



Technische Universität München

Fakultät für Mathematik
Lehrstuhl für Numerische Mathematik
(Prof. Dr. Barbara Wohlmuth)

Active Subspaces in Bayesian Inverse Problems

Mario Manuel Teixeira Parente, M.Sc.

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzende: Prof. Dr. Christina Kuttler

Prüfer der Dissertation:

1. Prof. Dr. Barbara Wohlmuth
2. Prof. Dr. Krzysztof Podgórski

Die Dissertation wurde am 27.05.2020 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 15.08.2020 angenommen.

Science, my lad, is made up of mistakes,
but they are mistakes which it is useful to
make, because they lead little by little to
the truth.

— **Jules Verne** (1828–1905)

Vieles hätte ich verstanden – hätte man es
mir nicht erklärt.

— **Stanisław Jerzy Lec** (1909–1966)

Most people say that it is the intellect
which makes a great scientist. They are
wrong: it is character.

— **Albert Einstein** (1879–1955)

Abstract

Bayesian inverse problems constitute an important methodological part of parameter studies in predictive modeling. Since high-dimensional parameter spaces often require a huge computational effort, approaches for cost reduction are of fundamental importance. For example, the costs can be substantially reduced by exploiting low-dimensional structure in the form of *active subspaces*. In this work, we derive new generalized bounds, develop an iterative algorithm, and demonstrate computational benefits with models from various applied disciplines.

Zusammenfassung

In der prädiktiven Modellierung stellen *Bayes'sche inverse Probleme* einen wichtigen methodischen Bestandteil von Parameterstudien dar. Da hierbei in hochdimensionalen Parameterräumen oft ein großer Rechenaufwand entsteht, sind Methodiken zur Kostenreduktion von fundamentaler Bedeutung. Eine Möglichkeit besteht darin, niederdimensionale Strukturen in der Form von *aktiven Unterräumen* auszunutzen. In dieser Arbeit werden neue verallgemeinerte Schranken hergeleitet, ein iterativer Algorithmus entwickelt und die Robustheit und Flexibilität der vorgestellten Methodik anhand von Modellen aus verschiedenen Anwendungsdisziplinen demonstriert.

Scientific contributions

Publications

• Included & first author

Authors: TP., M., Mattis, S., Gupta, S., Deusner, C., & Wohlmuth, B.
Title: Efficient Parameter Estimation for a Methane Hydrate Model with Active Subspaces
Journal: *Computational Geosciences*, 23(2):355–372, 2019
Reference: [256]
DOI: 10.1007/s10596-018-9769-x
Included in: Chapter 5

Authors: TP., M., Bittner, D., Mattis, S., Chiogna, G., & Wohlmuth, B.
Title: Bayesian Calibration and Sensitivity Analysis for a Karst Aquifer Model using Active Subspaces
Journal: *Water Resources Research*, 55(8):7086–7107, 2019
Reference: [258]
DOI: 10.1029/2019WR024739
Included in: Chapter 6

Authors: TP., M., Wallin, J., & Wohlmuth, B.
Title: Generalized Bounds for Active Subspaces
Journal: *Electronic Journal of Statistics*, 14(1):917–943, 2020
Reference: [257]
DOI: 10.1214/20-EJS1684
Included in: Section 3.3 and 3.4

• Not included

Authors: Bittner, D., TP., M., Mattis, S., Wohlmuth, B., & Chiogna, G.
Title: Identifying Relevant Hydrological and Catchment Properties in Active Subspaces: An Inference Study of a Lumped Karst Aquifer Model
Journal: *Advances in Water Resources*, 135:103472, 2020
Reference: [32]
DOI: 10.1016/j.advwatres.2019.103472

Talks, Conferences

• Speaker

Format: Invited oral presentation in a seminar
Name: Statistics seminar at the Department of Statistics, Lund University
Location: Lund, Sweden
Date: May 29, 2019
Talk: Active Subspaces in Bayesian Inverse Problems
Authors: **TP.**, **M.**, Mattis, S., & Wohlmuth, B.

Format: Oral presentation in a minisymposium at a conference
Name: SIAM Conference on Uncertainty Quantification (UQ20)
Minisymposium *UQ in Hydrology* (organized by UNMIX group)
Location: Garching near Munich
Date: March 24–27, 2020
Talk: Solving a Bayesian Inverse Problem for a Karst Aquifer Model with Active Subspaces
Authors: **TP.**, **M.**, Bittner, D., Mattis, S., Chiogna, G., & Wohlmuth, B.
Note: Our abstract was already accepted for the minisymposium.
Unfortunately, the conference was **canceled** on March 7, 2020 due to the outbreak of SARS-CoV-2.

• Contributor

Format: Contributed presentation at a conference
Name: SIAM Conference on Mathematical & Computational Issues in the Geosciences (GS19)
Location: Houston, Texas, USA
Date: March 11–14, 2019
Talk: Can a High Dimensional Parameter Space be Acceptable for a Lumped Hydrological Karst Model?
Authors: Chiogna, G., Bittner, D., **TP.**, **M.**, Mattis, S., & Wohlmuth, B.

Format: Contributed presentation at a conference
Name: EGU General Assembly 2019
Location: Vienna, Austria
Date: April 7–12, 2019
Talk: Parameter Dimension Reduction using the Active Subspace Method for a Lumped Karst Aquifer Model
Authors: Bittner, D., **TP.**, **M.**, Mattis, S., Wohlmuth, B., & Chiogna, G.

Format: Oral presentation at a conference
Name: 46th IAH Congress
Location: Malaga, Spain
Date: September 22–27, 2019
Talk: How to Model the Impacts of Land Use Changes in Karstic Environments: Model Development, Parameter Dimension Reduction and Uncertainty Quantification
Authors: Bittner, D., **TP.**, **M.**, Mattis, S., Wohlmuth, B., Disse, M., & Chiogna, G.

Preface

Journey as a PhD candidate

My supervisor Prof. Dr. Barbara Wohlmuth accepted me as a PhD candidate at the TUM Chair for Numerical Mathematics in late September 2016 such that I was able to start in mid-October 2016 with a third-party position.

In the first 8–10 months, I was looking for a topic which is 1) interesting enough to stay concentrated over a longer period, 2) challenging from a scientific and mathematical perspective, and 3) doable for me with my skills. Such a crucial decision is not easy to make, especially being a finished but only half-baked Master’s student. Since I was totally free in my choice, basically each project at the chair at that time could provide a potential topic. After some time, I decided for a topic at the interface of, say, statistics, scientific computing, mathematical modeling, and applications from other scientific disciplines like engineering or physics.

I have received a lot of support in getting familiar with the contents by my main technical mentor Steven (“Steve”) Mattis (former postdoc at TUM) who provided significant guidance and extraordinary patience for the first 2.5 years of my time at the chair; but also Jonas (“Jones”) Latz (former PhD student at TUM, now postdoc at University of Cambridge), with his broad but at the same time deep knowledge, has considerably contributed to my understanding.

Each PhD student at TUM is part of the TUM Graduate School and has to announce an official mentor who should be fairly independent of the structure at the department. In my case, it is Prof. Gerta Köster (University of Applied Sciences Munich) who agreed to take this responsibility. Although it was not necessary to get in contact with her to talk about serious issues, I know that she would have been open to meet and discuss at any time.

From July 2017 on, Barbara Wohlmuth offered a position of the department which also includes teaching and (IT) administrative tasks at the chair. Of course, I accepted since teaching can be an asset in many ways; also the IT administration that I have done together with Prof. Rainer Callies (TUM Chair for Numerical Mathematics) suited my competence due to my background as a software developer.

Teaching mainly encompassed preparing exercise sheets and giving tutorials for students. On the one hand, it was a welcome balancing factor to compensate frustration from cumbersome research topics and it often made me feel like having the opportunity of a direct influence on students and their mathematical education. On the other hand, (good) teaching is time consuming and it was sometimes difficult to make a compromise that satisfies both, teaching and research. Courses that I taught were Numerical Linear Algebra, Numerics of ODEs, and ODEs/PDEs for engineering students. Also, Barbara Wohlmuth and I have been part of a two week summer school in South-Tirol for students from different universities in September 2018,

where our group worked on several aspects of computer simulation and modeling.

In addition to the everyday life as a PhD student, I was allowed to take part at a long term (~ 1.5 years) teaching course offered by *TUM ProLehre* for teachers from the math department. I was really enjoying working with this group of PhD students and postdocs that also joined for the lessons. It has always been informative, educating, and fun, especially due to the huge effort of the organizers Rudolf Aichner, Emil Ratko-Dehnert, Vanessa Landgraf (former Krummeck), and Pit Forster.

In March 2018, I joined the *International Graduate School of Science and Engineering* (IGSSE) for the project *UNMIX* (Uncertainties due to boundary conditions in predicting mixing in groundwater) that was initiated in cooperation with the TUM Chair of Hydrology and River Basin Management and the Oden Institute Computational Hydraulics Group from the University of Texas at Austin (UT Austin). The IGSSE qualification program consists of several science-related and soft skill trainings. One mandatory element is an international research phase in which a PhD student is supposed to gain experience at another (or several) research institution(s) abroad for a longer period.

The first stay was at UT Austin for six weeks in February and March 2019 with my project partners Mónica Basilio Hazas, Pablo Merchan Rivera, and Daniel Bittner (all TUM Chair of Hydrology and River Basin Management). We were all able to get insight into the academic environment of a big U.S. American university and the scientific work of the research group lead by Prof. Clint Dawson. It was a personal challenge for me to stay away from home for a longer period, but, in retrospect, it was a great experience with several positive memories.

Shortly after coming home from Austin, in early April 2019, I was allowed to be part of hydrological field measurements in Trento (Italy). For me as a mathematician, it was very interesting to see how hydraulic heads are practically measured in cooperation and agreement with local authorities. So far, the measured quantities have only been of abstract mathematical type in my mind.

I did the second stay abroad at the Department of Statistics from Lund University in Lund (Sweden) under the supervision of Prof. Krzysztof Podgórski. It was split into two parts; the first was end of May 2019 and the second took place in August 2019. I learned a lot about statistical thinking and really enjoyed the time there, especially because I felt very welcome at the department over the whole period.

Science

The topic of my PhD project is (very) broadly speaking from the field of *Uncertainty Quantification* (UQ) and combines two facets of modern scientific computing. It investigates *Bayesian inverse problems*, i.e., the derivation and computation of (informed and informative) probability distributions on model inputs, and their low-dimensional structure in the form of linear subspaces gained by a technique called the *active subspace method*.

Many inverse problems of practical relevance have several input parameters and can suffer from the *curse of dimensionality* describing a substantial growth of computational complexity in the number of model inputs. This phenomenon can be tackled by reducing the actual dimension of the problem in the sense that computations focus only on its *intrinsic* structure.

The collection of my scientific contributions can be found in the preceding listings above. However, I would like to briefly mention (what I see as) the milestones of my work. The first publication, in cooperation with Shubhangi Gupta (former GEOMAR Helmholtz Centre for Ocean Research Kiel, now postdoc at TUM) and Christian Deusner (GEOMAR),

approximately solves a Bayesian inverse problem involving a complex model for the behavior of methane hydrates by investigating lower-dimensional active subspaces. We submitted a manuscript in January 2018 which was accepted in August 2018 by the Springer journal *Computational Geosciences*.

In the UNMIX project, we worked on a Bayesian calibration of a karst hydrological model, developed by Daniel Bittner and Gabriele Chiogna (TUM Chair of Hydrology and River Basin Management), paired with a *global sensitivity analysis*. A manuscript with our results was submitted in January 2019 and accepted in July 2019 by the Wiley journal *Water Resources Research*.

Both of the mentioned milestones forced me to work interdisciplinary which was interesting and educating on the one hand but on the other hand also challenging in the sense that communication across disciplines needs time, patience, and persistence.

In contrast, the third main contribution focused on a purely mathematical topic. Together with Jonas Wallin (Lund University) and the acknowledged support of Prof. Krzysztof Podgórski, we studied an interesting case in which we were able to prove that common error bounds of the active subspace method are not able to be obtained, but one rather has to fall back to *generalized* bounds. A corresponding manuscript was submitted in October 2019 and accepted in February 2020 by the *Electronic Journal of Statistics* (Institute of Mathematical Statistics, IMS).

With huge anticipation, I planned to take part at the *SIAM Conference on Uncertainty Quantification (UQ20)* on the TUM campus in Garching and contribute in the form of an already accepted talk in a minisymposium. The conference was scheduled for March 24–27, 2020 but unfortunately got canceled on March 7, 2020 due to the outbreak of SARS-CoV-2.

A final topic that I spent a lot of time and dedication on concerns an *iterative* construction of active subspaces that promise to fit better to a given Bayesian problem than subspaces computed with standard initial quantities. In particular, this can, for instance, be seen as a preconditioning step for further subsequent inference procedures in higher (but not full) dimensions. A corresponding computational experiment is conducted with an epidemiological model for the 2014 Ebola virus outbreak in West Africa. I want to sincerely thank Maria Vittoria Barbarossa (former Heidelberg University, now Frankfurt Institute for Advanced Studies), who is one of the modelers, for her patient correspondence and for providing subtle, crucial information about the model without which a successful implementation would not have been possible.

Infrastructure

All mentioned topics and projects involved a large amount of programming. The main part was developed in the form of scripts using the Python scripting language in version 3.7 [270], a documentation of which is available at <https://docs.python.org/3.7/>, and the Python distribution Anaconda [1].

In particular, I have made extensive use of popular Python libraries for scientific computing (NumPy [205], SciPy [271]), statistics/machine learning (scikit-learn [210]), data analysis (pandas [191]), and visualization (Matplotlib [131], seaborn [281]). For parallel computations, I employed the shell script launcher [5] which distributes batch jobs on a predefined number of processor cores. I want to express my reverence and thank all the developers of these high quality, freely available software packages and their (financial) supporters. In my opinion, their

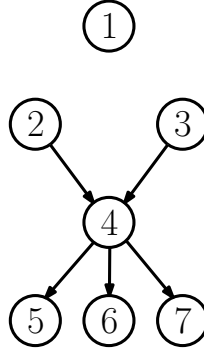


Figure 1: Dependencies among all chapters.

appreciation and contribution is highly underrated in the scientific community. My source code can be found in a repository at

<https://bitbucket.org/m-parente/uq-tools/>.

The methane hydrate model of Shubhangi Gupta is developed in the programming language C++ with the DUNE PDElab library [20, 76]. Daniel Bittner implemented the karst hydrological model “LuKARS” in the statistical scripting language R [10]. Wolfram Mathematica [285] notebooks and code to generate plots for the article with Jonas Wallin can be found at

<https://bitbucket.org/m-parente/asm-poincare-pub/>.

As a development environment, I used Microsoft’s highly extensible and customizable editor Visual Studio Code [3] and a corresponding Python extension [2].

The personal workstation that was kindly provided by the TUM Chair for Numerical Mathematics is a Dell Precision Tower 7810 equipped with an Intel(R) Xeon(R) CPU E5 (consisting of 8 logical cores each running at 3 GHz) and 32 GB RAM. The parallel computations involving the expensive methane hydrate model were conducted on one of the chair’s workstations which is also a Dell Precision Tower 7810 but equipped with two Intel(R) Xeon(R) CPU E5 (consisting of a total of 36 logical cores each running at 2.3 GHz) and 256 GB RAM.

This document, as well as most of my scientific notes and manuscripts, was created with the L^AT_EX distribution TeX Live [4] and the editor TeXstudio [6].

Structure

The content is logically divided into chapters, sections, subsections, and appendices; the depth of the subsequent table of contents is the section level. Definitions, assumptions, theorems, propositions, lemmata, and corollaries are all numbered within the same logical group on the section level. Equations, figures, and tables are numbered on chapter level. Algorithms are numbered consecutively in whole numbers.

It follows a description of the chapters’ contents. The dependencies among all chapters are illustrated in Fig. 1. **Chapter 1** provides an introduction to UQ by explaining different ways to define “uncertainty” and by describing several tasks that are commonly assigned to the field. In **Chapter 2**, we give a motivation for the consideration of inverse problems in a statistical setting and lay the mathematical foundation for the (approximate) solution of

Bayesian inverse problems. Active subspaces are introduced in the context of related dimension reduction approaches in **Chapter 3**. We study this technique from theoretical, as well as from practical perspectives. **Chapter 4** is concerned with bringing Bayesian inverse problems and the active subspace method formally and methodologically together. It forms the basis for subsequent investigations. An efficient parameter estimation via active subspaces involving the mentioned complex methane hydrate model is conducted in **Chapter 5**. **Chapter 6** describes our approach to calibrate the karst hydrological model from the UNMIX project involving a high-dimensional parameter space and provides a corresponding global sensitivity analysis. Finally, **Chapter 7** proposes an iterative scheme for the computation of active subspaces which is more robust to distributional assumptions on input parameters and thus can serve as a preconditioning/preprocessing step for further studies. **Appendix A** contains well-known, common results and definitions that are used and referred to throughout the document. **Appendix B** rigorously shows the invariance (w.r.t. orthogonal transformations) of Poincaré constants for multivariate normal distributions which is required in Ch. 3. **Appendix C** provides the model equations for the karst hydrological model “LuKARS” from Ch. 6. **Appendix D** includes a proof concerning a formal statement on the iterative algorithm from Ch. 7. The last two chapters contain a list of used **Notation** and all **References**.

Acknowledgments

It is now a good opportunity to say “thank you” to some persons since, first, without them I would not write this sentence, and, secondly, I missed the chance most probably several times during the last months and years.

I am so grateful to my family that they supported me throughout my whole lifetime (even if they are not suitable partners for scientific discussion). They do it in their own way and as it is possible for them, but the most important thing to me is that they put responsibility on my shoulders and trust me that I am able to form my own life and make my own decisions.

Fortunately, I have a large circle of good friends. I cannot name all of them, but I want to mention at least some of them. I am happy and lucky to have Bene as a friend. Since we met at Hochschule München, he is supporting me for almost ten years now with a lot of conversations of funny, philosophical, and scientific type. Fabi, Mali, Juls, Evi, and Olli, thank you for our friendship, your support, and for standing behind my back during challenging times. Also, our “Stammtisch” was very important for me to recharge my batteries and to put scientific thoughts aside for the moment. Thank you all for providing support not in a direct but all the more in an indirect way. There are so many others that I need to thank, but this would go beyond the scope and I hope that they forgive me.

Also, I would like to thank my supervisor Barbara Wohlmuth for her support during my time as a PhD student and the valuable scientific discussions we had.

Since the start at our chair, I benefited in so many ways from my colleagues, by scientific and personal discussion but also by sharing tasks. I already mentioned some of them above, but I really would like to thank everyone who contributed to my time as a PhD student. The most influential contributors were, in alphabetical order, Rudolf Aichner, Maria Vittoria Barbarossa, Nicole Barthel, Mónica Basilio Hazas, Florian Beiser, Daniel Bittner, Rainer Callies, Gabriele Chiogna, Daniel Drzisga, Marvin Fritz, Shubhangi Gupta, Thomas Horger, Markus Huber, Lorenz John, Brendan Keith, Ustim Khristenko, Tobias Köppl, Jo-Anna Küster, Vanessa Landgraf, Jonas Latz, Christian Ludwig, Giorgia Marcolini, Steven Mattis, Pablo Merchan Rivera, Markus Muhr, Thomas Nagler, Vanja Nikolić, Krzysztof Podgórski, Jenny Radeck,

Emil Ratko-Dehnert, Klaus-Dieter Reinsch, Laura Scarabosio, Daniel Schaden, Tanu Singh Cox, Piotr Swierczynski, Elisabeth Ullmann, Ettore Vidotto, Fabian Wagner, Jonas Wallin, Barbara Wohlmuth, Linus Wunderlich, and Olivier Zahm.

I sincerely apologize to everyone that I forgot to put in this list of contributors and hope that she or he forgives me.

Mario Teixeira Parente

Munich, Germany
September, 2020

Contents

Abstract	I
Scientific contributions	III
Preface	V
1 Quantification of uncertainties	1
1.1 Types of uncertainties	1
1.2 Some tasks in UQ	3
2 Bayesian inverse problems	9
2.1 Motivation	9
2.2 Mathematical setup	11
2.3 An example	16
2.4 Well-posedness	18
2.5 Sampling the posterior	20
3 Active subspaces	39
3.1 Idea	40
3.2 Common setup	43
3.3 Bounds for compactly supported and normal densities	48
3.4 Generalized bounds	50
3.5 Practical considerations	64
4 Active subspaces and Bayesian inverse problems	73
4.1 An example	75
4.2 MCMC in the active subspace	76
4.3 Construction of samples in the full space	79
4.4 Surrogate for the low-dimensional misfit	80
5 Parameter estimation for a methane hydrate model	83
5.1 Case study	84
5.2 Setup	91
5.3 Computational results	93

5.4	Discussion	98
6	Calibration and sensitivity analysis for a karst hydrological model	101
6.1	Case study	102
6.2	Setup	105
6.3	Global sensitivity analysis	107
6.4	Computational results	109
6.5	Discussion	113
7	Iterative active subspaces	117
7.1	Motivation	117
7.2	Iterative algorithm	119
7.3	Analysis with an illustrative case	119
7.4	Experiment with an Ebola spread model	122
7.5	Conclusions	133
A	Common results and definitions	137
A.1	General	137
A.2	Stochastics	137
B	Invariance of Poincaré constants for multivariate normal distributions	143
C	Model equations for LuKARS model	145
D	Pending proofs	147
D.1	Proof of Prop. 7.3.1	147
	Notation	151
	References	153

Quantification of uncertainties

The field of Uncertainty Quantification (UQ) is a recent, modern, and still emerging scientific discipline. Several attempts to “define” UQ had already been undertaken, but it seems that there is not a suitable unique definition every involved scientist agrees on. A possible reason for this type of ambiguity is the fact that scientists, who have backgrounds from different fields, may have different focuses or even disagree already on such a fundamental interpretation of what an *uncertainty* actually is; a topic that is discussed below in Sec. 1.1. UQ can, however, be seen as the interaction of fields like probability theory, statistics, mathematical modeling, experiments, numerical analysis, and computer simulations. Many models and predictions, e. g., from disciplines like medicine, geology, hydrology, engineering, and many more, are subject to uncertainties and, hence, there is a need for structured approaches on how to manage them within a suitable framework. For example, risk engineers need reliable guarantees for statistical properties of their models in order to justify a basis for important decisions. Since the mathematical formalism provides a natural context for a rigorous treatment of the quantification of uncertainties, UQ therefore utilizes concepts and methods from applied mathematics, e. g., from mathematical analysis, linear algebra, functional analysis, approximation theory, or perturbation theory.

