In that chapter, we propose incorporating constraints into the reference prior framework. These constraints are carefully designed to preserve objectivity as much as possible, while improving practical usability. Principally, our framework suggests a constraint that ensures the reference prior or its posterior will be proper. That proper characteristic of the posterior is essential for *a posteriori* inference based on posterior sampling, for instance. These works represent novel theoretical contributions that enrich the definition of priors built upon the foundation of reference priors, but tailored for real-world applicability.

Finally, in response to **question iii**, we address the challenge of solving the optimization problem that defines references priors. In this contribution, developed in chapter 6, we propose a method for approximating—potentially constrained—reference priors, even when their theoretical form is not known. As for the preceding chapter, this method embeds a new “blow” to the objectivity, in the sense that (i) it imposes the implicit parameterization of the prior through a neural network, and (ii) it does not guarantee that this parameterization will perfectly approximate the optimal (i.e., “most objective”) prior. Despite this, this approach, novel in both form and scope, represents a significant step toward enabling objective (or near-objective) Bayesian inference in practice. It also serves as the last link that connects the answers to the preceding questions within a tractable framework.

We therefore conclude that objectivity and usability are often in tension: quest for objectivity comes at a cost to practical application, and vice versa. However, we emphasize the capacity of the work presented in the part I of this manuscript to reconcile these two ideals as much as possible —generally by aiming to achieve the maximum possible practicality with the least possible sacrifice of objectivity.

## 11.3 On the contributions to the estimation of seismic fragility curves

The estimation of seismic fragility curves is a widely studied topic due to its concrete and impactful applications, making it both relevant and of crucial importance. In this thesis, we focused on the well-established probit-lognormal model, commonly used when observed data about structural responses are binary. The study of this model is a sufficiently rich topic to lead to significant contributions to our problem.

First, the construction of objective priors tailored for Bayesian estimation under this model had remained largely unexplored. We chose to address this gap through the scope of reference prior theory. Specifically, we derived, computed, and thoroughly analyzed the Jeffreys prior for the probit-lognormal model. This work, developed in [chapter 8](#) does not only answer **question iv**, but it also demonstrated that this prior performs significantly better in terms of both efficiency and accuracy compared to two serious competing approaches from the literature.

In parallel, we conducted a detailed study of the likelihood function of the model. Notably, the asymptotic behavior of this likelihood had not been analyzed in this form before. This analysis led us to identify and formalize the phenomenon of degeneracy, a critical issue that arises when the data provide insufficient information to the model. Degeneracy is frequently observed in estimates from many methods (such as Bayesian inference or maximum likelihood estimation with bootstrapping) and undermines the reliability of objective priors, among others. In response, we proposed a new construction of the prior, still supported by reference prior theory, and based on some of the theoretical contributions made in the [part I](#) of this manuscript. Indeed, the limitation introduced by degeneracy is directly related to the concerns raised in **question ii**. The application of this new methodology to the probit-lognormal model is developed in [chapter 9](#), thereby providing a complete response to **question v**.

Finally, we concluded our study by aiming to optimally integrate all sources of information available (*a priori* information and data-based information) for estimating fragility curves within the probit-lognormal framework. This way, in [chapter 10](#), we augmented prior information with an optimization of data-based information through a formal experimental design strategy. This methodology can be seen as a kind of “ultimate” approach to fragility curve estimation. Moreover, when used in conjunction with reference prior theory, it provides a valuable level of auditability in SPRA analyses. This work provides an answer that extends beyond **question vi**. Indeed, the proposed methodology handles many of the central issues identified throughout this research, and which concern the objectivity and robustness of the estimates, the ease of implementation and the achievement of minimal estimation bias. By an appropriate choice of data, our methodology makes the estimates robust. It allows to very effectively reduce the probability of likelihood degeneration for a given sample size. In the framework that we propose, we also found that, in doing so, the *a posteriori* estimation is very little influenced by the choice of the constraint that aims to make the prior proper. In other words, the methodology tends to reduce the “influence” of the prior on the estimations of the fragility curves. Finally, we observed that the methodology makes it possible to quickly achieve (i.e., with little data) the irreducible model bias. So, we provide a practical criterion and a threshold on the data size from which it becomes judicious to consider an alternative (typically non-parametric) model to conduct the inference.

## 11.4 General outlook

This thesis contributes not only to the theory of reference priors on one hand, and to the estimation of seismic fragility curves on the other, but more importantly, it highlights the connection between these two domains. The research carried out along each axis has fed into and enriched the other. This reflection is essential, as it underscores the value of technical motivations in guiding theoretical development, and conversely, the practical relevance of abstract considerations.

In our case, the application of reference priors to the estimation of seismic fragility curves is a problem with no precedent in the existing literature. Similarly, the development of reference prior theory driven by the specific context of fragility curve models is equally novel. These two parallel problems define the fundamental links between the chapters of this thesis, and have been addressed in tandem throughout the

work.

Specifically, the study of reference prior application began with the derivation of the Jeffreys prior for the probit-lognormal model, as presented in [chapter 8](#). Indeed, Jeffreys prior was already recognized in the literature as the solution to reference prior problem under general settings. With the idea to support even more objectivity of this prior beyond the historical framework of the theory, we led ourselves to revisit and generalize the foundational reference prior problem. This motivated the work carried out in [chapter 4](#), which validates the use of Jeffreys prior for conducting an objective Bayesian study.

Furthermore, the derivations performed in [chapter 8](#) revealed the limitations of Jeffreys prior when applied to the probit-lognormal model. These limitations concern the computational complexity and, more importantly, the improper nature of the posterior which can emerge when few data are available. The latter prompted a theoretical inquiry into how the reference prior framework might be adapted to avoid such drawbacks. This led to the development of the methods introduced in [chapter 5](#), which were directly applied in [chapter 9](#) to address and resolve those limitations.

[Chapter 6](#) offers a more open-ended contribution by proposing an approximation method for reference priors aimed at improving computational efficiency. While this method shows promise, it remains imperfect in terms of reliability. Therefore, although its potential was demonstrated in [chapter 9](#), it was not used in the “final” practical study conducted in [chapter 10](#).

Finally, [chapter 10](#) represents a sort of “ultimate” synthesis, combining all these elements, and including the optimization of data-based information in addition to prior information. This work resulted in a highly effective estimation methodology for fragility curves under the probit-lognormal model.

## 11.5 Perspectives and connections with other fields

The work conducted in this thesis opens several ways for future research and application. Three possible directions are proposed in what follows.

**Further development of the reference prior theory** First, it is possible to deepen the reference prior theory for objective Bayesian analysis. Reference priors are generally not unique, and this non-uniqueness becomes even more prominent when these priors are authorized to be defined conditionally on some constraints as proposed in the [part I](#) of this manuscript. In our opinion, the current theory remains limited by the lack of tools to compare and evaluate different priors systematically. This leads to several open questions: (i) Can we define a measure to compare the objectivity of priors in the context of reference prior theory? (ii) Could such a measure be extended to settings where the prior —and possibly even the posterior— is improper? (iii) Can we define the “most appropriate” way to constrain a prior, tailored to specific practical needs?

An answer to the first question necessitates further developments of the mutual information expression. It might be possible, for instance, to push further the current asymptotic form that it takes to determine and recover a “quantity of objectivity” that it measures.

The second question calls for a robust theoretical framework to formalize the use of improper priors in Bayesian modeling. This issue was partially addressed in the [part I](#) of this manuscript with a construction that aligns with that of Bioche (2015), for instance. However, a deeper treatment of the topological implications and their effect on reference prior theory remains to be explored.

The third question has a broader scope and can be seen as a natural application of answers to the first two.

**Application to the explainability of uncertainty in numerical models** Second, our coupling of reference priors with the estimation of seismic fragility curves in real-world scenarios has shown how this theoretical framework can anchor the explainability of outputs from numerical models. Indeed, “black-box” models —such as machine learning algorithms or numerical simulations— embed both aleatoric (data-related) and epistemic (model-related) uncertainties, as discussed in Hüllermeier and Waegeman, 2019 and in Chapter 1. As in the case of the probit-lognormal model used for seismic fragility curves estimation, explaining the

behavior of the “pipeline” requires quantifying those, and especially handling the second. Some works in the literature address the explainability of the uncertainties embedded in these models within a frequentist paradigm. We can cite Il Idrissi et al., 2024; Wimmer et al., 2023, which rely on sensitivity analysis tools to tackle the problem.

It is also possible to treat the problem under the Bayesian paradigm. This Bayesian path would involve defining appropriate priors over the model or its parameters and evaluating the posterior sensitivity of key quantities of interest to those prior assumptions. This path would raise the following questions that would leverage the definition of objective priors, as priors that are designed to let the data guide the posterior distribution: (i) What would be the appropriate prior to use in such context and how to approximate it efficiently? (ii) Would the result be sensitive to the prior? (iii) How to define and derive the aforementioned *a posteriori* sensitivity?

Naturally, research on these questions would also benefit from works on the perspectives mentioned earlier on the development of the reference prior theory.

**Towards better estimations of seismic fragility curves** Regarding the estimation of seismic fragility curves, the probit-lognormal model has been the main focus of this thesis, and its further refinement is not envisaged. However, their estimation using other frameworks, datasets and/or models remains open and remains a framework in which an objective study can be conducted.

First, one potential direction is to consider multiple IMs as inputs to the statistical model, allowing for a more comprehensive information on the seismic excitation on a structure. Second, instead of binary outcomes, continuous structural responses —such as EDPs— could be incorporated when available, since they embed richer information describing the impact of the input onto the structure. These directions have been explored using various models in existing literature. However, to our knowledge, none of these studies adopt a Bayesian approach that “optimally” allocates the information within the workflow (e.g., through the combination of an experimental design with a specifically tailored prior).

Lastly, we conclude by observing that different models naturally offer different fidelity levels in approximating the real fragility of the studied structure. As an example, we mention that we discuss in [appendix B](#) the use of a lower-fidelity model compared to the probit-lognormal model for fragility curves estimation. Combining multi-fidelity analyses, each bringing complementary information, could enhance the robustness and informativeness of the final fragility curve estimation.

## 11.6 Final words

In this thesis, we did not construct the definitive objective prior, nor did we estimate the most exact fragility curve for the mechanical structures studied. However, by addressing both of these questions individually and in close interaction, we have moved as close as possible to these two ideals.

We conclude by emphasizing that the new methodologies introduced in this work, along with their accompanying theoretical, formal, and experimental results, demonstrate a thoughtful balance between these two objectives. This balance has been key in providing the most effective response to answer the problem defined by this thesis.

For each result that forms the foundation of this conclusion, we insist that we have maintained a high level of rigor in our approach, in order to support the impact of the work to the problem context. Each individual contribution has been submitted (and in some cases already published) within the scientific community. These final words are thus backed by the seriousness and integrity with which this research has been conducted.



## APPENDICES



# A

## Influence of the IM choice on fragility curves estimated using a reference prior

This appendix is a postprint of the published work: Antoine Van Biesbroeck, Clément Gauchy, Cyril Feau, and Josselin Garnier (2023). “Influence of the choice of the seismic intensity measure on fragility curves estimation in a Bayesian framework based on reference prior”. In: *Proceedings of the 5th UNCECOMP Conference*, pp. 94–111. DOI: [10.7712/120223.10327.19899](https://doi.org/10.7712/120223.10327.19899)

**Abstract** The work presented in this appendix investigates the effect of the choice of the IM to Bayesian estimations of seismic fragility curves. As in the part II of this manuscript, we study the probit-lognormal model of fragility curves with datasets for which the structural responses are limited to binary outcomes (i.e., failures or non-failures). We implement a Bayesian approach that is based on an objective prior resulting from an approximation of the Jeffreys prior density. Considering two different IMs (PGA vs. PSA), we highlight that a more correlated IM to the structural response is more likely to yield degenerate scenarios. Such degeneracy compromises fragility curves estimations, either with Bayesian or classical frequentist methods. The consequences of these results on the estimates are thoroughly investigated on a case study, and compared for different approaches including ours. We precise that the case study considered here is identical only in terms of geometry to the piping system presented in chapter 7. However, a linear behavior of the structure is assumed here to accentuate the phenomena we wish to highlight, by increasing the correlation between its response and the IMs.

<b>A.1</b> Introduction . . . . .	<b>192</b>
<b>A.2</b> Bayesian problem . . . . .	<b>193</b>
<b>A.3</b> Reference prior theory . . . . .	<b>193</b>
<b>A.4</b> Jeffreys prior construction . . . . .	<b>194</b>
A.4.1 Jeffreys prior calculation . . . . .	194
A.4.2 IMs and practical implementation . . . . .	194
<b>A.5</b> Estimation tools, competing approaches and benchmarking metrics . . . . .	<b>195</b>
A.5.1 Fragility curves estimations via Monte-Carlo . . . . .	196
A.5.2 Fragility curves estimations in the Bayesian framework . . . . .	196
A.5.3 Competing approaches for performance evaluation . . . . .	196
A.5.4 Benchmarking metrics . . . . .	197
<b>A.6</b> Numerical application . . . . .	<b>198</b>
A.6.1 Presentation of the piping system and its correlation with the IMs . . . . .	198
A.6.2 Results and discussion . . . . .	199
<b>A.7</b> Conclusion . . . . .	<b>202</b>

## A.1 Introduction

Seismic fragility curves are key quantities of the Seismic Probabilistic Risk Assessment (SPRA) studies carried out on mechanical structures. They were introduced in the 1980s for safety studies of nuclear facilities (see e.g. Kennedy, Cornell, et al., 1980; Kennedy and Ravindra, 1984; Y.-J. Park, Hofmayer, and Chokshi, 1998; Kennedy, 1999; Cornell, 2004). They express the probability of failure of the mechanical structure conditional to a scalar value derived from the seismic ground motions —called Intensity Measure (IM)— such as the Peak Ground Acceleration (PGA) or the Pseudo-Spectral Acceleration (PSA) for a fixed frequency and damping. In practice, various data sources can be exploited to estimate fragility curves, namely: expert judgments supported by test data (Kennedy, Cornell, et al., 1980; Kennedy and Ravindra, 1984; Y.-J. Park, Hofmayer, and Chokshi, 1998; Zentner, Gündel, and Bonfils, 2017), experimental data (Y.-J. Park, Hofmayer, and Chokshi, 1998; Gardoni, Der Kiureghian, and Mosalam, 2002; Choe, Gardoni, and Rosowsky, 2007), results of damage collected on existing structures that have been subjected to an earthquake (Shinozuka, Feng, et al., 2000; Lallemand, Kiremidjian, and Burton, 2015; Straub and Der Kiureghian, 2008) and analytical results given by more or less refined numerical models using artificial or real seismic excitations (see e.g. Zentner, 2010; F. Wang and Feau, 2020; Mandal, Siddhartha Ghosh, and Pujari, 2016; Z. Wang, Pedroni, et al., 2018; Z. Wang, Zentner, and Zio, 2018; Zhao, Yu, and Mo, 2020). Parametric fragility curves were historically introduced in the SPRA framework because their estimates require small sample sizes. The probit-lognormal model has since become the most widely used model (see e.g. Shinozuka, Feng, et al., 2000; Lallemand, Kiremidjian, and Burton, 2015; Straub and Der Kiureghian, 2008; Zentner, 2010; F. Wang and Feau, 2020; Mandal, Siddhartha Ghosh, and Pujari, 2016; Z. Wang, Zentner, and Zio, 2018; Z. Wang, Pedroni, et al., 2018; Zhao, Yu, and Mo, 2020; Ellingwood, 2001; S.-H. Kim and Shinozuka, 2004; Mai, Konakli, and Sudret, 2017; Trevlopoulos, Feau, and Zentner, 2019; Katayama et al., 2021). Several strategies can be implemented to fit the median,  $\alpha$ , and the log standard deviation,  $\beta$ , of the model. Some of them are compared in Lallemand, Kiremidjian, and Burton, 2015 highlighting advantages and disadvantages. When the data is binary —i.e., when it indicates failure or not— Lallemand, Kiremidjian, and Burton (2015) recommended maximum likelihood estimation (MLE). When the data are independent, the bootstrap technique can be used to obtain confidence intervals relating to the size of the sample considered (e.g. Shinozuka, Feng, et al., 2000; Zentner, 2010; F. Wang and Feau, 2020).

Among the various other methods not mentioned in this short introduction, the Bayesian framework has recently become increasingly popular in seismic fragility analysis (see e.g. Gardoni, Der Kiureghian, and Mosalam, 2002; Z. Wang, Pedroni, et al., 2018; Katayama et al., 2021; Koutsourelakis, 2010; Damblin et al., 2014; Tadinada and A. Gupta, 2017; Kwag and A. Gupta, 2018; Jeon et al., 2019; Tabandeh, Asem, and Gardoni, 2020). It actually allows to solve irregularity issues encountered in the estimation of the parametric fragility curves. MLE-based methods can indeed lead to unrealistic or degenerate fragility curves such as unit step functions when the data availability is sparse. Those problems are especially encountered when resorting to complex and detailed modeling due to the calculation burden or when dealing with tests performed on shaking tables, etc. In earthquake engineering, Bayesian inference is often used to update probit-lognormal fragility curves obtained beforehand by various approaches, by assuming independent distributions for the prior values of  $\alpha$  and  $\beta$  such as log-normal distributions (see e.g. Tadinada and A. Gupta, 2017; Kwag and A. Gupta, 2018; Z. Wang, Pedroni, et al., 2018; Katayama et al., 2021; Straub and Der Kiureghian, 2008).

This work follows the one presented in chapter 8, which deals with Bayesian problems in which only few binary data are available. Using the reference prior theory as a support, we have proposed an objective approach to choose the prior and to simulate a posteriori fragility curves. This approach led to the Jeffreys prior and we have shown the robustness and advantages of the Jeffreys prior in terms of regularization (no irregular estimations) and stability (no outliers of the parameters) for fragility curves estimation under non-degenerate datasets. Since this prior depends only on the characteristics of the ground motion —the distribution of the IM of interest— its calculation is then suitable for any equipment of an industrial installation subjected to the same seismic hazard. So, in this work, we are interested in the influence of the choice of the IM —PGA vs. PSA— on the convergence of the estimates, for a given magnitude (M) - source-to-site distance (R) scenario and a given mechanical structure.

The appendix is organized as follows. Section A.2 presents the statement of the problem from the Bayesian

viewpoint. A review of the objective prior theory is presented in section A.3. The principal achievements of chapter 8 on which we rely are summarized in section A.4. Section A.5 is dedicated to reviewing estimation tools and benchmarking metrics used to support comparisons with classical approaches of the literature. They are implemented in section A.6 on a case study, a piping system. A conclusion is proposed in section A.7.

## A.2 Bayesian problem

A probit-lognormal model is often used to approximate fragility curves. In this model the probability of failure given the IM takes the following form:

$$P_f(a) = \mathbb{P}(\text{'failure'} | \text{IM} = a) = \Phi\left(\frac{\log a - \log \alpha}{\beta}\right), \quad (\text{A.1})$$

where  $\alpha, \beta \in (0, +\infty)$  are the two model parameters and  $\Phi$  is the cumulative distribution function of a standard Gaussian variable. In the following we denote  $\theta = (\alpha, \beta)$ . In the Bayesian point of view  $\theta$  is considered as a random variable. Its probability density function is denoted by  $\pi$  and called the prior density, it is supposed to be defined on a set  $\Theta \subset (0, +\infty)^2$ .

The statistical model consists in the observations of independent realizations  $(a_1, z_1), \dots, (a_k, z_k) \in \mathcal{A} \times \{0, 1\}$ ,  $\mathcal{A} \subset (0, \infty)$ ,  $k$  being the dataset size. For the  $i$ th seismic event,  $a_i$  is its observed IM and  $z_i$  is the observation of a failure ( $z_i$  is equal to one if failure has been observed during the  $i$ th seismic event and it is equal to zero otherwise). The joint probability density of the pair  $(a, z)$  conditionally to  $\theta$  has the form:

$$p(a, z | \theta) = h(a)\ell(z | a, \theta), \quad (\text{A.2})$$

where  $h(a)$  denotes the p.d.f. of the IM and  $\ell(z | a, \theta)$  is the density of a Bernoulli distribution whose parameter (the probability of failure) depends on  $a$  and  $\theta$  as expressed by eq. (A.1). The product of the conditional distributions  $\ell(z_i | a_i, \theta)$  is the likelihood of the model:

$$\ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta) = \prod_{i=1}^k \ell(z_i | a_i, \theta) = \prod_{i=1}^k \Phi\left(\frac{\log a_i - \log \alpha}{\beta}\right)^{z_i} \left(1 - \Phi\left(\frac{\log a_i - \log \alpha}{\beta}\right)\right)^{1-z_i}, \quad (\text{A.3})$$

denoting  $\mathbf{a}^k = (a_i)_{i=1}^k$ ,  $\mathbf{z}^k = (z_i)_{i=1}^k$ . The posterior density of  $\theta$  can be computed by Bayes theorem:

$$p(\theta | \mathbf{z}^k, \mathbf{a}^k) = \frac{\ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta)\pi(\theta)}{\int_{\Theta} \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta')\pi(\theta')d\theta'}. \quad (\text{A.4})$$

We recall that the likelihood is said to be degenerate in some cases: (i) when only failures or only non-failures are observed, and (ii) when an IM threshold allows separating failures from non-failures among the observations (i.e.,  $\exists a \in \mathcal{A}, \forall i, a_i < a \iff z_i = 0$ ). The degeneracy is defined in chapter 8 (section 8.5). In those cases, the likelihood decay rates make both priors considered in this study yielding improper posteriors.

## A.3 Reference prior theory

To choose a non-subjective prior, we consider as in chapter 8 a so-called reference prior. It consists in choosing the prior that maximizes the mutual information indicator  $\mathcal{I}^k$  which expresses the information provided by the data to the posterior, relatively to the prior. In other words, this criterion seeks the prior that maximizes the “learning” capacity from observations. We refer to the part I of the manuscript for more details. The mutual information indicator can be expressed as a function of the prior density:

$$\mathcal{I}^k(\pi) = \int_{(\mathcal{A} \times \{0, 1\})^k} D(p(\cdot | \mathbf{z}^k, \mathbf{a}^k) || \pi)p(\mathbf{z}^k, \mathbf{a}^k)d\lambda^{\otimes k}(\mathbf{a}^k, \mathbf{z}^k), \quad (\text{A.5})$$

where  $p(\mathbf{z}^k, \mathbf{a}^k) = \int_{\Theta} \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta')\prod_{l=1}^k h(a_l)\pi(\theta')d\theta'$  and the reference measure  $\lambda$  is the product of the Lebesgue measure over  $\mathcal{A}$  and the discrete measure  $\delta_0 + \delta_1$  over  $\{0, 1\}$ . The indicator in eq. (A.5) is based on a

## Appendix A. Influence of the IM choice on fragility curves estimated using a reference prior

divergence  $D$  between the posterior and the prior densities, which is known to numerically express this idea of the information provided by one distribution to another one. This divergence can be the Kullback-Leibler divergence or a  $\delta$ -divergence, for instance (see [chapter 3](#) and [chapter 4](#)).

A suitable definition of a reference prior is suggested as the solution of an asymptotic optimization of this mutual information metric. It has been proved that, under some mild assumptions which are satisfied in our framework, the Jeffreys prior, whose density is defined by

$$J(\theta) \propto \sqrt{|\det \mathcal{I}(\theta)|}, \quad (\text{A.6})$$

is the reference prior, with  $\mathcal{I}$  being the Fisher information matrix:

$$\mathcal{I}(\theta)_{i,j} = - \int_{\mathcal{A} \times \{0,1\}} \ell(z|a, \theta) \partial_{\theta_i \theta_j}^2 \log \ell(z|a, \theta) h(a) \lambda(da, dz). \quad (\text{A.7})$$

The Jeffreys prior is already well known in Bayesian theory for being invariant by a re-parametrization of the statistical model. This property is essential as it makes the choice of the model parameters  $\theta$  without any incidence on the resulting posterior.

## A.4 Jeffreys prior construction

### A.4.1 Jeffreys prior calculation

In a first step, we compute the Fisher information matrix  $\mathcal{I}(\theta)$  in our model. Here,  $\theta = (\alpha, \beta) \in \Theta$  and

$$\mathcal{I}(\theta)_{i,j} = - \int_{\mathcal{A} \times \{0,1\}} \ell(z|a, \theta) \partial_{\theta_i \theta_j}^2 \log p(z|a, \theta) h(a) \lambda(da, dz) \quad (\text{A.8})$$

for  $i, j \in \{1, 2\}$ , with  $\theta_1 = \alpha$ ,  $\theta_2 = \beta$ , i.e.,

$$\log \ell(z|a, \theta) = z \log \Phi \left( \frac{\log a - \log \alpha}{\beta} \right) + (1-z) \log \left( 1 - \Phi \left( \frac{\log a - \log \alpha}{\beta} \right) \right). \quad (\text{A.9})$$

From [chapter 8](#), the information matrix  $\mathcal{I}(\theta)$  is given by:

$$\mathcal{I}(\theta) = \begin{pmatrix} \frac{1}{\alpha^2 \beta^2} (A_{01} + A_{02}) & \frac{1}{\alpha \beta^3} (A_{11} + A_{12}) \\ \frac{1}{\alpha \beta^3} (A_{11} + A_{12}) & \frac{1}{\beta^4} (A_{21} + A_{22}) \end{pmatrix}, \quad (\text{A.10})$$

with

$$\begin{aligned} A_{01} &= \int_{\mathcal{A}} \frac{\Phi'(\gamma(a))^2}{\Phi(\gamma(a))} h(a) da, & A_{02} &= \int_{\mathcal{A}} \frac{\Phi'(\gamma(a))^2}{\Phi(-\gamma(a))} h(a) da, \\ A_{11} &= \int_{\mathcal{A}} \log \frac{a}{\alpha} \frac{\Phi'(\gamma(a))^2}{\Phi(\gamma(a))} h(a) da, & A_{12} &= \int_{\mathcal{A}} \log \frac{a}{\alpha} \frac{\Phi'(\gamma(a))^2}{\Phi(-\gamma(a))} h(a) da, \\ A_{21} &= \int_{\mathcal{A}} \log^2 \frac{a}{\alpha} \frac{\Phi'(\gamma(a))^2}{\Phi(\gamma(a))} h(a) da, & A_{22} &= \int_{\mathcal{A}} \log^2 \frac{a}{\alpha} \frac{\Phi'(\gamma(a))^2}{\Phi(-\gamma(a))} h(a) da, \end{aligned} \quad (\text{A.11})$$

and  $\gamma(a) = \beta^{-1} \log \frac{a}{\alpha}$ .