A main goal of scientists or engineers applying UQ is the *identification*, *quantification*, and *reduction* of uncertainties related to models, computer algorithms, experiments, and predictions of *quantities of interest* (QoI) [243]. This rather generally formulated problem can and must be subdivided into subtasks of which the most common are presented in Sec. 1.2.

For a more general overview and broader discussion of goals, fundamentals, and interpretations of UQ, see the books in [243, 253] and the review article of [204].

1.1 Types of uncertainties

The definition of the notion *uncertainty* is as fundamental as a definition can be for UQ. Although there exist several philosophical debates about a possible definition, we prefer to approach it by describing two types of common classifications for uncertainties.

The classes of *epistemic* (or systemic) and *aleatoric* (or statistical) uncertainties is well explained by a story about an important decision that Barack Obama had to make during his presidency in 2011 [93]. On April 29, 2011 Barack Obama had to decide whether or not to attack a compound Osama bin Laden was suspected to hide in. As Obama admitted

afterwards, this extraordinary situation put a lot of stress on him. This stress can be attributed to two different sources: First, although it has been a fact that either bin Laden was in the compound or not, Obama was still not certain about this fact. This type of uncertainty is called epistemic and is often described to be a *lack of knowledge*. Secondly, even if Obama had given the command for an attack, it was not certain that the operation would be successful. This uncertainty, not being a fact, is of statistical or probabilistic nature and thus called aleatoric (from the Latin word *ālea* meaning dice). It is still an open debate among philosophers and scientists from different fields if this classification is appropriate, or if it is even possible to uniquely attribute a particular uncertainty to one of the two mentioned categories. For example, “*an aleatory assessment that the military operation would succeed most of the time may also entail some epistemic uncertainty due to a lack of confidence in the assumptions underlying that assessment.*” [93, p. 2]

A second way of classification is more of conceptual type and seems to be more accessible compared to the two categories from above. In [204], all uncertainties appearing during the development and assessment of mathematical models in the context of UQ are assigned to one of the following classes of uncertainties:

- **Uncertainty in the system of reasoning or mathematical framework:** There are several frameworks and theories in which it is possible to express problems concerning the quantification of uncertainties. Examples are (different types of) probability theory or, according to [204], possibility theory, fuzzy sets, Dempster–Shafer theory, or interval arithmetic. It is, however, difficult to quantify this type uncertainty.
- **Data uncertainty:** Field experiments and measurements collecting data for, e. g., statistical inferences, are often subject to errors and uncertainties due to imprecise measurement devices or imperfect environmental conditions. Either there exist justified assumptions on the distribution of experimental errors, or particular models for this type of error are applied; see, e. g., [46, 236].
- **Model uncertainty:** A famous quote of George E.P. Box from 1978 in [37] is: “*All models are wrong, but some are useful.*” Almost always, building a model for a particular phenomenon automatically includes model errors that constitute a discrepancy in a model’s prediction and the actual phenomenon. This type of error, or uncertainty, is elusive and hard to grasp. However, there exist statistical approaches in the field of *model adequacy checking* on how to quantify model uncertainty; see, e. g., [197, Ch. 4].
- **Parameter uncertainty:** Many physical or mathematical models depend on parametric input values that directly impact the output of the model. Input parameters can be scalar- or function-valued and are often not clearly determined but uncertain. Examples are birth and death rates in a predator-prey model or initial/boundary conditions in a Navier-Stokes simulation. Most of the following content is dedicated to the statistical quantification of this type of uncertainty.
- **Discretization uncertainty:** A numerical implementation of models in efficient computer codes needs to discretize in space and time. For example, the numerical solution of a partial differential equation (PDE) is often computed on a discrete mesh in space and at discrete points in time. This uncertainty (or, more suitable, error) can be quantified by, e. g., a priori or a posteriori error estimates known from prominent discretization

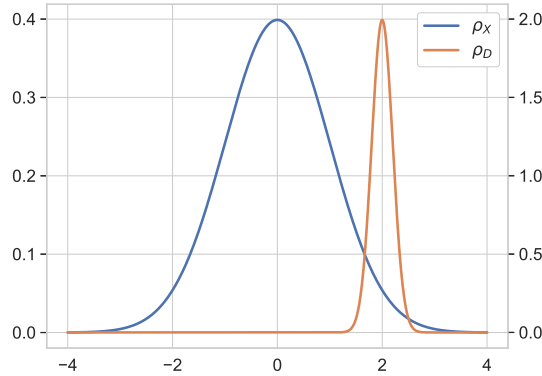


Figure 1.1: Forward propagation of a standard normal distribution with mean $\mu = 2$ and standard deviation $\sigma = 0.2$ through the map $\mathcal{G}(x) := 0.2x + 2$.

methods like the *finite difference method* [175] or the *finite element method* [39]. Traditional approaches employ deterministic error estimates, but there are recent advances in a probabilistic or Bayesian treatment of discretization errors; see, e. g., [124, 203].

1.2 Some tasks in UQ

UQ is a wide field with a lot of different problems. The variety is partly caused by the different types of uncertainties mentioned above. The following enumeration of common UQ tasks is certainly not complete but is supposed to give a broad overview making it easier to integrate the main investigations later. Note that these tasks mainly aim to solve problems related to parameter uncertainty. Uncertainty is, especially in this case, modeled by probability distributions known from statistics and probability theory.

In the following, let $\mathcal{G}(\mathbf{x})$ denote the output or QoI of a model or computation for some input parameter or variable \mathbf{x} .

Forward UQ

Forward uncertainty quantification comprises tasks like *uncertainty/forward propagation* and the computation of moments of QoIs.

Forward propagation deals with the computation or estimation of a probability distribution on the space of outputs. The (known or assumed) uncertainty of inputs \mathbf{x} propagates to an (unknown) uncertainty of corresponding outputs $\mathbf{d} := \mathcal{G}(\mathbf{x})$. Mathematically, given a distribution $\mathbf{P}_{\mathbf{X}}$ on the space of inputs, i. e., let $\mathbf{X} \sim \mathbf{P}_{\mathbf{X}}$ denote a random input variable, the task is to compute the so-called *push-forward distribution* of $\mathbf{D} := \mathcal{G}(\mathbf{X})$ which is $\mathbf{P}_{\mathbf{D}} := \mathbf{P}_{\mathbf{X}}(\mathcal{G}^{-1}(\cdot))$. For example, if we assume that our input $\mathbf{X} \sim \mathcal{N}(0, I)$ with values in \mathbf{R}^n , $n \in \mathbf{N}$, then this uncertainty propagates through the map $\mathcal{G}(\mathbf{x}) := \sigma\mathbf{x} + \boldsymbol{\mu}$, where $\boldsymbol{\mu} \in \mathbf{R}^n$ and $\sigma > 0$, to a normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix σI . Fig. 1.1 illustrates the case in 1D for $\sigma = 0.2$ and $\boldsymbol{\mu} = 2$. Of course, more complex problems do often not allow to analytically compute the propagated distribution. Hence, for such scenarios, more sophisticated methods must be applied; see, e. g., [172].

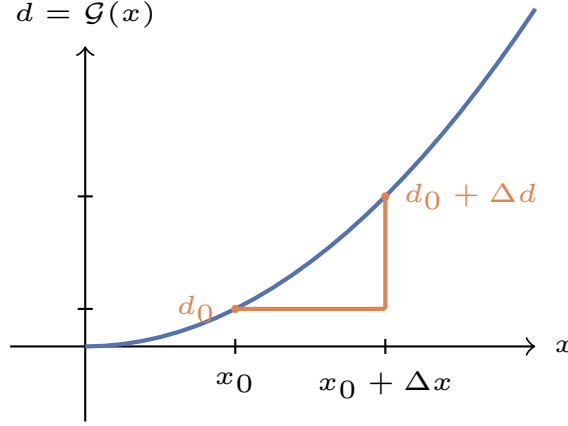


Figure 1.2: A perturbation in the input x_0 can have large effects on the corresponding output $d_0 = \mathcal{G}(x_0)$ depending on the derivative $d\mathcal{G}(x_0)/dx$.

Multilevel Monte Carlo methods are often applied for estimating the expectation of a computationally expensive QoI, i. e., $\mathbf{E}[\mathcal{G}] = \mathbf{E}[\mathcal{G}(\mathbf{X})]$, if the distribution of the input \mathbf{X} is given or can be assumed. Suppose there exists a sequence of approximations $\mathcal{G}_0, \dots, \mathcal{G}_L$ of \mathcal{G} (on so-called levels $0, \dots, L$) such that the evaluation of \mathcal{G}_{ℓ_1} is cheaper than that of \mathcal{G}_{ℓ_2} if $\ell_1 < \ell_2$ and such that $\mathcal{G}_L \rightarrow \mathcal{G}$ as $L \rightarrow \infty$. The linearity of the expectation operator allows to write $\mathbf{E}[\mathcal{G}_L]$ as a telescoping sum,

$$\mathbf{E}[\mathcal{G}_L] = \mathbf{E}[\mathcal{G}_0] + \sum_{\ell=1}^L \mathbf{E}[\mathcal{G}_\ell - \mathcal{G}_{\ell-1}]. \quad (1.1)$$

Since variances $\mathbf{Var}(\mathcal{G}_\ell - \mathcal{G}_{\ell-1}) \rightarrow 0$ as $\ell \rightarrow \infty$, it is not necessary to have a lot of (expensive) evaluations of \mathcal{G}_ℓ for larger ℓ which can considerably save computational costs. Only a few expensive samples have to be computed, but, on the other hand, even more so on lower (cheaper) levels are needed. More details, for example on convergence rates or the number of samples on each level, can be found in [51, 98].

Sensitivity analysis (SA)

Similar to forward UQ, SA studies how the uncertainties in the output can be assigned to different origins of uncertainties in inputs. For example, if a model is sensitive w.r.t. to some of its parameters, it is important to infer/know these parameters sufficiently accurately since, otherwise, the model output/prediction would not be reliable. There exist several approaches on how to define and quantify sensitivities. Here, we look at *local* and *global* sensitivity analysis.

Local SA uses partial derivatives of the forward model \mathcal{G} at a given nominal value \mathbf{x}_0 , i. e., we have to compute $\partial\mathcal{G}(\mathbf{x}_0)/\partial x_i$ for each $i = 1, \dots, n$. That means, if the local sensitivity of \mathcal{G} at \mathbf{x}_0 is large, a small perturbation or uncertainty in \mathbf{x}_0 can have large effects on the corresponding output $\mathcal{G}(\mathbf{x}_0)$; see Fig. 1.2 for illustration. In general, the nominal values, at which to compute the partial derivatives, are not known a priori. Also, local SA does neither intend to explore the whole input space nor systematically study perturbations in combinations of different input components.

For these reasons, there is a need for strategies and metrics analyzing global sensitivities. Here, “global” means that sensitivities are computed at several different locations to build an average sensitivity. These locations are sometimes chosen by a given underlying probability distribution. An overview of the most common techniques for global SA is provided in [232]. We present a particular approach, probably the most common in modern SA, called *variance-based SA* or *Sobol indices* [245, 246]. We assume w.l.o.g. that our forward map \mathcal{G} is real-valued and its inputs are independently uniformly distributed. In the vector-valued case, we can apply the following steps component-wise, and if the input space is not the unit cube, it can be transformed to be. Then, according to [245], $\mathcal{G}(\mathbf{X})$ can be decomposed in

$$\begin{aligned} \mathcal{G}(\mathbf{X}) = & \mathcal{G}_0 + \sum_{i=1}^n \mathcal{G}_i(X_i) + \sum_{\substack{i,j=1 \\ i < j}}^n \mathcal{G}_{i,j}(X_i, X_j) + \cdots \\ & + \mathcal{G}_{1,2,\dots,n}(X_1, \dots, X_n), \end{aligned} \quad (1.2)$$

where \mathcal{G}_0 is a constant and $\mathcal{G}_{i_1,\dots,i_s}$ are functions of X_{i_1}, \dots, X_{i_s} (for appropriate indices i_1, \dots, i_s , $s \geq 1$). A condition on every function $\mathcal{G}_{i_1,\dots,i_s}$ is that

$$\int_0^1 \mathcal{G}_{i_1,\dots,i_s}(x_{i_1}, \dots, x_{i_s}) dx_{i_k} = 0 \quad (1.3)$$

for each $k = 1, \dots, s$, which implies that all terms in (1.2) are orthogonal. Functions $\mathcal{G}_{i_1,\dots,i_s}$ are computed by taking suitable (conditional) expectations, e.g.,

$$\begin{aligned} \mathbf{E}[\mathcal{G}(\mathbf{X})] &= \mathcal{G}_0, \\ \mathbf{E}[\mathcal{G}(\mathbf{X}) | X_i] &= \mathcal{G}_0 + \mathcal{G}_i(X_i), \\ \mathbf{E}[\mathcal{G}(\mathbf{X}) | X_i, X_j] &= \mathcal{G}_0 + \mathcal{G}_i(X_i) + \mathcal{G}_j(X_j) + \mathcal{G}_{i,j}(X_i, X_j). \end{aligned} \quad (1.4)$$

That is, we have

$$\begin{aligned} \mathcal{G}_0 &= \mathbf{E}[\mathcal{G}(\mathbf{X})], \\ \mathcal{G}_i(x_i) &= \mathbf{E}[\mathcal{G}(\mathbf{X}) | X_i = x_i] - \mathcal{G}_0, \\ \mathcal{G}_{i,j}(x_i, x_j) &= \mathbf{E}[\mathcal{G}(\mathbf{X}) | X_i = x_i, X_j = x_j] - \mathcal{G}_i(x_i) - \mathcal{G}_j(x_j) - \mathcal{G}_0. \end{aligned} \quad (1.5)$$

The remaining functions are derived by analogous steps. Using the orthogonality of the terms in (1.2), the variance of $\mathcal{G}(\mathbf{X})$ can be decomposed similarly, i.e.,

$$\begin{aligned} \text{Var}(\mathcal{G}(\mathbf{X})) &= \mathbf{E}[\mathcal{G}(\mathbf{X})^2] - \mathcal{G}_0^2 \\ &= \sum_{i=1}^n V_i + \sum_{\substack{i,j=1 \\ i < j}}^n V_{i,j} + \cdots + V_{1,2,\dots,n}, \end{aligned} \quad (1.6)$$

where

$$\begin{aligned} V_i &= \text{Var}(\mathbf{E}[\mathcal{G}(\mathbf{X}) | X_i]), \\ V_{i,j} &= \text{Var}(\mathbf{E}[\mathcal{G}(\mathbf{X}) | X_i, X_j]) - V_i - V_j. \end{aligned} \quad (1.7)$$

The other terms in (1.6) are again defined analogously. Hence, (1.6) shows that the variance of $\mathcal{G}(\mathbf{X})$ can be decomposed into variances related to each input variable and interactions among them. There exist two common notions of sensitivity indices:

1. *first-order indices* [245],
2. *total-effect indices* [128].

A first-order index measures how much the variance caused by an interaction of certain input variables contributes to the total variance. This also includes variances caused by single (or marginal) input variables. For input variables X_{i_1}, \dots, X_{i_s} , it is defined by

$$S_{i_1, \dots, i_s} := \frac{V_{i_1, \dots, i_s}}{\text{Var}(\mathcal{G}(\mathbf{X}))}. \quad (1.8)$$

Note that

$$\sum_{i=1}^n S_i + \sum_{\substack{i,j=1 \\ i < j}}^n S_{i,j} + \dots + S_{1,2,\dots,n} = 1. \quad (1.9)$$

Since it can be computationally expensive to evaluate every single first-order index (the number of different first-order indices is $2^n - 1$), it is cheaper to compute the total-effect indices which measure sensitivities w.r.t. a certain input variable, i.e., these indices include the variance caused by the variable itself *and* each of the variances caused by its interactions. For an input variable X_i , the total-effect index is defined by

$$S_{\text{tot},i} := \frac{\mathbf{E}[\text{Var}(\mathcal{G}(\mathbf{X}) \mid \mathbf{X}_{\sim i})]}{\text{Var}(\mathcal{G}(\mathbf{X}))}, \quad (1.10)$$

where $\mathbf{X}_{\sim i}$ means all variables unless X_i . Note that there are only n total-effect indices to compute. Also, it holds that

$$\sum_{i=1}^n S_{\text{tot},i} \geq 1, \quad (1.11)$$

since variances caused by interactions are counted multiple times in this sum. For the computation of first-order and total-effect indices, we refer to [232, Ch. 4.6].

In Ch. 6, we will see that the *active subspace method* allows for another way to define a global sensitivity metric. In fact, these indices are comparable to Sobol indices as was shown in [53].

Data assimilation

Originally developed in the fields of weather forecasts, climate studies, and oceanographic sciences, *data assimilation* stands for the incorporation of noisy data in the estimation of the state of a dynamical, i.e., time-varying, system which describes the state's evolution or dynamics. For the specification of a mathematical setup, we follow [165] and only regard the discrete-time case although the continuous-time setting is also treated therein.

With indices j representing time, the dynamical state is modeled by a sequence of random variables $(\mathbf{S}_j)_{j \in \mathbf{N}_0}$ with values in \mathbf{R}^n , $n \in \mathbf{N}$, and the dynamics by a map $\Psi \in C(\mathbf{R}^n, \mathbf{R}^n)$, i.e.,

$$\begin{aligned} \mathbf{S}_{j+1} &= \Psi(\mathbf{S}_j) + \boldsymbol{\xi}_j, \quad j \in \mathbf{N}_0, \\ \mathbf{S}_0 &\sim \mathcal{N}(\mathbf{m}_0, C_0), \end{aligned} \quad (1.12)$$

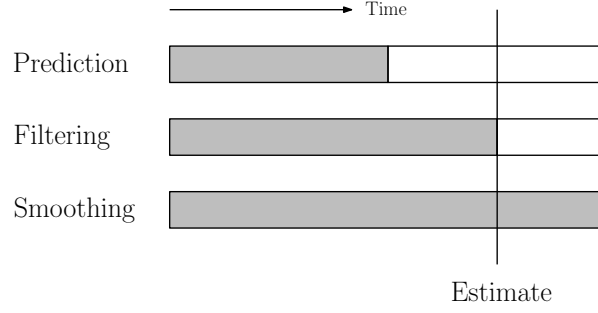


Figure 1.3: Depending on the relation of the time for the state estimate and the time data is available up to (gray), the corresponding tasks of estimation have different names in the literature.

where $(\xi_j)_{j \in \mathbf{N}_0}$ is a sequence of i.i.d. random variables modeling stochasticity in the dynamics (then called *stochastic dynamics*). The initial Gaussian distribution of \mathbf{S}_0 is assumed to have mean vector $\mathbf{m}_0 \in \mathbf{R}^n$ and covariance matrix $C_0 \in \mathbf{R}^{n \times n}$. For simplicity, it is assumed that all ξ_j are independent of \mathbf{S}_0 . If the dynamics is not stochastic, i.e., ξ_j is not used in (1.12), the dynamics is called *deterministic dynamics*.

Observable quantities of the system (QoIs), for which measurements are assumed to be available, are computed from the state and hence also evolve over time. Since measurements $(\mathbf{D}_j)_{j \in \mathbf{N}}$ for these QoIs (with values in \mathbf{R}^m , $m \in \mathbf{N}$) are often corrupted by noise, they are modeled by

$$\mathbf{D}_j = \mathcal{O}(\mathbf{S}_j) + \boldsymbol{\eta}_j, \quad j \in \mathbf{N}, \quad (1.13)$$

where $(\boldsymbol{\eta}_j)_{j \in \mathbf{N}}$ is a sequence of i.i.d. random variables representing observational noise and assumed to be independent of all \mathbf{S}_j and all ξ_j . Often, it is assumed that the noise is Gaussian. The map $\mathcal{O} \in C(\mathbf{R}^n, \mathbf{R}^m)$ is the so-called *observation operator* and maps an element from the state space to the data space.