The Jeffreys prior is known to be improper in numerous common cases (i.e., it cannot be normalized as a probability). In the present case, its asymptotic behavior is computed for different limits of  $\theta = (\alpha, \beta)$  in [chapter 8](#), which shows that it is indeed improper. Regarding the posterior, it is improper when the likelihood is degenerate, and proper otherwise.

### A.4.2 IMs and practical implementation

In this work we use  $10^5$  artificial seismic signals generated using the stochastic generator defined in Rezaeian, 2010 and implemented in Sainct et al., 2020 from 97 real accelerograms selected in the European Strong

Motion Database for  $5.5 \leq M \leq 6.5$  and  $R < 20$  km. Enrichment is not a necessity in the Bayesian framework—especially if a sufficient number of real signals is available—but it allows comparisons with the reference method of Monte-Carlo for simulation-based approaches as well as comparative studies of performance. For instance, fig. A.1 shows that the synthetic signals have the same PGA distribution as the real ones as well as the PSA which is here calculated at 5 Hz for 1% damping ratio (see section A.6 for justification). Moreover, the asymptotic expansions provided in chapter 8 give complementary and essential insight into the Jeffreys prior. They evince that its behavior in  $\alpha$  is similar to that of a log-normal distribution having the same median as that of the IM with a variance which is the sum of the variance of the IM and of a term which depends on  $\beta$ . Figure A.1 illustrates also this result for the two IMs.

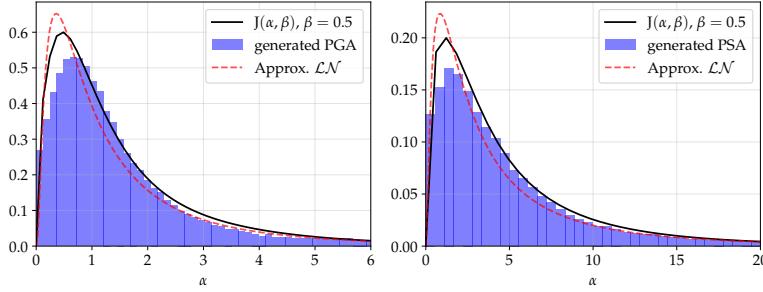


Figure A.1: Comparison of a sectional view of the Jeffreys prior density w.r.t.  $\alpha$  (black) with the approximated distributions of real accelerograms via Gaussian kernel estimation (red), the histograms of the generated signals (blue) and the log-normal approximations (purple) for the PGA (left) and for the PSA (right).

In practice, due to the use of Markov Chain Monte Carlo (MCMC) methods to sample the *a posteriori* distribution, the prior density must be evaluated (up to a multiplicative constant) many times in the calculations. Because of its computational complexity due to the integrals to be computed, we performed evaluations of the prior on an experimental design based on a fine-mesh grid of  $\Theta$  (here  $(0, +\infty)^2$ ) and to build an interpolated approximation of the Jeffreys prior density from this design. This strategy is more suitable for our numerical applications and very tractable because the domain  $\Theta$  is only two-dimensional. Figure A.2 shows plots of the Jeffreys prior densities.

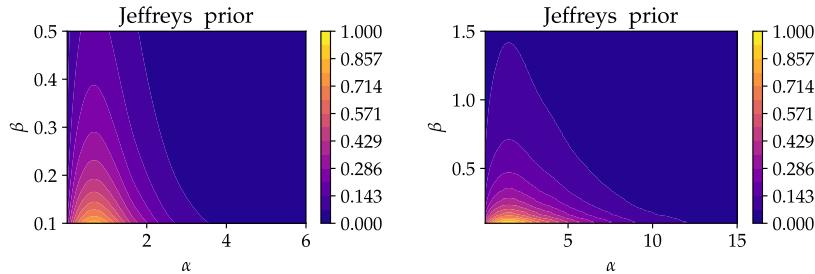


Figure A.2: The Jeffreys priors calculated from PGA (left) and PSA (right) on subdomains of  $\Theta = (0, +\infty)^2$ .

## A.5 Estimation tools, competing approaches and benchmarking metrics

In this section, we first present the Bayesian estimation tools and the Monte-Carlo reference method to which we refer to evaluate the relevance of the probit-lognormal model. Then, to evaluate the performance of the Jeffreys prior, we present two competing approaches that we implement. On the one hand, we apply the MLE method widely used in literature, coupled with a bootstrap technique. On the other hand, we apply a Bayesian technique implemented with the prior introduced by Straub and Der Kiureghian (2008). Performance evaluation metrics are then defined.

### A.5.1 Fragility curves estimations via Monte-Carlo

We assume that a validation dataset  $(\mathbf{a}^{\text{MC}}, \mathbf{z}^{\text{MC}}) = ((a_i^{\text{MC}})_{i=1}^{N^{\text{MC}}}, (z_i^{\text{MC}})_{i=1}^{N^{\text{MC}}})$  is available. For nonparametric estimations, good candidates are Monte-Carlo (MC) estimators which estimate the expected number of failures locally w.r.t. the IM. We first need to define a subdivision of the IM values and to estimate the failure probability on each of the sub-intervals. Regular subdivisions are not appropriate because the observed IMs are not uniformly distributed. We follow the suggestion by Trevlopoulos, Feau, and Zentner (2019) to take clusters of the IM using K-means. Given such  $N_c$  clusters  $(K_j)_{j=1}^{N_c}$ , the Monte-Carlo fragility curve estimated at the centroids  $(c_j)_{j=1}^{N_c}$  is expressed as

$$P_f^{\text{MC}}(c_j) = \frac{1}{n_j} \sum_{i, a_i^{\text{MC}} \in K_j} z_i^{\text{MC}}, \quad (\text{A.12})$$

where  $n_j = \text{Card}(i, a_i^{\text{MC}} \in K_j)$  is the size of cluster  $K_j$ . An asymptotic confidence interval for this estimator can also be derived using its Gaussian approximation. It is accepted that a MC-based fragility curve is a reference curve because it is not based on any assumption.

### A.5.2 Fragility curves estimations in the Bayesian framework

From eq. (A.4) *a posteriori* samples of  $\theta$  can be obtained by MCMC methods. We have implemented an adaptive Metropolis-Hastings (M-H) algorithm with Gaussian transition kernel and covariance adaptation (Haario, Saksman, and Tamminen, 2001). Such an algorithm allows simulating from a probability density known up to a multiplicative constant. The *a posteriori* samples of  $\theta$  can be used to define credible intervals for the probit-lognormal estimates of the fragility curves.

### A.5.3 Competing approaches for performance evaluation

**Multiple MLE by bootstrapping.** The best known parameter estimation method is the MLE, defined as the maximal argument of the likelihood derived from the observed data:

$$\hat{\theta}_k^{\text{MLE}} = \arg \max_{\theta \in \Theta} \ell_k(\mathbf{z}^k | \mathbf{a}^k, \theta). \quad (\text{A.13})$$

A common method for obtaining a wide range of  $\theta$  estimates is to compute multiple MLE by bootstrapping. Denoting the dataset size by  $k$ , bootstrapping consists in doing  $L$  independent draws with replacement of  $k$  items within the dataset. Those lead to  $L$  different likelihoods from the  $k$  initial observations, and so to  $L$  values of the estimator which can be averaged. In the context of fragility curves, this method is widespread (see e.g. Shinozuka, Feng, et al., 2000; Lallemand, Kiremidjian, and Burton, 2015; Gehl, Douglas, and Seyedi, 2015; Baker, 2015; F. Wang and Feau, 2020). The convergence of the MLE and the relevance of this method is stated in Vaart, 1992. Nevertheless the bootstrap method is often limited by the irregularity of the results for small values of  $k$  (see e.g. Zentner, Gündel, and Bonfils, 2017). In this context, the  $L$  values of  $\theta$  are used to define confidence intervals for the probit-lognormal estimates of the fragility curves.

**Prior suggested by Straub and Der Kiureghian (2008).** This prior, called SK prior, is defined as the product of a normal distribution for  $\ln(\alpha)$  and the improper distribution  $1/\beta$  for  $\beta$ , namely its density is:

$$\pi_{SK}(\theta) \propto \frac{1}{\alpha\beta} \exp\left(-\frac{(\log \alpha - \mu)^2}{2\sigma^2}\right). \quad (\text{A.14})$$

In Straub and Der Kiureghian, 2008 the parameters  $\mu$  and  $\sigma$  of the log-normal distribution are chosen to generate a non-informative prior. For a fair comparison with the approach proposed in this appendix, we decided to choose  $\mu$  and  $\sigma$  being equal to the mean and the standard deviation of the logarithm of the IM, whether the PGA or the PSA. This choice is consistent with the fact that the Jeffreys prior is similar to a log-normal distribution with these parameters (see fig. A.1). The prior density  $\pi_{SK}(\theta)$  is plotted in fig. A.3 for the two IMs.

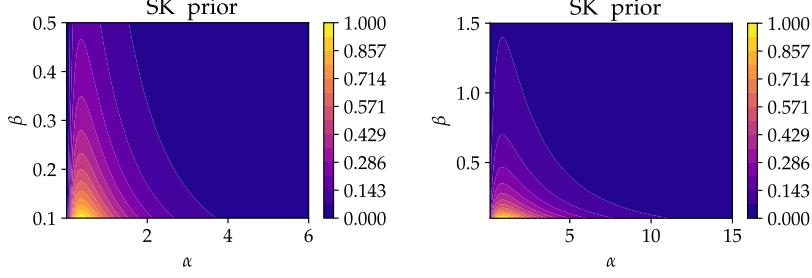


Figure A.3: Density of the prior suggested by Straub and Der Kiureghian (2008) scaled on log-normal estimations of the PGA (left) and of the PSA (right) distributions.

An analysis of the posterior which results from SK prior is given in chapter 8. It shows that the posterior is always improper. This statement jeopardizes the validity of *a posteriori* simulations using MCMC methods if we consider the whole domain  $\Theta = (0, +\infty)^2$ . This issue is nevertheless manageable with the consideration of a truncation w.r.t.  $\beta$ .

#### A.5.4 Benchmarking metrics

In order to evaluate the performance of the proposed approach, we consider two quantitative metrics which can be calculated for each of the methods described in section A.5.3. We consider the sample  $(\mathbf{z}^k, \mathbf{a}^k)$ . We denote by  $a \mapsto P_f^{|\mathbf{z}^k, \mathbf{a}^k}(a)$  the random process defined as the fragility curve conditional to the sample (the probability distribution of  $P_f^{|\mathbf{z}^k, \mathbf{a}^k}(a)$  is inherited from the *a posteriori* distribution of  $\theta$ ). For each value  $a$  the  $r$ -quantile of the random variable  $P_f^{|\mathbf{z}^k, \mathbf{a}^k}(a)$  is denoted by  $q_r^{|\mathbf{z}^k, \mathbf{a}^k}(a)$ . We define:

- The quadratic error:

$$\mathcal{E}^{|\mathbf{z}^k, \mathbf{a}^k} = \mathbb{E}[\|P_f^{|\mathbf{z}^k, \mathbf{a}^k} - P_f^{\text{ref}}\|_{L^2}^2 | \mathbf{z}^k, \mathbf{a}^k], \quad \|P\|_{L^2}^2 = \frac{1}{A_{\max}} \int_0^{A_{\max}} P(a)^2 da. \quad (\text{A.15})$$

$P_f^{\text{ref}}$  stands for non-parametric estimate of the fragility curve  $P_f^{\text{MC}}$  derived using the method described in section A.5.1 from a validation dataset.

- The  $1 - r$  square credibility width:

$$\mathcal{W}^{|\mathbf{z}^k, \mathbf{a}^k} = \|q_{1-r/2}^{|\mathbf{z}^k, \mathbf{a}^k} - q_{r/2}^{|\mathbf{z}^k, \mathbf{a}^k}\|_{L^2}^2. \quad (\text{A.16})$$

To estimate such variables, we simulate a set of  $L$  fragility curves  $(P_f^{\theta_i |\mathbf{z}^k, \mathbf{a}^k})_{i=1}^L$  where  $(\theta_i)_{i=1}^L$  is a sample of the *a posteriori* distribution of the model parameters obtained by MCMC. The empirical quantiles  $\hat{q}_r^{|\mathbf{z}^k, \mathbf{a}^k}(a)$  of  $(P_f^{\theta_i |\mathbf{z}^k, \mathbf{a}^k}(a))_{i=1}^L$  are approximations of the quantiles  $q_r^{|\mathbf{z}^k, \mathbf{a}^k}(a)$  of the random variable  $P_f^{|\mathbf{z}^k, \mathbf{a}^k}(a)$ . We derive:

- The approximated quadratic error:

$$\mathcal{E}^{|\mathbf{z}^k, \mathbf{a}^k} \approx \frac{1}{L} \sum_{i=1}^L \|P_f^{\theta_i |\mathbf{z}^k, \mathbf{a}^k} - P_f^{\text{MLE}}\|_{L^2}^2. \quad (\text{A.17})$$

- The approximated  $1 - r$  square credibility width:

$$\mathcal{W}^{|\mathbf{z}^k, \mathbf{a}^k} \approx \|\hat{q}_{1-r/2}^{|\mathbf{z}^k, \mathbf{a}^k} - \hat{q}_{r/2}^{|\mathbf{z}^k, \mathbf{a}^k}\|_{L^2}^2. \quad (\text{A.18})$$

The normalized  $L^2$  norms are normalized integrals over  $a \in [0, A_{\max}]$  which are approximated numerically using Simpson's interpolation on a regular subdivision  $0 = A_0 < \dots < A_p = A_{\max}$ . In the forthcoming section we use  $A_0 = 0$ ,  $A_{\max} = 24 \text{ m/s}^2$  for the PGA and  $A_{\max} = 50 \text{ m/s}^2$  for the PSA with  $p = 200$ .

For the MLE with bootstrapping, we can define a conditional quadratic error as in eq. (A.17) and conditional width of the  $1 - r$  confidence interval as in eq. (A.18) using a bootstrapped sample  $(\theta_i)_{i=1}^L$ .

## A.6 Numerical application

### A.6.1 Presentation of the piping system and its correlation with the IMs

This case study concerns a piping system that is a part of a secondary line of a French Pressurized Water Reactor. This piping system was studied, experimentally and numerically, as part of the ASG program (Touboul, Sollogoub, and Blay, 1999). Figure A.4 shows a view of the mock-up mounted on the shaking table Azalee of the EMSI laboratory of CEA Saclay whereas the Finite Element Model (FEM) —based on beam elements— is shown in fig. A.4-right. The latter has been implemented with the homemade FE code CAST3M (CEA, 2019) and has been validated via an experimental campaign.

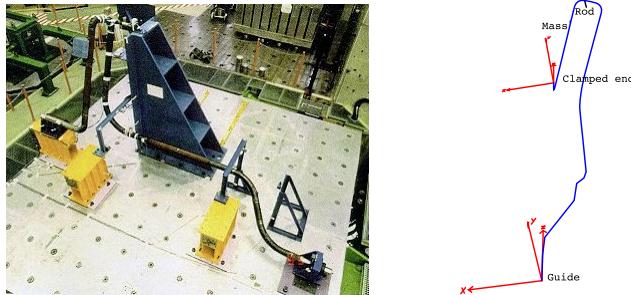


Figure A.4: (left) Overview of the piping system on the Azalee shaking table and (right) Mock-up FEM.

One end of the mock-up is clamped whereas the other is supported by a guide in order to prevent the displacements in the X and Y directions. Additionally, a rod is placed on the top of the specimen to limit the mass displacements in the Z direction (see fig. A.4-right). In the tests, the excitation is only imposed in the X direction. For this study, the artificial signals are filtered by a fictitious 2% damped linear single-mode building at 5 Hz, the first eigenfrequency of the 1% damped piping system. As failure criterion, we consider excessive out-of-plane rotation of the elbow located near the clamped end of the mock-up, as recommended in Touboul, Blay, et al., 2006. The critical rotation considered here is equal to 1.6°. This is the level quantile 90% of a sample of size  $10^5$  of numerical simulations carried out assuming a linear behavior of the piping system. A linear behavior is considered to simply highlight the influence of the choice of IM. Indeed we use on the one hand the PGA and on the other hand the PSA of the initial set of synthetic signals (i.e not filtered signals), calculated at 5 Hz and 1% damping ratio.

For the two IMs, fig. A.5 shows the comparisons between the reference MC-based fragility curves  $P_f^{\text{MC}}$  and their probit-lognormal estimates  $P_f^{\text{MLE}}$ , both estimated from a validation database of  $10^5$  simulations results. In both cases, the probit-lognormal fragility curves are good approximations of the reference ones.

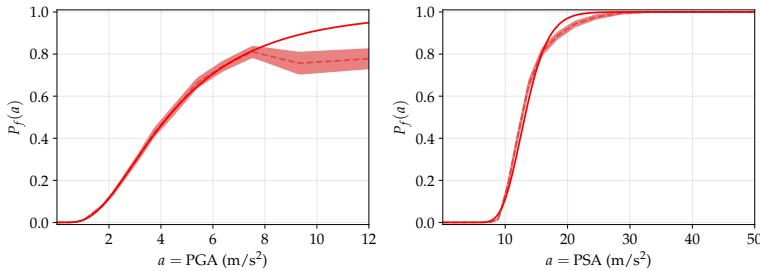


Figure A.5: Reference fragility curves  $P_f^{\text{MC}}$  compared with  $P_f^{\text{MLE}}$  computed using the full dataset generated ( $10^5$  items) for the PGA (left) and for the PSA (right).

**Correlation of the structure's response with the IMs** Figure A.6 shows that the PSA is clearly better correlated with the response of the structure than the PGA. This correlation is remarkable in fig. A.5 as well,

where the reference fragility curves  $P_f^{\text{MC}}$  (see section A.5.1) are plotted. Indeed, the PSA is a more discriminating indicator of the state of the structure than the PGA. This results in a “flatter” fragility curve when the PGA is used. In other words, with the PSA, the probability is greater that its values correspond to failure probabilities close to 1 or 0. As a consequence, random samples of PSAs are more likely to yield degenerate likelihoods, as shown in fig. A.7.

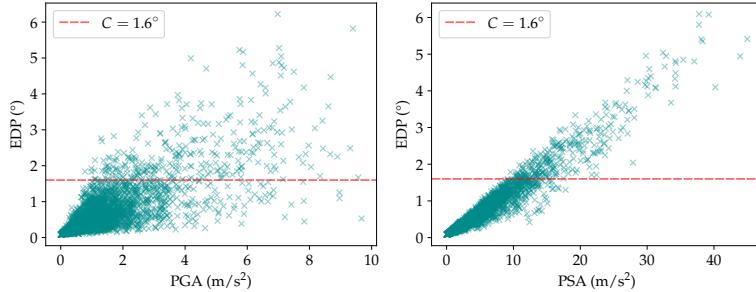


Figure A.6: Scatter plots of the out-of-plane elbow rotation as a function of (left) PGA and (right) PSA for a linear seismic behavior of the piping system.

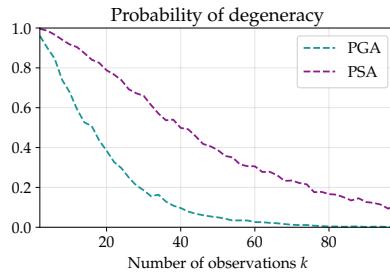


Figure A.7: Proportion of degeneracy yielded when drawing randomly a data sample in the database, as a function of the size  $k$  of the drawn sample.

## A.6.2 Results and discussion

Figures A.8 and A.9 present the results for the PGA and the PSA as IM respectively. Figures A.8-(a) and (b) (resp. figs. A.9-(a) and (b)) show the metrics  $\mathcal{E}|z^k, \alpha^k$ ,  $\mathcal{W}|z^k, \alpha^k$  for each of the three methods, for  $1 - r = 95\%$ , and for  $k$  varying from 5 to 100. Their averages and confidence intervals are derived from 200 draws of the datasets. Figures A.8-(c) and A.9-(c) present examples of fragility curves credibility (or confidence for the MLE) intervals for the three methods introduced in section A.5.3 for  $k = 30$ , in comparison with  $P_f^{\text{MC}}$ . Those are computed from generated pairs  $(\alpha, \beta)$  whose scatter plots are presented in figs. A.8-(d) and A.9-(d).

In addition, to get a better overview on the results, we also show in fig. A.10 the coefficients of variation of the two parameters  $(\alpha, \beta)$  as a function of both IM and  $k$ .

Generally speaking, these results show, as expected, that when the IM is more correlated to structural response, the differences between the methods are less marked, although in detail there may be some small differences depending on the sample size. Figure A.10 clearly shows that an IM more correlated to the response of the structure induces a lower variability of the estimate of the median of the probit-lognormal model. This is not the case for the log standard deviation whose estimate is affected by samples which are more degenerate with this kind of IM. Degenerate samples affect all methods to varying degrees but in all cases, the Jeffreys prior outperforms SK prior and MLE.

Whereas the SK prior is calibrated to look like the Jeffreys prior, figs. A.8-(d) and A.9-(d) show that many outliers —i.e., large values of the pair  $(\alpha, \beta)$ — are simulated with the SK prior. These values explain that the credible intervals of the fragility curves and the metrics  $\mathcal{E}|z^k, \alpha^k$  and  $\mathcal{W}|z^k, \alpha^k$  are larger with the SK prior. This observation is both supported by fig. A.10 and by the calculations provided in chapter 8 regarding the

## Appendix A. Influence of the IM choice on fragility curves estimated using a reference prior

asymptotic rates in  $\beta$ . Indeed, in chapter 8 is shown a better convergence of the Jeffreys prior toward 0 when  $\beta \rightarrow \infty$ . This better asymptotic behavior results in posteriors which happen to give a lower probability to outlier points —phenomenon particularly noticeable when the dataset is small— as well as to the weight of the likelihood within the posterior.

Although the intervals compared —that of the Bayesian framework and that of the MLE— are not of the same nature —credible interval for the first *versus* confidence interval for the second— these results clearly illustrate the advantage of the Bayesian framework over the MLE for small samples. Indeed, irregularities appear in the MLE method that are characterized by null estimates of  $\beta$ , which result (i) in important coefficient of variation of  $\beta$  (fig. A.10) and (ii) in “vertical” confidence intervals on fragility curves estimations (figs. A.8-(c) and A.9-(c)). When few failures are observed, some samples —both initial and bootstrapped samples— are degenerate, as explained earlier. As no prior is considered in the MLE-based approach, the likelihood can then be easily maximized with  $\beta = 0$ . In chapter 8, it is proven that such scenarios result in degenerate likelihood. This last statement is perceived better through figs. A.8-(d) and A.9-(d). The zero-degenerate  $\beta$  values that result from the MLE appear clearly. This leads to a confidence interval generally larger than the credible intervals, except for very small values of  $k$  ( $k \simeq 20$ ) when the IM is well correlated with the response of the structure, i.e. with the PSA. Indeed, with a perfect IM - which only exists if we know the structural response itself - the fragility curve is degenerate and has the form of a unit step function. The associated confidence interval is of null size, since in this case it does not require any sample to estimate the fragility curve. So, although apparently better, such confidence intervals are meaningless since they are based on degenerate estimates.

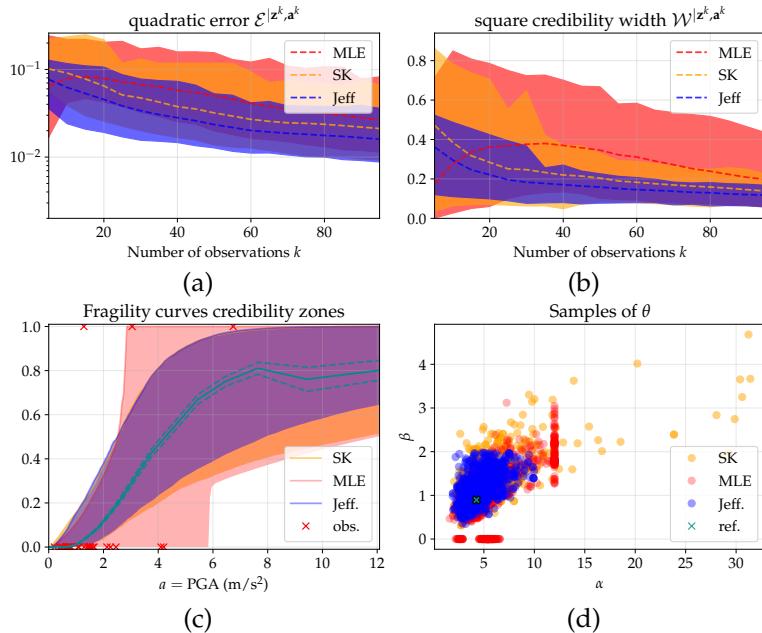


Figure A.8: (a) average values of  $E|z^k,a^k$  (dashed lines) surrounded by their 95%-confidence intervals; (b) average values of  $W|z^k,a^k$  (dashed lines) surrounded by their 95%-confidence intervals; (c) the shaded areas show the 95% credible (for Bayesian estimation) or confidence (for MLE) intervals resulting from a total of 5000 simulations of  $\theta$  using the three methods considered for  $k = 30$ ; (d) scatter plots of the generated  $\theta$  to estimate the fragility curves for the three methods (1000 points from the 5000  $\theta = (\alpha, \beta)$  generated are plotted); the green cross in (d) plots  $\theta^{\text{MLE}}$  which is derived from the validation dataset; the green line in (c) refers to  $P_f^{\text{MC}}$ , plotted along with its 95% intervals (dashed lines). Here the PGA is used as IM.