In this setup, the actual goal, from a Bayesian perspective, is to find the posterior distribution $\mathbf{P}_{\mathbf{S}_j | \mathbf{D}_{1:k}}(\cdot | \mathbf{d}_{1:k})$, $j, k \in \mathbf{N}$, i.e., we look for a distribution on the state space for the state at step j using measurement data $\mathbf{d}_{1:k} \in \mathbf{R}^m \times \cdots \times \mathbf{R}^m$, i.e., observations at time steps $1, \dots, k$. Depending on the relation between j and k , there are different notions for the corresponding task; see Tab. 1.1 or Fig. 1.3.

Table 1.1: The task of finding the posterior distribution $\mathbf{P}_{\mathbf{S}_j | \mathbf{D}_{1:k}}(\cdot | \mathbf{d}_{1:k})$ is named depending on the relation of j and k .

Relation	Notion
$j > k$	Prediction
$j = k$	Filtering
$j < k$	Smoothing

Remark. The Bayesian formulation of these problems also allows to incorporate the estimation of model parameters in this setting; see, e.g., [234]. Parameter studies are, however, not mentioned here but in the following subsection.

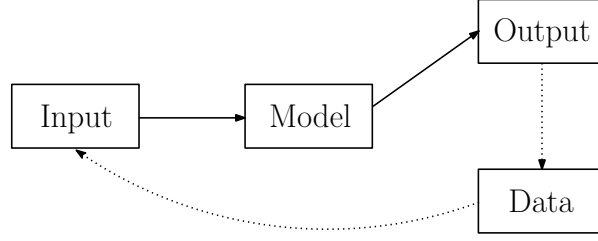


Figure 1.4: Illustration of the common setup in parameter calibration.

Much more mathematical details for filtering and smoothing, including their relation and well-posedness questions, can be found in [165, 234] and references therein.

The success story of filtering (based on a least squares approach) is due to the seminal article of Kalman in 1960 [148], who developed the so-called *Kalman filter* for linear problems, which impresses with its mathematical simplicity and generality. Kalman's approach was able to be easily transferred into the Bayesian context covered by [136]. Although a more comprehensive treatment of the most common algorithms and methodology is out of scope here, we refer to two important extensions of the classical Kalman filter: the *Extended Kalman filter* (for non-linear problems) [144] and the *Ensemble Kalman filter* (for larger problems by an efficient estimation of covariances) [88, 129].

Parameter calibration and validation

Given a (stationary) model $\mathcal{G} : \mathcal{X} \rightarrow \mathcal{D}$, that maps an element from a normed parameter space $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ into a normed space of QoIs $(\mathcal{D}, \|\cdot\|_{\mathcal{D}})$ (the data space), and experimental data $\mathbf{d} \in \mathcal{D}$, a common goal among engineering and applied disciplines is to find/infer a suitable set of parameters or a distribution on them that explains the measured data sufficiently well. This task is generally known as *parameter calibration* (also referred to as an *inverse problem*) and depicted in Fig. 1.4.

For a certain input parameter $\mathbf{x} \in \mathcal{X}$, the output $\mathcal{G}(\mathbf{x})$ is computed through the model and compared with experimental data via $\|\mathbf{d} - \mathcal{G}(\mathbf{x})\|_{\mathcal{D}}$. This information is used in turn to further continue and improve the inference of input parameters that are likely to explain the data. This gives a general circular inference process.

After calibration, the calibrated parameters \mathbf{x}^* should be tested on a new data set $\mathbf{d}' \in \mathcal{D}$ or, if possible, with a new QoI, i.e., a new forward operator $\mathcal{G}' : \mathcal{X} \rightarrow \mathcal{D}'$. For example, the difference $\|\mathbf{d}' - \mathcal{G}'(\mathbf{x}^*)\|_{\mathcal{D}'}$ could be studied and interpreted in a suitable manner depending on the context of the application. Other metrics are also possible. This general method is called *parameter validation*. It is important to note that validation is based on a purely new data set which is not supposed to be used for calibration purposes.

We keep the exposition on calibration and validation here rather simple because the next chapter, dedicated to present the Bayesian perspective on parameter calibration, is fairly comprehensive as it also explains the differences with other methods as, e.g., classical inverse problems.

Bayesian inverse problems

In the following, we present the formulation of inverse problems, particularly Bayesian inverse problems. This chapter constitutes a central role for the remaining results and discussions as it contains a consistent mathematical setup, basic theory, and examples that form a foundation for later chapters.

The structure is as follows. A comparison with classical inverse problems, as a motivation for Bayesian inverse problems, is provided in Sec. 2.1. Sec. 2.2 develops the basic mathematical setup for Bayesian inverse problems; an example for illustration is given in Sec. 2.3. Subsequently, well-posedness results are discussed in Sec. 2.4. Sec. 2.5 introduces strategies for posterior sampling that are rather basic but necessary to discuss in order to provide a self-contained exposition.

2.1 Motivation

Inverse problems frequently arise in applied sciences. They provide a context for the inference or reconstruction of input variables in a mathematical model which should explain given experimental data sufficiently well. Imagine, for example, an image that needs to be reconstructed but for which only a blurred version is available. Or, consider the average walking speed of a person in a model for pedestrian and evacuation dynamics. Pedestrian data that is measured, e. g., at a festival or at a supermarket, may be informative for the (distribution of the) walking speed of persons which itself is an input parameter for a corresponding model explaining the behavior of pedestrians in certain situations.

Before we concentrate on the Bayesian perspective on inverse problems, let us first regard *classical inverse problems* and see why a Bayesian formulation can be advantageous and attractive. Classical inverse problems require two objects: 1) Experimental data \mathbf{d} (also called observational or measured data) and 2) a model \mathcal{G} with a predefined input \mathbf{x} and output $\mathcal{G}(\mathbf{x})$ that is supposed to explain the measured data. The relationship between data, model, and input variables is given by

$$\mathbf{d} = \mathcal{G}(\mathbf{x}). \quad (2.1)$$

Note that \mathbf{d} and \mathcal{G} are given and \mathbf{x} is unknown. For a mathematical treatment of a classical inverse problem, we assume that both inputs and data are elements of Banach spaces. Hence, for Banach spaces $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{D}, \|\cdot\|_{\mathcal{D}})$, we have that $\mathbf{x} \in \mathcal{X}$, $\mathbf{d} \in \mathcal{D}$, and $\mathcal{G} : \mathcal{X} \rightarrow \mathcal{D}$.

The task in a classical inverse problem is to find $\mathbf{x} \in \mathcal{X}$ such that (2.1) holds. Let us look into questions regarding *well-posedness* of such a problem. The notion of well-posedness that we use here originates from an article of Hadamard in 1902 [115] and was set to the following definition during the last century.

Definition 2.1.1 (Well-posedness). A problem is called *well-posed* if and only if

1. a solution exists,
2. the solution is unique,
3. the solution depends continuously on the given context, i. e., a small perturbation in the given data does not cause a huge change in the solution.

Otherwise, the problem is called *ill-posed*.

Applying this definition to classical inverse problems, we immediately see that they are ill-posed in general, for example if $\mathbf{d} \notin \text{ran}(\mathcal{G})$. In such a case, we can reformulate the problem to finding a minimizer $\mathbf{x}^\dagger \in \mathcal{X}$ such that

$$\mathbf{x}^\dagger = \arg \min_{\mathbf{x} \in \mathcal{X}} \frac{1}{2} \|\mathbf{d} - \mathcal{G}(\mathbf{x})\|_{\mathcal{D}}^2. \quad (2.2)$$

However, it is known that (2.2) is still ill-posed since there might exist minimizing sequences $(\mathbf{x}_n)_{n \in \mathbb{N}}$ that do not converge in \mathcal{X} , the solution is not unique, or does not depend continuously on the data \mathbf{d} .

To reduce the effect of these issues, we can formulate a so-called *regularized* problem. There are several approaches on how such a regularization can be achieved, see e. g., [146, 253]. We present *Tikhonov-regularization* named after its originator A. N. Tikhonov [262, 263]. In this approach, the minimization functional gets an additional term, also called *penalty* or *penalization term*, that penalizes deviations from a given value $\mathbf{x}_0 \in \mathcal{E}$, where $(\mathcal{E}, \|\cdot\|_{\mathcal{E}})$ is a Banach space contained in \mathcal{X} that we need to introduce here in order to be mathematically correct in this abstract setting. The problem is thus modified to

$$\mathbf{x}^\dagger = \arg \min_{\mathbf{x} \in \mathcal{E}} \frac{1}{2} (\|\mathbf{d} - \mathcal{G}(\mathbf{x})\|_{\mathcal{D}}^2 + \|\mathbf{x} - \mathbf{x}_0\|_{\mathcal{E}}^2). \quad (2.3)$$

The penalty term $\|\mathbf{x} - \mathbf{x}_0\|_{\mathcal{E}}^2$ fulfills three objectives:

1. It adds a type of prior information to the problem specifying that we look for a solution “near” \mathbf{x}_0 . This could also make the problem uniquely solvable.
2. It avoids that the data \mathbf{d} is fit to closely. If this situation occurs, the continuous dependence on \mathbf{d} , which is one of the required properties of the well-posedness definition in Def. 2.1.1, can be circumvented.
3. It can help numerical optimization algorithms to be more stable since it avoids the solution \mathbf{x}^\dagger getting arbitrarily large in size.

However, although this type of regularization helps to reduce the issue of ill-posedness, it cannot completely avoid it.

The instability of classical inverse problems can also be understood intuitively. For example, since Newton, Leibniz, and others, many physical laws are described by differential equations. They contain not only aspects of *locality* in the sense that the dependence of a function, describing a physical phenomenon, and its derivatives at a given point is expressed based on physical conditions at this point, and only at this point. But also, physical laws are *causal* in nature, i. e., a later state of the described physical quantity depends on previous states. In contrast, inverse problems are most often *non-local* and *non-causal*. For non-causality, consider heat diffusion as an example. Small perturbations of the initial conditions disperse over time which makes their inference and reconstruction from temperatures observed at a final time rather difficult. We may find that starting the simulation with sufficiently different initial conditions lead to a similar final temperature within the accuracy of our measurements.

In order to avoid such unfavorable circumstances, it is important to incorporate as much *a priori* information (i. e., information we have *before* (*prior*) the observations) as possible since this increases the chance to get a well-posed problem. *Statistical inversion* (also known as *stochastic inversion* or *statistical/stochastic inference*), in a Bayesian context, provides a setting in which such prior information can be included in a proper and consistent way. This approach regards all involved quantities as random variables and defines a so-called *posterior distribution* giving a probabilistic description of unknown variables while taking observational data into account. It is also called *Bayesian updating* since the prior is “updated” to the posterior distribution.

The next sections in this chapter concentrate on stochastic inversion and a Bayesian viewpoint in particular. The topics covered there constitute the fundamentals for later approaches and discussions.

2.2 Mathematical setup

We mainly follow [73, 252] for the mathematical setup presented here but add some technical details.

Since later chapters only require a finite dimensional setting, we assume from now on that $\mathcal{X} = \mathbf{R}^n$ and $\mathcal{D} = \mathbf{R}^{n_d}$ with $n, n_d \in \mathbf{N}$. The starting point for Bayesian inverse problems is similar to the relation given in (2.1). The first difference in the Bayesian setting is that we can also include measurement errors, a type of uncertainty that was mentioned in Ch. 1 as data uncertainty. This results in the relation

$$\mathbf{D} = \mathcal{G}(\mathbf{x}) + \boldsymbol{\eta}, \quad (2.4)$$

where the observational noise $\boldsymbol{\eta} \sim \mathbf{P}_{\boldsymbol{\eta}}$ is modeled probabilistically. Note that the left hand side of (2.4) (the data) automatically becomes a random variable and is thus written with a capital letter. What makes the inference truly statistical is the fact that the variable \mathbf{x} can be also regarded as a random variable. Hence, (2.4) becomes

$$\mathbf{D} = \mathcal{G}(\mathbf{X}) + \boldsymbol{\eta}, \quad (2.5)$$

where all involved random variables are assumed to be continuous and elements of a common probability space $(\Omega, \mathcal{A}, \mathbf{P})$. In addition, we assume that \mathbf{X} and $\boldsymbol{\eta}$ are independent. Analogously to Sec. 2.1, \mathbf{X} , \mathbf{D} , and $\boldsymbol{\eta}$ are random variables with values in \mathcal{X} and \mathcal{D} , respectively, and \mathcal{G} is a map from \mathcal{X} to \mathcal{D} . As usual, we equip the spaces \mathcal{X} and \mathcal{D} with corresponding Borel σ -algebras $\mathcal{B}(\mathcal{X})$ and $\mathcal{B}(\mathcal{D})$, respectively, in order to get measurable spaces.

Remark. The forward operator \mathcal{G} does not necessarily need to be defined on the whole space \mathbf{R}^n . It can also be assumed that $\mathcal{X} \subseteq \mathbf{R}^n$ with $\mathbf{P}_{\mathbf{X}}(\mathcal{X}) = 1$, i. e., \mathcal{G} needs only to be defined on a set with $\mathbf{P}_{\mathbf{X}}$ -measure one.

Remark. We implicitly assume in (2.5) that the model \mathcal{G} is adequate for the data and can be evaluated exactly. Most often, especially when solving differential equations, this is not the case and thus we can only compute an approximation $\tilde{\mathcal{G}} \approx \mathcal{G}$ instead. Note that these two error sources were already mentioned in Ch. 1 as model and, respectively, discretization uncertainty. Including both errors, (2.5) would become

$$\mathbf{D} = \tilde{\mathcal{G}}(\mathbf{X}) + \varepsilon + \eta, \quad (2.6)$$

where the total error $\varepsilon := \mathcal{G}(\mathbf{X}) - \tilde{\mathcal{G}}(\mathbf{X})$ could also be modeled with a probability distribution. Actually, it is often modeled to be part of the measurement error η which can, however, be misleading. Indeed, it is difficult to find a probabilistic model for ε since it has non-zero mean and depends on the (random) input \mathbf{X} . For these reasons, we assume in the remainder that \mathcal{G} is adequate and exactly computable. It is, however, important to make such an assumption explicit.

In a Bayesian inverse problem, we are interested in the distribution of the random variable $\mathbf{X} | \mathbf{D} = \mathbf{d}$ for a given realization of the data $\mathbf{d} \in \mathcal{D}$. This distribution is called the *posterior distribution* (of \mathbf{X} given $\mathbf{D} = \mathbf{d}$). The distribution $\mathbf{P}_{\mathbf{X}}$, also called *prior distribution* in this context, represents the distributional information on \mathbf{X} *before* any observation of data. This information is updated by the knowledge contained in the data \mathbf{d} in a consistent way.

We get a more explicit expression of the posterior distribution by applying *Bayes' Theorem* (also known as *Bayes' rule*). Bayes' rule, in a first version, states that for two events $E_1, E_2 \in \mathcal{A}$ with $\mathbf{P}(E_1) > 0$, it holds that

$$\mathbf{P}(E_2 | E_1) = \frac{\mathbf{P}(E_1 | E_2)\mathbf{P}(E_2)}{\mathbf{P}(E_1)}. \quad (2.7)$$

This can be proven easily by using the definition of a conditional probability. A second version that is more useful for our purposes makes a similar statement about the involved distributions and their densities. For this, we need to develop a more formal context in order to make precise statements that are necessary for the remainder.

Let $\mathbf{P}_{\mathbf{X}} := \mathbf{P}(\mathbf{X} \in \cdot)$ and $\mathbf{P}_{\mathbf{D}} := \mathbf{P}(\mathbf{D} \in \cdot)$ denote probability distributions of \mathbf{X} and \mathbf{D} , respectively. Moreover, let the joint probability distribution for (\mathbf{X}, \mathbf{D}) be given by $\mathbf{P}_{\mathbf{X}, \mathbf{D}} := \mathbf{P}((\mathbf{X}, \mathbf{D}) \in \cdot)$. It can be shown (see, e. g., [30, p. 430] or [168]) that there exists a *Markov kernel* $K : \mathcal{D} \times \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ (see Def. A.2.12) such that

$$\mathbf{P}_{\mathbf{X}, \mathbf{D}}(A, C) = \int_C K(\mathbf{d}^*, A) \mathbf{P}_{\mathbf{D}}(d\mathbf{d}^*) \quad (2.8)$$

for all $A \in \mathcal{B}(\mathcal{X})$ and $C \in \mathcal{B}(\mathcal{D})$. For $\mathbf{d} \in \mathcal{D}$, the measure $K(\mathbf{d}, \cdot)$ is $\mathbf{P}_{\mathbf{D}}$ -a.s. unique and called the (*regular*) *conditional probability distribution* of \mathbf{X} given $\mathbf{D} = \mathbf{d}$. We denote it in the following by

$$\mathbf{P}_{\mathbf{X} | \mathbf{D}}(\cdot | \mathbf{d}) := \mathbf{P}(\mathbf{X} \in \cdot | \mathbf{D} = \mathbf{d}) := K(\mathbf{d}, \cdot). \quad (2.9)$$

The finite-dimensional setting allows us to assume that the distributions of random variables \mathbf{X} and \mathbf{D} , i. e., $\mathbf{P}_{\mathbf{X}}$ and $\mathbf{P}_{\mathbf{D}}$, have densities w.r.t. the Lebesgue measure which we

denote by $\rho_{\mathbf{X}}$ and $\rho_{\mathbf{D}}$, respectively. Let $\mathbf{x} \in \mathcal{X}$ and $\mathbf{d} \in \mathcal{D}$ be defined almost surely. We know that the joint density of $\mathbf{P}_{\mathbf{X},\mathbf{D}}$, denoted by $\rho_{\mathbf{X},\mathbf{D}}$, satisfies

$$\rho_{\mathbf{X}}(\mathbf{x}) = \int_{\mathcal{D}} \rho_{\mathbf{X},\mathbf{D}}(\mathbf{x}, \mathbf{d}^*) \, \mathrm{d}\mathbf{d}^* \quad \text{and} \quad \rho_{\mathbf{D}}(\mathbf{d}) = \int_{\mathcal{X}} \rho_{\mathbf{X},\mathbf{D}}(\mathbf{x}^*, \mathbf{d}) \, \mathrm{d}\mathbf{x}^*. \quad (2.10)$$

This leads to the fact that we can represent the conditional distribution $\mathbf{P}_{\mathbf{X}|\mathbf{D}}$ by its Lebesgue density $\rho_{\mathbf{X}|\mathbf{D}}$ explicitly. Indeed, it holds that

$$\rho_{\mathbf{X}|\mathbf{D}}(\mathbf{x}|\mathbf{d}) = \begin{cases} \rho_{\mathbf{X},\mathbf{D}}(\mathbf{x}, \mathbf{d}) / \rho_{\mathbf{D}}(\mathbf{d}) & \text{if } \rho_{\mathbf{D}}(\mathbf{d}) > 0, \\ 0 & \text{otherwise.} \end{cases} \quad (2.11)$$

Note that we can use an analogous construction to get an explicit Lebesgue density $\rho_{\mathbf{D}|\mathbf{X}}$ for $\mathbf{P}_{\mathbf{D}|\mathbf{X}}$, i. e., we have

$$\rho_{\mathbf{D}|\mathbf{X}}(\mathbf{d}|\mathbf{x}) = \begin{cases} \rho_{\mathbf{X},\mathbf{D}}(\mathbf{x}, \mathbf{d}) / \rho_{\mathbf{X}}(\mathbf{x}) & \text{if } \rho_{\mathbf{X}}(\mathbf{x}) > 0, \\ 0 & \text{otherwise.} \end{cases} \quad (2.12)$$

With this setup given, we can formulate a version of Bayes' Theorem that is suitable for our context; see, e. g., [164, Thm. 2.5].

Theorem 2.2.1 (Bayes' Theorem). *Let $\mathbf{d} \in \mathcal{D}$ be defined $\mathbf{P}_{\mathbf{D}}$ -almost surely. If $\rho_{\mathbf{D}|\mathbf{X}}(\mathbf{d}|\cdot)$ is in $L^1(\mathcal{X}, \mathbf{P}_{\mathbf{X}})$ and strictly positive, then it holds that $\rho_{\mathbf{D}}(\mathbf{d}) > 0$ and*

$$\rho_{\mathbf{X}|\mathbf{D}}(\mathbf{x}|\mathbf{d}) = \frac{\rho_{\mathbf{D}|\mathbf{X}}(\mathbf{d}|\mathbf{x}) \rho_{\mathbf{X}}(\mathbf{x})}{\rho_{\mathbf{D}}(\mathbf{d})} \quad (2.13)$$

for λ -a. e. $\mathbf{x} \in \mathcal{X}$.