As proven in chapter 8, in the Bayesian context, the same degenerate samples also produce degeneracy but “less marked” than for MLE, as the phenomena is regularized by the prior distribution. As the results show, this affects SK prior more than Jeffreys one. This is observed, in particular, when the IM is very well correlated with the response of the structure since, in this case, it is more probable to obtain this kind of

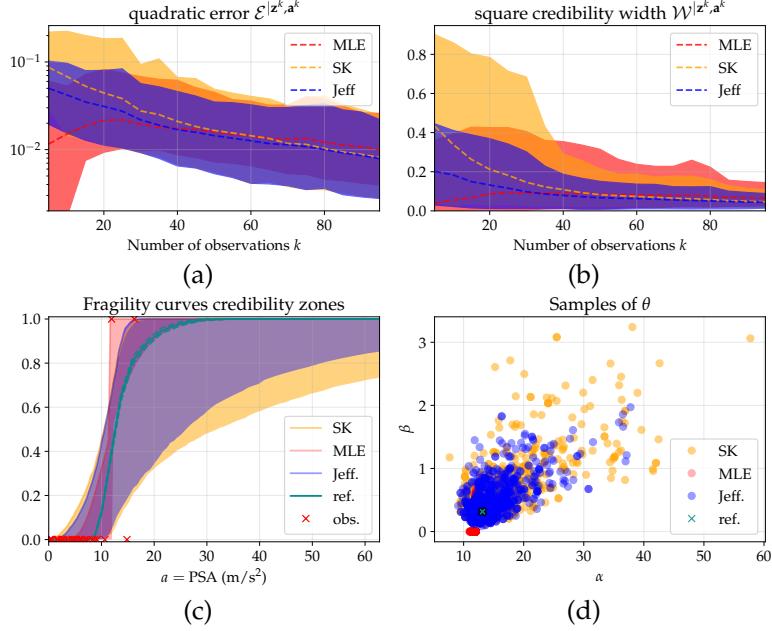


Figure A.9: Same as in fig. A.8, but here the PSA is used as IM.

degenerate samples. Therefore, when  $k < 40$ , the credible intervals are slightly larger with the PSA than with the PGA. This is also confirmed by the results shown in fig. A.10. So, counter-intuitively, when very few data are available, a less well-specified problem from the point of view of the choice of the IM leads to better convergences of the estimates, because it produces fewer degenerate samples. However, this remains confined to very small sample sizes and therefore cannot be considered representative. Note that, in all cases, the Jeffreys prior outperforms SK prior and MLE.

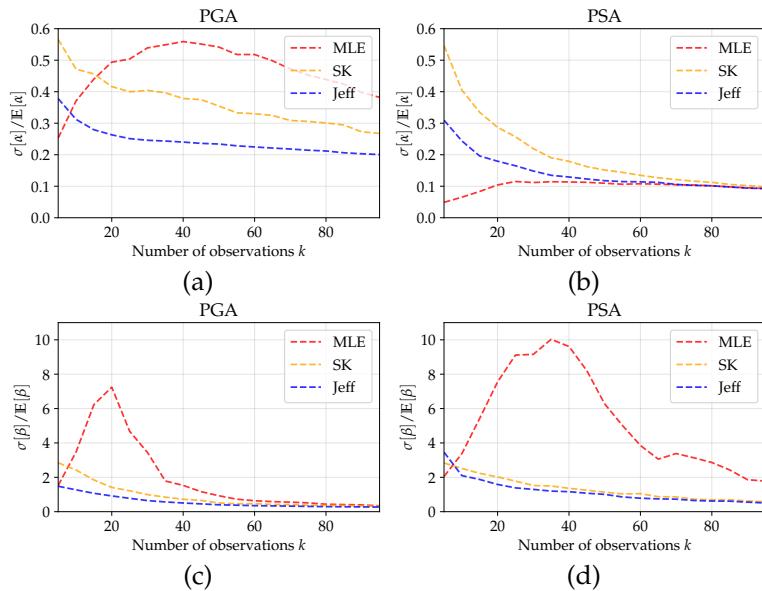


Figure A.10: Average coefficient of variation of  $\alpha$  (a-b) and of  $\beta$  (c-d) for the PGA and the PSA. For each value of  $k$ , 200 samples of size  $k$  have been used to compute the average values of the resulting 200 coefficients of variation from 5000 estimates of  $\theta$  each.

## A.7 Conclusion

Assessing the seismic fragility of Structures and Components when little data is available is a daunting task. The Bayesian framework is known to be efficient for this kind of problems. Nevertheless, the choice of the prior remains tricky because it has a non-negligible influence on the *a posteriori* distribution and therefore on the estimation of any quantity of interest linked to the fragility curves.

Using the reference prior theory to define an objective prior, we have derived the Jeffreys prior for the probit-lognormal model with binary data which indicate the state of the structure (e.g. failure or non-failure). In doing so, this prior is completely defined, it does not depend on an additional subjective choice.

Since this prior depends only on the characteristics of the ground motion —the distribution of the IM of interest— its calculation is then suitable for any equipment of an industrial installation subjected to the same seismic hazard. So, in this work, we were interested in the influence of the choice of the IM —PGA vs. PSA— on the convergence of the estimates, for a given (M,R) seismic scenario and a given mechanical structure, namely a piping system.

The results show, as expected, that when the IM is more correlated with the structural response, the differences between methods are less marked, although in detail there may be some small differences depending on sample size, due to possible degenerate samples. These results testify to the fact that an IM more correlated to the response of the structure essentially induces a lower variability of the estimate of the median of the probit-lognormal model. This is not the case for the log standard deviation whose estimate is affected by samples which are more degenerate with this kind of IM. Such degeneracy affects all methods, however, in all cases, the Jeffreys prior outperforms the classical approaches of the literature both in terms of accuracy and stability (absence of outliers when sampling the *a posteriori* distribution of the parameters).

# B

## Design of experiments for a low fidelity model of seismic fragility curves

This appendix is a postprint of the accepted work: Antoine Van Biesbroeck, Clément Gauchy, Cyril Feau, and Josselin Garnier (2025a). “Design of experiment based on a low fidelity model for seismic fragility estimation”. In: *ESAIM: ProcS* (to appear). URL: <https://hal.science/hal-04719458v1>

**Abstract** In this appendix, we propose an efficient modeling for the estimation of seismic fragility curves in the Bayesian context that is not based on the common probit-lognormal model. We implement instead a low fidelity model of the structure’s response to the ground motion signal and an objective prior. In this work, we do not limit the knowledge about the state of the structure to a binary outcome. The analytical expression of our modeling allows fast generation of estimates. Also, the representative bias arisen by the modeling choice is partly handled with a design of experiments methodology. Finally, our approach is evaluated on a real case study. Our results prove its ability to satisfactorily overcome the irreducible bias when coupled with the design of experiments we propose. However, they also highlight the strong limitations of such simple model, even if it takes richer information about the structural responses as inputs than only binary outcomes.

<b>B.1</b>	<b>Introduction</b>	204
<b>B.2</b>	<b>Low fidelity model for fragility curves</b>	205
B.2.1	Linear regression model	205
B.2.2	Likelihood	205
B.2.3	Prior and posterior	206
<b>B.3</b>	<b>Sensitivity index for design of experiments</b>	207
B.3.1	A review of the global sensitivity analysis	207
B.3.2	Sequential planning of experiments via global sensitivity index maximization	207
<b>B.4</b>	<b>Numerical application</b>	208
B.4.1	Case study presentation	208
B.4.2	Benchmarking metrics	209
B.4.3	Numerical results	210
<b>B.5</b>	<b>Discussion</b>	211
<b>B.6</b>	<b>Conclusion</b>	212
<b>B.7</b>	<b>Details regarding the construction of z</b>	212

## B.1 Introduction

The probabilistic seismic risk assessment framework (SPRA) introduced in the 1980s for the nuclear industry is based on the estimation of seismic fragility curves, for the structures and components (SCs) of interest (Kennedy, Cornell, et al., 1980; Kennedy and Ravindra, 1984; Y.-J. Park, Hofmayer, and Chokshi, 1998; Kennedy, 1999; Cornell, 2004). These curves are defined as the conditional probability that an engineering demand parameter (EDP)—such as the interstory drift ratio—exceeds a limit threshold, given a scalar value derived from the seismic ground motion and called intensity measure (IM). The IM can be for instance the peak ground acceleration (PGA) or a pseudo-spectral acceleration (PSA) evaluated for a given frequency and damping ratio (Ciano, Gioffrè, and Grigoriu, 2020; Saint et al., 2020; Ciano, Gioffrè, and Grigoriu, 2022). As explained in Cornell, 2004, it is therefore assumed that the seismic hazard, on a given site, can be reduced to such a single indicator.

Practitioners have several data sources at their disposal to estimate fragility curves, namely: expert judgments supported by test data (Kennedy, Cornell, et al., 1980; Kennedy and Ravindra, 1984; Y.-J. Park, Hofmayer, and Chokshi, 1998; Zentner, Gündel, and Bonfils, 2017), experimental data (Y.-J. Park, Hofmayer, and Chokshi, 1998; Gardoni, Der Kiureghian, and Mosalam, 2002; Choe, Gardoni, and Rosowsky, 2007), results of damage collected on existing structures that have been subjected to earthquakes (Shinozuka, Feng, et al., 2000; Lallement, Kiremidjian, and Burton, 2015; Straub and Der Kiureghian, 2008) as well as analytical results given by more or less refined numerical models using synthetic or real seismic excitations (Zentner, 2010; F. Wang and Feau, 2020; Mandal, Siddhartha Ghosh, and Pujari, 2016; Z. Wang, Zentner, and Zio, 2018; Z. Wang, Pedroni, et al., 2018; Zhao, Yu, and Mo, 2020). Over the years, many methods have been developed to estimate these curves (Shinozuka, Feng, et al., 2000; Lallement, Kiremidjian, and Burton, 2015; Zentner, Gündel, and Bonfils, 2017). Nowadays, even though machine learning techniques are becoming very popular (J. Park and Towashiraporn, 2014; Seo and Linzell, 2013; Gidaris, Taflanidis, and Mavroeidis, 2015; Z. Wang, Pedroni, et al., 2018; Saint et al., 2020), parametric fragility curves historically introduced in the SPRA framework are ubiquitous in practice and the log-normal model is the most widely used model due to its proven ability to handle limited data (Shinozuka, Feng, et al., 2000; Lallement, Kiremidjian, and Burton, 2015; Straub and Der Kiureghian, 2008; Zentner, 2010; F. Wang and Feau, 2020; Mandal, Siddhartha Ghosh, and Pujari, 2016; Hariri-Ardebili and Saouma, 2016; Z. Wang, Zentner, and Zio, 2018; Z. Wang, Pedroni, et al., 2018; Zhao, Yu, and Mo, 2020; Ellingwood, 2001; S.-H. Kim and Shinozuka, 2004; Mai, Konakli, and Sudret, 2017; Trevlopoulos, Feau, and Zentner, 2019; Katayama et al., 2021).

Different strategies can be implemented to estimate the parameters that define the fragility curve in the log-normal model. Among these we distinguish the Bayesian framework (Gardoni, Der Kiureghian, and Mosalam, 2002; Z. Wang, Pedroni, et al., 2018; Katayama et al., 2021; Koutsourelakis, 2010; Damblin et al., 2014; Tadinada and A. Gupta, 2017; Kwag and A. Gupta, 2018; Jeon et al., 2019; Tabandeh, Asem, and Gardoni, 2020; Lee, Kwag, and Ju, 2023). This framework is interesting because it allows to solve the irregularity issues encountered when few data are available. This occurs with the widely used maximum likelihood estimation coupled with a bootstrap technique to estimate a confidence interval when the data are binary, that is, when they represent the failing or non-failing state of the structure (as in the studies conducted in the part II of this manuscript). In practice, binary data are encountered when dealing with tests performed on shaking tables for instance.

In earthquake engineering, within the SPRA framework, Bayesian inference is often used to update log-normal fragility curves obtained beforehand by various approaches, by assuming independent distributions for the prior values of the parameters, such as log-normal distributions for instance (Tadinada and A. Gupta, 2017; Kwag and A. Gupta, 2018; Z. Wang, Pedroni, et al., 2018; Katayama et al., 2021; Straub and Der Kiureghian, 2008). In this thesis, based on the reference prior theory, we proposed the use of an objective prior, in order to remove any subjectivity that could legitimately lead to inevitable open questions on the influence of the *a priori* on the quantities of interest. In all these approaches, the use of Markov chain Monte Carlo (MCMC) methods is nevertheless necessary to sample the *a posteriori* distribution of the parameters, which can prove cumbersome to implement, particularly if we want a rapid first estimate of a fragility curve with limited data.

We circumvent this problem in this work by proposing an effective approach for the estimation of fragility

curves, which avoids the use of the MCMC method. We rely on a low-fidelity linear model between the logarithm of the Engineering Demand Parameter and the one of the IM (Lallemand, Kiremidjian, and Burton, 2015; Hariri-Ardebili and Saouma, 2016; Zentner, Gündel, and Bonfils, 2017; Swarup Ghosh and Chakraborty, 2020). Supported by the Bayesian framework, our model benefits from a fully analytical form; that former allows an efficient implementation and a solution for fast generation of estimates with limited data. The reliability of the Bayesian scheme w.r.t. its prior choice is answered as well with the derivation of an objective prior derived on the basis of the reference prior theory (see the part I of this manuscript). Finally, since the low-fidelity model is a linear model, we also propose a sequential planning of experiments strategy to minimize the representation bias. The design we suggest relies on the maximization of the information brought by the observation of a new data item onto the posterior distribution. That one is measured through global sensitivity indices described in Da Veiga, 2015.

The remainder of this appendix is organized as follows: the statement of our low-fidelity modeling strategy for the fragility curves estimation in a Bayesian framework is presented in section B.2. After a brief review devoted to global sensitivity analysis, we describe in section B.3 our design for a sequential planning of experiments, taking the global sensitivity indices as a support. Section B.4 is dedicated to the implementation of our methodology on a case study from the nuclear industry. Finally, section B.5 which precedes the conclusion offers a discussion on the performance of our method.

## B.2 Low fidelity model for fragility curves

### B.2.1 Linear regression model

The fragility curve which we seek to estimate is defined by

$$P_f(a) = \mathbb{P}(\text{EDP} > C \mid \text{IM} = a). \quad (\text{B.1})$$

Up to a refinement of a multiplicative constant in the definition of the engineering demand parameter (EDP),  $C$  can be supposed to be equal to 1 in what follows. The EDP is supposed to be correlated with the intensity measure (IM) as follows

$$\log \text{EDP} = \rho \log \text{IM} + \epsilon, \quad (\text{B.2})$$

where the random variable  $\epsilon$  follows the distribution  $\mathcal{N}(\mu, \sigma^2)$ . Here  $\rho$  is supposed unknown, as well as  $\mu$  and  $\sigma$ .

### B.2.2 Likelihood

We have at our disposal a data-set composed by the tuples  $(a_i, y_i)_{i=1}^k$ ,  $a_i \in \mathcal{A} \subset (0, \infty)$  denoting the measured IM from the  $i$ -th seismic ground motion signal and  $y_i \in \mathcal{Y} \subset (0, \infty)$  denoting the structure's EDP. Conditionally to the parameter  $\theta = (\rho, \mu, \sigma)$ , we suppose the observations to be independent and identically distributed. As  $a_i$  follows a distribution assumed to admit a density  $a \mapsto h(a)$  w.r.t. the Lebesgue measure and to be independent of  $\theta$ , the distribution of the  $y_i$  is known conditionally to  $(a_i, \theta)$ :

$$\log y_i | a_i, \theta \sim \mathcal{N}(\rho \log a_i + \mu, \sigma^2). \quad (\text{B.3})$$

The likelihood for the parameter  $\theta$  is therefore:

$$\ell_k^0(\hat{\mathbf{y}}, \hat{\mathbf{a}} | \theta) = \prod_{i=1}^k \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\hat{y}_i - \rho \hat{a}_i - \mu)^2}{2\sigma^2}\right) h(a_i), \quad (\text{B.4})$$

where  $\hat{y}_i$  (resp.  $\hat{a}_i$ ) denotes  $\log y_i$  (resp.  $\log a_i$ ) and  $\hat{\mathbf{y}}$  (resp.  $\hat{\mathbf{a}}$ ) denotes the vector  $(\log y_i)_{i=1}^k$  (resp.  $(\log a_i)_{i=1}^k$ ).

This likelihood introduces a challenge due to the lack of clear separation among the three parameters that constitute  $\theta$ . Within the Bayesian framework, which we develop later, this challenge could result in a posterior distribution that is hardly tractable. We address this issue by introducing the following quantities:

$$\rho_k = \frac{\text{Cov}_k(\hat{\mathbf{y}}, \hat{\mathbf{a}})}{\text{Var}_k \hat{\mathbf{a}}}, \quad \mathbf{z} = \tilde{\mathbf{P}}_{\mathbf{a}}^\top (\hat{\mathbf{y}} - \rho_k \hat{\mathbf{a}}), \quad (\text{B.5})$$

## Appendix B. Design of experiments for a low fidelity model of seismic fragility curves

where  $\text{Cov}_k$ ,  $\text{Var}_k$  respectively denote the empirical covariance and variance, and the matrix  $\tilde{P}_{\mathbf{a}}$  is defined in section B.7. Their conditional distributions are given by:

$$\rho_k | \mathbf{a}, \theta \sim \mathcal{N} \left( \rho, \frac{\sigma^2}{k \text{Var}_k \hat{\mathbf{a}}} \right), \quad (\text{B.6})$$

$$\mathbf{z} | \mathbf{a}, \theta \sim \mathcal{N}(\mu \tilde{P}_{\mathbf{a}}^\top \mathbf{1}, \sigma^2 D); \quad D = \text{diag} \left( 1, \dots, 1, \frac{\hat{\mathbf{a}}^\top \hat{\mathbf{a}}}{k \text{Var}_k \hat{\mathbf{a}}} \right), \quad (\text{B.7})$$

where  $\mathbf{a}$  is the vector  $(a_i)_{i=1}^k$ ,  $\mathbf{1}$  denotes the vector of  $\mathbb{R}^k$  composed only of ones, and  $\text{diag}(\lambda_1, \dots, \lambda_{k-1})$  refers to the diagonal matrix of  $\mathbb{R}^{(k-1) \times (k-1)}$  whose diagonal coefficients are the  $(\lambda_i)_{i=1}^{k-1}$ .

Let us denote by  $(w_i)_{i=1}^k$  the orthogonal columns of the matrix  $P_{\mathbf{a}}$  defined in section B.7. Noticing  $\rho_k = w_k^\top \hat{\mathbf{y}}$  and  $\mathbf{z}$  is spanned by the  $(w_i)_{i=1}^{k-1}$ , we deduce that  $\mathbf{z}$  and  $\rho_k$  are independent conditionally to  $(\mathbf{a}, \theta)$ , and that the knowledge of  $(\mathbf{z}, \rho_k, \mathbf{a})$  is equivalent to the one of  $(\mathbf{y}, \mathbf{a})$  or  $(\hat{\mathbf{y}}, \hat{\mathbf{a}})$ . Thus, the likelihood issued from the observations of  $(\mathbf{z}, \rho_k, \mathbf{a})$  is

$$\ell_k(\mathbf{z}, \rho_k, \mathbf{a} | \theta) = h(\mathbf{a}) \frac{\|\hat{\mathbf{a}} - \bar{\mathbf{a}}\|}{\sqrt{2\pi \hat{\mathbf{a}}^\top \hat{\mathbf{a}} \sigma}} \exp \left( -\frac{(z_{k-1} - \mu \sqrt{k})^2}{2\sigma^2 \frac{\hat{\mathbf{a}}^\top \hat{\mathbf{a}}}{\|\hat{\mathbf{a}} - \bar{\mathbf{a}}\|^2}} \right) \exp \left( -\frac{(\rho_k - \rho)^2}{2\sigma^2 / (k \text{Var}_k \hat{\mathbf{a}})} \right) \prod_{i=1}^{k-2} \frac{1}{\sqrt{2\pi \sigma}} \exp \left( -\frac{z_i^2}{2\sigma^2} \right), \quad (\text{B.8})$$

where  $\bar{\mathbf{a}} = \frac{1}{k} \sum_{i=1}^k \log a_i$ , and  $h(\mathbf{a}) = \prod_{i=1}^k h(a_i)$ .

### B.2.3 Prior and posterior

Within a Bayesian context, the parameter of interest  $\theta$  is itself a random variable taking values in a space  $\Theta \subset \mathbb{R}^2 \times (0, \infty)$  and following a distribution called the prior. We take as a support the reference prior theory (see the part I of this manuscript) to justify the choice of the Jeffreys prior for  $\theta$ , derived from the likelihood expressed in eq. (B.8). Conditionally to  $\mathbf{a}$ , that former is a Gaussian density, making the associated Fisher information matrix being:

$$\mathcal{I}(\theta) = \int_{\mathcal{A}^k} \begin{pmatrix} k \frac{\|\hat{\mathbf{a}} - \bar{\mathbf{a}}\|^2}{\sigma^2 \hat{\mathbf{a}}^\top \hat{\mathbf{a}}} & 0 & 0 \\ 0 & k \frac{2}{\sigma^2} & 0 \\ 0 & 0 & \frac{k \text{Var}_k \hat{\mathbf{a}}}{\sigma^2} \end{pmatrix} \prod_{i=1}^k h(a_i) da_i. \quad (\text{B.9})$$

The Jeffreys' prior being defined as the one whose density  $J$  w.r.t. the Lebesgue measure is proportional to  $\sqrt{|\mathcal{I}(\theta)|}$ , we obtain

$$J(\theta) \propto \frac{1}{\sigma^3}. \quad (\text{B.10})$$

Finally, the posterior distribution of  $\theta$  is given by its density, which is proportional to the product of the likelihood (from eq. (B.8)) with the prior:

$$p(\theta | \mathbf{z}, \mathbf{a}, \rho_k) \propto \frac{1}{\sigma^{k+3}} \exp \left( -\frac{\sum_{i=1}^{k-2} z_i^2}{2\sigma^2} \right) \exp \left( -\frac{k \|\hat{\mathbf{a}} - \bar{\mathbf{a}}\|^2 (z_{k-1} k^{-1/2} - \mu)^2}{2\sigma^2 \hat{\mathbf{a}}^\top \hat{\mathbf{a}}} \right) \exp \left( -\frac{(\rho_k - \rho)^2}{2\sigma^2 / (k \text{Var}_k \hat{\mathbf{a}})} \right). \quad (\text{B.11})$$

We recognize the above as a product of square inverse gamma distributions. More precisely,  $\sigma^{-2}$  follows a gamma distribution, and  $\mu$  and  $\rho$  follow independent Gaussian distributions conditionally to  $\sigma$ .

This posterior allows to elucidate the distribution of what expresses the fragility curve:

$$P_f(a) | \theta \sim \mathbb{P}(\hat{y} > 0 | \hat{a}, \theta) = \Phi \left( \frac{\rho \log a + \mu}{\sigma} \right), \quad (\text{B.12})$$

with  $\Phi$  being the c.d.f. of a standard Gaussian distribution.

Its distribution is known *a posteriori*, given that  $\frac{\rho \log a + \mu}{\sigma}$ , conditionally to  $(a, \mathbf{a}, \mathbf{y})$  for any  $a \in \mathcal{A}$ , is distributed as the sum of a variable with Gaussian distribution and the square root of a variable with Gamma distribution (both variables being independent):

$$\frac{\rho \log a + \mu}{\sigma} | a, \mathbf{a}, \mathbf{y} \sim \mathcal{N} \left( 0, \frac{\log^2 a}{k \text{Var}_k \hat{\mathbf{a}}} + \frac{\hat{\mathbf{a}}^\top \hat{\mathbf{a}}}{k \|\hat{\mathbf{a}} - \bar{\mathbf{a}}\|^2} \right) + \left( \rho_k \log a + \frac{z_{k-1}}{\sqrt{k}} \right) \Gamma^{1/2}(\tilde{c}, \tilde{d}), \quad (\text{B.13})$$

$$\tilde{c} = k/2, \quad \tilde{d} = \frac{1}{2} \sum_{i=1}^{k-2} z_i^2. \quad (\text{B.14})$$

## B.3 Sensitivity index for design of experiments

### B.3.1 A review of the global sensitivity analysis

Global sensitivity analysis (GSA) is a cornerstone of uncertainty quantification studies of computer simulators. It aims at quantifying how the uncertainties within the observed output of a model are influenced by the uncertainties of one or several of its inputs (Iooss and Lemaître, 2015). More formally, in classical GSA settings, a system outputs an observed variable  $Y$ , supposed to be a function of input variables  $Y = \eta(X_1, \dots, X_p)$ , where the input  $X_i$ 's are assumed to follow a known distribution and to be mutually independent. Since the first indices introduced by Sobol' (1993), GSA's tools measure statistically how  $Y$  is impacted by one or some of the  $X_i$  (Da Veiga et al., 2021). Global sensitivity indices (Da Veiga, 2015) are quantities whose class regroups a large range of these tools. According to their definition, letting  $D$  be a dissimilarity measure between probability distributions, the impact of input  $X_i$  onto the output  $Y$  can be derived as

$$S_i = \mathbb{E}_{X_i} [D(\mathbb{P}_Y || \mathbb{P}_{Y|X_i})], \quad (\text{B.15})$$

where  $\mathbb{P}_Y$  is the distribution of  $Y$ ,  $\mathbb{P}_{Y|X_i}$  is the conditional distribution of  $Y$  given  $X_i$ . The choice of  $D$  can depend on the expected properties. A classical example is to set  $D(P||Q) = \|\mathbb{E}_{X \sim P}[X] - \mathbb{E}_{X \sim Q}[X]\|^2$  which gives the un-normalized Sobol' index (Sobol', 1993).