Let us look again at the terms involved in (2.13):

- $\rho_{\mathbf{X}|\mathbf{D}}(\mathbf{x}|\mathbf{d})$: Density of posterior distribution $\mathbf{P}_{\mathbf{X}|\mathbf{D}}(\cdot|\mathbf{d})$.
- $\rho_{\mathbf{D}|\mathbf{X}}(\mathbf{d}|\mathbf{x})$: This expression is called the *likelihood function* (in \mathbf{x}), or just *likelihood*. Note that it is not a probability density function in \mathbf{x} but in \mathbf{d} for λ -a. e. $\mathbf{x} \in \mathcal{X}$. Let $\mathbf{x} \in \mathcal{X}$ such that $\rho_{\mathbf{X}}(\mathbf{x}) > 0$. Using the relation above in (2.5), we get that

$$\rho_{\mathbf{D}|\mathbf{X}}(\mathbf{d}|\mathbf{x}) = \rho_{\boldsymbol{\eta}}(\mathbf{d} - \mathcal{G}(\mathbf{x})), \quad (2.14)$$

where $\rho_{\boldsymbol{\eta}}$ is the Lebesgue density of the noise distribution $\mathbf{P}_{\boldsymbol{\eta}}$. The likelihood is commonly written as

$$\rho_{\boldsymbol{\eta}}(\mathbf{d} - \mathcal{G}(\mathbf{x})) = \exp(-\Phi(\mathbf{x}, \mathbf{d})) \quad (2.15)$$

for the *negative log-likelihood* Φ (also called *potential*). It is also known as *data misfit function* and denoted by

$$f^{\mathbf{d}}(\mathbf{x}) := \Phi(\mathbf{x}, \mathbf{d}). \quad (2.16)$$

- $\rho_{\mathbf{X}}(\mathbf{x})$: Density of the distribution of \mathbf{X} , i. e., of $\mathbf{P}_{\mathbf{X}}$. This density (or distribution) is called prior density (or prior distribution). It should contain all information that is available prior to any data measurements.

- $\rho_D(\mathbf{d})$: Density of the distribution of \mathbf{D} , i.e., of \mathbf{P}_D . It is the normalizing constant for the posterior distribution, also called *model evidence*. Thus, it holds that

$$\rho_D(\mathbf{d}) = \int_{\mathcal{X}} \rho_{D|X}(\mathbf{d}|\mathbf{x}^*) \rho_X(\mathbf{x}^*) d\mathbf{x}^* \quad (2.17)$$

for λ -a.e. $\mathbf{d} \in \mathcal{D}$. Note that $\rho_D(\mathbf{d})$ has to be greater than zero since the ratio in (2.13) would be undefined otherwise. This case is, however, excluded by the assumptions of Bayes' Theorem (Thm. 2.2.1). For example, more exotic noise models like uniform or Dirac type distributions do not satisfy these assumptions and could therefore cause such problems.

Remark. Bayes' Theorem can be applied in a more abstract setting including infinite-dimensional parameter and data spaces under rather weak assumptions. For details, we refer to [73, 252, 253]. However, since the theory then gets more technical and we are not regarding infinite-dimensional spaces in this text, we focus instead on the finite-dimensional case.

In the remainder, we abbreviate the notation of the prior and posterior distribution and their densities with

$$\mu_0 := \mathbf{P}_X, \quad \rho_0 := \rho_X \quad (2.18)$$

and

$$\mu^{\mathbf{d}} := \mathbf{P}_{X|D}(\cdot|\mathbf{d}) = \mathbf{P}(X \in \cdot | D = \mathbf{d}), \quad \rho^{\mathbf{d}} := \rho_{X|D}(\cdot|\mathbf{d}) \quad (2.19)$$

for $\mathbf{d} \in \mathcal{D}$. Summarizing, for the density of the posterior distribution it holds that

$$\begin{aligned} \rho^{\mathbf{d}}(\mathbf{x}) &= \frac{\rho_{\eta}(\mathbf{d} - \mathcal{G}(\mathbf{x})) \rho_0(\mathbf{x})}{\rho_D(\mathbf{d})} \\ &\propto \rho_{\eta}(\mathbf{d} - \mathcal{G}(\mathbf{x})) \rho_0(\mathbf{x}). \end{aligned} \quad (2.20)$$

for λ -a.e. $\mathbf{x} \in \mathcal{X}$ and $\mathbf{d} \in \mathcal{D}$.

In the following, we assume that the observational noise is Gaussian distributed with zero mean and covariance matrix $\Gamma \in \mathbf{R}^{n_d \times n_d}$. Then, we get that

$$\rho_{\eta}(\mathbf{d} - \mathcal{G}(\mathbf{x})) \propto \exp \left(-\frac{1}{2} \|\Gamma^{-1/2}(\mathbf{d} - \mathcal{G}(\mathbf{x}))\|_2^2 \right), \quad (2.21)$$

i.e., $f^{\mathbf{d}}(\mathbf{x}) = \frac{1}{2} \|\mathbf{d} - \mathcal{G}(\mathbf{x})\|_{\Gamma}^2$ in (2.16) with $\|\cdot\|_{\Gamma} := \|\Gamma^{-1/2} \cdot\|_2$. Note that the expression on the right hand side in (2.21) is defined *for every* $\mathbf{x} \in \mathcal{X}$ and $\mathbf{d} \in \mathcal{D}$ (and not only λ -a.s.).

Remark. From a statistical perspective, we perform *parametric inference* in a Bayesian setting, also called *Bayesian inference*, with the statistical model

$$(\mathcal{S}, \{\mathbf{P}_{\theta}\}_{\theta \in \Theta}) := (\mathcal{D}, \{\mathcal{N}(\mathcal{G}(\mathbf{x}), \Gamma)\}_{\mathbf{x} \in \mathcal{X}}), \quad (2.22)$$

where \mathcal{S} is the set of possible observations and $\{\mathbf{P}_{\theta}\}_{\theta \in \Theta}$ is a *parametrized* family of distributions with parameter θ from a set Θ .

If we additionally suppose that the prior distribution is also Gaussian distributed with mean $\mathbf{m}_0 \in \mathbf{R}^n$ and covariance matrix $\Sigma_0 \in \mathbf{R}^{n \times n}$, the corresponding posterior density becomes

$$\rho^{\mathbf{d}}(\mathbf{x}) \propto \exp \left(-\frac{1}{2} \|\mathbf{d} - \mathcal{G}(\mathbf{x})\|_{\Gamma}^2 - \frac{1}{2} \|\mathbf{x} - \mathbf{m}_0\|_{\Sigma_0}^2 \right). \quad (2.23)$$

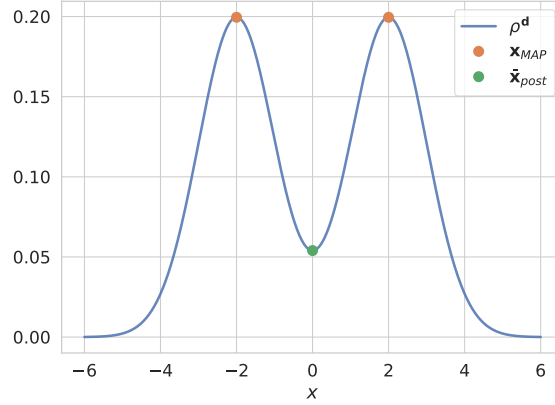


Figure 2.1: Multimodal posterior distribution with a non-unique MAP estimator and a mean value as a local minimum.

Note that the expression in the exponential is the same (up to the sign) as the objective of the optimization for a Tikhonov-regularized classical inverse problem in (2.3). That is, the minimizer \mathbf{x}^\dagger from (2.3) maximizes the posterior density from (2.23).

Such a point estimate for the input variable \mathbf{x} is generally called a *maximum a posteriori estimator* (or MAP estimator), i. e.,

$$\mathbf{x}_{\text{MAP}} := \arg \max_{\mathbf{x} \in \mathcal{X}} \rho^{\mathbf{d}}(\mathbf{x}). \quad (2.24)$$

It is also a *mode* of the posterior distribution. Also, we can regard the mean value of the posterior distribution, i. e.,

$$\bar{\mathbf{x}}_{\text{post}} := \mathbf{E}[\mathbf{X} \mid \mathbf{D} = \mathbf{d}]. \quad (2.25)$$

However, note that these estimators can be misleading in general (non-unique MAP estimator, mean value as a local minimum) as the situation in Fig. 2.1 shows.

Another alternative point estimate from frequentist statistics is the *maximum likelihood estimator* (MLE) [202, 228]. This approach is not using Bayes' Theorem at all, and thus no prior information, but instead it is interested in the estimate of a point of maximum likelihood \mathbf{x}_{MLE} , i. e.,

$$\mathbf{x}_{\text{MLE}} := \arg \max_{\mathbf{x} \in \mathcal{X}} \rho_{\mathbf{D}|\mathbf{X}}(\mathbf{d}|\mathbf{x}) = \arg \min_{\mathbf{x} \in \mathcal{X}} \Phi(\mathbf{x}, \mathbf{d}) \quad (2.26)$$

for given $\mathbf{d} \in \mathcal{D}$.

Related literature

As mentioned, the main references for our setting are [73, 252]. Further general introductions to the Bayesian approach to statistical inverse problems are provided in [23, 146].

The discretization error ε from (2.6) and the noise term $\boldsymbol{\eta}$ can also be modeled separately; see, e. g., [146, Ch. 7] and [104, 132, 151, 206, 208]. It is also possible to model this error type by extending the variable \mathbf{X} to include model aspects as, e. g., missing physics; see [63].

We also give some references for *large-scale* computations in Bayesian inverse problems. For highly expensive models, linear approximations of the forward operator \mathcal{G} are often computed in

order to make the resulting Bayesian inverse problem computationally tractable. A framework for corresponding infinite-dimensional inferences is provided in [42, 44]. In [89], the authors suggest algorithms using low-rank Hessian approximations of the data misfit function $f^{\mathbf{d}}$ for linear problems. For general large-scale problems, [43] proposes an adaptive algorithm employing Gaussian processes [283] for a Hessian-based approximation of posterior densities.

Finally, we refer to [280] for Bayesian inverse problems involving priors that rely on the ℓ_1 norm of the parameters including distributions with regularity less than Gaussian distributions allowing for, e. g., discontinuities and blockiness.

2.3 An example

A first example should provide some intuition for the behavior of the involved quantities in Bayesian inverse problems. We regard a basic situation in which the parameter-to-QoI operator \mathcal{G} is linear, i. e., $\mathcal{G}(\mathbf{x}) := A\mathbf{x}$ for some matrix $A \in \mathbf{R}^{n_d \times n}$, and the prior as well as the observational noise is Gaussian, i. e., $\mathbf{X} \sim \mathcal{N}(\mathbf{m}_0, \Sigma_0)$ and $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \Gamma)$ for covariance matrices $\Sigma_0 \in \mathbf{R}^{n \times n}$ and $\Gamma \in \mathbf{R}^{n_d \times n_d}$. Remember that

$$\mathbf{D}|\mathbf{X} = \mathbf{x} \sim \mathcal{N}(A\mathbf{x}, \Gamma) \quad (2.27)$$

according to (2.4). Also, note that the random vector (\mathbf{X}, \mathbf{D}) is jointly Gaussian (which is not true in general for *dependent* Gaussian random variables). This can be seen, for example, by computing the joint density of (\mathbf{X}, \mathbf{D}) . Indeed, if we define $\mathbf{b} := (\mathbf{x}, \mathbf{d})$, we get

$$\begin{aligned} \rho_{\mathbf{X}, \mathbf{D}}(\mathbf{x}, \mathbf{d}) &= \rho_{\mathbf{D}|\mathbf{X}}(\mathbf{d}|\mathbf{x})\rho_{\mathbf{X}}(\mathbf{x}) \\ &\propto \exp\left(-\frac{1}{2}\|\mathbf{d} - A\mathbf{x}\|_{\Gamma}^2 - \frac{1}{2}\|\mathbf{x} - \mathbf{m}_0\|_{\Sigma_0}^2\right) \\ &= \exp\left(-\frac{1}{2}\psi(\mathbf{x}, \mathbf{d})\right) \end{aligned} \quad (2.28)$$

for

$$\begin{aligned} \psi(\mathbf{x}, \mathbf{d}) &= \|(-A \ I)\mathbf{b}\|_{\Gamma}^2 + \|(I \ 0)(\mathbf{b} - \mathbf{b}_0)\|_{\Sigma_0}^2 \\ &= \mathbf{b}^{\top}\Sigma_1\mathbf{b} + (\mathbf{b} - \mathbf{b}_0)^{\top}\Sigma_2(\mathbf{b} - \mathbf{b}_0), \end{aligned} \quad (2.29)$$

where

$$\Sigma_1 := (-A \ I)^{\top}\Gamma^{-1}(-A \ I), \quad \Sigma_2 := (I \ 0)^{\top}\Sigma_0^{-1}(I \ 0), \quad (2.30)$$

and $\mathbf{b}_0 := (\mathbf{m}_0, *)$. The $*$ means that it does not matter which values are at its place; see (2.29). Completing the squares with the help of Lem. A.2.5 gives

$$\psi(\mathbf{x}, \mathbf{d}) = (\mathbf{b} - M^{-1}\mathbf{v})^{\top}M(\mathbf{b} - M^{-1}\mathbf{v}) + R, \quad (2.31)$$

where

$$M := \Sigma_1 + \Sigma_2, \quad \mathbf{v} := \Sigma_2\mathbf{b}_0, \quad (2.32)$$

and R is a constant value independent of \mathbf{b} . Note that M is a square matrix with full rank and is thus invertible. Hence, $\rho_{\mathbf{X}, \mathbf{D}}$ can be written as a Gaussian density implying that (\mathbf{X}, \mathbf{D}) is jointly Gaussian.

Using basic properties of covariance operators (see Def. A.2.1) and the independence of \mathbf{X} and $\boldsymbol{\eta}$, we get that

$$\begin{aligned}\mathbf{Cov}(\mathbf{D}) &= \mathbf{Cov}(A\mathbf{X}) + \mathbf{Cov}(\boldsymbol{\eta}) = A\Sigma_0A^\top + \Gamma, \\ \mathbf{Cov}(\mathbf{X}, \mathbf{D}) &= \mathbf{Cov}(\mathbf{X}, A\mathbf{X} + \boldsymbol{\eta}) = \mathbf{Cov}(\mathbf{X}, A\mathbf{X}) + \mathbf{Cov}(\mathbf{X}, \boldsymbol{\eta}) = \Sigma_0A^\top, \\ \mathbf{Cov}(\mathbf{D}, \mathbf{X}) &= \mathbf{Cov}(\mathbf{X}, \mathbf{D})^\top = A\Sigma_0.\end{aligned}\tag{2.33}$$

That is, we have that $(\mathbf{X}, \mathbf{D}) \sim \mathcal{N}(\bar{\mathbf{m}}, \bar{\Sigma})$ with

$$\bar{\mathbf{m}} = (\mathbf{m}_0, A\mathbf{m}_0) \quad \text{and} \quad \bar{\Sigma} = \begin{pmatrix} \Sigma_0 & \Sigma_0A^\top \\ A\Sigma_0 & A\Sigma_0A^\top + \Gamma \end{pmatrix}.\tag{2.34}$$

By Thm. A.2.6, it follows that $\mathbf{X}|\mathbf{D} = \mathbf{d} \sim \mathcal{N}(\mathbf{m}, \Sigma)$ with

$$\begin{aligned}\mathbf{m} &= \mathbf{m}_0 + \Sigma_0A^\top(\Gamma + A\Sigma_0A^\top)^{-1}(\mathbf{d} - A\mathbf{m}_0), \\ \Sigma &= \Sigma_0 - \Sigma_0A^\top(\Gamma + A\Sigma_0A^\top)^{-1}A\Sigma_0 \\ &= (\Sigma_0^{-1} + A^\top\Gamma^{-1}A)^{-1}.\end{aligned}\tag{2.35}$$

The last equation is an application of the *Woodbury matrix identity* (Lem. A.1.2), also known as *Sherman–Morrison–Woodbury formula*.

If $\mathbf{d} = A\mathbf{m}_0$, the posterior mean \mathbf{m} is not updated what makes sense intuitively. For the posterior covariance matrix Σ , the prior covariance matrix Σ_0 is updated by subtracting a symmetric positive semi-definite matrix which is “informative” for the posterior in the sense that (co-)variances get smaller. Also, if the noise in the data gets larger, then, by the last line in (2.35), we see that Σ stays more similar to Σ_0 . It is also interesting to note that the posterior covariance Σ does not depend on the realization of data \mathbf{d} that we condition on. This is a special property of conditional Gaussian distributions; see Thm. A.2.6.

Let us continue with a special case to gain even more intuition. Since we want to regard a one-dimensional case, we denote all variables with normal (thin) letters and denote variances with σ^2 and γ^2 instead of Σ and Γ , respectively. So, suppose we have $n = n_{\mathbf{d}} = 1$, $X \sim \mathcal{N}(0, 1)$, $\eta \sim \mathcal{N}(0, \gamma^2)$, $A = a \in \mathbf{R}$, and $d = ax_0$ for some $x_0 \in \mathbf{R}$. Recall that the relationship between data and parameters is

$$D = aX + \eta.\tag{2.36}$$

By (2.35), we get that

$$m = \frac{a^2x_0}{\gamma^2 + a^2} \quad \text{and} \quad \sigma^2 = 1 - \frac{a^2}{\gamma^2 + a^2} = \frac{\gamma^2}{\gamma^2 + a^2}.\tag{2.37}$$

For $a = 0.2$, $x_0 = 2$, and multiple $\gamma^2 \in \{0.01, 0.02, 0.05, 0.1\}$, we get posterior distributions depicted in Fig. 2.2. We see that smaller noise levels cause more concentrated posterior distributions with means approaching the “true” value x_0 . This matches the intuition since smaller noises mean more accurate data, and more accurate data mean that we can be more certain about our estimation.

However, if we transfer this simple example into a more general one with higher dimensions, the results can change. In general, the so-called *small noise limit* of the posterior depends on whether the problem is under- or overdetermined and on the rank of the linear operator A . If we

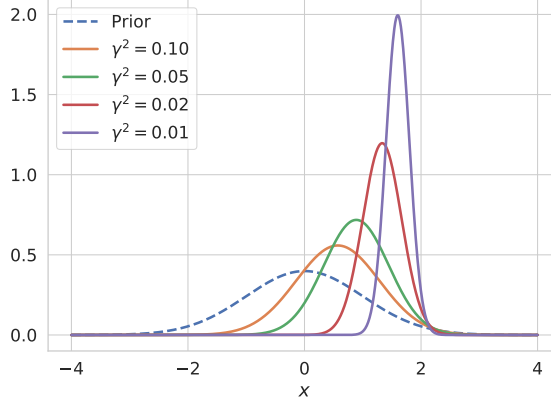


Figure 2.2: In the described one-dimensional case, smaller noises give a more concentrated and informed posterior distribution.

assume that A has full rank, then in the small noise limit the prior can play an important role in the (frequently occurring) underdetermined case, whereas in the overdetermined case the prior loses his influence on the posterior which would become a Dirac measure in this situation. For more details that also include all the assumptions, we refer to [252, Example 2.1, 2.2 and Thm. 2.3, 2.5].

2.4 Well-posedness

In Sec. 2.1, we mentioned that classical inverse problems and also their regularization are generally ill-posed, i.e., one or more points listed in Def. 2.1.1, which are necessary for a problem in order to be well-posed, are not fulfilled. This fact changes for Bayesian inverse problems. Since it is not a point estimate for the parameters that we are looking for but rather a (posterior) distribution on the space of parameters, the corresponding problem seems to contain a sort of “natural regularity.” Indeed, inverse problems formulated in a Bayesian setting, as we did in Sec. 2.2, can be shown to be well-posed in the sense of Def. 2.1.1.

In finite dimensions, only the assumptions of Bayes’ Theorem (Thm. 2.2.1) are necessary in order to have a unique (more precisely, λ -a.s. unique) posterior density $\rho_{\mathbf{X}|\mathbf{D}}(\mathbf{x}|\mathbf{d})$. These include, for example, a positive density of the data \mathbf{d} that we condition on, i.e., $\rho_{\mathbf{D}}(\mathbf{d}) > 0$. However, the continuous dependence of the posterior w.r.t. the data (measured with the Hellinger distance; see Def. A.2.8) requires additional assumptions on the forward operator \mathcal{G} .

Assumption 2.4.1 ([252, Assumption 2.7]). The forward operator $\mathcal{G} : \mathcal{X} \rightarrow \mathcal{D}$ satisfies the following properties:

- (i) For every $\varepsilon > 0$, there is an $M = M(\varepsilon) \in \mathbf{R}$ such that for all $\mathbf{x} \in \mathcal{X}$ it holds that

$$\|\mathcal{G}(\mathbf{x})\|_{\Gamma} \leq \exp(\varepsilon\|\mathbf{x}\|_{\mathcal{X}}^2 + M). \quad (2.38)$$

- (ii) For every $r > 0$, there is a $K = K(r) > 0$ such that for all $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{X}$ with

$$\max\{\|\mathbf{x}_1\|_{\mathcal{X}}, \|\mathbf{x}_2\|_{\mathcal{X}}\} < r \quad (2.39)$$

it holds that

$$\|\mathcal{G}(\mathbf{x}_1) - \mathcal{G}(\mathbf{x}_2)\|_{\Gamma} \leq K \|\mathbf{x}_1 - \mathbf{x}_2\|_{\mathcal{X}}. \quad (2.40)$$

The first assumption requires \mathcal{G} to satisfy a type of local bound, whereas the second one is a Lipschitz property. Applying [252, Corollary 4.4], we have the following continuity result.

Theorem 2.4.2. *Assume $\mathcal{G} : \mathcal{X} \rightarrow \mathcal{D}$ satisfies Assumption 2.4.1. Also, assume that the prior μ_0 is Gaussian. Then, the posterior $\mu^{\mathbf{d}}$, defined in (2.19), is Lipschitz continuous in the data w.r.t. the Hellinger distance, i. e., if $\mu^{\mathbf{d}}$ and $\mu^{\mathbf{d}'}$ denote two posterior measures corresponding to $\mathbf{d}, \mathbf{d}' \in \mathcal{D}$, respectively, then for every $r > 0$ there is a constant $C = C(r) > 0$ such that for all \mathbf{d}, \mathbf{d}' with $\max\{\|\mathbf{d}\|_{\Gamma}, \|\mathbf{d}'\|_{\Gamma}\} < r$ it holds that*

$$d_{\text{Hell}}(\mu^{\mathbf{d}}, \mu^{\mathbf{d}'}) \leq C \|\mathbf{d} - \mathbf{d}'\|_{\Gamma}. \quad (2.41)$$

Proof. We refer to the proof of a more general result for infinite-dimensional parameter and data spaces; see [252, Thm. 4.2]. ■

Remark. Note that Thm. 2.4.2 gives *Lipschitz* continuity in the data, which is a rather strong property. Therefore, the corresponding proof uses a tail property of the prior distribution μ_0 according to Fernique's theorem (Thm. A.2.7). The assumption on the prior μ_0 to be Gaussian, which is sufficient for satisfying the mentioned tail property, can actually be generalized. In fact, each distribution with tails not heavier than Gaussian tails can be used as a prior such that the results of Thm. 2.4.2 still hold. In particular, this includes any distribution with compact support as, e. g., uniform distributions.

Although it is not part of the original well-posedness definition given in Def. 2.1.1, we also comment on a stability property of posterior measures. Suppose that there is a sequence of forward operators $(\mathcal{G}_{\ell})_{\ell \in \mathbf{N}}$ converging to a limit \mathcal{G} in some sense to be described later. A natural question is whether the sequence of corresponding posterior distributions $(\mu_{\ell}^{\mathbf{d}})_{\ell \in \mathbf{N}}$ is also converging to a limit $\mu^{\mathbf{d}}$ that is related to \mathcal{G} (for a suitable fixed $\mathbf{d} \in \mathcal{D}$). That is, we want to translate the approximation of \mathcal{G} into a statement regarding the corresponding approximation of $\mu^{\mathbf{d}}$.

Recall that \mathcal{G} and hence $\Phi(\mathbf{x}, \mathbf{d}) := \frac{1}{2} \|\mathbf{d} - \mathcal{G}(\mathbf{x})\|_{\Gamma}^2$ induce a posterior distribution $\mu^{\mathbf{d}}$ with density

$$\rho^{\mathbf{d}}(\mathbf{x}) \propto \exp(-\Phi(\mathbf{x}, \mathbf{d})) \rho_0(\mathbf{x}). \quad (2.42)$$

For an approximation of the exact forward operator \mathcal{G} , denoted by \mathcal{G}_{ℓ} , we define

$$\Phi_{\ell}(\mathbf{x}, \mathbf{d}) := \frac{1}{2} \|\mathbf{d} - \mathcal{G}_{\ell}(\mathbf{x})\|_{\Gamma}^2. \quad (2.43)$$

According to (2.42), Φ_{ℓ} induces a posterior distribution $\mu_{\ell}^{\mathbf{d}}$ with density

$$\rho_{\ell}^{\mathbf{d}}(\mathbf{x}) \propto \exp(-\Phi_{\ell}(\mathbf{x}, \mathbf{d})) \rho_0(\mathbf{x}). \quad (2.44)$$

The following theorem shows that the approximation quality of \mathcal{G}_{ℓ} translates to the posterior approximation quality in the Hellinger distance; see [252, Corollary 4.9].

Theorem 2.4.3. *Let $\ell \in \mathbf{N}$. Assume that the posterior densities given in (2.42) and (2.44) are well-defined and that for any $\varepsilon > 0$ there is $K' = K'(\varepsilon) > 0$ such that*

$$|\mathcal{G}(\mathbf{x}) - \mathcal{G}_\ell(\mathbf{x})| \leq K' \exp(\varepsilon \|\mathbf{x}\|_{\mathcal{X}}^2) \psi(\ell) \quad (2.45)$$

for all $\mathbf{x} \in \mathcal{X}$, where $\psi : \mathbf{N} \rightarrow \mathbf{R}$ such that $\psi(\ell) \rightarrow 0$ as $\ell \rightarrow \infty$. If, additionally, \mathcal{G} and \mathcal{G}_ℓ satisfy Assumption 2.4.1(i) uniformly in ℓ , there is a constant $C' > 0$, independent of ℓ , such that

$$d_{\text{Hell}}(\mu^{\mathbf{d}}, \mu_\ell^{\mathbf{d}}) \leq C' \psi(\ell). \quad (2.46)$$

Remark. Consequently, if the sequence of forward operators $(\mathcal{G}_\ell)_{\ell \in \mathbf{N}}$ converge to \mathcal{G} in the sense made precise in (2.45), then the corresponding posterior approximations also converge to the posterior distribution $\mu^{\mathbf{d}}$ corresponding to \mathcal{G} . In fact, this result can be derived as a corollary from a stronger result involving infinite-dimensional data; see [252, Thm. 4.6].

Related literature

To study well-posedness of a Bayesian inverse problem as it is done above, it is necessary to decide for a certain concept of continuity that is supposed to be investigated. The conventional setting from [73, 252], which we also apply above, is based on Lipschitz continuity which might however be difficult to verify in practice.

The author of [164] studies well-posedness regarding the standard form of continuity and also provides a proper measure-theoretic setup which is used above for Thm. 2.2.1 (Bayes' Theorem). Note that respective continuity results do not need to assume a tail property on the prior distribution according to Fernique's theorem as above. However, these weaker assumptions “only” yield *continuity* instead of *Lipschitz continuity* in the data as above.

Furthermore, local Lipschitz stability of Bayesian inverse problems is investigated in [248].

2.5 Sampling the posterior

In most cases, it is not trivial to create samples that follow a posterior distribution in Bayesian inverse problems. In contrast to the example discussed in Sec. 2.3, the posterior distribution is often not cheaply available, i. e., the posterior density

$$\rho^{\mathbf{d}}(\mathbf{x}) \propto \exp\left(-\frac{1}{2}\|\mathbf{d} - \mathcal{G}(\mathbf{x})\|_{\Gamma}^2\right) \rho_0(\mathbf{x}) \quad (2.47)$$

can be computationally expensive to evaluate. Note that most of the computational cost usually lies in the execution of the model which is hidden in the parameter-to-QoI map \mathcal{G} .

Many algorithms, methods, and techniques for the generation of posterior samples were proposed in recent decades. In Subsec. 2.5.1, we first look at rather straightforward techniques that underlie simple ideas, are easy to implement, and are suitable to gain intuition for sampling from a distribution. As we will see, these fundamental approaches are too primitive for more computationally expensive and high-dimensional Bayesian inverse problems. Therefore, we subsequently investigate *Markov chain Monte Carlo* (MCMC) methods in more depth in Subsec. 2.5.2 since we will use a particular one in later chapters.

2.5.1 Fundamental strategies

We first describe a class of strategies that sample from a different distribution than the one of actual interest. The constructed samples are then *corrected* in order to follow the target distribution. For example, *rejection sampling* (see, e. g., [184, Sec. 2.2] or [222, Sec. 2.3]) and *importance sampling* (see, e. g., [184, Sec. 2.5] or [222, Sec. 3.3]) belong to this type of class.

Rejection sampling

In rejection sampling, we create samples following a distribution that is easier, or cheaper, to sample and subsequently *reject* or *accept* these samples according to a certain probability involving the density of the actual target distribution.

Assume that we are given a target density ρ and that we want to create samples from the distribution induced by ρ . The simple idea behind rejection sampling is based on the identity

$$\rho(\mathbf{x}) = \int_0^{\rho(\mathbf{x})} 1 \, d\mathbf{u} = \int_0^\infty \underbrace{\mathbf{1}_{\{\mathbf{u}' \mid 0 < \mathbf{u}' < \rho(\mathbf{x})\}}(\mathbf{u})}_{=:h(\mathbf{x}, \mathbf{u})} \, d\mathbf{u} \quad (2.48)$$

for all $\mathbf{x} \in \mathcal{X}$. Note that this identity allows to interpret ρ as the marginal density of the uniform distribution under the area of the curve of ρ (denoted by $h(\mathbf{x}, \mathbf{u})$). Thus, we can generate a sample according to ρ by sampling uniformly from the area under ρ , i. e., $\{(\mathbf{x}, \mathbf{u}) \mid 0 < \mathbf{u} < \rho(\mathbf{x})\}$, and then keeping the \mathbf{x} part of the pairs.

In general, it is not easy to sample uniformly from the mentioned area. However, we can take a second density q , the so-called *proposal density* inducing a *proposal distribution* from which we can easily sample, such that its area is wrapping (up to a multiplicative constant) the area under ρ . Uniform samples from the area under q are then rejected or accepted according to their position relative to the curve of ρ .

The idea is illustrated in Fig. 2.3 showing a target density ρ of a Beta(2,2) distribution and a scaled uniform density q around ρ . This example, however, uses the bounded support of the target density. In fact, it is sufficient to find a proposal density q and a constant $M > 1$ such that $\rho(\mathbf{x}) < Mq(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{X}$.

The steps for rejection sampling can be found in Alg. 1. It can be shown that they refer back to the situation in (2.48) and Fig. 2.3. The algorithm is justified by the following theorem.

Algorithm 1: Rejection sampling

Input : Target distribution with density ρ , proposal density q , and constant $M > 1$ such that $\rho(\mathbf{x}) < Mq(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{X}$.

Output: Samples following the distribution induced by ρ .

We generate a sample according to the target density ρ by the following two steps.

1. Draw a sample \mathbf{Y} from the proposal distribution induced by q .
2. Accept \mathbf{Y} with probability

$$\frac{\rho(\mathbf{Y})}{Mq(\mathbf{Y})}, \quad (2.49)$$

otherwise go back to step 1.

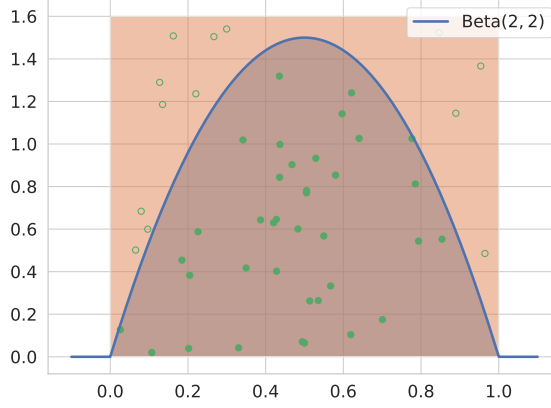


Figure 2.3: Illustration of the idea behind rejection sampling. The filled green dots are accepted uniform samples under the area of the density of a Beta(2, 2) distribution with bounded support in $[0, 1]$. The orange area shows the area under the curve of a scaled uniform (proposal) density on $[0, 1]$.

Theorem 2.5.1. *Given a target density ρ , a proposal density q , and a constant $M > 1$ such that $\rho(\mathbf{x}) < Mq(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{X}$, Alg. 1 generates samples from the target distribution induced by ρ .*

Proof. Define

$$Z := \text{Bernoulli}\left(\frac{\rho(\mathbf{Y})}{Mq(\mathbf{Y})}\right). \quad (2.50)$$

So, we have that $\{\mathbf{Y} \text{ is accepted}\} = \{Z = 1\}$ and $\{\mathbf{Y} \text{ is rejected}\} = \{Z = 0\}$. For the sake of clearness of this proof, we set $\rho_{\mathbf{Y}} := q$, denote the density of Z with ρ_Z , and the joint density of (\mathbf{Y}, Z) with $\rho_{\mathbf{Y}, Z}$.

We have to show that

$$\mathbf{P}(\mathbf{Y} \in A \mid \mathbf{Y} \text{ is accepted}) = \mathbf{P}(\mathbf{Y} \in A \mid Z = 1) = \int_A \rho(\mathbf{y}) \, d\mathbf{y} \quad (2.51)$$

for each $A \in \mathcal{B}(\mathcal{X})$. Let $A \in \mathcal{B}(\mathcal{X})$. Using the definition of conditional probabilities, we get that

$$\mathbf{P}(\mathbf{Y} \in A \mid Z = 1) = \frac{\mathbf{P}(\mathbf{Y} \in A, Z = 1)}{\mathbf{P}(Z = 1)}. \quad (2.52)$$

We compute

$$\begin{aligned} \mathbf{P}(\mathbf{Y} \in A, Z = 1) &= \int_A \rho_{\mathbf{Y}, Z}(\mathbf{y}, 1) \, d\mathbf{y} \\ &= \int_A \rho_{\mathbf{Y}}(\mathbf{y}) \rho_{Z \mid \mathbf{Y}}(1 \mid \mathbf{y}) \, d\mathbf{y} \\ &= \int_A q(\mathbf{y}) \frac{\rho(\mathbf{y})}{Mq(\mathbf{y})} \, d\mathbf{y} \\ &= \frac{1}{M} \int_A \rho(\mathbf{y}) \, d\mathbf{y}. \end{aligned} \quad (2.53)$$

Also,

$$\begin{aligned}\mathbf{P}(Z = 1) &= \mathbf{P}(\mathbf{Y} \in \mathcal{X}, Z = 1) \\ &= \frac{1}{M} \int_{\mathcal{X}} \rho(\mathbf{y}) \, d\mathbf{y} = \frac{1}{M}.\end{aligned}\tag{2.54}$$

Hence, we get the equality from (2.51) by combining (2.53) and (2.54). \blacksquare

Remark. The steps of the proof can also be carried out if the target density ρ is only known up to a constant, i. e., we only know $\bar{\rho}$, where $\rho = C\bar{\rho}$, $C > 0$, provided it still holds that $\bar{\rho}(\mathbf{x}) < Mq(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{X}$. The acceptance probability is then

$$\frac{\bar{\rho}(\mathbf{Y})}{Mq(\mathbf{Y})}.\tag{2.55}$$

Importance sampling

Importance sampling, on the other hand, corrects samples from a simpler distribution not by accepting or rejecting but by *(re-)weighting*. We are again given a target density ρ inducing a distribution which is difficult to sample, and a proposal density q inducing a proposal distribution which is simple to sample.

For a random variable \mathbf{X} distributed according to the distribution induced by ρ , importance sampling is based on the identity

$$\mathbf{P}(\mathbf{X} \in A) = \int_A \rho(\mathbf{x}) \, d\mathbf{x} = \int_A \frac{\rho(\mathbf{x})}{q(\mathbf{x})} q(\mathbf{x}) \, d\mathbf{x} = \int_A w(\mathbf{x}) q(\mathbf{x}) \, d\mathbf{x}\tag{2.56}$$

for arbitrary $A \in \mathcal{B}(\mathcal{X})$ and $w(\mathbf{x}) := \rho(\mathbf{x})/q(\mathbf{x})$. Note that $q(\mathbf{x})$ has to be greater than zero for λ -a.e. $\mathbf{x} \in \mathcal{X}$ with $\rho(\mathbf{x}) > 0$.

Denoting

$$\mathbf{E}_{\pi}[g(\mathbf{X})] := \int_{\mathcal{X}} g(\mathbf{x}) \pi(\mathbf{x}) \, d\mathbf{x}\tag{2.57}$$

for a measurable function g and a probability density π , we can derive a generalization of the identity from (2.56). Indeed, it holds that

$$\mathbf{E}[h(\mathbf{X})] = \int_{\mathcal{X}} h(\mathbf{x}) \rho(\mathbf{x}) \, d\mathbf{x} = \int_{\mathcal{X}} h(\mathbf{x}) w(\mathbf{x}) q(\mathbf{x}) \, d\mathbf{x} = \mathbf{E}_q[h(\mathbf{X})w(\mathbf{X})]\tag{2.58}$$

for a real-valued ρ -integrable function h . Here, we need to assume that $q(\mathbf{x}) > 0$ for λ -a.e. $\mathbf{x} \in \mathcal{X}$ with $h(\mathbf{x})\rho(\mathbf{x}) \neq 0$.

For random variables $\mathbf{X}_1, \dots, \mathbf{X}_N$ which are independently identically distributed (i.i.d.) according to q , by the strong law of large numbers (Thm. A.2.11) we get that

$$\frac{1}{N} \sum_{i=1}^N h(\mathbf{X}_i) w(\mathbf{X}_i) \xrightarrow[N \rightarrow \infty]{a.s.} \mathbf{E}_q[h(\mathbf{X})w(\mathbf{X})] = \mathbf{E}[h(\mathbf{X})]\tag{2.59}$$

provided that $\mathbf{E}_q[|h(\mathbf{X})w(\mathbf{X})|]$ exists. This means that we can estimate $\mu := \mathbf{E}[h(\mathbf{X})]$ by

$$\hat{\mu}_N := \frac{1}{N} \sum_{i=1}^N h(\mathbf{X}_i) w(\mathbf{X}_i).\tag{2.60}$$

However, note that the weights $w(\mathbf{X}_i)$, $i = 1, \dots, N$, do not sum up to N in general. Hence, we might want to use a normalized version $\hat{\mu}_w$ of $\hat{\mu}_N$ to estimate μ with

$$\hat{\mu}_w := \frac{1}{\sum_{i=1}^N w(\mathbf{X}_i)} \sum_{i=1}^N h(\mathbf{X}_i) w(\mathbf{X}_i). \quad (2.61)$$

The steps of importance sampling are summarized in Alg. 2.

Algorithm 2: Importance sampling

Input : Function h , target density ρ , and proposal density q such that the assumptions from above are fulfilled.

Output: Estimator for $\mathbf{E}_\rho[h(\mathbf{X})]$.

- 1 Draw samples $\mathbf{X}_1, \dots, \mathbf{X}_N$ from q .
 - 2 Set $w(\mathbf{X}_i) = \rho(\mathbf{X}_i)/q(\mathbf{X}_i)$ for $i = 1, \dots, N$.
 - 3 Return either $\hat{\mu}_N$ (2.60) or $\hat{\mu}_w$ (2.61).
-

Statistical properties like bias and variance of the two described estimators $\hat{\mu}_N$ and $\hat{\mu}_w$ are proven in the following theorem.

Analogously to the definition of \mathbf{E}_π above, we set

$$\mathbf{Var}_\pi(Z_1) := \mathbf{E}_\pi[(Z_1 - \mathbf{E}_\pi[Z_1])^2] \quad (2.62)$$

and

$$\mathbf{Cov}_\pi(Z_1, Z_2) := \mathbf{E}_\pi[(Z_1 - \mathbf{E}_\pi[Z_1])(Z_2 - \mathbf{E}_\pi[Z_2])] \quad (2.63)$$

for random variables Z_1 and Z_2 .

Theorem 2.5.2 (Bias and variance of importance sampling, [184, Sec. 2.5]). *The following identities hold:*

a)

$$\mathbf{E}_q[\hat{\mu}_N] = \mu, \quad (2.64)$$

b)

$$\mathbf{Var}_q(\hat{\mu}_N) = \frac{1}{N} \mathbf{Var}_q(h(\mathbf{X})w(\mathbf{X})), \quad (2.65)$$

c)

$$\begin{aligned} \mathbf{E}_q[\hat{\mu}_w] &= \mu + \frac{1}{N} \left(\mu \mathbf{Var}_q(w(\mathbf{X})) - \mathbf{Cov}_q(w(\mathbf{X}), h(\mathbf{X})w(\mathbf{X})) \right. \\ &\quad \left. + \mu^2 \mathbf{Var}_q(w(\mathbf{X})) \right) + O(N^{-2}), \end{aligned} \quad (2.66)$$

d)

$$\begin{aligned} \mathbf{Var}_q(\hat{\mu}_w) &= \frac{1}{N} \left(\mathbf{Var}_q(h(\mathbf{X})w(\mathbf{X})) - 2\mu \mathbf{Cov}_q(w(\mathbf{X}), h(\mathbf{X})w(\mathbf{X})) \right. \\ &\quad \left. + \mu^2 \mathbf{Var}_q(w(\mathbf{X})) \right) + O(N^{-2}). \end{aligned} \quad (2.67)$$

The theorem above has some implications. First, we note that the estimator $\hat{\mu}_N$ is unbiased whereas $\hat{\mu}_w$ is biased; see Thm. 2.5.2a) and c). However, $\hat{\mu}_w$ might have a smaller variance than $\hat{\mu}_N$; see Thm. 2.5.2b) and d). Secondly, we can see that the normalized estimator $\hat{\mu}_w$ has another advantage over $\hat{\mu}_N$. The estimator $\hat{\mu}_w$ requires that we know the target density ρ only up to a constant, i. e., it suffices to know $\bar{\rho}$ such that $\rho = C\bar{\rho}$ for some constant $C > 0$. Indeed, for $w(\mathbf{X}_i) = C\bar{\rho}(\mathbf{X}_i)/q(\mathbf{X}_i)$ we compute that

$$\begin{aligned}\hat{\mu}_w &= \frac{1}{\sum_{i=1}^N \frac{C\bar{\rho}(\mathbf{X}_i)}{q(\mathbf{X}_i)}} \sum_{i=1}^N h(\mathbf{X}_i) \frac{C\bar{\rho}(\mathbf{X}_i)}{q(\mathbf{X}_i)} \\ &= \frac{1}{\sum_{i=1}^N \frac{\bar{\rho}(\mathbf{X}_i)}{q(\mathbf{X}_i)}} \sum_{i=1}^N h(\mathbf{X}_i) \frac{\bar{\rho}(\mathbf{X}_i)}{q(\mathbf{X}_i)},\end{aligned}\tag{2.68}$$

i. e., we only need $\bar{\rho}$ to compute $\hat{\mu}_w$. This can be a huge advantage, especially in Bayesian contexts where the normalization constant of a posterior distribution (see (2.17)) is often not available due to intractable computational costs.