### B.3.2 Sequential planning of experiments via global sensitivity index maximization

Following the idea of Da Veiga (2015), a judicious data acquisition strategy would be to minimize the sensitivity that the posterior would get from the observations. This way, the IM  $a_{k+1}$  that has to be chosen for the next simulation which would output an EDP  $y_{k+1}$ , after having observed  $(\mathbf{y}, \mathbf{a}) = (y_i, a_i)_{i=1}^k$  is the one such that the following index is maximized:

$$\mathbb{E}_{y_{k+1}|a_{k+1}, \mathbf{y}, \mathbf{a}} [D(\mathbb{P}_{\theta|y, \mathbf{a}} || \mathbb{P}_{\theta|y_{k+1}, a_{k+1}, \mathbf{y}, \mathbf{a}})], \quad (\text{B.16})$$

where  $\mathbb{E}_{y_{k+1}|a_{k+1}, \mathbf{y}, \mathbf{a}}$  is the expectation with respect to the distribution of  $y_{k+1}$  given  $a_{k+1}, \mathbf{y}, \mathbf{a}$ . Within the GSA's scope, this makes the next experiment being chosen as the one such that the resulting observation of the structure's response provides the most impact onto the parameter of interest  $\theta$ . Sequentially, the observations are chosen to maximize the evolution of the posterior distribution. We invite one to notice that this viewpoint joins the reference prior theory one. Indeed, the reference prior is built to be the one such that the posterior distribution is expected to evolve the most from the prior.

If the relation between the logarithm of the EDP and that of the IM is "very close" to a linear relation, eq. (B.16) is sufficient to improve the learning of the fragility curve. In other words, a strategy based on this equation makes it possible to sufficiently explore the space of the IMs, in order to maximize their empirical variance and thus reduce the *a posteriori* variance of the estimation of the fragility curve, all things being equal (cf. eq. (B.13)). Note that there is no mathematical proof of this in this appendix, but it has been tested numerically.

In practice, since the linear model is expected to be biased, the way to reduce the bias is to localize the learning, even if it is not optimal with respect to the *a posteriori* variance of the estimation of the fragility

curve. In our work, the locality of interest corresponds to the values of IMs for which the fragility curve evolves “significantly” from 0 to 1. For this reason, we propose a refinement of the data acquisition strategy of eq. (B.16) which includes the researched information (failure or non-failure for a given  $a_{k+1}$ ):

$$\mathbb{E}_{s_{k+1}|a_{k+1}, \mathbf{y}, \mathbf{a}}[D(\mathbb{P}_{\theta|\mathbf{y}, \mathbf{a}} || \mathbb{P}_{\theta|s_{k+1}, a_{k+1}, \mathbf{y}, \mathbf{a}})], \quad (\text{B.17})$$

with  $s_{k+1} = \mathbb{1}_{\hat{y}_{k+1} > 0}$ .

As a dissimilarity measure, we suggest the following, defined as a Sobol’ index of the fragility curve:

$$\begin{aligned} D(\mathbb{P}_{\theta|\mathbf{y}, \mathbf{a}} || \mathbb{P}_{\theta|s_{k+1}, a_{k+1}, \mathbf{y}, \mathbf{a}}) &= \left\| \mathbb{E}[P_f | \mathbf{y}, \mathbf{a}] - \mathbb{E}[P_f | s_{k+1}, a_{k+1}, \mathbf{y}, \mathbf{a}] \right\|_{L^2}^2 \\ &= \int_{\mathcal{A}} \left| \mathbb{E}[P_f(a) | \mathbf{y}, \mathbf{a}] - \mathbb{E}[P_f(a) | s_{k+1}, a_{k+1}, \mathbf{y}, \mathbf{a}] \right|^2 da, \end{aligned} \quad (\text{B.18})$$

where for any  $a \in \mathcal{A}$ ,  $P_f(a) = \Phi((\rho \log a + \mu)/\sigma)$  inherits from the distribution of  $\theta$ . Conditionally to  $(\mathbf{y}, \mathbf{a})$ , its distribution has been elucidated in section B.2.3. Also,

$$p(\theta | s_{k+1}, a_{k+1}, \mathbf{y}, \mathbf{a}) = \frac{p(s_{k+1} | a_{k+1}, \mathbf{y}, \mathbf{a}, \theta)}{\mathbb{E}[p(s_{k+1} | a_{k+1}, \mathbf{y}, \mathbf{a}, \theta) | \mathbf{y}, \mathbf{a}]} p(\theta | \mathbf{y}, \mathbf{a}) \quad (\text{B.19})$$

with  $p(s_{k+1} | a_{k+1}, \mathbf{y}, \mathbf{a}, \theta) = P_f(a_{k+1})^{s_{k+1}} (1 - P_f(a_{k+1}))^{1-s_{k+1}}$ . Thus, samples of  $\theta$  conditionally to  $(\mathbf{y}, \mathbf{a})$  allow the approximation of both expectations in eq. (B.18) by Monte-Carlo averages. The integrals in  $a$  are estimated by Simpson’s rule. In the following example, a regular subdivision of  $\mathcal{A} = [0, A_{\max}]$  is suggested (see section B.4.2).

The calculation of this index necessitates several initial observations. Actually, the derivations of both the likelihood and the posterior, as outlined in sections B.2.2 and B.2.3, require that  $k > 2$  and  $a_1 \neq a_2$  (refer to section B.7). In our experiments, we randomly select  $k_0 = 3$  initial seismic signals with distinct IMs from their original distribution. The planning of experiments strategy is then sequentially implemented to select subsequent IM values by maximizing the numerical approximation of the index expressed in eq. (B.17). The optimization in one dimension is carried out using the BFGS algorithm.

## B.4 Numerical application

### B.4.1 Case study presentation

This case study concerns the seismic behavior of a piping system forming part of the secondary line of a French pressurized water reactor. Figure B.1 presents a perspective of the mock-up positioned on the Azalee shaking table at the EMSI laboratory of CEA/Saclay. Simultaneously, fig. B.1-right depicts the finite element model (FEM), employing beam elements and implemented through the proprietary FE code CAST3M (CEA, 2019). The validation of the FEM was carried out thanks to an experimental campaign described in Touboul, Sollogoub, and Blay, 1999.

The mock-up comprises a carbon steel TU42C pipe with an outer diameter of 114.3 mm, a thickness of 8.56 mm, and a 0.47 elbow characteristic parameter. This pipe, filled with water without pressure, includes three elbows, with a valve-mimicking mass of 120 kg, constituting over 30% of the mock-up’s total mass. One end of the mock-up is clamped, while the other is guided to restrict displacements in the X and Y directions. Additionally, a rod is positioned atop the specimen to limit mass displacements in the Z direction (refer to Figure 1-right). During testing, excitation was applied exclusively in the X direction.

The numerous simulations carried out for this case study were obtained with artificial seismic signals generated with the stochastic generator proposed by Rezaeian (2010). This generator implemented in Saint et al., 2020 was calibrated from 97 real accelerograms selected in the European Strong Motion Database for a magnitude  $M$  such that  $5.5 \leq M \leq 6.5$ , and a source-to-site distance  $R < 20$  km (Ambraseys et al., 2000). Note that enrichment is not a necessity in the Bayesian framework —especially if a sufficient number of real signals is available— but it allows comparative performance studies, such as those presented in this work.

As in practice the piping system is located in a building, the artificial signals were filtered using a fictitious 2% damped linear single-mode building at 5 Hz, which corresponds to the first eigenfrequency of the 1% damped piping system. The chosen failure criterion is based on the assessment of excessive out-of-plane rotation of the elbow near the clamped end of the mock-up, following the recommendation in Touboul, Blay, et al., 2006. The chosen IM is the PSA which is calculated here at 5 Hz for a damping ratio of 1%.

In order to evaluate the effectiveness of the proposed method, we considered the nonlinear seismic behavior of the piping system. Regarding the nonlinear constitutive law of the material, a bilinear law exhibiting kinematic hardening was used to reproduce the overall nonlinear behavior of the mock-up with satisfactory agreement compared to the results of the seismic tests (Touboul, Sollogoub, and Blay, 1999).

In this work, the critical rotation threshold is set at  $C = 4.1^\circ$ , representing the 90%-level quantile derived from a sample of  $10^4$  nonlinear numerical simulations.

Finally, the fragility curve that we will call “reference” in the following was obtained by Monte-Carlo averages on clusters of the IM using the K-means algorithm, following the suggestion of Trevlopoulos, Feau, and Zentner (2019), from the  $10^4$  data that we dispose. The estimation procedure is presented in chapter 7. In this method, the average goes along its confidence intervals; in the computation of the metrics that are suggested in the next section, the average is considered as the reference.

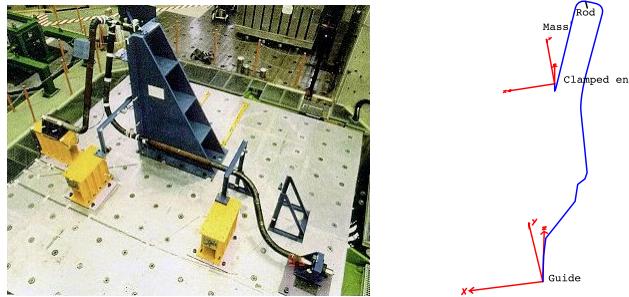


Figure B.1: (left) Overview of the piping system on the Azalee shaking table and (right) Mock-up FEM.

## B.4.2 Benchmarking metrics

In order to evaluate the effects that our planning of experiments method has over the fragility curve estimates, we consider four quantitative metrics described hereafter. Those are later implemented on the *a posteriori* estimates conditional to two types of dataset: one derived from our planning of experiments methodology, and one without. Considering a sample  $(\mathbf{a}, \mathbf{y})$ , we denote by  $a \mapsto P_f^{|\mathbf{a}, \mathbf{y}}(a)$  the random process defined as the fragility curve conditionally to the sample.  $P_f^{|\mathbf{a}, \mathbf{y}}(a) = \Phi((\rho \log a + \mu)/\sigma)$  inherits from the *a posteriori* distribution of  $\theta$ . For each value  $a$  the  $r$ -quantile of the random variable  $P_f^{|\mathbf{a}, \mathbf{y}}(a)$  is denoted by  $q_r^{|\mathbf{a}, \mathbf{y}}(a)$ , and its median is denoted by  $m^{|\mathbf{a}, \mathbf{y}}(a)$ . Also, we take into account the reference fragility curve  $a \mapsto P_f^{\text{ref}}(a)$ , evoked in section B.4.1; and we consider a bounded set  $\mathcal{A} = [0, A_{\max}]$  for the IM, the truncation is set to the maximal IM within the database of  $10^4$  seismic signals we have at disposal for this work, as disclosed in section B.4.1. We define:

- The square bias to the median:  $\mathcal{B}^{|\mathbf{a}, \mathbf{y}} = \|m^{|\mathbf{a}, \mathbf{y}} - P_f^{\text{ref}}\|_{L^2}^2$ ; where  $\|P\|_{L^2}^2 = \frac{1}{A_{\max}} \int_0^{A_{\max}} P(a)^2 da$ .
- The quadratic error:  $\mathcal{E}^{|\mathbf{a}, \mathbf{y}} = \mathbb{E} \left[ \|P_f^{|\mathbf{a}, \mathbf{y}} - P_f^{\text{ref}}\|_{L^2}^2 | \mathbf{a}, \mathbf{y} \right]$ .
- The  $1 - r$ -square credibility width:  $\mathcal{W}^{|\mathbf{a}, \mathbf{y}} = \|q_{1-r/2}^{|\mathbf{a}, \mathbf{y}} - q_{r/2}^{|\mathbf{a}, \mathbf{y}}\|_{L^2}^2$ .
- The  $1 - r$ -coverage probability:  $\mathcal{P}^{|\mathbf{a}, \mathbf{y}} = \frac{1}{A_{\max}} \int_0^{A_{\max}} \mathbb{1}_{P_f^{\text{ref}}(a) \in [q_{1-r/2}^{|\mathbf{a}, \mathbf{y}}(a), q_{r/2}^{|\mathbf{a}, \mathbf{y}}(a)]} da$ .

For the forthcoming implementation of these metrics, numerous *a posteriori* samples of the process  $P_f^{|\mathbf{a},\mathbf{y}}$  are generated from their known distribution (see eq. (B.11)) and serve the computation of the medians, quantiles and means through Monte-Carlo derivations. The integrals are approximated numerically from Simpsons' interpolation on sub-intervals of regular size  $0 = A_0 < \dots < A_p = A_{\max}$ . In our computations, we use  $A_{\max} = 55 \text{ m/s}^2$ , and  $p = 200$ .

### B.4.3 Numerical results

Figure B.2 shows examples of fragility curve estimations based on different dataset sizes. The results presented in fig. B.2-(c) come from our planning-of-experiments (PE) methodology while the results presented in figs. B.2-(a) and B.2-(b) come from independent random samples, with IMs that have been drawn w.r.t. their original standard distribution or w.r.t. a uniform distribution. These qualitative results clearly illustrate the contribution of our methodology.

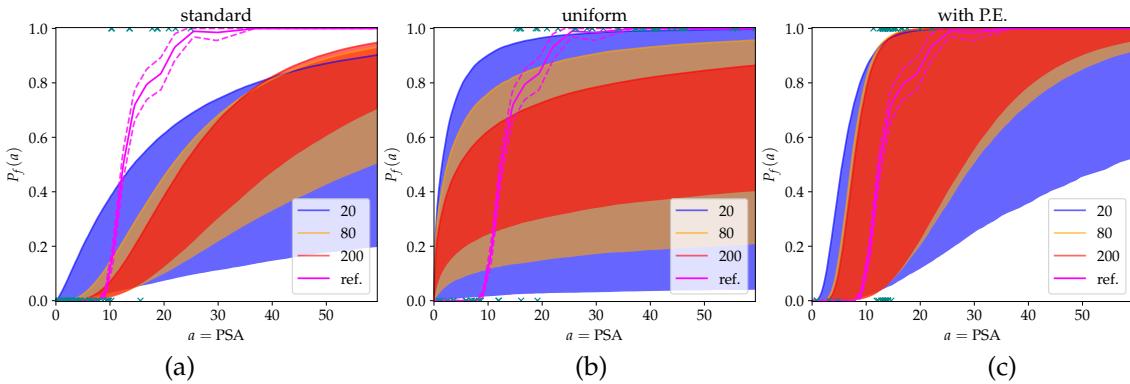


Figure B.2: Examples of fragility curve estimations for different number of observations, with the PSA considered as IM. Are plotted the 95% credibility interval for 3 datasets of respective sizes 20 (blue), 80 (orange), and 200 (red); the reference curve  $P_f^{\text{ref}}$  is drawn in magenta and is accompanied with the confidence interval of the procedure in dashed lines. The observations are chosen (a) w.r.t. the standard distribution of the IM; or (b) w.r.t. a uniform distribution on  $[0, A_{\max}]$ ; or (c) using our planning of experiments method. The green crosses represent 50 pairs  $(a_i, \mathbb{1}_{y_i > C})$  drawn for each method.

When samples are randomly drawn from the original IM distribution or from a uniform distribution, the results show a rapid convergence of the estimates —in the sense that the associated credibility intervals decrease rapidly— towards biased estimations of the fragility curves. Conversely, when the samples come from the PE methodology, the bias decreases but the convergence is slower.

These observations are confirmed on a larger scale by the results presented in figs. B.3 and B.4. These are issued from computations of the metrics  $\mathcal{B}^{|\mathbf{a},\mathbf{y}}$ ,  $\mathcal{E}^{|\mathbf{a},\mathbf{y}}$ ,  $\mathcal{W}^{|\mathbf{a},\mathbf{y}}$  and  $\mathcal{P}^{|\mathbf{a},\mathbf{y}}$  described in section B.4.2, for various observation sets  $(\mathbf{a}, \mathbf{y})$ . They compare the performances of our PE methodology with the two methods that are based on independently drawn observations: the one that involves IM samples drawn from their standard distribution, and the one that involves IM samples drawn from a uniform distribution over the range  $[0, A_{\max}]$ .

Figure B.3 shows empirical comparisons of the bias and of the quadratic error between the method involving a design of experiments and without. These two results clearly illustrate that the PE approach outperforms the standard and the uniform approaches. Although the quadratic error is strongly related to the variance of the estimates it is significantly offset by the fact that the bias is smaller with the PE approach. Indeed, fig. B.4-left illustrates that the 95%-square credibility width is smaller with the standard and uniform approaches than with the PE-based approach. This good result nevertheless masks a lack of robustness of the standard or uniform approaches since the estimate turns out to be strongly biased, as shown in fig. B.4-right. This figure shows indeed the coverage probability for the both methods, as a function of the dataset size. It measures the average inclusion of reference fragility curves in posterior credibility intervals.

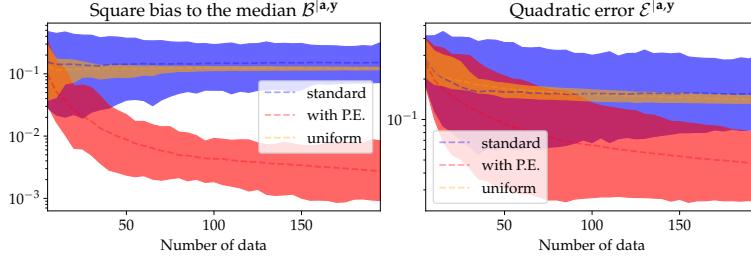


Figure B.3: Confidence intervals and means w.r.t. the  $(\mathbf{a}, \mathbf{y})$  for (left) the square bias to the median  $\mathcal{B}^{|\mathbf{a}, \mathbf{y}|}$  and (right) the quadratic error  $\mathcal{E}^{|\mathbf{a}, \mathbf{y}|}$ ; as a function of the number of observations. For each value of  $k = 5, 10, \dots, 200$ , a number of  $L = 200$  dataset have been drawn following the standard distribution of the IM firstly (for the blue curves), following a uniform distribution on  $[0, A_{\max}]$  secondly (for the orange curves), and following the planning of experiments method thirdly (for the red curves).

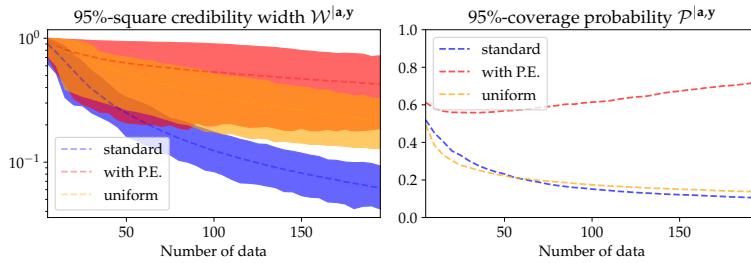


Figure B.4: (left) 95%-confidence intervals and means w.r.t.  $(\mathbf{a}, \mathbf{y})$  for the 95%-square credibility width  $\mathcal{W}^{|\mathbf{a}, \mathbf{y}|}$ , as a function of the number of observations. (right) mean w.r.t.  $(\mathbf{a}, \mathbf{y})$  for the 95%-coverage probability  $\mathcal{P}^{|\mathbf{a}, \mathbf{y}|}$ , as a function of the number of observations. For each value of  $k = 5, 10, \dots, 200$ ,  $L = 200$  datasets have been drawn following the standard distribution of the IM firstly (for the blue curves), following a uniform distribution on  $[0, A_{\max}]$  secondly (for the orange curves), and following the planning of experiments method thirdly (for the red curves).

## B.5 Discussion

The results presented in the previous section clearly illustrate the superiority of the PE-based approach over the standard and uniform approaches. Similar results, not presented here for the sake of brevity, and obtained with the PGA as IM, as well as with other types of structures, also confirm these results.

The strength of the approach proposed in this work is its completely analytical nature, which avoids the use of MCMC methods for the *a posteriori* estimation of fragility curves. To do this, however, it is necessary to assume that the logarithm of the EDP evolves linearly as a function of the logarithm of the IM.

So, to solve this problem, we are faced with a contradiction. In order to satisfy the linearity assumption, it is necessary, on the one hand, that the learning zone is local, that is to say restricted to the vicinity of the IMs for which the fragility curve evolves significantly from 0 to 1. On the other hand, eq. (B.13) shows that an empirical variance of the IMs that is too small is penalizing from the point of view of the variance of the estimation of the fragility curves. The variance of the latter is in fact inversely proportional to that of the IMs considered for learning.

As the numerical results show, the proposed learning method localizes the learning in the area of interest and significantly reduces the model bias. As a result, this is accompanied by a slight reduction in the size of the credibility interval with the number of training data.

This therefore suggests that the proposed method is effective for samples with limited size (for instance, given the results presented in section B.4.3, an appropriate limit could be a sample size smaller than 100 for the case study treated in this appendix). Beyond that, given the cost associated with each training data, it seems preferable to move towards less constrained and more sophisticated methods.

## B.6 Conclusion

Assessing the seismic fragility of structures and components when few data are available is a challenging task and the Bayesian framework is known to be effective for these types of problems.

In this work we proposed an efficient Bayesian methodology whose strength lies in its fully analytical nature, which avoids the use of MCMC methods for the *a posteriori* estimation of fragility curves. The effectiveness of the method comes from the assumption of linearity between the logarithm of the EDP and that of the IM of interest. As this hypothesis implies a model bias in most practical cases, we proposed a strategy in order to minimize this bias, by concentrating the learning in the vicinity of the IMs for which the fragility curve evolves significantly from 0 to 1.

The numerical results clearly illustrate the superiority of the proposed approach over an approach without a learning strategy. They emphasize the robustness of a design of experiments which is based on a sensitivity analysis of the posterior distribution. Such construction is not limited to the modeling we derive in this work in particular, and could still be adapted to another to increase its learning abilities. They also suggest that the proposed method is effective for a limited sample size (about 100 in our settings). Beyond that, given the cost associated with each training data, it seems preferable to move towards less constrained and more sophisticated methods, in order to more effectively minimize both the biases and the variance of the estimates.

For practitioners, this method therefore constitutes a rapid and robust tool for first estimates of fragility curves in a context where the datasets are of limited size.

## B.7 Details regarding the construction of $\mathbf{z}$

Conditionally to  $(\mathbf{a}, \theta)$ , we derive the distribution of  $\hat{\mathbf{y}} - \rho_k \hat{\mathbf{a}}$ :

$$\hat{\mathbf{y}} - \rho_k \hat{\mathbf{a}} | \mathbf{a}, \theta \sim \mathcal{N}(\mu \mathbf{1}, \sigma^2 U_{\mathbf{a}}), \quad \text{with } U_{\mathbf{a}} = I - \frac{\hat{\mathbf{a}} (\frac{1}{2} \hat{\mathbf{a}} - \bar{\mathbf{a}})^\top}{k \text{Var}_k \hat{\mathbf{a}}} - \frac{(\frac{1}{2} \hat{\mathbf{a}} - \bar{\mathbf{a}}) \hat{\mathbf{a}}^\top}{k \text{Var}_k \hat{\mathbf{a}}}, \quad \bar{\mathbf{a}} = \frac{1}{k} \sum_{i=1}^k \log a_i, \quad (\text{B.20})$$

where  $\mathbf{1}$  denotes the vector of  $\mathbb{R}^k$  which contains only ones. Below is suggested a diagonalization of the matrix  $U_{\mathbf{a}}$ : we define  $P_{\mathbf{a}}$  such that  $P_{\mathbf{a}}^\top P_{\mathbf{a}} = I$  and  $P_{\mathbf{a}}^\top U_{\mathbf{a}} P_{\mathbf{a}}$  is a diagonal matrix. To define  $\mathbf{z}$ , we denote by  $\tilde{P}_{\mathbf{a}}$  the matrix in  $\mathbb{R}^{k \times (k-1)}$  composed by the  $k-1$  first columns of  $P_{\mathbf{a}}$ . In what follows, we assume that  $k > 2$  and that the coordinates of  $\mathbf{a}$  are not all identical.

The matrix  $U_{\mathbf{a}}$  takes the form of  $I - uv^\top - vu^\top$  for some vectors  $u$  and  $v$  of  $\mathbb{R}^k$ , which are linearly independent. It is clear that the diagonalization of  $U_{\mathbf{a}}$  is linked with the one of  $V = uv^\top + vu^\top$ .

First of all notice that  $v^\perp$  and  $u^\perp$  are two different hyperplanes because of the linear independence of  $u$  and  $v$ . That makes  $u^\perp \cap v^\perp$  a subspace of dimension  $k-2$ . Therefore, as one would notice that  $u^\perp \cap v^\perp \subset \ker V$  and  $\text{im } V \subset \text{Span}(u, v)$ , the converse inclusions stand.

This way, while 0 is the first eigenvalue of  $V$  with rank  $k-2$ , an other eigenvalue  $r$  must admit eigenvectors in  $\text{Span}(u, v)$ , which should make the system

$$\begin{cases} r\gamma &= \gamma v^\top u + \delta v^\top v \\ r\delta &= \gamma u^\top u + \delta u^\top v \end{cases} \quad (\text{B.21})$$

admitting an infinity of solutions w.r.t.  $(\gamma, \delta)$ . Equivalently, its determinant must be null which lead to the two solutions

$$r = v^\top u \pm \|v\| \|u\|. \quad (\text{B.22})$$

As  $u$  and  $v$  are linearly independent, the equation above defines two different eigenvalues  $r_+$  and  $r_-$ , both of rank 1. Let  $r$  be one of those, a resolution of the equation system (B.21) gives that the eigenspace associated with  $r$  is  $\text{Span}(v^\top vu + (r - v^\top u)v)$ .