A final important question is whether there exists an *optimal* proposal density q^* in the sense that q^* minimizes the variance of the estimators. For the unbiased estimator $\hat{\mu}_N$, it was shown in [145] that there exists such an optimal proposal density.

Theorem 2.5.3 (Optimal proposal density for $\hat{\mu}_N$). *The proposal density*

$$q^*(\mathbf{x}) := \frac{|h(\mathbf{x})|\rho(\mathbf{x})}{\int_{\mathcal{X}} |h(\mathbf{x}')|\rho(\mathbf{x}') d\mathbf{x}'},\tag{2.69}$$

for $\mathbf{x} \in \mathcal{X}$, is optimal, i. e., it minimizes $q \mapsto \mathbf{Var}_q(\hat{\mu}_N)$.

Proof. We refer to [145]. ■

Remark. The optimality result is rather formal and not directly applicable in practice. In fact, we would need to know the normalization constant from (2.69) which is (up to the absolute value in the integral) exactly the quantity that we want to estimate. So, since the proposal distribution was assumed to produce cheap samples, the optimal proposal density from above is not usable. However, an important implication of this result is that we should choose a proposal density that is similar in shape to $\mathbf{x} \mapsto |h(\mathbf{x})|\rho(\mathbf{x})$.

The two described sampling strategies are rather fundamental and only served as introductory examples to get a feeling what it means to create samples of an arbitrary distribution. There exist several versions and combinations among these two and other strategies. We refer to [184, 222] for additional approaches and their details.

However, many fundamental sampling methods share a disadvantage that prevents their application in more complex scenarios. They suffer from the so-called *curse of dimensionality*, i. e., from a huge growth of complexity in the number of unknowns, i. e., in the dimension of the problem; see, e. g., [176]. That is, if we want to create samples in, for example, $n = 1000$ dimensions with one of the described sampling approaches, we would need a very large amount of proposal samples in order to sample from the correct high-dimensional distribution. If we aim to construct samples from a posterior distribution, then every proposal also means an execution of the (potentially expensive) forward model, which might be rather costly; hence, these methods are not appropriate for advanced applications. For this reason, we introduce sampling strategies that have shown to be more suitable for purposes of larger scale, *Markov chain Monte Carlo* (MCMC) methods.

2.5.2 Markov chain Monte Carlo (MCMC)

Markov chain Monte Carlo (MCMC) methods [184, 222] also aim to produce samples following a given target density or target distribution. These approaches construct a Markov chain, according to a prescribed transition function, from which the components, or states, are then taken as samples from the target distribution. Before we introduce and discuss two of the most common MCMC algorithms, the *Random Walk Metropolis algorithm* and the *Metropolis–Hastings algorithm*, we give a proper theoretical probabilistic setup regarding Markov chains, i. e., we recall a couple of definitions and fundamental theoretical results. This is necessary in order to be able to provide evidence for the applicability and usability of the two mentioned algorithms.

Markov chains

The class of *Markov chains* is a particular class of *discrete-time stochastic processes* [30, 79, 221] that possess the so-called *memorylessness* property, also known as *Markov property*. The states of a stochastic process with this property do *only* depend on the previous state and *not* on the path that lead to the previous state. This idea is formalized in the following definition.

Definition 2.5.4 (Markov chain). A stochastic process $(\mathbf{X}_j)_{j \in \mathbf{N}_0}$, consisting of random variables \mathbf{X}_j with values in $\mathcal{X} \subseteq \mathbf{R}^n$, is called a (discrete-time continuous-space) *Markov chain* if

$$\mathbf{P}(\mathbf{X}_{j+1} \in A \mid \mathbf{X}_0 = \mathbf{x}_0, \dots, \mathbf{X}_j = \mathbf{x}_j) = \mathbf{P}(\mathbf{X}_{j+1} \in A \mid \mathbf{X}_j = \mathbf{x}_j) \quad (2.70)$$

for all $j \in \mathbf{N}_0$, $A \in \mathcal{B}(\mathcal{X})$, and $\mathbf{x}_0, \dots, \mathbf{x}_j \in \mathcal{X}$.

Such a chain is also called a *general state space chain* since it can take on an uncountable number of states. If \mathcal{X} was a finite set (equipped with a suitable σ -algebra), we would speak of a *finite state space chain*. We present most of the theory for general state space chains, since the algorithms that we apply construct Markov chains with values in \mathbf{R}^n .

We consider only *time-homogeneous*, or *stationary*, Markov chains which means that the conditional probabilities, and hence the transitions between states, do not depend on the time index, i. e., it holds that

$$\mathbf{P}(\mathbf{X}_{j+1} \in A \mid \mathbf{X}_j = \mathbf{x}) = \mathbf{P}(\mathbf{X}_{j+m+1} \in A \mid \mathbf{X}_{j+m} = \mathbf{x}) \quad (2.71)$$

for each pair of indices $j, m \in \mathbf{N}_0$, measurable set $A \in \mathcal{B}(\mathcal{X})$, and state $\mathbf{x} \in \mathcal{X}$.

It is particularly convenient to describe the transitions between states in a Markov chain by a so-called *kernel function*. In fact, the law of a Markov chain for any finite number of steps can be specified by a starting distribution μ , i. e., $\mathbf{X}_0 \sim \mu$, and a *Markov kernel* $K : \mathcal{X} \times \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ (see Def. A.2.12) which does not depend on the time index due to the time-homogeneity. Indeed, we have that

$$\begin{aligned} \mathbf{P}(\mathbf{X}_0 \in A_0, \mathbf{X}_1 \in A_1, \dots, \mathbf{X}_j \in A_j) &= \\ &= \int_{A_0} \int_{A_1} \dots \int_{A_{j-1}} K(\mathbf{x}_{j-1}, A_j) K(\mathbf{x}_{j-2}, d\mathbf{x}_{j-1}) \dots K(\mathbf{x}_0, d\mathbf{x}_1) \mu(d\mathbf{x}_0) \end{aligned} \quad (2.72)$$

for all sets $A_0, A_1, \dots, A_j \in \mathcal{B}(\mathcal{X})$. Reversely, we can construct a Markov chain for any Markov kernel K and starting distribution μ [221]. In other words, the kernel K describes the transition

probabilities in the Markov chain in the sense that

$$K(\mathbf{x}, A) = \mathbf{P}(\mathbf{X}_{j+1} \in A \mid \mathbf{X}_j = \mathbf{x}) \quad (2.73)$$

for all $\mathbf{x} \in \mathcal{X}$, $A \in \mathcal{B}(\mathcal{X})$, and $j \in \mathbf{N}_0$.

It will be useful to define a *p-step ahead* ($p \in \mathbf{N}$) conditional distribution, i. e., we consider for arbitrary $j \in \mathbf{N}_0$, $\mathbf{x}_j \in \mathcal{X}$, $p \in \mathbf{N}$, and $A \in \mathcal{B}(\mathcal{X})$,

$$\begin{aligned} \mathbf{P}(\mathbf{X}_{j+p} \in A \mid \mathbf{X}_j = \mathbf{x}_j) &= \\ &= \int_{\mathcal{X}} \cdots \int_{\mathcal{X}} K(\mathbf{x}_{j+p-1}, A) K(\mathbf{x}_{j+p-2}, d\mathbf{x}_{j+p-1}) \cdots K(\mathbf{x}_j, d\mathbf{x}_{j+1}) \\ &= \int_{\mathcal{X}^{p-1}} K(\mathbf{x}_{j+p-1}, A) \prod_{i=j}^{j+p-2} K(\mathbf{x}_i, d\mathbf{x}_{i+1}) \\ &=: K^p(\mathbf{x}_j, A). \end{aligned} \quad (2.74)$$

Note that we have $K^1 = K$ according to (2.73). The *p-step ahead* conditional distribution $K^p(\mathbf{x}, A)$ specifies how likely it is for the chain to get into a set $A \in \mathcal{B}(\mathcal{X})$ in exactly p steps starting from some point \mathbf{x} .

The kernel K can also be described, or induced, by a *transition function* k . In the literature, these two terms (kernel and transition function) are often used interchangeably since they are equivalent.

Definition 2.5.5 (Transition function). A *transition function* is a map

$$k : (\mathcal{X} \times \mathcal{X}, \mathcal{B}(\mathcal{X}) \otimes \mathcal{B}(\mathcal{X})) \rightarrow \mathbf{R}_{\geq 0} \quad (2.75)$$

that induces a Markov kernel $K : \mathcal{X} \times \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ by

$$K(\mathbf{x}, A) = \int_A k(\mathbf{x}, \mathbf{y}) d\mathbf{y} \quad (2.76)$$

for all $\mathbf{x} \in \mathcal{X}$ and $A \in \mathcal{B}(\mathcal{X})$.

Remark. (i) Note that, for $\mathbf{x} \in \mathcal{X}$, $k(\mathbf{x}, \cdot)$ is the (Lebesgue) density of the distribution $K(\mathbf{x}, \cdot)$. (ii) In the following, we assume that a transition function k is, by notation, associated with the Markov kernel K by (2.76).

A natural question at this point is if the marginal distributions of the Markov chain get *stationary* after a certain amount of steps. In other words, we want to know if there exists a distribution π and an index j such that

$$\mathbf{X}_j \sim \pi \Rightarrow \mathbf{X}_{j+1} \sim \pi. \quad (2.77)$$

Definition 2.5.6 (Invariant/stationary distribution). A distribution π is *invariant* or *stationary* for a Markov kernel K if

$$\pi(A) = \int_{\mathcal{X}} K(\mathbf{x}, A) \pi(d\mathbf{x}) \quad (2.78)$$

for each $A \in \mathcal{B}(\mathcal{X})$.

To ensure the uniqueness of an invariant distribution, we require a certain sufficient property.

Definition 2.5.7 (Irreducibility). A Markov chain with kernel K is called ψ -irreducible for a distribution ψ if for all $\mathbf{x} \in \mathcal{X}$ and sets $A \in \mathcal{B}(\mathcal{X})$ with $\psi(A) > 0$, there exists some $p \in \mathbf{N}$ such that

$$K^p(\mathbf{x}, A) > 0. \quad (2.79)$$

A Markov chain is called *irreducible* if there exists a distribution ψ such that the chain is ψ -irreducible.

A Markov chain being ψ -irreducible means that we can reach any set $A \in \mathcal{B}(\mathcal{X})$ with $\psi(A) > 0$ in a *finite* number of steps from an *arbitrary* starting point. In particular, this property implies uniqueness.

Theorem 2.5.8 (Uniqueness of invariant distribution, [222, Ch. 6]). *If an irreducible Markov chain admits an invariant distribution π , then π is unique.*

The existence of an invariant distribution is implied by the Markov chain, or, more exactly, the associated transition function, satisfying a certain sufficient condition.

Theorem 2.5.9 (Detailed balance condition). *If a transition function k satisfies the detailed balance condition for a distribution π with Lebesgue density ρ , i. e.,*

$$\rho(\mathbf{x})k(\mathbf{x}, \mathbf{x}') = \rho(\mathbf{x}')k(\mathbf{x}', \mathbf{x}) \quad (2.80)$$

for all $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$, then π is the invariant distribution of the Markov chain associated with k .

Proof. The result follows by noting that, for any $A \in \mathcal{B}(\mathcal{X})$, it holds that

$$\begin{aligned} \int_{\mathcal{X}} K(\mathbf{x}, A) \pi(d\mathbf{x}) &= \int_{\mathcal{X}} \int_A k(\mathbf{x}, \mathbf{x}') d\mathbf{x}' \pi(d\mathbf{x}) \\ &= \int_{\mathcal{X}} \int_A k(\mathbf{x}, \mathbf{x}') \rho(\mathbf{x}) d\mathbf{x}' d\mathbf{x} \\ &= \int_{\mathcal{X}} \int_A k(\mathbf{x}', \mathbf{x}) \rho(\mathbf{x}') d\mathbf{x}' d\mathbf{x} \\ &= \int_A \underbrace{\int_{\mathcal{X}} k(\mathbf{x}', \mathbf{x}) d\mathbf{x}}_{=K(\mathbf{x}', \mathcal{X})=1} \pi(d\mathbf{x}') \\ &= \pi(A). \end{aligned} \quad (2.81)$$

The change of the order of the integrals is justified by the joint measurability of k and Fubini's theorem. ■

Remark. (i) The detailed balance condition is also sufficient for a Markov chain (associated with k) to be *reversible*; see Def. A.2.13. (ii) The detailed balance condition is *not* a necessary condition for π to be an invariant distribution.

The detailed balance condition asks for a kind of symmetry in the sense that getting to a state \mathbf{x} and transitioning to a next state \mathbf{x}' is equally likely as getting to \mathbf{x}' and transitioning to \mathbf{x} . In addition to the irreducibility property, it is this condition that we will check later for the transition functions specified by the particular MCMC algorithms.

Note that, so far, we only know that there exists a unique invariant distribution for an irreducible Markov chain satisfying the detailed balance condition. It remains to show that a Markov chain is indeed *converging* to this distribution in a prescribed sense. In order to ensure this convergence, we additionally need to introduce two sufficient properties in the next two definitions.

Definition 2.5.10 (Positive Harris chain). A Markov chain is called *positive Harris* if

1. there is a distribution ψ such that the chain is ψ -irreducible,
2. it admits an invariant distribution,
3. for all $A \in \mathcal{B}(\mathcal{X})$ with $\psi(A) > 0$, A is *Harris recurrent*, i. e., the probability of the Markov chain returning infinitely many times to A starting from \mathbf{x} is one for each $\mathbf{x} \in A$.

Definition 2.5.11 ((A)periodicity for finite state space chains). Let \mathcal{X} be a finite set and regard (finite state space) chains with values in \mathcal{X} . The natural number

$$\tau_{\mathbf{x}} := \gcd(\{m \in \mathbf{N} \mid \mathbf{P}(\mathbf{X}_m = \mathbf{x} \mid \mathbf{X}_0 = \mathbf{x}) > 0\}) \quad (2.82)$$

is called the *period* of the state $\mathbf{x} \in \mathcal{X}$, provided the above set is not empty. If $\tau_{\mathbf{x}} = 1$, the state \mathbf{x} is called *aperiodic*, otherwise *periodic* with period $\tau_{\mathbf{x}}$.

A Markov chain with values in \mathcal{X} is called *aperiodic* if each state in \mathcal{X} is aperiodic.

Remark. In fact, if one state from \mathcal{X} is aperiodic for an *irreducible* (finite state space) Markov chain, then *all* states in \mathcal{X} are aperiodic.

Note that this definition of (a)periodicity is for *finite* state space chains. A corresponding generalization for *general* state space chains exists with so-called *(d-)cycles* (see [222, Subsec. 6.3.3] or [193, Subsec. 5.4.3]) but is rather technical and exceeds the scope of this section.

We now have all the ingredients for a convergence result of aperiodic and positive Harris Markov chains.

Theorem 2.5.12 (Convergence of aperiodic and positive Harris chains, [193, Thm. 13.3.3]). *If a Markov chain is aperiodic and positive Harris, it converges to its invariant distribution π in the total variation norm (see Def. A.2.9) for any initial distribution μ , i. e.,*

$$\lim_{p \rightarrow \infty} \left\| \int_{\mathcal{X}} K^p(\mathbf{x}, \cdot) \mu(d\mathbf{x}) - \pi \right\|_{\text{TV}} = 0. \quad (2.83)$$

The remainder of this section is devoted to present the two mentioned MCMC algorithms. We start with the Random Walk Metropolis (RWM) algorithm [192] which is actually a special case of the Metropolis–Hastings (MH) algorithm [123] described later.

Random Walk Metropolis algorithm

The aim is still to compute samples that follow a given target distribution π with density ρ . The RWM algorithm was proposed by Metropolis et al. in [192]. Starting from an arbitrary point in \mathcal{X} , it successively constructs elements of a Markov chain by *proposing* a potentially new state \mathbf{X}' that is *accepted* or *rejected* according to a certain acceptance probability involving the actual target density.

The proposed new state, called the *proposal*, is sampled from a distribution, the *proposal distribution*, that is symmetric around the current state of the Markov chain. The formal steps of this approach are given in Alg. 3. Note that the proposal distribution, and hence the

Algorithm 3: Random Walk Metropolis algorithm

Input : Starting point $\mathbf{X}_0 := \mathbf{x}_0 \in \mathcal{X}$, target distribution π with density ρ , symmetric distribution \mathbf{Q} .

Output: Samples approximately following π .

Iterating for $j = 1, 2, \dots$, the j -th step is:

1. Draw $\varepsilon \sim \mathbf{Q}$ independently of \mathbf{X}_{j-1} and define the proposal

$$\mathbf{X}' = \mathbf{X}_{j-1} + \varepsilon. \quad (2.84)$$

2. Compute acceptance probability

$$\alpha(\mathbf{X}'|\mathbf{X}_{j-1}) := \min \left\{ 1, \frac{\rho(\mathbf{X}')}{\rho(\mathbf{X}_{j-1})} \right\}. \quad (2.85)$$

3. Set $\mathbf{X}_j = \mathbf{X}'$ with probability $\alpha(\mathbf{X}'|\mathbf{X}_{j-1})$, otherwise set $\mathbf{X}_j = \mathbf{X}_{j-1}$.
-

Markov kernel corresponding to the resulting Markov chain, is not depending on the current step, i. e., it is time-homogeneous.

In (2.85), the acceptance probability is defined in terms of the ratio $\rho(\mathbf{X}')/\rho(\mathbf{X}_{j-1})$ which implies that the target density has to be known only up to a constant. This is particularly advantageous when applying the RWM algorithm in the context of Bayesian inverse problems. Indeed, the normalizing constant of the posterior density from (2.20), which can be computationally expensive or even intractable, does not need to be calculated.

After accepting or rejecting the proposal, the elements of the constructed Markov chain are then taken as samples of the target distribution π . Investigations that theoretically justify this algorithm are carried out later in the context of the more general MH algorithm.

Since we will use the RWM algorithm in later chapters, we should discuss different aspects as performance measures, how they can be controlled, and well-known pitfalls and their diagnostics.

The main performance indicator for a (realization of a) Markov chain is its *effective sample size* (ESS). For example, in theory, given *independent* samples, the central limit theorem bounds uncertainty in the estimates with the number of samples. If we have *dependent* or *correlated* samples, this number is replaced by the ESS denoted by N_{eff} . The number N_{eff} of uncorrelated samples has the same estimation potential as the number of all correlated

samples.

Definition 2.5.13 (Effective sample size [40]). Let $(\mathbf{x}_n)_{n=1,\dots,N}$ be a realization of a Markov chain with $N \in \mathbf{N}$ elements. The *effective sample size in the i -th component* of the realization, i. e., of $((\mathbf{x}_n)_i)_{n=1,\dots,N}$, is defined as

$$N_{\text{eff},i} := \frac{N}{1 + 2 \sum_{\tau=1}^{\infty} r^{(i)}(\tau)}, \quad (2.86)$$

where $r^{(i)}(\tau)$ denotes the *auto-correlation with lag $\tau \in \mathbf{N}$* in the i -th component. We conservatively define the *effective sample size* of the (whole) realization as

$$N_{\text{eff}} := \min_{i=1,\dots,n} N_{\text{eff},i}. \quad (2.87)$$

In practice, $N_{\text{eff},i}$ is often not analytically available and thus has to be estimated. The infinite sum from above is approximated by $\sum_{\tau=1}^{K_i} r^{(i)}(\tau)$, $K_i \in \mathbf{N}$, such that $r^{(i)}(\tau)$ is negligible for $\tau > K_i$.

For a definition of auto-correlation, the resulting Markov chain is viewed as a component-wise *discrete-time signal*.

Definition 2.5.14 (Auto-correlation function of a discrete-time signal). The *auto-correlation function* of a real-valued discrete-time signal $(x_n)_{n=1,\dots,N}$ with $N \in \mathbf{N}$ elements is defined as

$$r(\tau) = r_x(\tau) := \sum_{n=0}^{N-\tau} x_n x_{n+\tau}, \quad (2.88)$$

where $\tau \in \mathbf{N}$ denotes the *lag*.

The auto-correlation function can be computed efficiently using *fast Fourier transforms*; see [38].

Note that the ESS increases when the auto-correlations decrease. The auto-correlation is mainly affected by the choice of the *proposal variance*, i. e., the variance of the proposal distribution, which also directly controls the *acceptance rate*, i. e., the ratio of accepted and total samples. That is, the goal is to find an “optimal” proposal variance (or, equivalently, acceptance rate) that minimizes the auto-correlation of the chain.

On the one hand, we could choose a large proposal variance since this would decrease auto-correlation if samples are accepted. However, on the other hand, if it is too large, the proposals get rejected and the chain does not move which in turn increases auto-correlation. This ambivalence is a well-known conflict of interest. For general target distributions, it is difficult to find the optimal proposal variance; but there are results that yield useful heuristics.

Suppose we want to create samples from a one-dimensional standard normal distribution $\mathcal{N}(0, 1)$ using the RWM algorithm and study the influence of the proposal variance σ^2 when choosing $\mathbf{Q} = \mathcal{N}(0, \sigma^2)$. A first guess for a corresponding optimal value would be $\sigma^2 = 1$. However, it turns out that this is not the case. Fig. 2.4 shows sample paths of four Markov chains with different proposal variances $\sigma^2 \in \{0.1^2, 1, 2.38^2, 5^2\}$. We see that the number of rejected proposals (orange circles) increases with an increasing proposal variance. But, at the same time, we also recognize a higher auto-correlation caused by smaller proposal variances.

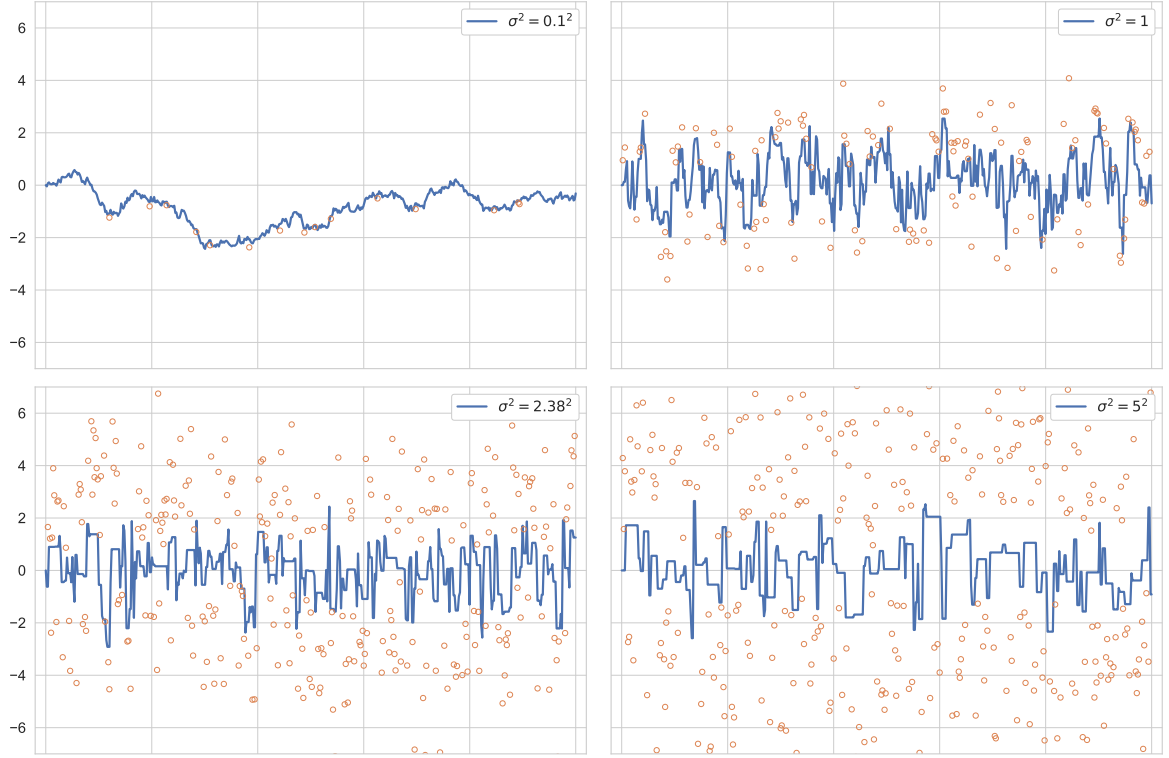


Figure 2.4: Sample paths of four Markov chains with different proposal variances. The orange circles depict rejected samples.

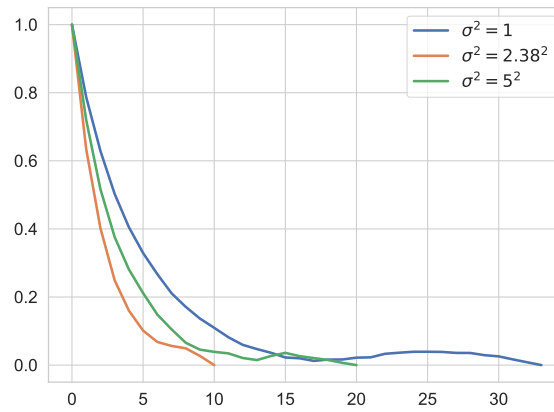


Figure 2.5: Auto-correlation functions of three Markov chains constructed with the RWM algorithm to sample from $\mathcal{N}(0, 1)$ using different proposal variances σ^2 . The plot confirms the result of [96] proving that $\sigma_{\text{opt}}^2 = 2.38^2$ in this example.

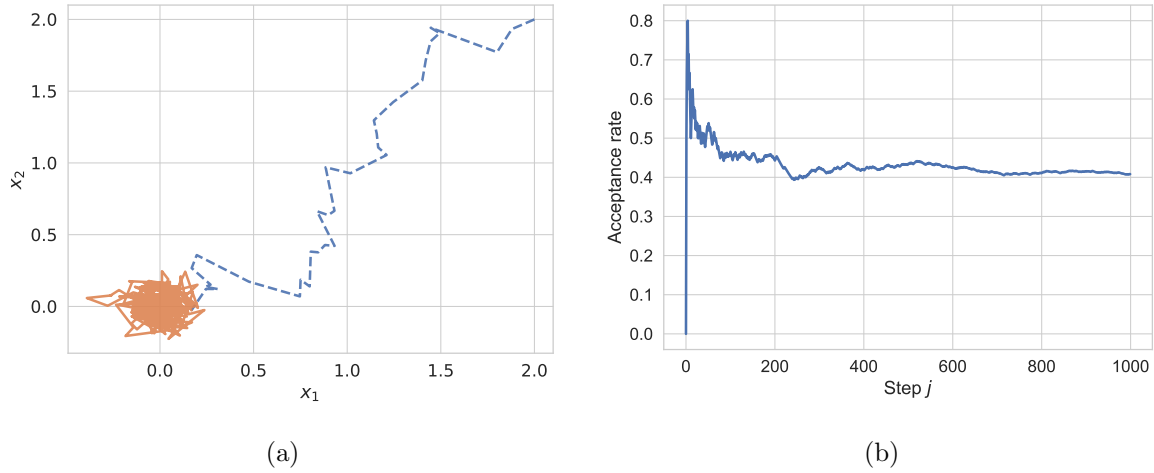


Figure 2.6: (a) The plot shows a Markov chain starting at $\mathbf{x}_0 = (2, 2)^\top$, a point with small probability. During the burn-in period (dashed blue lines), the chain is moving towards a region with higher probability (solid orange lines). (b) We see how the corresponding acceptance rate evolves. It takes about 200 steps for the chain to stabilize its rate of acceptance which is also an indicator of the burn-in period.

It was shown in [96] that the optimal variance for this example is $\sigma_{\text{opt}}^2 = 2.38^2$. This is supported by Fig. 2.5 which depicts the auto-correlation functions corresponding to the proposal variances $\sigma^2 \in \{1, 2.38^2, 5^2\}$.

Under an i.i.d. assumption on the target density, the optimal acceptance rate for the RWM algorithm was proven to be 0.234 [223]. More recent investigations can avoid the restrictive i.i.d. assumption and generalize this optimality result to a larger class of target distributions [289]. However, [22] shows that these results are not the end of the story and demonstrates that the optimal acceptance rate can be smaller in more general cases.

A common rule of thumb is that a “good” acceptance rate lies around 0.2 – 0.4. In any case, a practitioner, who would like to check the quality of a certain realization of a Markov chain, is recommended to calculate its ESS.

There is one additional pitfall that we need to comment on since it can highly mislead the interpretation and diagnostics of resulting Markov chains. Thm. 2.5.12 from above gives a convergence result for Markov chains under certain assumptions. We will see below that a Markov chain generated by the RWM algorithm (or, more general, the MH algorithm) satisfies these assumptions and, hence, converges. This is, however, an *asymptotic* result. Unfortunately, there is not much known about the *pre-asymptotic* period of a Markov chain in general. In our case, it might take some steps for the Markov chain to get into regions with higher probability in order to produce samples that follow the target distribution. That means that very often the first elements of the Markov chain are not sampled from the correct distribution and thus should not be taken as samples for a subsequent estimation of some quantity.

This period, that a Markov chain needs to produce valuable samples, is called the *burn-in period*. Fig. 2.6 illustrates that a Markov chain might need an initial amount of steps not only to reach a region with higher probability but also to get a “stable” acceptance rate. In practice, it is not easy to determine the end of the burn-in period. Heuristically, we can remove the

first $N_{\text{burn-in}} \in \mathbf{N}$ samples (with $N_{\text{burn-in}} < N$) from the chain and use the remaining elements as the final set of samples. This is often done in a generous way, i.e., there are scenarios with $N_{\text{burn-in}} \approx 10^5$ or even more depending on the concrete application and its complexity.

As a concluding remark, the RWM algorithm is often also subject to the curse of dimensionality, meaning that a lot of steps are necessary in order to get a sufficiently large ESS, since the symmetric proposal distribution cannot exploit much structure of an arbitrary target distribution. For example, a target distribution can be asymmetric or informative for only a subset of components or linear combinations of them, i.e., informative on a low-dimensional subspace. As we will see in Ch. 4, we nevertheless employ this algorithm in our applications for reasons becoming clear later. However, the next paragraphs discuss a popular generalization of the algorithm and provide theoretical results that justify its usage.

Metropolis–Hastings algorithm

Hastings [123] generalizes the RWM algorithm to the MH algorithm in the sense that the proposal distribution can be an arbitrary Markov kernel. This allows a lot of flexibility that can be used to create more informative proposals following a distribution similar to the target.

Choosing the proposal distribution, however, remains to be an art. If it is done in an advantageous way, then the resulting algorithm yields better mixing (see, e.g., the *Metropolis-adjusted Langevin algorithm* (MALA) [24, 224, 226]) or even circumvents the curse of dimensionality (see, e.g., the *preconditioned Crank-Nicolson algorithm* [64, 118]). For an overview and history of MCMC methods and their convergence properties, we refer to [27].

The formal steps of the MH algorithm, which are similar to the RWM algorithm (Alg. 3), are made precise in Alg. 4. The only difference to the RWM algorithm is the more flexi-

Algorithm 4: Metropolis–Hastings algorithm

Input : Starting point $\mathbf{X}_0 := \mathbf{x}_0 \in \mathcal{X}$, target distribution π with density ρ , Markov kernel $\mathbf{Q} : \mathcal{X} \times \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ such that $\mathbf{Q}(\mathbf{x}, \cdot)$ has density $q(\cdot|\mathbf{x})$ for each $\mathbf{x} \in \mathcal{X}$.

Output: Samples approximately following π .

Iterating for $j = 1, 2, \dots$, the j -th step is:

1. Draw proposal $\mathbf{X}' \sim \mathbf{Q}(\mathbf{X}_{j-1}, \cdot)$.
2. Compute acceptance probability

$$\alpha(\mathbf{X}'|\mathbf{X}_{j-1}) := \min \left\{ 1, \frac{\rho(\mathbf{X}')}{\rho(\mathbf{X}_{j-1})} \frac{q(\mathbf{X}_{j-1}|\mathbf{X}')}{q(\mathbf{X}'|\mathbf{X}_{j-1})} \right\}. \quad (2.89)$$

3. Set $\mathbf{X}_j = \mathbf{X}'$ with probability $\alpha(\mathbf{X}'|\mathbf{X}_{j-1})$, otherwise set $\mathbf{X}_j = \mathbf{X}_{j-1}$.
-

ble choice of the proposal distribution as a Markov kernel \mathbf{Q} and the adjusted acceptance probability in (2.89). For $\mathbf{x} \in \mathcal{X}$, let $q(\cdot|\mathbf{x})$ denote the (Lebesgue) density of the probability measure $\mathbf{Q}(\mathbf{x}, \cdot)$, i.e.,

$$\frac{d\mathbf{Q}(\mathbf{x}, \cdot)}{d\lambda} = q(\cdot|\mathbf{x}). \quad (2.90)$$

In the following, a Markov chain generated by the MH algorithm is called a “Metropolis–Hastings chain”, or just “MH chain”.

Let us define

$$a(\mathbf{x}_j) := \int_{\mathcal{X}} \alpha(\mathbf{x}'|\mathbf{x}_j)q(\mathbf{x}'|\mathbf{x}_j) d\mathbf{x}' \quad (2.91)$$

for any $\mathbf{x}_j \in \mathcal{X}$ as the probability that the algorithm accepts a proposal given that $\mathbf{X}_j = \mathbf{x}_j$, $j \in \mathbf{N}_0$. Then, we can derive the following transition function of a MH chain.

Proposition 2.5.15 (Transition function of the MH algorithm). *The transition function of (the Markov kernel of) a MH chain is*

$$k(\mathbf{x}_j, \mathbf{x}_{j+1}) = \alpha(\mathbf{x}_{j+1}|\mathbf{x}_j)q(\mathbf{x}_{j+1}|\mathbf{x}_j) + (1 - a(\mathbf{x}_j))\delta_{\mathbf{x}_j}(\mathbf{x}_{j+1}) \quad (2.92)$$

for any $\mathbf{x}_j, \mathbf{x}_{j+1} \in \mathcal{X}$, where $\delta_{\mathbf{x}_j}$ denotes the Dirac measure at \mathbf{x}_j .

Remark. Note that we make use of an abuse of notation here. In fact, there is no density function $\delta_{\mathbf{y}}(\mathbf{x})$, $\mathbf{x} \in \mathcal{X}$, for the Dirac measure $\delta_{\mathbf{y}}(A)$, $A \in \mathcal{B}(\mathcal{X})$. However, we can always equivalently argue with the corresponding Markov kernel K induced by k , where everything is defined rigorously, and so we accept this abuse of notation for the sake of convenience and clearness.

Proof. Let $A \in \mathcal{B}(\mathcal{X})$ and $\mathbf{x}_j \in \mathcal{X}$. Note that

$$\mathbf{P}(\mathbf{X}_{j+1} \in A, \mathbf{X}' \text{ is accepted} | \mathbf{X}_j = \mathbf{x}_j) = \int_A \alpha(\mathbf{x}'|\mathbf{x}_j)q(\mathbf{x}'|\mathbf{x}_j) d\mathbf{x}' \quad (2.93)$$

and

$$\begin{aligned} \mathbf{P}(\mathbf{X}_{j+1} \in A, \mathbf{X}' \text{ is rejected} | \mathbf{X}_j = \mathbf{x}_j) &= \\ &= \mathbf{P}(\mathbf{X}_{j+1} \in A | \mathbf{X}' \text{ is rejected}, \mathbf{X}_j = \mathbf{x}_j) \cdot \mathbf{P}(\mathbf{X}' \text{ is rejected} | \mathbf{X}_j = \mathbf{x}_j) \\ &= \mathbf{1}_A(\mathbf{x}_j) \cdot (1 - a(\mathbf{x}_j)) \\ &= \int_A (1 - a(\mathbf{x}_j)) \delta_{\mathbf{x}_j}(d\mathbf{x}_{j+1}). \end{aligned} \quad (2.94)$$

The result (2.92) follows by

$$\begin{aligned} \mathbf{P}(\mathbf{X}_{j+1} \in A | \mathbf{X}_j = \mathbf{x}_j) &= \mathbf{P}(\mathbf{X}_{j+1} \in A, \mathbf{X}' \text{ is accepted} | \mathbf{X}_j = \mathbf{x}_j) \\ &\quad + \mathbf{P}(\mathbf{X}_{j+1} \in A, \mathbf{X}' \text{ is rejected} | \mathbf{X}_j = \mathbf{x}_j). \end{aligned} \quad (2.95)$$

■

The existence of an invariant distribution of a MH chain is implied by the transition function satisfying the detailed balance condition. In particular, the following result also implies that the invariant distribution is actually equal to the target distribution π .

Theorem 2.5.16. *The transition function of a MH chain from (2.92) satisfies the detailed balance condition (2.80) for the target distribution π with density ρ .*

Proof. It suffices to show that $(\mathbf{x}_j, \mathbf{x}_{j+1}) \mapsto \alpha(\mathbf{x}_{j+1}|\mathbf{x}_j)q(\mathbf{x}_{j+1}|\mathbf{x}_j)$ satisfies the detailed balance condition for π since the steps in the proof of Thm. 2.5.9 can then be carried out similarly to show that π is the invariant distribution.

For arbitrary $\mathbf{x}_j, \mathbf{x}_{j+1} \in \mathcal{X}$, we compute

$$\begin{aligned} \alpha(\mathbf{x}_{j+1}|\mathbf{x}_j)q(\mathbf{x}_{j+1}|\mathbf{x}_j)\rho(\mathbf{x}_j) &= \min \left\{ 1, \frac{\rho(\mathbf{x}_{j+1})}{\rho(\mathbf{x}_j)} \frac{q(\mathbf{x}_j|\mathbf{x}_{j+1})}{q(\mathbf{x}_{j+1}|\mathbf{x}_j)} \right\} q(\mathbf{x}_{j+1}|\mathbf{x}_j)\rho(\mathbf{x}_j) \\ &= \min \{ \rho(\mathbf{x}_j)q(\mathbf{x}_{j+1}|\mathbf{x}_j), \rho(\mathbf{x}_{j+1})q(\mathbf{x}_j|\mathbf{x}_{j+1}) \} \\ &= \min \left\{ \frac{\rho(\mathbf{x}_j)}{\rho(\mathbf{x}_{j+1})} \frac{q(\mathbf{x}_{j+1}|\mathbf{x}_j)}{q(\mathbf{x}_j|\mathbf{x}_{j+1})}, 1 \right\} q(\mathbf{x}_j|\mathbf{x}_{j+1})\rho(\mathbf{x}_{j+1}) \\ &= \alpha(\mathbf{x}_j|\mathbf{x}_{j+1})q(\mathbf{x}_j|\mathbf{x}_{j+1})\rho(\mathbf{x}_{j+1}). \end{aligned} \tag{2.96}$$

■

The π -irreducibility would give the uniqueness of π as the invariant distribution. An immediate sufficient condition for π -irreducibility is positivity of the conditional density q , i. e.,

$$q(\mathbf{x}'|\mathbf{x}_j) > 0 \quad \text{for every } \mathbf{x}_j, \mathbf{x}' \in \text{supp}(\rho), \tag{2.97}$$

since then every subset of the support of ρ with positive Lebesgue measure can be reached in a single step.

In fact, irreducibility of a MH chain implies Harris recurrence (point 3 in Def. 2.5.10) which is needed to show convergence.

Lemma 2.5.17 ([222, Lem. 7.3]). *If a MH chain is π -irreducible, then it is Harris recurrent.*

Hence, with the condition (2.97), we get a π -irreducible, Harris-recurrent MH chain that admits an invariant distribution, i. e., the chain is positive Harris according to Def. 2.5.10.

We still require aperiodicity to apply the convergence result from Thm. 2.5.12. A MH chain is ensured to be aperiodic if the probability that the chain remains in the same state is positive, i. e.,

$$\mathbf{P}(X_{j+1} = X_j) > 0 \tag{2.98}$$

for all $j \in \mathbf{N}_0$. In particular, this is the case if the probability of accepting a new proposal is less than one, that is, if

$$\mathbf{P}\left(\rho(\mathbf{X}')q(\mathbf{X}'|\mathbf{X}_j) < \rho(\mathbf{X}_j)q(\mathbf{X}_j|\mathbf{X}')\right) < 1 \tag{2.99}$$

for each $j \in \mathbf{N}_0$.

However, the two conditions from (2.97) and (2.99), yielding convergence, seem to be rather restrictive. Although most of the common proposal distributions, like multivariate Gaussian distributions, fulfill these conditions, there is a weaker and more general condition that gives π -irreducibility and aperiodicity at the same time. In fact, it is sufficient to require that there exists a uniform radius for a neighborhood of the states $\mathbf{x} \in \mathcal{X}$ such that the probability of transitioning within this neighborhood is uniformly bounded from below. It is easy to imagine that this property allows the Markov chain to reach every state starting from an arbitrary starting point in a finite number of steps, hence yielding irreducibility.

Lemma 2.5.18 ([222, Lem. 7.6]). *Let ρ be bounded and positive on every compact subset of its connected support $\text{supp}(\rho)$. If there exist real numbers $\varepsilon, \delta > 0$ such that*

$$q(\mathbf{x}'|\mathbf{x}_j) > \varepsilon \quad \text{if } \|\mathbf{x}_j - \mathbf{x}'\|_2 < \delta \quad (2.100)$$

for each pair $\mathbf{x}_j, \mathbf{x}' \in \mathcal{X}$, then the corresponding MH chain is π -irreducible and aperiodic.

Summarizing, the MH algorithm (and particularly the RWM algorithm) yields a Markov chain that converges to its unique invariant target distribution π under rather mild assumptions on the proposal distribution (or Markov kernel) \mathbf{Q} .