Coming back to  $U_{\mathbf{a}}$ , the arguments above show that the eigenvalues of  $U_{\mathbf{a}}$  are  $1$ ,  $1 - r_-$  and  $1 - r_+$  with respective ranks  $k - 2$ ,  $1$  and  $1$ , and with:

$$r_+ = 1, \quad r_- = \frac{-k\hat{\mathbf{a}}^2}{\|\hat{\mathbf{a}} - \bar{\mathbf{a}}\|^2}, \quad (\text{B.23})$$

because

$$\|\hat{\mathbf{a}}\|^2 \left\| \frac{1}{2} \hat{\mathbf{a}} - \bar{\mathbf{a}} \right\|^2 = \sum_{i=1}^k \hat{a}_i^2 \left( \frac{1}{4} \sum_{i=1}^k \hat{a}_i^2 + \left( \sum_{i=1}^k \hat{a}_i^2 \right) \frac{1}{k} \left( \sum_{i=1}^k \hat{a}_i \right)^2 - \left( \sum_{i=1}^k \hat{a}_i^2 \right) \frac{1}{k} \left( \sum_{i=1}^k \hat{a}_i \right)^2 \right) = \frac{1}{4} \left( \sum_{i=1}^k \hat{a}_i^2 \right)^2. \quad (\text{B.24})$$

Now, let us choose  $w_1, \dots, w_{k-2}$  an orthonormal basis of  $\hat{\mathbf{a}}^\perp \cap (\frac{1}{2}\hat{\mathbf{a}} - \bar{\mathbf{a}})^\perp$ . Let us define

$$w_{k-1} = \frac{1}{\sqrt{k}} \mathbf{1}, \quad w_k = \frac{\hat{\mathbf{a}} - \bar{\mathbf{a}}}{\|\hat{\mathbf{a}} - \bar{\mathbf{a}}\|}. \quad (\text{B.25})$$

We remind  $\mathbf{1}$  is the vector whose coordinates are ones. Therefore, denoting  $P_{\mathbf{a}}$  the matrix whose columns are the  $w_i$ ,  $i = 1, \dots, k$ ; it comes  $P_{\mathbf{a}}^\top P_{\mathbf{a}} = I$  and

$$P_{\mathbf{a}}^\top U_{\mathbf{a}} P_{\mathbf{a}} = \text{diag} \left( 1, \dots, 1, \frac{\hat{\mathbf{a}}^\top \hat{\mathbf{a}}}{\|\hat{\mathbf{a}} - \bar{\mathbf{a}}\|^2}, 0 \right). \quad (\text{B.26})$$

The complete definition of  $P_{\mathbf{a}}$  depends on the chosen orthonormal basis  $w_1, \dots, w_{k-2}$  of  $\hat{\mathbf{a}}^\perp \cap (\frac{1}{2}\hat{\mathbf{a}} - \bar{\mathbf{a}})^\perp$ . In practice, we proceed as follows to construct it, starting from  $w_{k-1} = (w_{k-1}^{(j)})_{j=1}^k$  and  $w_k = (w_k^{(j)})_{j=1}^k$  as defined in eq. (B.25). We denote by  $e_i$  the canonical vectors of  $\mathbb{R}^k$  (the  $j$ th coordinate of  $e_i$  is equal to  $0$  iff  $j \neq i$ ). As the coordinates of  $\mathbf{a}$  are not all the same, there exist  $j, p$  such that  $w_k^{(j)} \neq w_k^{(p)}$  (in practice, we select the minimal  $j$  and the minimal  $p$  such that this property is verified). Thus, we can show that the vectors  $w_k, w_{k-1}, (e_i)_{i \neq j, p}$  form a basis of  $\mathbb{R}^k$  by computing their determinant:

$$\det(w_k, w_{k-1}, (e_i)_{i \neq j, p}) = w_k^{(j)} w_{k-1}^{(p)} (-1)^{j+p+1} - w_{k-1}^{(j)} w_k^{(p)} (-1)^{j+p+1} \neq 0, \quad (\text{B.27})$$

as  $w_{k-1}^{(j)} = w_{k-1}^{(p)}$ . Eventually, the family of the  $w_i$ ,  $i = 1, \dots, k$  is the result of the Gram-Schmidt process applied to  $u_k, \dots, u_1 = w_k, w_{k-1}, (e_i)_{i \neq j, p}$ :

$$w_{k-i} = \frac{\tilde{w}_{k-i}}{\|\tilde{w}_{k-i}\|}, \quad \tilde{w}_{k-i} = u_{k-i} - \sum_{j < i} w_{k-j}^\top u_{k-i} \cdot w_{k-j}, \quad (\text{B.28})$$

Note that this process leaves the expressions of  $w_k$  and  $w_{k-1}$  unchanged. The family  $w_1, \dots, w_{k-2}$  thus forms a basis of  $\text{Span}(w_k, w_{k-1})^\perp = \hat{\mathbf{a}}^\perp \cap (\frac{1}{2}\hat{\mathbf{a}} - \bar{\mathbf{a}})^\perp$ .

We invite the reader to notice that in any way, the construction of those  $k - 2$  first columns of  $P_{\mathbf{a}}$  has no influence on the resulting posterior distribution of interest (given by eq. (B.11)). Indeed, that latter only involves the expressions of  $z_{k-1}$  and of  $\sum_{i=1}^{k-2} z_i^2$ . Concerning the first one, it is equal to  $w_{k-1}^\top (\hat{\mathbf{y}} - \rho_k \hat{\mathbf{a}})$ , and concerning the second one, it is equal to:

$$\begin{aligned} \sum_{i=1}^{k-2} z_i^2 &= \sum_{i=1}^{k-2} |w_i^\top (\hat{\mathbf{y}} - \rho_k \hat{\mathbf{a}})|^2 = \sum_{i=1}^k |w_i^\top (\hat{\mathbf{y}} - \rho_k \hat{\mathbf{a}})|^2 - |w_{k-1}^\top (\hat{\mathbf{y}} - \rho_k \hat{\mathbf{a}})|^2 - |w_k^\top (\hat{\mathbf{y}} - \rho_k \hat{\mathbf{a}})|^2 \\ &= \|\hat{\mathbf{y}} - \rho_k \hat{\mathbf{a}}\|^2 - |w_{k-1}^\top (\hat{\mathbf{y}} - \rho_k \hat{\mathbf{a}})|^2 - |w_k^\top (\hat{\mathbf{y}} - \rho_k \hat{\mathbf{a}})|^2. \end{aligned}$$



# C

## Non-asymptotic reference priors, a brief study

This appendix compiles some theoretical developments that were conducted during my end-of-study internship that took place at CEA Saclay in 2021.

**Abstract** This short appendix complements [chapter 3](#) as it provides a further study of the original definition of the mutual information. In this work, we are interested about the maximization of the mutual information when the number of observations is fixed. Our results take the form of implicit formulations of what we call “non-asymptotic reference priors”. Two case studies are considered for deriving those, one involving the least possible assumption about the prior, and the other incorporating linear constraints to the framework. This study provides further insights on the expression of the reference priors as well as the optimal characteristic of the Jeffreys prior in the theory.

<a href="#">C.1 Motivations and context</a> . . . . .	<a href="#">215</a>
<a href="#">C.2 Bayesian framework and mutual information</a> . . . . .	<a href="#">216</a>
<a href="#">C.3 Non-asymptotic reference priors</a> . . . . .	<a href="#">216</a>
<a href="#">C.4 Discussion and link with asymptotic reference priors</a> . . . . .	<a href="#">219</a>

### C.1 Motivations and context

Reference prior theory has been built on the idea of constructing priors whose influence onto the posterior would be minimized, in order to let the latter being informed by the data in priority. The aim of that idea is to qualify the resulting prior as “objective”.

To achieve that goal, instead of seeking to maximize the expected divergence between the prior and the posterior (i.e., the mutual information), the authors who developed the theory (Bernardo, [1979b](#); Bernardo and Smith, [1994](#)) suggested maximizing the asymptotic value of that quantity as the number of observations tends to infinity (see [chapter 3](#) for a complete review of the theory). Indeed, it is seen as an issue that the mutual information—and its maximal argument—depends on the number of observations  $k$ , since the “objective” prior should, by nature, be adequate for any data samples, no matter their size. According to Bernardo and Smith ([1994](#)), when the number of observations  $k$  is fixed, the distribution of the vector of the data is not fully described. They advance that, denoting  $\mathcal{I}^k$  the mutual information when  $k$  data items are observed, the limit  $\mathcal{I}^\infty$  (if it exists) measures the knowledge missing from the prior.

However, there exist few elements that support formally this asymptotic definition of the reference prior in the literature. While there is consensus on the definition of the reference prior proposed by Bernardo ([1979b](#))

that consists in asymptotically maximizing the mutual information, we are interested in the expression of the prior that maximizes it non-asymptotically. This study is motivated by the two following thoughts: (i) in a case where the number of observations represents a fundamental element of the problem, a prior that takes it into account could be valued; (ii) expressing the non-asymptotic reference prior can help to understand or to express the asymptotic reference priors (i.e., the reference priors as they are originally defined).

We suggest in this short appendix two main results that express implicitly the non-asymptotic reference priors. First, we recall briefly the framework of the reference prior theory in the next section, then our results are developed in section C.3. Finally, we discuss our results in section C.4, and we elucidate their link with classical asymptotic reference priors.

## C.2 Bayesian framework and mutual information

We remind that the original framework of the reference prior theory is comprehensively detailed in chapter 3. We consider a Bayesian framework: observations  $\mathbf{Y}_k \in \mathcal{Y}^k$  follow conditionally to  $T = \theta \in \Theta$  the distribution  $\mathbb{P}_{\mathbf{Y}_k|\theta} = \mathbb{P}_{Y|\theta}^{\otimes k}$ . The marginal distribution is denoted  $\mathbb{P}_{\mathbf{Y}_k}$ , the prior distribution  $\Pi$ . We suppose that the model admits a likelihood denoted  $\ell_k$  w.r.t. a measure  $\mu$  on  $\mathcal{Y}$ . We also suppose that the prior admits a density  $\pi$  w.r.t. a measure  $\nu$  on  $\Theta$ . The marginal density is denoted  $p_{\mathbf{Y}_k}$ . Finally, given a sample of observations  $\mathbf{y} \in \mathcal{Y}^k$ , the posterior distribution (resp. density) is denoted  $\mathbb{P}_{T|\mathbf{y}}$  (resp.  $p(\cdot|\mathbf{y})$ ).

Under those settings, if  $\Pi$  is proper, the mutual information given  $k$  observations is defined as

$$\mathcal{I}^k(\Pi) = \mathbb{E}_{\mathbf{Y}_k \sim \mathbb{P}_{\mathbf{Y}_k}} [\text{KL}(\mathbb{P}_{T|\mathbf{Y}_k} || \Pi)], \quad \text{with} \quad \text{KL}(\mathbb{P}_{T|\mathbf{Y}_k} || \Pi) = \int_{\Theta} \log \left( \frac{p(\theta|\mathbf{Y}_k)}{\pi(\theta)} \right) p(\theta|\mathbf{Y}_k) d\nu(\theta). \quad (\text{C.1})$$

## C.3 Non-asymptotic reference priors

In this appendix, we propose to define a non-asymptotic reference prior as a maximal argument of the mutual information, as expressed in the definition below.

**Definition C.1** (Non-asymptotic reference prior). Let  $\mathcal{P} \subset \mathcal{M}^\nu$  be a set of proper priors. A prior  $\Pi \in \mathcal{P}$  is called a non-asymptotic reference prior over  $\mathcal{P}$  if

$$\Pi \in \arg \max_{P \in \mathcal{P}} \mathcal{I}(P) \quad (\text{C.2})$$

where  $\bar{P}$  denotes  $P(\cdot|\Theta)$ , i.e., the renormalized probability distribution associated to  $P$ .

$\Pi$  is said to be unique if for any other non-asymptotic reference prior  $\Pi'$  over  $\mathcal{P}$ ,  $\Pi \simeq \Pi'$ .

As a comparison with the original definition of reference priors (which we refer as asymptotic reference priors in this appendix), a non-asymptotic reference prior must be proper. Indeed, the proper characteristic of  $\Pi$  is essential for the mutual information to be well-defined. We recall that, in this thesis, a proper prior is a measure that has a finite total mass, which does not necessarily equal 1 (see chapter 3 (section 3.4.1) for more details).

The result below is a contribution of ourselves. It gives an implicit expression of the non-asymptotic reference prior among a non-restrictive set of priors

**Theorem C.1.** *We assume that  $\Theta$  is a compact subset of  $\mathbb{R}^d$ , and that there exist  $0 < l_1 < l_2$  such that  $\ell_k(\mathbf{y}|\theta) \in [l_1, l_2]$  for any  $\mathbf{y} \in \mathcal{Y}^k$ ,  $\theta \in \Theta$ . Let  $\mathcal{P} = \{\Pi \in \mathcal{M}_C^\nu, \Pi > 0, \Pi(\Theta) < \infty\}$  the set of continuous and positive proper priors. Then there exists a unique non-asymptotic reference prior over  $\mathcal{P}$ . Its density  $\pi$  verifies*

$$\pi \propto f_\pi \quad \text{with} \quad f_\pi(\theta) = \exp \left( \int_{\mathcal{Y}^k} \ell_k(\mathbf{y}|\theta) \log \frac{\ell_k(\mathbf{y}|\theta)\pi(\theta)}{\int_{\Theta} \ell_k(\mathbf{y}|\theta')\pi(\theta')d\nu(\theta')} d\mu^{\otimes k}(\mathbf{y}) \right). \quad (\text{C.3})$$

Following the idea provided in chapter 5, we suggest also the study of non-asymptotic reference prior under constraints. The result below considers constraints that take the form of linear constraints.

**Theorem C.2.** Under the assumptions of theorem C.1, consider  $(g_j)_{j=1}^q$  a family of functions in  $\mathcal{R}_{\mathcal{C}^b}$  that is such that  $g_0, \dots, g_q$  are linearly independent, where  $g_0 := \theta \mapsto 1$ . Let  $c_1, \dots, c_q \in \mathbb{R}$ , and define  $\mathcal{P}' = \{\Pi \in \mathcal{P}, \forall j, \int_{\Theta} g_j d\Pi = c_j\}$ . Then there exists a unique non-asymptotic reference prior over  $\mathcal{P}'$ . Its density  $\tilde{\pi}$  verifies

$$\tilde{\pi} \propto f_{\tilde{\pi}} \cdot \exp \left( \sum_{j=1}^q \lambda_j g_j \right), \quad (\text{C.4})$$

for some  $\lambda_1, \dots, \lambda_q \in \mathbb{R}$ .

*Proof of theorem C.1.*

Let  $P \in \mathcal{P}$ , we denote  $p$  its renormalized density, i.e.,  $P(B|\Theta) = \int_B p(\theta) d\nu(\theta)$ . The quantity  $\mathcal{J}^k(\bar{P})$  can be written as a function of  $p$ :

$$\mathcal{J}^k(\bar{P}) = \hat{I}(p) := \int_{\mathcal{Y}^k} \int_{\Theta} \ell_k(\mathbf{y}|\theta) p(\theta) \log \frac{\ell_k(\mathbf{y}|\theta)}{\int_{\Theta} \ell_k(\mathbf{y}|\theta') p(\theta') d\nu(\theta')} d\nu(\theta) d\mu^{\otimes k}(\mathbf{y}) \quad (\text{C.5})$$

Therefore maximizing  $P \mapsto \mathcal{J}^k(\bar{P})$  over  $\mathcal{P}$  is equivalent to maximize  $\hat{I}$  over  $\mathcal{R}_{\mathcal{C}^b}^* \cap \{p, \int_{\Theta} p d\nu = 1\}$ , where  $\mathcal{R}_{\mathcal{C}^b}^* = \{p \in \mathcal{R}_{\mathcal{C}^b}, p > 0\}$ . In addition,  $\hat{I}$  can be expressed as  $\hat{I} = \hat{I}_1 + \hat{I}_2$  where

$$\begin{aligned} \hat{I}_1(p) &= \int_{\mathcal{Y}^k} \int_{\Theta} \ell_k(\mathbf{y}|\theta) p(\theta) \log \ell_k(\mathbf{y}|\theta) d\nu(\theta) d\mu^{\otimes k}(\mathbf{y}) \\ \hat{I}_2(p) &= \int_{\mathcal{Y}^k} \psi \left( \int_{\Theta} \ell_k(\mathbf{y}|\theta) p(\theta) d\nu(\theta) \right) d\mu^{\otimes k}(\mathbf{y}) \end{aligned}$$

with  $\psi : x \mapsto -x \log(x)$ , which is a strictly concave function. Reminding that  $\mathcal{R}_{\mathcal{C}^b}^*$  is convex and considering  $p \neq q \in \mathcal{R}_{\mathcal{C}^b}^*, \lambda \in ]0, 1[$ , we have:

$$\begin{aligned} \hat{I}_2(\lambda p + (1-\lambda)q) &= \int_{\mathcal{Y}^k} \psi \left( \lambda \int_{\Theta} \ell_k(\mathbf{y}|\theta) p(\theta) d\nu(\theta) + (1-\lambda) \int_{\Theta} \ell_k(\mathbf{y}|\theta) q(\theta) d\nu(\theta) \right) d\mu^{\otimes k}(\mathbf{y}) \\ &> \int_{\mathcal{Y}^k} \left[ \lambda \psi \left( \int_{\Theta} \ell_k(\mathbf{y}|\theta) p(\theta) d\nu(\theta) \right) + (1-\lambda) \psi \left( \int_{\Theta} \ell_k(\mathbf{y}|\theta) q(\theta) d\nu(\theta) \right) \right] d\mu^{\otimes k}(\mathbf{y}) \quad (\text{C.6}) \\ &> \lambda \hat{I}_2(p) + (1-\lambda) \hat{I}_2(q), \end{aligned}$$

hence  $\hat{I}_2$  is strictly concave. Since  $\hat{I}_1$  is a linear function, we deduce that  $\hat{I}$  is strictly concave on  $\mathcal{R}_{\mathcal{C}^b}^*$ .

We call  $E$  the set of  $\nu$ -a.e. continuous and locally bounded functions. Note that  $E$  provided with the supremum norm  $\|f\| = \sup_{\Theta} |f|$  is a Banach set whose  $U = \mathcal{R}_{\mathcal{C}^b}^* \cap \{p, \int_{\Theta} p d\nu \in (1-\xi, 1+\xi)\} \xi \in (0, 1)$  is an open and convex subset as  $\Theta$  is supposed compact. Next part of the proof will be dedicated to show that  $\hat{I}$  is differentiable on  $U$ .

As a consequence of the assumption on  $\ell_k$ ,  $|\log \ell_k(\cdot|\cdot)|$  is uniformly bounded on  $\mathcal{Y}^k \times \Theta$ . Then for any  $p \in U$  we have:

$$\begin{aligned} |\hat{I}_1(p)| &\leq \int_{\mathcal{Y}^k} \int_{\Theta} \ell_k(\mathbf{y}|\theta) d\nu(\theta) d\mu^{\otimes k}(\mathbf{y}) \sup_{\mathbf{y}, \theta} |\log \ell_k(\mathbf{y}|\theta)| \|p\| \\ &\leq \nu(\Theta) \sup_{\mathbf{y}, \theta} |\log \ell_k(\mathbf{y}|\theta)| \|p\|, \end{aligned} \quad (\text{C.7})$$

which implies that  $\hat{I}_1$  is a differentiable linear form. Its derivative in  $p$  is  $d\hat{I}_1(p) = \hat{I}_1$ .

We consider  $0 < \tilde{l}_1 < (1-\xi)l_1\nu(\Theta)$  and  $\tilde{l}_2 > (1+\xi)l_2\nu(\Theta)$  and write  $\hat{I}_2(p) = \Phi_1(p \times G(p))$ , with

## Appendix C. Non-asymptotic reference priors, a brief study

$G := C - \varphi \circ \Phi_2$ , and where  $\Phi_1, \Phi_2, \varphi$  and  $C$  are defined as follows:

$$1. \quad \Phi_2 \begin{cases} U & \longrightarrow \overset{\circ}{\mathcal{C}}(\mathcal{Y}^k \times \Theta, (\tilde{l}_1, \tilde{l}_2)) \\ p & \longmapsto [(t, \theta) \mapsto \int_{\Theta} \ell_k(\mathbf{y}|\tilde{\theta}) p(\tilde{\theta}) d\nu(\tilde{\theta})] \end{cases},$$

where  $\overset{\circ}{\mathcal{C}}(A, B)$  denotes the interior of  $\mathcal{C}(A, B) := \{f : A \mapsto B, \text{continuous}\}$ ;

$$2. \quad \varphi \begin{cases} \overset{\circ}{\mathcal{C}}(\mathcal{Y}^k \times \Theta, (\tilde{l}_1, \tilde{l}_2)) & \longrightarrow \overset{\circ}{\mathcal{C}}(\mathcal{Y}^k \times \Theta, (\log \tilde{l}_1, \log \tilde{l}_2)) \\ q & \longmapsto \log \circ q \end{cases};$$

3.  $C$  is the constant of  $\overset{\circ}{\mathcal{C}}(\mathcal{Y}^k \times \Theta, (\log \tilde{l}_1, \log \tilde{l}_2))$  that equals  $\varphi(\ell_k)$ ;

$$4. \quad \Phi_1 \begin{cases} \overset{\circ}{\mathcal{C}}(\mathcal{Y}^k \times \Theta, [\hat{l}_1, \hat{l}_2]) & \longrightarrow \mathbb{R} \\ r & \longmapsto \int_{\mathcal{Y}^k} \int_{\Theta} r(t, \theta) \ell_k(\mathbf{y}|\theta) d\nu(\theta) d\mu^{\otimes k}(\mathbf{y}) \end{cases}$$

where  $0 < \hat{l}_1 < \min(\log \tilde{l}_1, l_1)$  and  $\hat{l}_2 > \max(\log \tilde{l}_2, l_2)$ . We justify hereafter the correct definition of  $\Phi_2$ . Let  $p$  be in  $\mathcal{P}_{\mathcal{C}^b}^*$ . As  $\ell_k$  is continuous and uniformly bounded,  $p$  is bounded, and  $\nu(\Theta) < \infty$  we obtain that  $\Phi_2(p)$  is continuous, moreover, we have

$$(1 - \xi)l_1\nu(\Theta) < \Phi_2(p) < (1 + \xi)l_2\nu(\Theta) \quad (\text{C.8})$$

this way  $\Phi_2(p) \in \mathcal{C}(\mathcal{Y}^k \times \Theta, ((1 - \xi)l_1\nu(\Theta), (1 + \xi)l_2\nu(\Theta)))$  which is included in  $\overset{\circ}{\mathcal{C}}(\mathcal{Y}^k \times \Theta, (\tilde{l}_1, \tilde{l}_2))$ .

Now we will show that all these functions are differentiable.  $\Phi_2$  is linear, and for any  $p \in U$ ,  $\|\Phi_2(p)\| \leq \|p\|l_1\nu(\Theta)$  thus it is continuous and differentiable.  $\Phi_1$  is also linear and for any  $r \in \overset{\circ}{\mathcal{C}}(\mathcal{Y}^k \times \Theta, (\hat{l}_1, \hat{l}_2))$ ,  $\|\Phi_1(r)\| \leq \|r\|\nu(\Theta)$ , which states its differentiability. As the application  $(q, p) \mapsto qp$  is a continuous and differentiable bi-linear application, last task we have to do is to demonstrate the differentiability of  $\varphi$ .

We consider  $q \in \overset{\circ}{\mathcal{C}}(\mathcal{Y}^k \times \Theta, (\tilde{l}_1, \tilde{l}_2))$ . Let  $\varepsilon > 0$ , there exists  $\eta > 0$  such that for any  $x \in (-\eta, \eta)$ ,

$$|\log(1+x) - x| \leq \varepsilon|x| \inf_{\mathcal{Y}^k \times \Theta} |p|$$

Therefore if we consider  $\tau$  such that  $\mathcal{B}(q, \tau) \subset \overset{\circ}{\mathcal{C}}(\mathcal{Y}^k \times \Theta, (\tilde{l}_1, \tilde{l}_2))$  and  $\tau / \inf_{\mathcal{Y}^k \times \Theta} |q| < \eta$  it comes for any  $\mathbf{y}, \theta$  and  $h \in \mathcal{B}(q, \tau)$ ,

$$\left| \varphi(q+h) - \varphi(q) - \frac{h}{q} \right|(\mathbf{y}, \theta) = \left| \log \left( 1 + \frac{h(\mathbf{y}, \theta)}{q(\mathbf{y}, \theta)} \right) - \frac{h(\mathbf{y}, \theta)}{q(\mathbf{y}, \theta)} \right| \leq \varepsilon|h(\mathbf{y}, \theta)|$$

Thus  $\left\| \varphi(q+h) - \varphi(q) - \frac{h}{q} \right\| \leq \varepsilon\|h\|$

which shows that  $\varphi$  is differentiable with  $d\varphi(q)h = h/q$ .

Finally,  $\hat{l}_2$  is differentiable. Going back to  $\hat{l}$ , we have shown that is it differentiable. Moreover, as  $q \mapsto d\varphi(q)$  is continuous on  $\overset{\circ}{\mathcal{C}}(\mathcal{Y}^k \times \Theta, (\tilde{l}_1, \tilde{l}_2))$ , we deduce that  $\hat{l}$  continuously differentiable on  $U$ . We compute bellow its differentiate:

$$\begin{aligned} d\hat{l}(p)h &= \Phi_1(hG(p) - pd\varphi(\Phi_2(p))\Phi_2(h)) \\ &= \int_{\mathcal{Y}^k} \int_{\Theta} \ell_k(\mathbf{y}|\theta) \left( h(\theta) \log \frac{\ell_k(\mathbf{y}|\theta)}{\int_{\Theta} \ell_k(\mathbf{y}|\tilde{\theta}) p(\tilde{\theta}) d\nu(\tilde{\theta})} - p(\theta) \frac{\int_{\Theta} \ell_k(\mathbf{y}|\tilde{\theta}) h(\tilde{\theta}) d\nu(\tilde{\theta})}{\int_{\Theta} \ell_k(\mathbf{y}|\tilde{\theta}) p(\tilde{\theta}) d\nu(\tilde{\theta})} \right) d\nu(\theta) d\mu^{\otimes k}(\mathbf{y}) \end{aligned} \quad (\text{C.9})$$

The maximal argument point we are looking for is the  $\arg \max$  of  $\hat{l}$  on  $U$  under the constraint  $p \in F = g^{-1}(\{0\})$ ,  $g(p) := \int_{\Theta} pd\nu - 1$ .  $g$  is continuous and then  $\mathcal{C}^1$  on  $U$  with  $dg(p) = h \mapsto \int_{\Theta} hd\nu$  that is a surjective function for any  $p \in F$ .  $F$  is convex and  $\hat{l}$  is strictly concave on  $F$  and bounded from above. This

way,  $\hat{I}$  admits a unique maximal argument  $\pi$  over  $F$ . According to the Lagrange multipliers theorem, there exists  $\lambda \in \mathbb{R}$  such that  $\pi$  is a critical point of the function  $\mathcal{L} = \hat{I} - \lambda g$  i.e.

$$d\mathcal{L}(\pi) = d\hat{I}(\pi) - \lambda dg(\pi) = 0 \quad (\text{C.10})$$

This implies that, for any  $h \in E$ , the expression derived in eq. (C.9) is equal to  $\lambda \int_{\Theta} h(\theta) d\nu(\theta)$ . Thus, for any  $\theta \in \Theta$ ,

$$\lambda = \int_{\mathcal{Y}^k} \ell_k(\mathbf{y}|\theta) \left( \log \frac{\ell_k(\mathbf{y}|\theta)}{\int_{\Theta} \ell_k(\mathbf{y}|\tilde{\theta}) \pi(\tilde{\theta}) d\nu(\tilde{\theta})} - \frac{\int_{\Theta} \ell_k(\mathbf{y}|\tilde{\theta}) \pi(\tilde{\theta}) d\nu(\tilde{\theta})}{\int_{\Theta} \ell_k(\mathbf{y}|\tilde{\theta}) \pi(\tilde{\theta}) d\nu(\tilde{\theta})} \right) d\mu^{\otimes k}(\mathbf{y}) \quad (\text{C.11})$$

expressing the exponential of last equation and multiplying by  $\pi(\theta)$  leads to what follows.

$$\begin{aligned} \pi(\theta) e^{\lambda+1} &= \exp \left[ \int_{\mathcal{Y}^k} \ell_k(\mathbf{y}|\theta) \log \frac{\ell_k(\mathbf{y}|\theta)}{\int_{\Theta} \ell_k(\mathbf{y}|\tilde{\theta}) \pi(\tilde{\theta}) d\nu(\tilde{\theta})} d\mu^{\otimes k}(\mathbf{y}) + \log \pi(\theta) \right] \\ &= \exp \left[ \int_{\mathcal{Y}^k} \ell_k(\mathbf{y}|\theta) \log \frac{\ell_k(\mathbf{y}|\theta)}{\int_{\Theta} \ell_k(\mathbf{y}|\tilde{\theta}) \pi(\tilde{\theta}) d\nu(\tilde{\theta})} d\mu^{\otimes k}(\mathbf{y}) + \int_{\mathcal{Y}^k} \ell_k(\mathbf{y}|\theta) \log \pi(\theta) d\mu^{\otimes k}(\mathbf{y}) \right] \end{aligned} \quad (\text{C.12})$$

by using the fact that  $\int_{\mathcal{Y}^k} \ell_k(\mathbf{y}|\theta) d\mu^{\otimes k}(\mathbf{y}) = 1$ . Finally, we obtain,

$$\pi(\theta) e^{\lambda+1} = \exp \left[ \int_{\mathcal{Y}^k} \ell_k(\mathbf{y}|\theta) \log \frac{\ell_k(\mathbf{y}|\theta) \pi(\theta)}{\int_{\Theta} \ell_k(\mathbf{y}|\tilde{\theta}) \pi(\tilde{\theta}) d\nu(\tilde{\theta})} d\mu^{\otimes k}(\mathbf{y}) \right]. \quad (\text{C.13})$$

□

*Proof of theorem C.2.* Using the notations introduced in the preceding proof, we are looking for a maximizer  $\tilde{\pi} \in U$  of  $\hat{I}$  that satisfies the equality constraints  $\int_{\Theta} g_j(\theta) \tilde{\pi}(\theta) d\nu(\theta) = c_j$  for any  $j \in \{0, \dots, q\}$ . These constraints are linear and continuous letting the strict concavity of  $\hat{I}$  ensuring the existence of a unique solution  $\tilde{\pi}$ . Since we remain under the assumptions of the Lagrange multipliers theorem, we obtain that there exist  $\lambda, \lambda_1, \dots, \lambda_q \in \mathbb{R}$  such that

$$d\hat{I}(\tilde{\pi})h - \lambda \int_{\Theta} h d\nu + \sum_{j=1}^q \lambda_j \int_{\Theta} g_j h d\nu = 0 \quad (\text{C.14})$$

for any  $h \in E$ . Thus, for any  $\theta \in \Theta$ ,

$$\lambda - \sum_{j=1}^q \lambda_j g_j(\theta) = \int_{\mathcal{Y}^k} \ell_k(\mathbf{y}|\theta) \log \frac{\ell_k(\mathbf{y}|\theta)}{\int_{\Theta} \ell_k(\mathbf{y}|\tilde{\theta}) \tilde{\pi}(\tilde{\theta}) d\nu(\tilde{\theta})} d\mu^{\otimes k}(\mathbf{y}) - 1. \quad (\text{C.15})$$

Applying exponential function to the above and multiplying by  $\tilde{\pi}$  leads to the result as:

$$\tilde{\pi}(\theta) e^{\lambda+1} e^{-\sum_{j=1}^q \lambda_j g_j(\theta)} = \exp \left[ \int_{\mathcal{Y}^k} \ell_k(\mathbf{y}|\theta) \log \frac{\ell_k(\mathbf{y}|\theta) \tilde{\pi}(\theta)}{\int_{\Theta} \ell_k(\mathbf{y}|\tilde{\theta}) \tilde{\pi}(\tilde{\theta}) d\nu(\tilde{\theta})} d\mu^{\otimes k}(\mathbf{y}) \right]. \quad (\text{C.16})$$

□

## C.4 Discussion and link with asymptotic reference priors

The results provided by theorems C.1 and C.2 provide implicit expressions of non-asymptotic reference priors. The resolution of these equations to issue an explicit formulation of the priors remains complex. However, further work could study the approximation of the priors, as the solutions of fixed-point problems.

Furthermore, one should notice that the implicit formulations actually express the priors as a function of their posterior. This remark motivates their asymptotic study since, as  $k \rightarrow \infty$ , the dependence of the posterior in the prior should vanish. We develop this heuristic by taking the Bernstein-von Mises theorem as a support. Under appropriate assumptions, this theorem states that the posterior density is asymptotically close to a Gaussian density whose mean is a frequentist asymptotically sufficient and consistent estimator,  $\hat{\theta}_k$ , and whose variance is  $\frac{1}{\sqrt{k}}\mathcal{I}(\theta)^{-1/2}$ , where  $\mathcal{I}(\theta)$  denotes the Fisher information matrix. Considering this statement, we write

$$\log p(\theta|\mathbf{y}) \simeq \log \frac{k}{2\pi} - \frac{k}{2}\|\mathcal{I}(\theta)^{1/2}(\theta - \hat{\theta}_k)\|^2 + \frac{1}{2}\log |\det \mathcal{I}(\theta)|. \quad (\text{C.17})$$

Using that, under minimal assumptions,  $\mathbb{E}_{\mathbf{Y}_k \sim \mathbb{P}_{\mathbf{Y}_k|\theta}} k\|\mathcal{I}(\theta)^{1/2}(\theta - \hat{\theta})\|^2$  tends toward 0 as  $k \rightarrow \infty$ , we obtain that

$$f_\pi(\theta) \simeq \frac{k}{2\pi} \exp \log \frac{1}{2} |\det \mathcal{I}(\theta)|, \quad (\text{C.18})$$

where  $f_\pi$  is the function expressed in theorem C.1. Finally, within this heuristic, denoting  $J$  the Jeffreys prior density:  $J(\theta) = \sqrt{|\det \mathcal{I}(\theta)|}$ , we obtain that, asymptotically,

1. the solution of theorem C.1 becomes proportional to the Jeffreys prior,
2. the solution of theorem C.2 becomes proportional to  $\tilde{\pi}^\infty$  defined as a modified Jeffreys prior:

$$\tilde{\pi}^\infty(\theta) = J(\theta) \cdot \exp \left( \sum_{j=1}^q \lambda_j g_j \right). \quad (\text{C.19})$$

We conclude by highlighting that the expression of the functional  $f_\pi$  in theorem C.1 echoes the one of the functional that formally tends to the asymptotic reference priors in the theorem stated by Berger, Bernardo, and Sun (2009) (see chapter 3, section 3.3, where this theorem is expressed). In their theorem, they leverage the statement according to which the posterior should become independent of the prior asymptotically. Indeed, they replace in  $f_\pi$  the posterior that is derived from  $\pi$  by the posterior that is derived from another suitable prior.



# Ma thèse en quelques pages

En quoi consiste cette thèse ? C'est la question à laquelle on cherche à répondre ici, en seulement quelques pages et trois minutes de lecture. Cette annexe propose alors un résumé simple du contexte, des enjeux et du travail apporté au cours de cette thèse.

<b>La thèse en quelques mots, et quelques phrases</b>	<b>221</b>
<b>Initiation à la réflexion sur l'a priori Bayésien</b>	<b>222</b>
<b>Le lien avec les séismes et les centrales nucléaires dans tout ça</b>	<b>222</b>
<b>Concrètement, qu'est ce qu'on a fait, qu'est-ce qu'on en conclut ?</b>	<b>223</b>

## La thèse en quelques mots, et quelques phrases

La France est le pays qui admet le plus grand ratio de centrales nucléaires par habitants. Le nucléaire représente alors aujourd'hui une industrie importante et essentielle à notre pays. Il s'agit également d'une industrie hautement sensible, qui se doit d'être dimensionnée pour résister à des événements exceptionnels. En effet, on se souvient tous de l'incident survenu à Fukushima, au Japon, en 2011 : un tsunami est venu s'abattre sur la centrale nucléaire de Fukushima-Daiichi provoquant la fusion du cœur et une irradiation de la zone et des populations alentours. Les études et analyses sur cette catastrophe ont démontré que les signaux sismiques (le séisme ayant causé le tsunami) qui ont été pris en compte pour dimensionner la résistance de la centrale ont été sous évalués par les autorités japonaises.

En France, le rapport à la sûreté des installations nucléaires se construit autour d'un dialogue entre EDF, qui est l'exploitant du parc nucléaire civil, et l'autorité publique de sûreté nucléaire et de radioprotection (ASN), qui doit fixer et veiller au respect des normes de sûreté nucléaires. Dans ce dialogue entre les experts de ces deux partis, le CEA intervient avec une place ambivalente ; tout d'abord au même titre qu'EDF, le CEA est responsable d'installations nucléaires (de recherche), qu'il opère. Aussi, il participe à des études d'expertise d'équipements nucléaires d'EDF.

Suite à l'incident de Fukushima, l'ASN a imposé en France des normes encore plus strictes qu'auparavant relatives à la robustesse des équipements nucléaires devant le risque sismique. C'est pourquoi, au CEA, la recherche sur l'évolution des méthodes d'études sismiques de sûreté est un enjeu permanent. C'est dans ce cadre que s'inscrit ma thèse.

La robustesse des équipements reste encore aujourd'hui complexe à évaluer, même avec les outils scientifiques actuels. De plus, il est essentiel d'introduire dans ces études la notion d'aléa présente dans le séisme et dans la réponse de l'installation, ces éléments comportant de nombreuses inconnues aléatoires. Pour étudier mathématiquement cet aléatoire et ses effets, on s'intéresse dans cette thèse à la méthode qui s'appelle l'inférence bayésienne. Cette méthode admet cependant une limite qui provient du choix dit "a

## \* Ma thèse en quelques pages

"priori" qui constitue un vaste sujet de recherche mathématique dont on parle plus en détail dans la section suivante.

L'objectif affiché est de développer cette méthode mathématique dans l'idée de produire des démonstrations robustes et auditables de fragilité (ou de non fragilité) des équipements mécaniques qui composent une centrale nucléaire.

## Initiation à la réflexion sur l'a priori Bayésien

Imaginez un dé, simple et à six faces. Je le lance et vous demande, selon vous, quelle est la probabilité que le résultat soit 6. J'imagine que beaucoup diront une chance sur six, en partant du principe qu'il peut s'agir de chaque face du dé avec la même probabilité. Mettons alors maintenant que je vous donne une information : le chiffre sur lequel le dé est tombé est pair. Et alors, à nouveau, je vous demande la probabilité que ce chiffre soit 6. Si vous êtes bien partis du principe initial selon lequel chaque face pouvait tomber de manière équiprobable, alors vous me répondrez sûrement une chance sur trois. Puisque si le résultat est pair il ne peut s'agir que de 2, 4 ou 6. Ce que vous avez donc fait, c'est que vous êtes partis d'une probabilité initiale que vous croyez et que vous avez mis à jour après observation d'une partie du réel. Mais tout cela, c'est sans compter votre voisin ou la personne derrière vous qui aurait pu vous dire "Non. Son dé, je l'ai vu, pour moi il est truqué, je connais ça. Il y a deux fois plus de chances pour que le chiffre tombe sur 1, 2 ou 3." Ou bien "les dés comme le sien je connais, j'en ai fabriqué plein, il n'y a pas plus d'une chance sur quarante-huit qu'il sorte un 6." Finalement, ce problème revient exactement à ce que l'on appelle l'inférence bayésienne. En partant d'un *a priori*, sur la distribution des probabilités des différentes faces du dé, vous avez mis à jour la probabilité d'obtenir 6 après connaissance d'information sur le réel (le chiffre est pair). Cette mise à jour et le résultat un tiers que vous avez déterminé constituent en un sens l'*a posteriori*.



Mais alors que croire, car si vous aviez écouté l'une des deux expertises qui proposait chacune son propre *a priori*, vous seriez tombé sur un résultat final différent de probabilité d'obtenir 6 qu'en considérant votre jugement initial, que vous trouviez sûrement le plus logique et le plus objectif<sup>1</sup> en l'absence de connaissance ou d'information sur le dé. Le problème du choix *a priori* est alors celui de pouvoir définir un *a priori* objectif lorsque l'on n'a pas ou l'on ne croit pas en de telles expertises.

Ce problème est celui qui est au centre de ma thèse. Je cherche à définir et étudier mathématiquement ce qu'est un *a priori* objectif et comment il est possible de l'employer dans notre contexte de travail.

## Le lien avec les séismes et les centrales nucléaires dans tout ça

Dans cette thèse, le dé peut alors représenter une centrale, ou un équipement de cette centrale. Et plutôt qu'un dé à six faces, on imagineraît plutôt une structure empilée de dés à dix-huit mille faces (voire même une infinité). Au lieu d'être lancés, les dés bougent à la suite d'un tremblement de terre. Et on se demandera alors quelle est la probabilité que la structure s'effondre.

<sup>1</sup>Pour rassurer le lecteur, dans ce cas précis du dé son intuition est la bonne, puisque le prior le plus objectif mathématiquement est bien celui qui donne la même probabilité à chaque face du dé.



Voici en fig. a ci-dessous un exemple concret de "dés" que l'on étudie dans cette thèse. Il ne s'agit pas de toute une centrale, mais d'un équipement, en l'occurrence, un circuit de tuyauterie, qui provient d'un réacteur nucléaire en France. Sur cette photo, la tuyauterie se trouve sur une table vibrante du CEA qui reproduit les mouvements du sol lors d'un séisme afin d'observer le comportement du système dans ces conditions.

De la même manière qu'on observait une partie du réel (le chiffre est pair) dans l'exemple du dé de la section précédente, on observe ici une partie du comportement mécanique du système lorsqu'il est soumis à certains signaux sismiques envisageables. Avec la méthode bayésienne, on veut se servir de cette information pour déduire la probabilité de défaillance du système. C'est là qu'on peut comprendre que le choix de l'*a priori* est critique et essentiel dans une telle étude. En effet, afin d'être convaincus du résultat obtenu relativement à la fragilité d'un tel système et de convaincre également de ce résultat, la démarche entière doit être auditable et exempte de toute subjectivité introduite. Comme dit plus haut, mon travail a alors été de réfléchir à des méthodes efficaces et intelligibles pour définir mathématiquement des *a priori* objectifs pour estimer de manière robuste et auditable la fragilité sismique de tels équipements. Mathématiquement parlant, on cherche à construire de tels *a priori* en s'appuyant sur et en développant une théorie qui s'appelle la théorie des priors de référence.

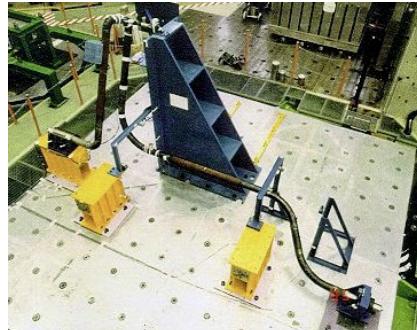


Figure a: Section de tuyauterie provenant du circuit de refroidissement d'une centrale nucléaire en France.

## Concrètement, qu'est ce qu'on a fait, qu'est-ce qu'on en conclut ?

Le premier axe de travail majeur de cette thèse a été de construire une définition de ce qu'est mathématiquement, en inférence bayésienne, un prior objectif en partant de la théorie existante des priors de référence. Cette définition, qui va un peu plus loin que la théorie initiale, a donné lieu à un des résultats théoriques principaux de ma thèse : je démontre que le prior objectif est dans la plupart des cas celui qui s'appelle "prior de Jeffreys".

Ensuite, pour aller plus loin, j'ai travaillé et défini des méthodes pour adapter ce prior objectif de Jeffreys dans des cas où sa forme théorique est trop complexe pour être utilisée en pratique.

Ces résultats ont constitué un apport direct au second acte de ma thèse, qui fut consacré à l'implémentation de ces priors théoriques dans nos cas d'étude de fragilité sismique d'équipements nucléaires. J'ai étudié principalement l'interaction du prior objectif avec la modélisation mathématique de la fragilité sismique afin de mieux la comprendre et la quantifier. En plus de cela, j'ai cherché à proposer diverses techniques qui ont pour

### \* *Ma thèse en quelques pages*

but d'améliorer l'inférence bayésienne pour donner lieu à la meilleure estimation avec le moins de données possible. Nos résultats, principalement issus de l'application de nos méthodes aux circuits de tuyauterie présentés en section précédente, ont démontré une bonne performance de nos méthodologies.

Au final, ces études ont donné lieu à une estimation de fragilité sismique que l'on considère comme étant robuste. Cette dernière est issue d'une méthodologie auditable grâce à l'emploi d'un prior objectif particulièrement construit pour de tels cas pratiques.

# Bibliography

- Agrell, Christian and Kristina Rognlien Dahl (2021). "Sequential Bayesian optimal experimental design for structural reliability analysis". In: *Statistics and Computing* 31.3, p. 27. doi: [10.1007/s11222-021-10000-2](https://doi.org/10.1007/s11222-021-10000-2).
- Ambraseys, Nicholas N. et al. (2000). "Dissemination of European StrongMotion Data". CD-ROM collection. European Commission, Directorate-General XII, Environmental and Climate Programme, ENV4-CT97-0397, Brussels, Belgium.
- Arbel, Julyan et al. (2023). "Bayes in action in deep learning and dictionary learning". In: *ESAIM: ProcS* 74, pp. 90–107. doi: [10.1051/proc/202374090](https://doi.org/10.1051/proc/202374090).
- Aronszajn, Nachman (1950). "Theory of reproducing kernels". In: *Transactions of the American Mathematical Society* 68.3, pp. 337–404. doi: [10.1090/S0002-9947-1950-0051437-7](https://doi.org/10.1090/S0002-9947-1950-0051437-7).
- ASN (2002). *Règle Fondamentale de Sûreté : Développement et utilisation des études probabilistes de sûreté*. RFS 2002-01. Autorité de sûreté nucléaire. URL: <https://www ASN.fr/reglementation/rfs/rfs-relatives-aux-rep/rfs-2002-1-du-26-12-2002>.
- Azzimonti, Dario, David Ginsbourger, Clément Chevalier, Julien Bect, and Yann Richet (2021). "Adaptive Design of Experiments for Conservative Estimation of Excursion Sets". In: *Technometrics* 63.1, pp. 13–26. doi: [10.1080/00401706.2019.1693427](https://doi.org/10.1080/00401706.2019.1693427).
- Bach, Francis (2023). "Sum-of-Squares Relaxations for Information Theory and Variational Inference". arXiv: 2206.13285. doi: [10.48550/arXiv.2206.13285](https://doi.org/10.48550/arXiv.2206.13285).
- Baillie, Nils, Antoine Van Biesbroeck, Cyril Feau, and Clément Gauchy (2025). "Bayesian estimation of seismic fragility curves based on variational reference priors using neural networks". In: *Proceedings of the 6th UNCECOMP Conference*. URL: <https://2025.uncecomp.org/proceedings/pdf/21225.pdf>.
- Baillie, Nils, Antoine Van Biesbroeck, and Clément Gauchy (2025). "Variational inference for approximate objective priors using neural networks". arXiv:2502.02364. doi: [10.48550/arXiv.2502.02364](https://doi.org/10.48550/arXiv.2502.02364).
- Baker, Jack W. (2015). "Efficient analytical fragility function fitting using dynamic structural analysis". In: *Earthquake Spectra* 31.1, pp. 579–599. doi: [10.1193/021113EQS025M](https://doi.org/10.1193/021113EQS025M).
- Basir, Shamsulhaq and Inanc Senocak (2023). "An adaptive augmented Lagrangian method for training physics and equality constrained artificial neural networks". arXiv:2306.0490. doi: [10.48550/arXiv.2306.04904](https://doi.org/10.48550/arXiv.2306.04904).
- Bect, Julien, François Bachoc, and David Ginsbourger (2019). "A supermartingale approach to Gaussian process based sequential design of experiments". In: *Bernoulli* 25.4, pp. 2883–2919. doi: [10.3150/18-BEJ1074](https://doi.org/10.3150/18-BEJ1074).
- Berger, James O. (1990). "Robust Bayesian analysis: sensitivity to the prior". In: *Journal of Statistical Planning and Inference* 25.3, pp. 303–328. doi: [10.1016/0378-3758\(90\)90079-A](https://doi.org/10.1016/0378-3758(90)90079-A).
- Berger, James O. and José M. Bernardo (1992a). "On the development of Reference Priors". In: *Bayesian Statistics* 4. Oxford: Oxford University Press, pp. 35–60.
- (1992b). "Ordered group reference priors with application to the multinomial problem". In: *Biometrika* 79.1, pp. 25–37. doi: [10.1093/biomet/79.1.25](https://doi.org/10.1093/biomet/79.1.25).
- Berger, James O., José M. Bernardo, and Dongchu Sun (2009). "The formal definition of reference priors". In: *The Annals of Statistics* 37.2, pp. 905–938. doi: [10.1214/07-AOS587](https://doi.org/10.1214/07-AOS587).
- (2015). "Overall Objective Priors". In: *Bayesian Analysis* 10.1, pp. 189–221. doi: [10.1214/14-BA915](https://doi.org/10.1214/14-BA915).
- Berger, James O. and Dongchu Sun (2008). "Objective priors for the bivariate normal model". In: *The Annals of Statistics* 36.2, pp. 963–982. doi: [10.1214/07-AOS501](https://doi.org/10.1214/07-AOS501).

## Bibliography

- Bernardo, José M. (1979a). "Expected Information as Expected Utility". In: *The Annals of Statistics* 7.3, pp. 686–690. doi: [10.1214/aos/1176344689](https://doi.org/10.1214/aos/1176344689).
- (1979b). "Reference Posterior Distributions for Bayesian Inference". In: *Journal of the Royal Statistical Society. Series B* 41.2, pp. 113–147. doi: [10.1111/j.2517-6161.1979.tb01066.x](https://doi.org/10.1111/j.2517-6161.1979.tb01066.x).
- (2005). "Reference Analysis". In: *Bayesian Thinking*. Ed. by D. K. Dey and C. R. Rao. Vol. 25. Handbook of Statistics. ISSN: 0169-7161. Elsevier, pp. 17–90. doi: [10.1016/S0169-7161\(05\)25002-2](https://doi.org/10.1016/S0169-7161(05)25002-2).
- Bernardo, José M. and Adrian F. M. Smith (1994). *Bayesian Theory*. 1st. Wiley Series in Probability and Statistics. Wiley.
- Bernier, Carl and Jamie E. Padgett (2019). "Fragility and risk assessment of aboveground storage tanks subjected to concurrent surge, wave, and wind loads". In: *Reliability Engineering & System Safety* 191, p. 106571. doi: [10.1016/j.ress.2019.106571](https://doi.org/10.1016/j.ress.2019.106571).
- Beylat, Delphine (2020). "Contribution à l'étude du comportement dynamique aléatoire de structures libres et empilées sous séisme". PhD thesis. Université Clermont Auvergne. URL: <https://theses.hal.science/tel-03159237>.
- Bioche, Christèle (2015). "Approximation de lois impropre et applications". Theses. Université Blaise Pascal - Clermont-Ferrand II. URL: <https://theses.hal.science/tel-01308523>.
- Bioche, Christèle and Pierre Druilhet (2016). "Approximation of improper priors". In: *Bernoulli* 22.3. doi: [10.3150/15-bej708](https://doi.org/10.3150/15-bej708).
- Bodnar, Olha and Clemens Elster (2014). "Analytical derivation of the reference prior by sequential maximization of Shannon's mutual information in the multi-group parameter case". In: *Journal of Statistical Planning and Inference* 147, pp. 106–116. doi: [10.1016/j.jspi.2013.11.003](https://doi.org/10.1016/j.jspi.2013.11.003).
- Bousquet, Nicolas (2024a). "Contributions to the statistical quantification of uncertainties affecting the use of numerical models". Accreditation to supervise research. Sorbonne université. URL: <https://perso.lpsm.paris/~bousquet/hdr/hdr.pdf>.
- (2024b). "Discussion of "Specifying prior distributions in reliability applications": Towards new formal rules for informative prior elicitation?" In: *Applied Stochastic Models in Business and Industry* 40.1, pp. 92–102. doi: [10.1002/asmb.2794](https://doi.org/10.1002/asmb.2794).
- Buratti, Nicola, Fabio Minghini, Elena Ongaretto, Marco Savoia, and Nerio Tullini (2017). "Empirical seismic fragility for the precast RC industrial buildings damaged by the 2012 Emilia (Italy) earthquakes". In: *Earthquake Engineering & Structural Dynamics* 46.14, pp. 2317–2335. doi: [10.1002/eqe.2906](https://doi.org/10.1002/eqe.2906).
- Castillo, Ismaël, Johannes Schmidt-Hieber, and Aad van der Vaart (2015). "Bayesian linear regression with sparse priors". In: *The Annals of Statistics* 43.5, pp. 1986–2018. doi: [10.1214/15-AOS1334](https://doi.org/10.1214/15-AOS1334).
- CEA (2019). "CAST3M". URL: <http://www-cast3m.cea.fr/>.
- Chang, J. T. and D. Pollard (1997). "Conditioning as disintegration". In: *Statistica Neerlandica* 51.3, pp. 287–317. doi: [10.1111/1467-9574.00056](https://doi.org/10.1111/1467-9574.00056).
- Chen, Ming-Hui, Dipak K. Dey, Peter Müller, Dongchu Sun, and Keying Ye (2010). "Objective Bayesian Inference with Applications". In: *Frontiers of Statistical Decision Making and Bayesian Analysis: In Honor of James O. Berger*. New York, NY: Springer New York, pp. 31–68. doi: [10.1007/978-1-4419-6944-6\\_2](https://doi.org/10.1007/978-1-4419-6944-6_2).
- Chen, Ming-Hui and Joseph G. Ibrahim (2006). "The relationship between the power prior and hierarchical models". In: *Bayesian Analysis* 1.3, pp. 551–574. doi: [10.1214/06-BA118](https://doi.org/10.1214/06-BA118).
- Chen, Ming-Hui, Joseph G. Ibrahim, and Sungduk Kim (2008). "Properties and Implementation of Jeffreys's Prior in Binomial Regression Models". In: *Journal of the American Statistical Association* 103.484, pp. 1659–1664. doi: [10.1198/016214508000000779](https://doi.org/10.1198/016214508000000779).
- Choe, Do-Eun, Paolo Gardoni, and David Rosowsky (2007). "Closed-Form Fragility Estimates, Parameter Sensitivity, and Bayesian Updating for RC Columns". In: *Journal of Engineering Mechanics* 133.7, pp. 833–843. doi: [10.1061/\(ASCE\)0733-9399\(2007\)133:7\(833\)](https://doi.org/10.1061/(ASCE)0733-9399(2007)133:7(833)).
- Chu, John T. (1955). "On Bounds for the Normal Integral". In: *Biometrika* 42.1, pp. 263–265. doi: [10.2307/2333443](https://doi.org/10.2307/2333443).
- Ciano, Matteo, Massimiliano Gioffrè, and Mircea Dan Grigoriu (2020). "The role of intensity measures on the accuracy of seismic fragilities". In: *Probabilistic Engineering Mechanics* 60, p. 103041. doi: [10.1016/j.probengmech.2020.103041](https://doi.org/10.1016/j.probengmech.2020.103041).

- (2022). “A novel approach to improve accuracy in seismic fragility analysis: The modified intensity measure method”. In: *Probabilistic Engineering Mechanics* 69, p. 103301. doi: [10.1016/j.probengmech.2022.103301](https://doi.org/10.1016/j.probengmech.2022.103301).
- Clarke, Bertrand S. and Andrew R. Barron (1990). “Information-theoretic asymptotics of Bayes methods”. In: *IEEE Transactions on Information Theory* 36.3, pp. 453–471. doi: [10.1109/18.54897](https://doi.org/10.1109/18.54897).
- (1994). “Jeffreys’ prior is asymptotically least favorable under entropy risk”. In: *Journal of Statistical Planning and Inference* 41.1, pp. 37–60. doi: [10.1016/0378-3758\(94\)90153-8](https://doi.org/10.1016/0378-3758(94)90153-8).
- Clarke, Bertrand S. and Subhashis Ghosal (2010). “Reference priors for exponential families with increasing dimension”. In: *Electronic Journal of Statistics* 4, pp. 737–780. doi: [10.1214/10-EJS569](https://doi.org/10.1214/10-EJS569).
- Clarke, Bertrand S. and Dongchu Sun (1997). “Reference Priors under the Chi-Squared Distance”. In: *Sankhyā: The Indian Journal of Statistics* 59.2, pp. 215–231. doi: [10.2307/25051152](https://doi.org/10.2307/25051152).
- Consonni, Guido, Dimitris Fouskakis, Brunero Liseo, and Ioannis Ntzoufras (2018). “Prior Distributions for Objective Bayesian Analysis”. In: *Bayesian Analysis* 13.2, pp. 627–679. doi: [10.1214/18-BA1103](https://doi.org/10.1214/18-BA1103).
- Cornell, C. Allin (1968). “Engineering seismic risk analysis”. In: *Bulletin of the Seismological Society of America* 58.5, pp. 1583–1606. doi: [10.1785/BSSA0580051583](https://doi.org/10.1785/BSSA0580051583).
- (2004). “Hazard, ground motions and probabilistic assessment for PBSD”. In: *Proceedings of the International Workshop on Performance-Based Seismic Design - Concepts and Implementation*. Bled, Slovenia: PEER Center, University of California, CA, pp. 39–52.
- Csiszár, Imre (1967). “Information-type measures of difference of probability distributions and indirect observation”. In: *Studia Scientiarum Mathematicarum Hungarica* 2, pp. 229–318.
- D’Andrea, Amanda M. E. et al. (2021). “Objective bayesian analysis for multiple repairable systems”. In: *PLOS ONE* 16, pp. 1–19. doi: [10.1371/journal.pone.0258581](https://doi.org/10.1371/journal.pone.0258581).
- Da Veiga, Sébastien (2015). “Global sensitivity analysis with dependence measures”. In: *Journal of Statistical Computation and Simulation* 85.7, pp. 1283–1305. doi: [10.1080/00949655.2014.945932](https://doi.org/10.1080/00949655.2014.945932).
- Da Veiga, Sébastien, Fabrice Gamboa, Bertrand Iooss, and Clémentine Prieur (2021). *Basics and Trends in Sensitivity Analysis*. 1st ed. Computational Science & Engineering. Philadelphia, PA: Society for Industrial and Applied Mathematics. doi: [10.1137/1.9781611976694](https://doi.org/10.1137/1.9781611976694).
- Damblin, Guillaume, Merlin Keller, Alberto Pasanisi, Pierre Barbillon, and Éric Parent (2014). “Approche décisionnelle bayésienne pour estimer une courbe de fragilité”. In: *Journal de la Société Française de Statistique* 155.3, pp. 78–103. URL: <https://hal.archives-ouvertes.fr/hal-01545648>.
- Datta, Gauri Sankar and Malay Ghosh (1996). “On the invariance of noninformative priors”. In: *The Annals of Statistics* 24.1, pp. 141–159. doi: [10.1214/aos/1033066203](https://doi.org/10.1214/aos/1033066203).
- Davis, Philip J. (1959). “Leonhard Euler’s Integral: A Historical Profile of the Gamma Function: In Memoriam: Milton Abramowitz”. In: *The American Mathematical Monthly* 66.10, pp. 849–869.
- Ellingwood, Bruce R. (1990). “Validation studies of seismic PRAs”. In: *Nuclear Engineering and Design* 123.2, pp. 189–196. doi: [10.1016/0029-5493\(90\)90237-R](https://doi.org/10.1016/0029-5493(90)90237-R).
- (2001). “Earthquake risk assessment of building structures”. In: *Reliability Engineering & System Safety* 74.3, pp. 251–262. doi: [10.1016/S0951-8320\(01\)00105-3](https://doi.org/10.1016/S0951-8320(01)00105-3).
- EPRI (2011). *Advanced Methods for Determination of Seismic Fragilities: Seismic Fragilities Using Scenario Earthquakes*. 1002988. Electric Power Research Institute. URL: <https://www.epri.com/research/products/1022995>.
- (2013). *Seismic Probabilistic Risk Assessment Implementation Guide*. 3002000709. Electric Power Research Institute. URL: <https://www.epri.com/research/products/3002000709>.
- Galharret, Jean-Michel and Anne Philippe (2023). “Bayesian analysis for mediation and moderation using g-priors”. In: *Econometrics and Statistics* 27, pp. 161–172. doi: [10.1016/j.ecosta.2021.12.009](https://doi.org/10.1016/j.ecosta.2021.12.009).
- Gao, Yansong, Rahul Ramesh, and Pratik Chaudhari (2022). “Deep Reference Priors: What is the best way to pretrain a model?” In: *Proceedings of the 39th International Conference on Machine Learning*. Vol. 162. Proceedings of Machine Learning Research. PMLR, pp. 7036–7051. URL: <https://proceedings.mlr.press/v162/gao22d.html>.
- Gardoni, Paolo, Armen Der Kiureghian, and Khalid M. Mosalam (2002). “Probabilistic Capacity Models and Fragility Estimates for Reinforced Concrete Columns based on Experimental Observations”. In: *Journal of Engineering Mechanics* 128.10, pp. 1024–1038. doi: [10.1061/\(ASCE\)0733-9399\(2002\)128:10\(1024\)](https://doi.org/10.1061/(ASCE)0733-9399(2002)128:10(1024)).

## Bibliography

- Gauchy, Clément, Cyril Feau, and Josselin Garnier (2021). "Importance sampling based active learning for parametric seismic fragility curve estimation". arXiv:2109.04323. doi: [10.48550/arxiv.2109.04323](https://doi.org/10.48550/arxiv.2109.04323).
- (2024). "Uncertainty quantification and global sensitivity analysis of seismic fragility curves using kriging". In: *International Journal for Uncertainty Quantification* 14.4, pp. 39–63. doi: [10.1615/Int.J.UncertaintyQuantification.2023046480](https://doi.org/10.1615/Int.J.UncertaintyQuantification.2023046480).
- Gautschi, Walter (1959). "Some Elementary Inequalities Relating to the Gamma and Incomplete Gamma Function". In: *Journal of Mathematics and Physics* 38.1, pp. 77–81. doi: [10.1002/sapm195938177](https://doi.org/10.1002/sapm195938177).
- Gehl, Pierre, John Douglas, and Darius M. Seyed (2015). "Influence of the Number of Dynamic Analyses on the Accuracy of Structural Response Estimates". In: *Earthquake Spectra* 31.1, pp. 97–113. doi: [10.1193/102912EQS320M](https://doi.org/10.1193/102912EQS320M).
- Gelman, Andrew, Aki Vehtari, et al. (2020). "Bayesian Workflow". arXiv:2011.01808. doi: [10.48550/arXiv.2011.01808](https://doi.org/10.48550/arXiv.2011.01808).
- Gelman, Andrew and Yuling Yao (2020). "Holes in Bayesian statistics". In: *Journal of Physics G: Nuclear and Particle Physics* 48.1, p. 014002. doi: [10.1088/1361-6471/abc3a5](https://doi.org/10.1088/1361-6471/abc3a5).
- Ghobarah, Ahmed (2001). "Performance-based design in earthquake engineering: state of development". In: *Engineering Structures* 23.8, pp. 878–884. doi: [10.1016/S0141-0296\(01\)00036-0](https://doi.org/10.1016/S0141-0296(01)00036-0).
- Ghosh, Malay, Victor Mergel, and Ruitao Liu (2011). "A general divergence criterion for prior selection". In: *Annals of the Institute of Statistical Mathematics* 63.1, pp. 43–58. doi: [10.1007/s10463-009-0226-4](https://doi.org/10.1007/s10463-009-0226-4).
- Ghosh, Swarup and Subrata Chakraborty (2020). "Seismic fragility analysis of structures based on Bayesian linear regression demand models". In: *Probabilistic Engineering Mechanics* 61, p. 103081. doi: [10.1016/j.probengmech.2020.103081](https://doi.org/10.1016/j.probengmech.2020.103081).
- Gidaris, Ioannis, Alexandros A. Taflanidis, and George P. Mavroelidis (2015). "Kriging metamodeling in seismic risk assessment based on stochastic ground motion models". In: *Earthquake Engineering & Structural Dynamics* 44.14, pp. 2377–2399. doi: [10.1002/eqe.2586](https://doi.org/10.1002/eqe.2586).
- Giovenale, Paolo, C. Allin Cornell, and Luis Esteva (2004). "Comparing the adequacy of alternative ground motion intensity measures for the estimation of structural responses". In: *Earthquake Engineering & Structural Dynamics* 33.8, pp. 951–979. doi: [10.1002/eqe.386](https://doi.org/10.1002/eqe.386).
- Gretton, Arthur, Karsten M. Borgwardt, Malte J. Rasch, Bernhard Schölkopf, and Alexander Smola (2012). "A Kernel Two-Sample Test". In: *Journal of Machine Learning Research* 13.25, pp. 723–773. URL: <http://jmlr.org/papers/v13/gretton12a.html>.
- Grigoriu, Mircea Dan and Alin Radu (2021). "Are seismic fragility curves fragile?" In: *Probabilistic Engineering Mechanics* 63, p. 103115. doi: [10.1016/j.probengmech.2020.103115](https://doi.org/10.1016/j.probengmech.2020.103115).
- Gu, Mengyang (2019). "Jointly Robust Prior for Gaussian Stochastic Process in Emulation, Calibration and Variable Selection". In: *Bayesian Analysis* 14.3, pp. 857–885. doi: [10.1214/18-BA1133](https://doi.org/10.1214/18-BA1133).
- Gu, Mengyang and James O. Berger (2016). "Parallel partial Gaussian process emulation for computer models with massive output". In: *The Annals of Applied Statistics* 10.3, pp. 1317–1347. doi: [10.1214/16-AOAS934](https://doi.org/10.1214/16-AOAS934).
- Gupta, Mayetri and Joseph G. Ibrahim (2009). "An information matrix prior for Bayesian Analysis in Generalized Linear Models with high dimensional data". In: *Statistica Sinica* 19.4, pp. 1641–1663. URL: <https://www.jstor.org/stable/24308922>.
- Haario, Heikki, Eero Saksman, and Johanna Tamminen (2001). "An Adaptive Metropolis Algorithm". In: *Bernoulli* 7.2, pp. 223–242. doi: [10.2307/3318737](https://doi.org/10.2307/3318737).
- Hariri-Ardebili, Mohammad A. and Victor E. Saouma (2016). "Probabilistic seismic demand model and optimal intensity measure for concrete dams". In: *Structural Safety* 59, pp. 67–85. doi: [10.1016/j.strusafe.2015.12.001](https://doi.org/10.1016/j.strusafe.2015.12.001).
- Hashimoto, Shintaro (2021). "Reference priors via alpha-divergence for a certain non-regular model in the presence of a nuisance parameter". In: *Journal of Statistical Planning and Inference* 213, pp. 162–178. doi: [10.1016/j.jspi.2020.11.007](https://doi.org/10.1016/j.jspi.2020.11.007).
- Hüllermeier, Eyke and Willem Waegeman (2019). "Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods". In: *Machine Learning* 110, pp. 457–506. doi: [10.1007/s10994-021-05946-3](https://doi.org/10.1007/s10994-021-05946-3).
- IAEA (2015). *The Fukushima Daiichi Accident. Technical Volume 2/5. Safety Assessment.* Vol. 2. 5 vols. Vienna (Austria): International Atomic Energy Agency. 186 pp.

- (2020). *Probabilistic Safety Assessment for Seismic Events*. TECDOC Series 1937. Vienna (Austria): International Atomic Energy Agency. URL: <https://www.iaea.org/publications/14744/probabilistic-safety-assessment-for-seismic-events>.
- Il Idrissi, Marouane, Nicolas Bousquet, Fabrice Gamboa, Bertrand Iooss, and Jean-Michel Loubes (2024). “Quantile-constrained Wasserstein projections for robust interpretability of numerical and machine learning models”. In: *Electronic Journal of Statistics* 18.2, pp. 2721–2770. DOI: [10.1214/24-EJS2268](https://doi.org/10.1214/24-EJS2268).
- Iooss, Bertrand (2009). “Contributions au traitement des incertitudes en modélisation numérique : propagation d’ondes en milieu aléatoire et analyse statistique d’expériences simulées”. Accreditation to supervise research. Université Paul Sabatier - Toulouse III. URL: <https://theses.hal.science/tel-00360995>.
- Iooss, Bertrand and Paul Lemaître (2015). “A Review on Global Sensitivity Analysis Methods”. In: *Uncertainty Management in Simulation-Optimization of Complex Systems: Algorithms and Applications*. Ed. by Gabriella Dellino and Carlo Meloni. Boston, MA: Springer US, pp. 101–122. DOI: [10.1007/978-1-4899-7547-8\\_5](https://doi.org/10.1007/978-1-4899-7547-8_5).
- Ito, Kiyosi and Henry P. McKean (1974). *Diffusion processes and their sample paths*. Berlin: Springer-Verlag.
- Jeffreys, Harold (1946). “An invariant form for the prior probability in estimation problems”. In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 186.1007, pp. 453–461. DOI: [10.1098/rspa.1946.0056](https://doi.org/10.1098/rspa.1946.0056).
- Jeon, Jong-Su, Sujith Mangalathu, Junho Song, and Reginald Desroches (2019). “Parameterized Seismic Fragility Curves for Curved Multi-frame Concrete Box-Girder Bridges Using Bayesian Parameter Estimation”. In: *Journal of Earthquake Engineering* 23.6, pp. 954–979. DOI: [10.1080/13632469.2017.1342291](https://doi.org/10.1080/13632469.2017.1342291).
- Kallioinen, Noa, Topi Paananen, and Paul-Christian Bürkner (2023). “Detecting and diagnosing prior and likelihood sensitivity with power-scaling”. In: *Statistics and Computing* 34.57, pp. 1573–1375. DOI: [10.1007/s11222-023-10366-5](https://doi.org/10.1007/s11222-023-10366-5).
- Kass, Robert E. and Larry A. Wasserman (1996). “The Selection of Prior Distributions by Formal Rules”. In: *Journal of the American Statistical Association* 91.435, pp. 1343–1370. DOI: [10.1080/01621459.1996.10477003](https://doi.org/10.1080/01621459.1996.10477003).
- Katayama, Yoshifumi, Yasuki Ohtori, Toshiaki Sakai, and Hitoshi Muta (2021). “Bayesian-estimation-based method for generating fragility curves for high-fidelity seismic probability risk assessment”. In: *Journal of Nuclear Science and Technology* 58.11, pp. 1220–1234. DOI: [10.1080/00223131.2021.1931517](https://doi.org/10.1080/00223131.2021.1931517).
- Kennedy, Robert P. (1999). “Risk based seismic design criteria”. In: *Nuclear Engineering and Design* 192.2, pp. 117–135. DOI: [10.1016/S0029-5493\(99\)00102-8](https://doi.org/10.1016/S0029-5493(99)00102-8).
- Kennedy, Robert P., C. Allin Cornell, Robert D. Campbell, Stan J. Kaplan, and Harold F. Perla (1980). “Probabilistic seismic safety study of an existing nuclear power plant”. In: *Nuclear Engineering and Design* 59.2, pp. 315–338. DOI: [10.1016/0029-5493\(80\)90203-4](https://doi.org/10.1016/0029-5493(80)90203-4).
- Kennedy, Robert P. and Mayasandra K. Ravindra (1984). “Seismic fragilities for nuclear power plant risk studies”. In: *Nuclear Engineering and Design* 79.1, pp. 47–68. DOI: [10.1016/0029-5493\(84\)90188-2](https://doi.org/10.1016/0029-5493(84)90188-2).
- Khansefid, Ali, Seyed Mahmoudreza Yadollahi, Francesca Taddei, and Gerhard Müller (2023). “Fragility and comfortability curves development and seismic risk assessment of a masonry building under earthquakes induced by geothermal power plants operation”. In: *Structural Safety* 103, p. 102343. DOI: [10.1016/j.strusafe.2023.102343](https://doi.org/10.1016/j.strusafe.2023.102343).
- Kiani, Jalal, Charles Camp, and Shahram Pezeshk (2019). “On the application of machine learning techniques to derive seismic fragility curves”. In: *Computers & Structures* 218, pp. 108–122. DOI: [10.1016/j.compstruc.2019.03.004](https://doi.org/10.1016/j.compstruc.2019.03.004).
- Kim, Sang-Hoon and Masanobu Shinozuka (2004). “Development of fragility curves of bridges retrofitted by column jacketing”. In: *Probabilistic Engineering Mechanics* 19.1, pp. 105–112. DOI: [10.1016/j.probengmech.2003.11.009](https://doi.org/10.1016/j.probengmech.2003.11.009).
- Kingma, Diederik P. and Max Welling (2014). “Auto-Encoding Variational Bayes”. In: *Proceedings of the 2nd International Conference on Learning Representations (ICLR)*. DOI: [10.48550/arXiv.1312.6114](https://arxiv.org/abs/1312.6114).
- (2019). “An Introduction to Variational Autoencoders”. In: *Foundations and Trends in Machine Learning* 12.4, pp. 307–392. DOI: [10.1561/2200000056](https://doi.org/10.1561/2200000056).
- Kobyzev, Ivan, Simon J. D. Prince, and Marcus A. Brubaker (2021). “Normalizing Flows: An Introduction and Review of Current Methods”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43.11, pp. 3964–3979. DOI: [10.1109/TPAMI.2020.2992934](https://doi.org/10.1109/TPAMI.2020.2992934).
- Kolmogorov, Andreï N. (1933). *Foundations of the Theory of Probability*. 2nd. Dover Books on Mathematics. Dover Publications.

## Bibliography

- Koutsourelakis, P. Steve (2010). "Assessing structural vulnerability against earthquakes using multi-dimensional fragility surfaces: A Bayesian framework". In: *Probabilistic Engineering Mechanics* 25.1, pp. 49–60. DOI: [10.1016/j.probengmech.2009.05.005](https://doi.org/10.1016/j.probengmech.2009.05.005).
- Kristan, Matej, Aleš Leonardis, and Danijel Skočaj (2011). "Multivariate online kernel density estimation with Gaussian kernels". In: *Pattern Recognition. Semi-Supervised Learning for Visual Content Analysis and Understanding* 44.10, pp. 2630–2642. DOI: [10.1016/j.patcog.2011.03.019](https://doi.org/10.1016/j.patcog.2011.03.019).
- Kwag, Shinyoung and Abhinav Gupta (2018). "Computationally efficient fragility assessment using equivalent elastic limit state and Bayesian updating". In: *Computers & Structures* 197, pp. 1–11. DOI: [10.1016/j.compstruc.2017.11.011](https://doi.org/10.1016/j.compstruc.2017.11.011).
- Lafferty, John D. and Larry A. Wasserman (2001). "Iterative Markov Chain Monte Carlo Computation of Reference Priors and Minimax Risk". In: *Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence (UAI)*. Ed. by Jack S. Breese and Daphne Koller. Morgan Kaufmann, pp. 293–300. DOI: [10.48550/arXiv.1301.2286](https://doi.org/10.48550/arXiv.1301.2286).
- Laguerre, Marc-Ansy, Mohammad Salehi, and Reginald Desroches (2024). "Empirical Fragility Analysis of Haitian Reinforced Concrete and Masonry Buildings". In: *Buildings* 14.3, p. 792. DOI: [10.3390/buildings-14030792](https://doi.org/10.3390/buildings-14030792).
- Lallemand, David, Anne S. Kiremidjian, and Henry Burton (2015). "Statistical procedures for developing earthquake damage fragility curves". In: *Earthquake Engineering & Structural Dynamics* 44.9, pp. 1373–1389. DOI: [10.1002/eqe.2522](https://doi.org/10.1002/eqe.2522).
- Lartaud, Paul, Philippe Humbert, and Josselin Garnier (2025). "Sequential design for surrogate modeling in Bayesian inverse problems". arXiv:2402.16520. DOI: [10.48550/arXiv.2402.16520](https://doi.org/10.48550/arXiv.2402.16520).
- Le, Tri Minh (2014). "The formal definition of reference priors under a general class of divergence". PhD thesis. University of Missouri-Columbia. DOI: [10.32469/10355/44185](https://doi.org/10.32469/10355/44185).
- Lee, Sangwoo, Shinyoung Kwag, and Bu-seog Ju (2023). "On efficient seismic fragility assessment using sequential Bayesian inference and truncation scheme: A case study of shear wall structure". In: *Computers & Structures* 289, p. 107150. DOI: [10.1016/j.compstruc.2023.107150](https://doi.org/10.1016/j.compstruc.2023.107150).
- Lehmann, Erich L. (1999). *Elements of Large-Sample Theory*. 1st ed. Springer Texts in Statistics. Springer New York, NY. DOI: [10.1007/b98855](https://doi.org/10.1007/b98855).
- Liang, Feng, Rui Paulo, German Molina, Merlise A. Clyde, and James O. Berger (2008). "Mixtures of g Priors for Bayesian Variable Selection". In: *Journal of the American Statistical Association* 103.481, pp. 410–423. DOI: [10.1198/016214507000001337](https://doi.org/10.1198/016214507000001337).
- Lindley, David V. (1956). "On a Measure of the Information Provided by an Experiment". In: *The Annals of Mathematical Statistics* 27.4, pp. 986–1005. DOI: [10.1214/aoms/1177728069](https://doi.org/10.1214/aoms/1177728069).
- Liu, Ruitao, Arijit Chakrabarti, Tapas Samanta, Jayanta K. Ghosh, and Malay Ghosh (2014). "On Divergence Measures Leading to Jeffreys and Other Reference Priors". In: *Bayesian Analysis* 9.2, pp. 331–370. DOI: [10.1214/14-BA862](https://doi.org/10.1214/14-BA862).
- Luco, Nicolas and C. Allin Cornell (2007). "Structure-Specific Scalar Intensity Measures for Near-Source and Ordinary Earthquake Ground Motions". In: *Earthquake Spectra* 23.2, pp. 357–392. DOI: [10.1193/1.2723158](https://doi.org/10.1193/1.2723158).
- MacKay, David J. C. (2003). *Information Theory, Inference, and Learning Algorithms*. Copyright Cambridge University Press.
- Mackie, Kevin and Božidar Stojadinović (2001). "Probabilistic Seismic Demand Model for California Highway Bridges". In: *Journal of Bridge Engineering* 6.6, pp. 468–481. DOI: [10.1061/\(ASCE\)1084-0702\(2001\)6:6\(468\)](https://doi.org/10.1061/(ASCE)1084-0702(2001)6:6(468)).
- Mai, Chu, Katerina Konakli, and Bruno Sudret (2017). "Seismic fragility curves for structures using non-parametric representations". In: *Frontiers of Structural and Civil Engineering* 11.2, pp. 169–186. DOI: [10.1007/s11709-017-0385-y](https://doi.org/10.1007/s11709-017-0385-y).
- Mai, Chu, Minas. D. Spiridonakos, Eleni Chatzi, and Bruno Sudret (2016). "Surrogate modeling for stochastic dynamical systems by combining nonlinear autoregressive with exogenous input models and polynomial chaos expansions". In: *International Journal for Uncertainty Quantification* 6, pp. 313–339. DOI: [10.1615/Int.J.UncertaintyQuantification.2016016603](https://doi.org/10.1615/Int.J.UncertaintyQuantification.2016016603).
- Mandal, Tushar K., Siddhartha Ghosh, and Nikil N. Pujari (2016). "Seismic fragility analysis of a typical Indian PHWR containment: Comparison of fragility models". In: *Structural Safety* 58, pp. 11–19. DOI: [10.1016/j.strusafe.2015.08.003](https://doi.org/10.1016/j.strusafe.2015.08.003).

- Mikkola, Petrus et al. (2023). "Prior Knowledge Elicitation: The Past, Present, and Future". In: *Bayesian Analysis*, pp. 1–33. DOI: [10.1214/23-BA1381](https://doi.org/10.1214/23-BA1381).
- Minka, Tom (2005). *Divergence Measures and Message Passing*. MSR-TR-2005-173. Microsoft Ltd, p. 17. URL: <https://www.microsoft.com/en-us/research/publication/divergence-measures-and-message-passing/>.
- Mironov, Ilya (2017). "Rényi Differential Privacy". In: *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*, pp. 263–275. DOI: [10.1109/CSF.2017.11](https://doi.org/10.1109/CSF.2017.11).
- Mitropoulou, Chara Ch. and Manolis Papadrakakis (2011). "Developing fragility curves based on neural network IDA predictions". In: *Engineering Structures* 33.12, pp. 3409–3421. DOI: [10.1016/j.engstruct.2011.07.005](https://doi.org/10.1016/j.engstruct.2011.07.005).
- Muré, Joseph (2018). "Objective Bayesian analysis of Kriging models with anisotropic correlation kernel". PhD thesis. Université Sorbonne Paris Cité. URL: <https://theses.hal.science/tel-02184403v1>.
- Nalisnick, Eric and Padhraic Smyth (2017). "Learning Approximately Objective Priors". In: *Proceedings of the 33rd Conference on Uncertainty in Artificial Intelligence (UAI)*. Association for Uncertainty in Artificial Intelligence (AUAI). DOI: [10.48550/arXiv.1704.01168](https://doi.org/10.48550/arXiv.1704.01168).
- Natarajan, Ranjini and Robert E. Kass (2000). "Reference Bayesian Methods for Generalized Linear Mixed Models". In: *Journal of the American Statistical Association* 95.449, pp. 227–237. DOI: [10.1080/01621459.2000.10473916](https://doi.org/10.1080/01621459.2000.10473916).
- Neyman, Jerzy and Elizabeth L. Scott (1948). "Consistent Estimates Based on Partially Consistent Observations". In: *Econometrica* 16.1, pp. 1–32. DOI: [10.2307/1914288](https://doi.org/10.2307/1914288).
- Nguyen, XuanLong, Martin J. Wainwright, and Michael I. Jordan (2009). "On surrogate loss functions and f-divergences". In: *The Annals of Statistics* 37.2, pp. 876–904. DOI: [10.1214/08-AOS595](https://doi.org/10.1214/08-AOS595).
- NoCEDAL, Jorge and Stephen J. Wright (2006). "Numerical optimization". In: *Springer Series in Operations Research and Financial Engineering*. Section: Penalty and Augmented Lagrangian Methods. Springer New York, pp. 497–528. DOI: [10.1007/978-0-387-40065-5\\_17](https://doi.org/10.1007/978-0-387-40065-5_17).
- Noh, Hae Young, David Lallement, and Anne S. Kiremidjian (2014). "Development of empirical and analytical fragility functions using kernel smoothing methods". In: *Earthquake Engineering & Structural Dynamics* 44.8, pp. 1163–1180. DOI: [10.1002/eqe.2505](https://doi.org/10.1002/eqe.2505).
- Nott, David J., Xueou Wang, Michael Evans, and Berthold-Georg Englert (2020). "Checking for Prior-Data Conflict Using Prior-to-Posterior Divergences". In: *Statistical Science* 35.2, pp. 234–253. DOI: [10.1214/19-STS731](https://doi.org/10.1214/19-STS731).
- NRC (1983). *PRA Procedures Guide: A Guide to the Performance of Probabilistic Risk Assessments for Nuclear Power Plants*. NUREG/CR-2300. Nuclear Regulatory Commission. URL: <https://www.nrc.gov/reading-rm/doc-collections/nuregs/contract/cr2300/vol1/index.html>.
- Papamakarios, George, Eric Nalisnick, Danilo Jimenez Rezende, Shakir Mohamed, and Balaji Lakshminarayanan (2021). "Normalizing Flows for Probabilistic Modeling and Inference". In: *Journal of Machine Learning Research* 22.57, pp. 1–64. URL: <http://jmlr.org/papers/v22/19-1028.html>.
- Park, Joonam and Peeranan Towashiraporn (2014). "Rapid seismic damage assessment of railway bridges using the response-surface statistical model". In: *Structural Safety* 47, pp. 1–12. DOI: [10.1016/j.strusafe.2013.10.001](https://doi.org/10.1016/j.strusafe.2013.10.001).
- Park, Young-Ji, Charles H. Hofmayer, and Nilesh C. Chokshi (1998). "Survey of seismic fragilities used in PRA studies of nuclear power plants". In: *Reliability Engineering & System Safety* 62.3, pp. 185–195. DOI: [10.1016/S0951-8320\(98\)00019-2](https://doi.org/10.1016/S0951-8320(98)00019-2).
- Paulo, Rui (2005). "Default priors for Gaussian processes". In: *The Annals of Statistics* 33.2, pp. 556–582. DOI: [10.1214/009053604000001264](https://doi.org/10.1214/009053604000001264).
- Pérez, José M. and James O. Berger (2002). "Expected-posterior prior distributions for model selection". In: *Biometrika* 89.3, pp. 491–512. DOI: [10.1093/biomet/89.3.491](https://doi.org/10.1093/biomet/89.3.491).
- Picard-Weibel, Antoine and Benjamin Guedj (2022). "On change of measure inequalities for f-divergences". arXiv:2202.05568. DOI: [10.48550/arXiv.2202.05568](https://doi.org/10.48550/arXiv.2202.05568).
- Radu, Alin and Mircea Dan Grigoriu (2018). "An earthquake-source-based metric for seismic fragility analysis". In: *Bulletin of Earthquake Engineering* 16. DOI: [10.1007/s10518-018-0341-9](https://doi.org/10.1007/s10518-018-0341-9).
- Renyi, Alfred (1970). *Foundations of Probability*. Dover books on mathematics. Dover Publications.

## Bibliography

- Rezaeian, Sanaz (2010). "Stochastic Modeling and Simulation of Ground Motions for Performance-Based Earthquake Engineering". PhD thesis. UC Berkeley. URL: <https://escholarship.org/uc/item/51m3669m>.
- Rezaeian, Sanaz and Armen Der Kiureghian (2008). "A stochastic ground motion model with separable temporal and spectral nonstationarities". In: *Earthquake Engineering & Structural Dynamics* 37.13, pp. 1565–1584. DOI: [10.1002/eqe.831](https://doi.org/10.1002/eqe.831).
- (2010). "Simulation of synthetic ground motions for specified earthquake and site characteristics". In: *Earthquake Engineering & Structural Dynamics* 39.10, pp. 1155–1180. DOI: [10.1002/eqe.997](https://doi.org/10.1002/eqe.997).
- Robert, Christian (2007). *The Bayesian Choice*. 2nd ed. Texts in Statistics. Springer.
- Roger, Mathias (2020). "Le séisme, la centrale et la règle : instaurer et maintenir la robustesse des installations nucléaires en France". PhD thesis. Université Paris Cité. URL: <https://theses.hal.science/tel-03414838>.
- Rubio, Francisco J. and Mark F. J. Steel (2014). "Inference in Two-Piece Location-Scale Models with Jeffreys Priors". In: *Bayesian Analysis* 9.1, pp. 1–22. DOI: [10.1214/13-BA849](https://doi.org/10.1214/13-BA849).
- Saint, Rémi, Cyril Feau, Jean-Marc Martinez, and Josselin Garnier (2020). "Efficient methodology for seismic fragility curves estimation by active learning on Support Vector Machines". In: *Structural Safety* 86, p. 101972. DOI: [10.1016/j.strusafe.2020.101972](https://doi.org/10.1016/j.strusafe.2020.101972).
- Schölkopf, Bernhard and Alexander J. Smola (2001). *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. The MIT Press. DOI: [10.7551/mitpress/4175.001.0001](https://doi.org/10.7551/mitpress/4175.001.0001).
- Seo, Junwon and Daniel G. Linzell (2013). "Use of response surface metamodels to generate system level fragilities for existing curved steel bridges". In: *Engineering Structures* 52, pp. 642–653. DOI: [10.1016/j.engstruct.2013.03.023](https://doi.org/10.1016/j.engstruct.2013.03.023).
- Shinozuka, Masanobu and George Deodatis (1988). "Stochastic process models for earthquake ground motion". In: *Probabilistic Engineering Mechanics* 3.3, pp. 114–123. DOI: [10.1016/0266-8920\(88\)90023-9](https://doi.org/10.1016/0266-8920(88)90023-9).
- Shinozuka, Masanobu, Maria Qing Feng, Jongheon Lee, and Toshihiko Naganuma (2000). "Statistical Analysis of Fragility Curves". In: *Journal of Engineering Mechanics* 126.12, pp. 1224–1231. DOI: [10.1061/\(ASCE\)0733-9399\(2000\)126:12\(1224\)](https://doi.org/10.1061/(ASCE)0733-9399(2000)126:12(1224)).
- Simpson, Daniel, Håvard Rue, Andrea Riebler, Thiago G. Martins, and Sigrunn H. Sørbye (2017). "Penalising Model Component Complexity: A Principled, Practical Approach to Constructing Priors". In: *Statistical Science* 32.1, pp. 1–28. DOI: [10.1214/16-STS576](https://doi.org/10.1214/16-STS576).
- Sobol', Ilya M. (1993). "Sensitivity estimates for non linear mathematical models". In: *Mathematical Modeling and Computer Experiments* 1, pp. 407–414.
- Spitzner, Dan J. (2011). "Neutral-data comparisons for Bayesian testing". In: *Bayesian Analysis* 6.4, pp. 603–638. DOI: [10.1214/11-BA623](https://doi.org/10.1214/11-BA623).
- Straub, Daniel and Armen Der Kiureghian (2008). "Improved seismic fragility modeling from empirical data". In: *Structural Safety* 30.4, pp. 320–336. DOI: [10.1016/j.strusafe.2007.05.004](https://doi.org/10.1016/j.strusafe.2007.05.004).
- Sudret, Bruno (2007). "Uncertainty propagation and sensitivity analysis in mechanical models - Contributions to structural reliability and stochastic spectral methods". Accreditation to supervise research. Université Blaise Pascal - Clermont II. URL: <https://ethz.ch/content/dam/ethz/special-interest/baug/ibk/risk-safety-and-uncertainty-dam/publications/reports/HDRSudret.pdf>.
- Tabandeh, Armin, Pouyan Asem, and Paolo Gardoni (2020). "Physics-based probabilistic models: Integrating differential equations and observational data". In: *Structural Safety* 87, p. 101981. DOI: [10.1016/j.strusafe.2020.101981](https://doi.org/10.1016/j.strusafe.2020.101981).
- Tadinada, Sashi Kanth and Abhinav Gupta (2017). "Structural fragility of T-joint connections in large-scale piping systems using equivalent elastic time-history simulations". In: *Structural Safety* 65, pp. 49–59. DOI: [10.1016/j.strusafe.2016.12.003](https://doi.org/10.1016/j.strusafe.2016.12.003).
- Taraldsen, Gunnar and Bo Henry Lindqvist (2016). "Conditional probability and improper priors". In: *Communications in Statistics - Theory and Methods* 45.17, pp. 5007–5016. DOI: [10.1080/03610926.2014.935430](https://doi.org/10.1080/03610926.2014.935430).
- Touboul, Françoise, Nadine Blay, Pierre Sollogoub, and Stéphane Chapuliot (2006). "Enhanced seismic criteria for piping". In: *Nuclear Engineering and Design* 236.1, pp. 1–9. DOI: [10.1016/j.nucengdes.2005.07.002](https://doi.org/10.1016/j.nucengdes.2005.07.002).
- Touboul, Françoise, Pierre Sollogoub, and Nadine Blay (1999). "Seismic behaviour of piping systems with and without defects: experimental and numerical evaluations". In: *Nuclear Engineering and Design* 192.2, pp. 243–260. DOI: [10.1016/S0029-5493\(99\)00111-9](https://doi.org/10.1016/S0029-5493(99)00111-9).

- Trevlopoulos, Konstantinos, Cyril Feau, and Irmela Zentner (2019). "Parametric models averaging for optimized non-parametric fragility curve estimation based on intensity measure data clustering". In: *Structural Safety* 81, p. 101865. doi: [10.1016/j.strusafe.2019.05.002](https://doi.org/10.1016/j.strusafe.2019.05.002).
- Vaart, Aad van der (1992). *Asymptotic statistics*. 1st ed. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.
- Van Biesbroeck, Antoine (2024a). "Generalized mutual information and their reference priors under Csizar f-divergence". arXiv:2310.10530. doi: [10.48550/arXiv.2310.10530](https://doi.org/10.48550/arXiv.2310.10530).
- (2024b). "Properly constrained reference priors decay rates for efficient and robust posterior inference". arXiv:2409.13041. doi: [10.48550/arXiv.2409.13041](https://doi.org/10.48550/arXiv.2409.13041).
- Van Biesbroeck, Antoine, Cyril Feau, and Josselin Garnier (2025). "Design of experiments for efficient and conform Bayesian learning of seismic fragility curves". In: *Proceedings of the 28th conference on Structural Mechanics in Reactor Techology (SMiRT)*.
- Van Biesbroeck, Antoine, Clément Gauchy, Cyril Feau, and Josselin Garnier (2023). "Influence of the choice of the seismic intensity measure on fragility curves estimation in a Bayesian framework based on reference prior". In: *Proceedings of the 5th UNCECOMP Conference*, pp. 94–111. doi: [10.7712/120223.10327.19899](https://doi.org/10.7712/120223.10327.19899).
- (2024). "Reference prior for Bayesian estimation of seismic fragility curves". In: *Probabilistic Engineering Mechanics* 76, p. 103622. doi: [10.1016/j.probengmech.2024.103622](https://doi.org/10.1016/j.probengmech.2024.103622).
- (2025a). "Design of experiment based on a low fidelity model for seismic fragility estimation". In: *ESAIM: ProcS* (to appear). URL: <https://hal.science/hal-04719458v1>.
- (2025b). "Robust a posteriori estimation of probit-lognormal seismic fragility curves via sequential design of experiments and constrained reference prior". arXiv:2503.07343. doi: [10.48550/arXiv.2503.07343](https://doi.org/10.48550/arXiv.2503.07343).
- Villemonteix, Julien, Emmanuel Vazquez, and Eric Walter (2009). "An informational approach to the global optimization of expensive-to-evaluate functions". In: *Journal of Global Optimization* 44.4, pp. 509–534. doi: [10.1007/s10898-008-9354-2](https://doi.org/10.1007/s10898-008-9354-2).
- Wang, Fan and Cyril Feau (2020). "Influence of Input Motion's Control Point Location in Nonlinear SSI Analysis of Equipment Seismic Fragilities: Case Study on the Kashiwazaki-Kariwa NPP". In: *Pure and Applied Geophysics* 177, pp. 2391–2409. doi: [10.1007/s00024-020-02467-3](https://doi.org/10.1007/s00024-020-02467-3).
- Wang, Zhiyi, Nicola Pedroni, Irmela Zentner, and Enrico Zio (2018). "Seismic fragility analysis with artificial neural networks: Application to nuclear power plant equipment". In: *Engineering Structures* 162, pp. 213–225. doi: [10.1016/j.engstruct.2018.02.024](https://doi.org/10.1016/j.engstruct.2018.02.024).
- Wang, Zhiyi, Irmela Zentner, and Enrico Zio (2018). "A Bayesian framework for estimating fragility curves based on seismic damage data and numerical simulations by adaptive neural networks". In: *Nuclear Engineering and Design* 338, pp. 232–246. doi: [10.1016/j.nucengdes.2018.08.016](https://doi.org/10.1016/j.nucengdes.2018.08.016).
- Wimmer, Lisa, Yusuf Sale, Paul Hofman, Bernd Bischl, and Eyke Hüllermeier (2023). "Quantifying aleatoric and epistemic uncertainty in machine learning: Are conditional entropy and mutual information appropriate measures?" In: *Proceedings of the Thirty-Ninth Conference on Uncertainty in Artificial Intelligence*. Vol. 216. PMLR, pp. 2282–2292. URL: <https://proceedings.mlr.press/v216/wimmer23a.html>.
- Winkelbauer, Andreas (2014). "Moments and Absolute Moments of the Normal Distribution". arXiv:1209.4340. doi: [10.48550/arXiv.1209.4340](https://doi.org/10.48550/arXiv.1209.4340).
- Xie, Qun and Andrew R. Barron (1997). "Minimax redundancy for the class of memoryless sources". In: *IEEE Transactions on Information Theory* 43.2, pp. 646–657. doi: [10.1109/18.556120](https://doi.org/10.1109/18.556120).
- Zeidler, Eberhard (2012). *Applied functional Analysis*. 1st. Applied Mathematical Sciences. Springer New York, NY. doi: [10.1007/978-1-4612-0821-1](https://doi.org/10.1007/978-1-4612-0821-1).
- Zentner, Irmela (2010). "Numerical computation of fragility curves for NPP equipment". In: *Nuclear Engineering and Design* 240.6, pp. 1614–1621. doi: [10.1016/j.nucengdes.2010.02.030](https://doi.org/10.1016/j.nucengdes.2010.02.030).
- Zentner, Irmela, Max Gundel, and Nicolas Bonfils (2017). "Fragility analysis methods: Review of existing approaches and application". In: *Nuclear Engineering and Design* 323, pp. 245–258. doi: [10.1016/j.nucengdes.2016.12.021](https://doi.org/10.1016/j.nucengdes.2016.12.021).
- Zentner, Irmela and Fabrice Poirion (2012). "Enrichment of seismic ground motion databases using Karhunen-Loeve expansion". In: *Earthquake Engineering & Structural Dynamics* 41.14, pp. 1945–1957. doi: [10.1002/eqe.2166](https://doi.org/10.1002/eqe.2166).
- Zerva, Aspasia (1988). "Seismic source mechanisms and ground motion models". In: *Probabilistic Engineering Mechanics* 3.2, pp. 64–74. doi: [10.1016/0266-8920\(88\)90017-3](https://doi.org/10.1016/0266-8920(88)90017-3).

- Zhao, Chnfeng, Na Yu, and Yilung Mo (2020). "Seismic fragility analysis of AP1000 SB considering fluid-structure interaction effects". In: *Structures* 23, pp. 103–110. doi: [10.1016/j.istruc.2019.11.003](https://doi.org/10.1016/j.istruc.2019.11.003).
- Zhu, Huaiyu and Richard Rohwer (1995). *Information geometric measurements of generalisation*. Monograph NCRG/4350. Place: Birmingham, UK Num Pages: 403709 Publisher: Aston University. Neural Computing Research Group. URL: <https://publications.aston.ac.uk/id/eprint/507/>.
- Zhu, Xujia, Marco Broccardo, and Bruno Sudret (2023). "Seismic fragility analysis using stochastic polynomial chaos expansions". In: *Probabilistic Engineering Mechanics* 72, p. 103413. doi: [10.1016/j.probengmech.2023.103413](https://doi.org/10.1016/j.probengmech.2023.103413).









**Titre:** Théorie des priors de référence étendue pour une inférence objective et pratique, application à l'estimation robuste et auditable de courbes de fragilité sismique

**Mots clés:** Prior objectif, Analyse bayésienne, Aléa sismique, Courbe de fragilité, Quantification des incertitudes

**Résumé:** La théorie des priors de référence fournit un cadre approprié à une inférence bayésienne objective, puisqu'elle vise à minimiser la subjectivité introduite et à permettre aux informations issues des données d'orienter la distribution des estimations. Pour cette raison, l'application de cette théorie à l'estimation des courbes de fragilité sismique est particulièrement pertinente. En effet, ces courbes sont des éléments essentiels des études sismiques probabilistes de sûreté ; elles expriment la probabilité de défaillance d'une structure mécanique en fonction d'indicateurs définissant des scénarios sismiques. Puisqu'elles informent des décisions critiques en matière de sécurité des infrastructures, une auditabilité complète de l'approche qui conduit aux estimations de ces courbes est nécessaire.

Cette thèse étudie l'interaction entre la théorie des priors de référence et l'estimation des courbes de fragilité sismique, apportant des contributions originales dans ces deux domaines. Tout d'abord, nous complétons les fondements théoriques des priors de référence en développant de nouvelles constructions de ceux-ci. Notre objectif est de soutenir leur objectivité tout en améliorant leur applicabilité pratique. Nos résultats prennent la forme de contributions théoriques dans ce domaine qui sont basées sur une définition généralisée de

l'information mutuelle. Nos approches abordent les principaux problèmes des priors de référence, à savoir le caractère impropre de leur distribution ou de leur distribution *a posteriori*, et leur formulation complexe pour une utilisation pratique.

Ensuite, nous revisitons l'estimation des courbes de fragilité sismique basée sur le modèle probit-lognormal dans un contexte où les données sont particulièrement rares. Notre objectif est de réaliser une estimation bayésienne des courbes de fragilité qui tire parti de l'optimisation de toutes sortes d'informations, y compris l'information *a priori*, afin de fournir des estimations robustes et auditables. Nos résultats mettent en évidence les limites et les irrégularités du modèle et proposent des méthodes qui fournissent des estimations précises et efficaces des courbes. Les évaluations de nos approches sont réalisées sur différents cas d'étude issus de l'industrie nucléaire.

Cette thèse établit un lien fort entre ces deux domaines. L'application aux courbes de fragilité sismique a non seulement motivé les développements théoriques mais leur a aussi directement profité, produisant finalement un cadre d'estimation plus robuste, interprétable et vérifiable.

**Title:** Extended reference prior theory for objective and practical inference, application to robust and auditable seismic fragility curves estimation

**Keywords:** Objective prior, Bayesian analysis, Seismic hazard, Fragility curve, Uncertainties quantification

**Abstract:** Reference prior theory provides a principled framework for objective Bayesian inference, aiming to minimize subjective input and allow data-based information to drive the estimates distribution. For this reason, the application of this theory to the estimation of seismic fragility curves is particularly relevant. Indeed, these curves are essential elements of seismic probabilistic risk assessment studies; they express the probability of failure of a mechanical structure as a function of indicators that define seismic scenarios. Since they inform critical decisions in infrastructure safety, a complete auditability of the pipeline that leads to the estimates of these curves is required.

This thesis investigates the interplay between reference prior theory and seismic fragility curves estimation, yielding original contributions in these two domains. First, we complement the theoretical foundations of reference priors by developing novel constructions of them. Our goal is to support their objectivity while improving their practical applicability. Our results take the form of theoretical contributions in this domain that are based on a general-

ized definition of the mutual information. Our approaches tackle the principal issues of reference priors, namely their improper characteristic or that of their posterior, and their complex formulation for practical use.

Second, we revisit the estimation of seismic fragility curves based on the prominent probit-lognormal model in a context where the data are particularly sparse. Our goal is to conduct a Bayesian estimation of seismic fragility curves that leverages the optimization of every sort of information, including the *a priori* one, in order to provide estimates that are robust and auditable. Our results highlight the limitations and irregularities of the model and propose methods that provide accurate and efficient estimates of the curves. The evaluations of our approaches are carried out on different case studies taken from the nuclear industry.

This thesis builds a strong link between these two domains. The application to seismic fragility curves not only motivated theoretical developments but also directly benefited them, ultimately producing a more robust, interpretable, and verifiable estimation framework.