Referring back to the Bayesian context, the MH algorithm can be applied to create samples of the posterior by running the algorithm for

$$\pi = \mu^d \quad (2.101)$$

and a suitable proposal distribution \mathbf{Q} . As already indicated, the proposal distribution is the component which mainly determines the performance (burn-in period, auto-correlation, and convergence) of the overall algorithm. In the case of high-dimensional posterior distributions involving potentially expensive models, it is especially important to find proposal distributions that are dimension-robust and similar to the posterior.

Our approach is, however, different since we do not solve a certain Bayesian inverse problem in the full space but employ a dimension reduction technique which allows to run a Markov chain in a space of reduced dimension. This improves the performance without the need of a complex proposal distribution. Actually, we use a *symmetric* proposal distribution, i. e., we run the RWM algorithm. The mentioned technique is explained in the next chapter, Ch. 3. Its connection to Bayesian inverse problems is then presented in Ch. 4.

Related literature

As mentioned, the standard MCMC algorithms presented above are rather elementary. When it comes to posterior sampling, there are more sophisticated approaches, a few of which we refer to in this subsection.

A concept called *delayed rejection* (DR) [107, 194, 261] waits until a (first-stage) rejected proposal is finally rejected to avoid that it and its corresponding expensive forward computations are rejected “accidentally.” The *Delayed Rejection Adaptive Metropolis* (DRAM) algorithm [111] combines DR with the *Adaptive Metropolis* (AM) algorithm [112, 113]. Here, adaptivity means that the covariance matrix of the Gaussian proposal distribution is selected on the basis of the whole history of the Markov chain. A discussion of examples of adaptive MCMC methods is provided in [225].

Contrarily, *delayed acceptance* waits until a (first-stage) accepted proposal is finally accepted. Corresponding approaches and applications comprise, e. g., [66, 67, 105, 219, 239].

The authors of [50] demonstrate how to formulate an accelerated MCMC algorithm when a cheap approximation or surrogate model is available which can be especially beneficial in the context of expensive PDE models.

When dealing with parameterized PDE models, one way to construct a surrogate or reduced-order model and, hence, speed up computations is to build a so-called *reduced basis* [229]. In [71], a data-driven reduced basis, combined with a delayed acceptance approach, is used to define an MCMC algorithm that adaptively refines the reduced basis in areas with high posterior probability.

We indicated that a particular problem of MCMC methods is their deteriorating behavior in high dimensions. In order to avoid bad mixing times and high auto-correlations, e. g., when refining a time or spatial mesh in a PDE model, we mentioned that it is necessary to choose a good proposal distribution. For this, the authors of [186] build a local Gaussian approximation based on gradient and Hessian information of the data misfit function to propose samples with higher posterior probability. Other similar approaches that also exploit second order curvature information are given in [100, 166].

An MCMC algorithm for infinite-dimensional inference is proposed in [64], where carefully chosen time meshes of stochastic dynamical systems are used to result in a dimension-robust speed of convergence. Another dimension-robust MCMC method from [48] relies on hierarchical models and samplers for latent Gaussian processes. In [201], formulating a “localized” problem, i. e., enforcing local structure of the prior distribution and the observations, speeds up MCMC while retaining similar posterior moments. More recently, [264] studies a MALA-within-Gibbs sampler for high-dimensional distributions with sparse conditional structure from theoretical and practical viewpoints.

Finally, recent approaches exploit so-called *transport maps* [188] for an adaptive Metropolis–Hastings algorithm [209] as well as to construct more informative proposal distributions, which use a low-fidelity (cheap) model, for “preconditioning” an MCMC algorithm that evaluates the high-fidelity (expensive) model [211].

Active subspaces

Reference. The content of Sec. 3.3 and 3.4 as well as the definition of Poincaré constants as a supremum in Sec. 3.2 is based on our article in [257].

The previous chapter introduced a context for Bayesian inverse problems in a self-contained manner. We do the same in this chapter for a dimension reduction technique that also builds a basis for later chapters. In particular, we need the contents of this chapter for the study of low-dimensional structures in Bayesian inverse problems which is introduced in Ch. 4.

Many modern computational problems, e. g., in engineering disciplines or earth sciences, consist of computationally expensive physical models with a large number of unknown parameters which often makes them subject to the *curse of dimensionality*, a phenomenon characterized by a huge growth of complexity in the number of unknowns [265, 266], i. e., in the dimension of the problem.

Dimension reduction techniques are a common way to make otherwise intractable problems solvable. For a given function $f : \mathcal{X} \rightarrow \mathbf{R}$, $\mathcal{X} \subseteq \mathbf{R}^n$, we want to find maps $g : \mathcal{Y} \rightarrow \mathbf{R}$, $\mathcal{Y} \subseteq \mathbf{R}^k$, $k \leq n$, and $h : \mathcal{X} \rightarrow \mathcal{Y}$ such that

$$f \approx g \circ h. \quad (3.1)$$

Depending on the type of the map h , we distinct between three (nested) classes of dimension reduction approaches.

The smallest class, called *index-based* or *subset-based*, consists of techniques that choose a certain subset of input variables that are most important to reconstruct the behavior of the original function f . They can also be called *coordinate-aligned*. This class consists of sensitivity analysis methods [232] (see Sec. 1.2) and a variety of other approaches described in, e. g., [282].

The next level is called *subspace-based* dimension reduction. In this class, h is a linear function, represented by a matrix $A^\top \in \mathbf{R}^{k \times n}$, that maps a high-dimensional vector to a low-dimensional one. Note that the columns of the matrix A span a subspace, consisting of important *directions*, on which most of the behavior of f can be explained. That is, we are interested in *linear combinations* of input variables instead of just a subset of them. The *active subspace method* (ASM), or short, *active subspaces* [52, 54], which we explain in detail below, belongs to this type of dimension reduction.

Finally, the most general class allows h to be a function of arbitrary type and is called *nonlinear dimension reduction*; see, e. g., [171] for a variety of corresponding approaches. Tech-

niques and also intuition for this general class enormously differ from their linear counterpart. For this reason, we do not cover them in this text.

As already mentioned, we are interested in subspace-based dimension reduction approaches, particularly in ASM. Since these techniques allow for linear combinations of variables and not only for certain coordinates, they promise to be applicable for many practical situations. This advantage comes, however, with two drawbacks compared to subset-based approaches as described in [52, Ch. 2]. First, it is more difficult to interpret results from a practical perspective. For a practitioner, a result of the type “Parameter 1 is important, but parameter 2 can be neglected” is easy to interpret and evaluate. On the other hand, saying “the combination of $0.4 \times$ parameter 1 and $0.6 \times$ parameter 2 is dominating” is more difficult to interpret, especially if the parameters appearing in important combinations do not have an obvious relationship or different units. Secondly, fixing a coordinate while varying others is simpler than fixing or varying linear combinations. We have to be careful here, especially if the domain for the input variables is bounded.

In Sec. 3.1, this chapter starts with an explanation of the idea behind subspace-based dimension reduction which also includes a short description of related approaches. We continue in Sec. 3.2 by giving a formal setup of ASM and derive first upper bounds for the mean square error of a given function of interest and a low-dimensional approximation constructed by ASM. Concrete bounds for compactly supported and multivariate normal distributions are provided in Sec. 3.3. Subsequently, we derive *generalized bounds* for more general situations and show that a practitioner needs to be careful when using distributions with exponential tails (Sec. 3.4). Eventually, we discuss how to perform ASM in practical situations in Sec. 3.5 and thus prepare for the application of ASM in the context of Bayesian inverse problems described in Ch. 4.

3.1 Idea

Imagine that we are standing in the middle of a landscape with transparent mountains described by a function of interest $f : \mathcal{X} \rightarrow \mathbf{R}$, $\mathcal{X} \subseteq \mathbf{R}^n$, and want to turn around until we do not see much of variation in height of the landscape anymore. Once we find such a position, we know that the corresponding direction can be rather neglected for further analyses since it provides little information about the function f . At the same time, we might also look for (orthogonal) directions in which the landscape is varying much more. Of course, we hope that the amount of negligible directions is (much) larger than the amount of important ones to allow for a reasonable reduction of the dimensionality.

In mathematical words, consider, for example, the function f to be a so-called *ridge function* [216], i. e.,

$$f(\mathbf{x}) = g(A^\top \mathbf{x}) \quad (3.2)$$

for each $\mathbf{x} \in \mathcal{X}$, a *profile function* $g : \mathcal{Y} \rightarrow \mathbf{R}$, $\mathcal{Y} \subseteq \mathbf{R}^k$, $k \leq n$, and a matrix $A \in \mathbf{R}^{n \times k}$ with full rank. Choosing $\mathbf{x} \in \mathcal{X}$ and $\mathbf{v} \in \ker(A^\top)$ such that $\mathbf{x} + \mathbf{v} \in \mathcal{X}$, we see that

$$f(\mathbf{x} + \mathbf{v}) = g(A^\top(\mathbf{x} + \mathbf{v})) = g(A^\top \mathbf{x}) = f(\mathbf{x}). \quad (3.3)$$

That is, the ridge function f does not change along the null space of A^\top and thus has an intrinsic dimension of k instead of n .

Fig. 3.1 depicts two two-dimensional examples of ridge functions. The left plot (Fig. 3.1a) shows the contour lines of the function $f_1(x_1, x_2) := \exp(0.7x_1 + 0.3x_2)$ which was already

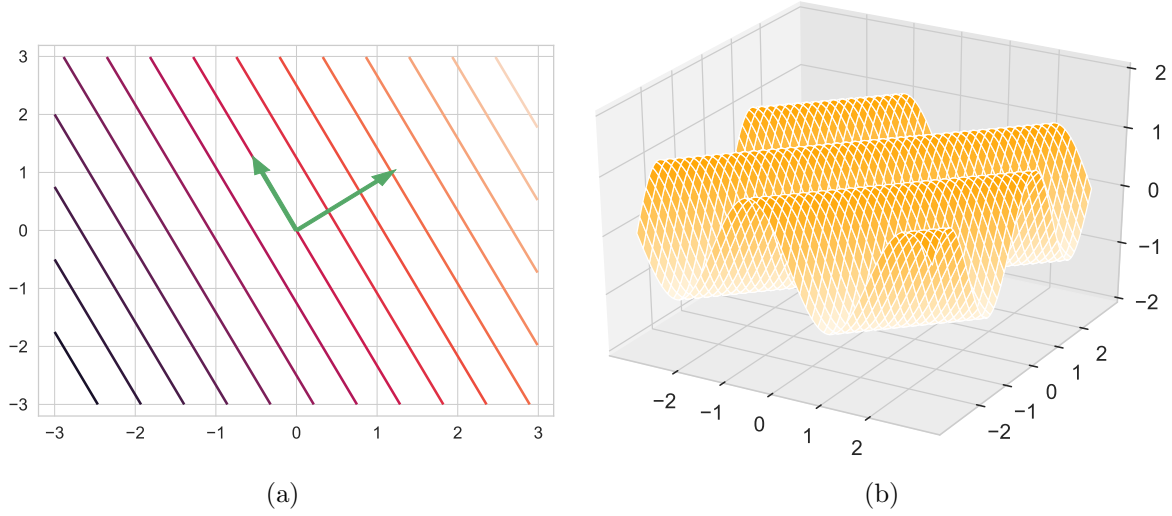


Figure 3.1: (a) Contour plot of the ridge function $f_1(x_1, x_2) := \exp(0.7x_1 + 0.3x_2)$. (b) 3D surface plot of the ridge function $f_2(x_1, x_2) := \sin(-2x_1 + 2x_2)$. Both cases demonstrate that a ridge function contains directions along which the function is constant.

used as an example in [52, 54]. Regarding (3.2), we have $g(y) = \exp(y)$ and $A^\top = \begin{pmatrix} 0.7 & 0.3 \end{pmatrix}$ in this case. The right part (Fig. 3.1b) plots the 3D surface of the function $f_2(x_1, x_2) := \sin(-2x_1 + 2x_2)$ for which it holds that $g(y) = \sin(y)$ and $A^\top = \begin{pmatrix} -2 & 2 \end{pmatrix}$. We can see in both cases that there is always one direction along which the function is constant and a perpendicular direction along which all of the variation occurs.

In practice, we cannot assume that a given function f is a ridge function. However, we can try to *approximate* f by a ridge function. This process is called *ridge approximation*. That is, our aim is to find a ridge profile g and a matrix A such that

$$f(\mathbf{x}) \approx g(A^\top \mathbf{x}) \quad (3.4)$$

for each $\mathbf{x} \in \mathcal{X}$. The quality behind “ \approx ” above is determined by the actual approximation approach and will be clearer later. Note that, if the columns of A are orthogonal, then this process can also be interpreted as a rotation of the coordinate system such that f can then be represented low-dimensionally in the rotated coordinate system.

If a technique finds a ridge function as an approximation to a general function f , it can be verified by, for example, a *sufficient summary plot* [58]. A sufficient summary plot depicts a finite amount of function evaluations from f on the axes representing the important directions. That is, the samples are projected on the important subspace (spanned by the columns of A) and plotted together with the corresponding function value. Of course, this is only possible if the important subspace has dimension 1 or 2. Fig. 3.2 shows sufficient summary plots for the ridge function f_2 from above. The left part (Fig. 3.2a) plots the function on the original coordinate axis x_1 and does not show a one-dimensional behavior as expected. However, the right plot (Fig. 3.2b) shows the function values on the important axis as a one-dimensional representation. We will use this type of plot later to demonstrate low-dimensionality in more complex scenarios.

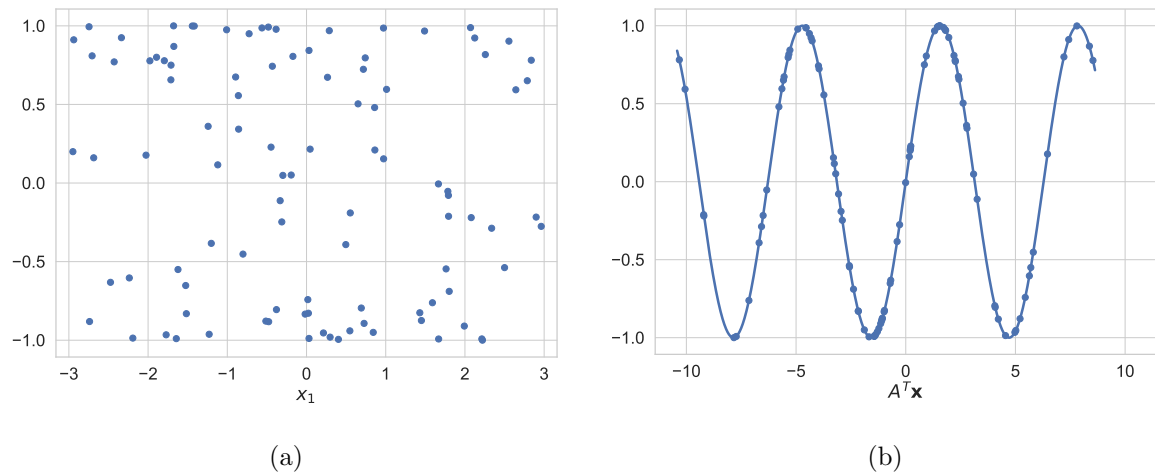


Figure 3.2: Sufficient summary plots for the ridge function f_2 from above. The samples are randomly picked. We see that f_2 does not have a one-dimensional behavior along the original coordinate axis x_1 (a), but along the important axis (b).

Related literature

Let us conclude this section with a short literature review emphasizing the two main branches, or perspectives, of subspace-based dimension reduction: statistics and approximation theory.

In the context of statistics, important subspaces are exploited to reduce the dimension of random variables and, particularly, of regression models [58, 59]. This process is called *sufficient dimension reduction* and, in the context of regression, aims to find linear combinations, represented by a rectangular matrix A (see above), of predictors \mathbf{x} such that the conditional distribution of the regression function, given $A^T \mathbf{x}$, is equal to the conditional distribution, given all predictors \mathbf{x} . It has motivated several advances and techniques in the field like *sliced inverse regression* (SIR) [61, 179, 279], *sliced average variance estimation* (SAVE) [57], and *contour regression* (CR) [178, 177]; see also, e.g., [60, 180, 287]. For a detailed review, we refer to [185].

Ridge approximation, already mentioned above, is the approximation theoretic approach to subspace-based dimension reduction. For example, a gradient-free compressed sensing method for ridge approximation is investigated in [92]. Two other techniques based on *Gaussian Process Regression* (GPR) [283], i.e., fitting a low-dimensional Gaussian process to given function evaluations, are discussed in [238, 267], where the former describes an algorithm that alternates the optimization of the important subspace spanned by A (see above) and the hyperparameters of the Gaussian process, and the latter regards the important subspace itself as a hyperparameter of the Gaussian process and bases the decision for the dimension of the subspace, i.e., the natural number k , on a Bayesian information criterion [237, 284]. The authors of [292] provide an extension of gradient-based ridge approximation, as it is used in this chapter, for vector-valued functions.

Note that, despite the differences of these two perspectives, there is a connection as was shown in [103]. For SIR and SAVE, [103, Thm. 6] proves that if and only if there exists a conditional independence of inputs and outputs of a function from the approximation setting, then the function is indeed a ridge function.

For dimension reduction in Bayesian and statistical inverse problems, we refer to [56, 69, 181, 293] at this point and come back to them later in Ch. 4.

Finally, the PhD thesis of Russi in [230] demonstrates that using a function's gradient at randomly picked locations and computing a singular value decomposition with them gives directions along which the function is only mildly changing, on average. The term “active subspace” is actually adopted from this thesis since the idea behind ASM is similar. In the following, we concentrate on ASM for which we give a formal setup in the next section.

3.2 Common setup

Recall that we want to approximate a given function of interest $f : \mathcal{X} \rightarrow \mathbf{R}$, $\mathcal{X} \subseteq \mathbf{R}^n$, by a ridge function, i. e., find a ridge profile g and a matrix $A \in \mathbf{R}^{n \times k}$, $k \leq n$, such that

$$f(\mathbf{x}) \approx g(A^\top \mathbf{x}) \quad (3.5)$$

for each $\mathbf{x} \in \mathcal{X}$. For the setup and description of ASM, we follow the two main sources [52, 54].

Consider a random variable \mathbf{X} distributed according to $\mathbf{P}_{\mathbf{X}}$ with (Lebesgue) density $\rho_{\mathbf{X}}$. Define $\mathcal{X} := \text{supp}(\rho_{\mathbf{X}}) \subseteq \mathbf{R}^n$ as the support of $\rho_{\mathbf{X}}$. Furthermore, assume that $f \in C^1(\mathcal{X}, \mathbf{R})$ with partial derivatives that are square-integrable w.r.t. $\mathbf{P}_{\mathbf{X}}$. We denote the gradient of f w.r.t. \mathbf{x} by $\nabla^{\mathbf{x}} f = \nabla f \in \mathbf{R}^n$. The starting point for ASM to investigate important and dominant directions of f is the matrix

$$\begin{aligned} C &:= \mathbf{E}[\nabla f(\mathbf{X}) \nabla f(\mathbf{X})^\top] \\ &= \int_{\mathcal{X}} \nabla f(\mathbf{x}) \nabla f(\mathbf{x})^\top \rho_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}, \end{aligned} \quad (3.6)$$

which is the average outer product of the gradient of f with itself. Note that this matrix is similar to a covariance matrix of the random vector $\nabla f(\mathbf{X})$ and actually equal to the auto-correlation matrix (see Def. A.2.2) of $\nabla f(\mathbf{X})$.

Since $C \in \mathbf{R}^{n \times n}$ is symmetric, we can choose orthonormal eigenvectors $\mathbf{w}_i \in \mathbf{R}^n$, $i = 1, \dots, n$, and the positive definiteness of C gives that corresponding eigenvalues λ_i are non-negative. Hence, we can get the eigendecomposition

$$C =: W \Lambda W^\top, \quad (3.7)$$

where

$$W =: \begin{pmatrix} | & & | \\ \mathbf{w}_1 & \cdots & \mathbf{w}_n \\ | & & | \end{pmatrix} \in \mathbf{R}^{n \times n} \quad (3.8)$$

is orthogonal and $\Lambda =: \text{diag}(\lambda_1, \dots, \lambda_n)$ with $\lambda_1 \geq \dots \geq \lambda_n \geq 0$.

The reason why we build and eigendecompose C according to (3.6) and (3.7) becomes clear by the equality

$$\lambda_i = \mathbf{w}_i^\top C \mathbf{w}_i = \mathbf{E}[(\mathbf{w}_i^\top \nabla f(\mathbf{X}))^2] \quad (3.9)$$

holding for each $i = 1, \dots, n$. This means that an eigenvalue of C indicates the *averaged* sensitivity of the function f in the direction of the corresponding eigenvector. In the extreme case that there exists a zero eigenvalue, i. e., $\lambda_i = 0$ for some i , f does not change along the

direction of \mathbf{w}_i at all; that is, f is a ridge function. We emphasize that all the sensitivities of f described by (3.9) are *averaged* sensitivities (weighted w.r.t. \mathbf{P}_X).

The equality in (3.9) also suggests to study f on a different coordinate system that is aligned with the sensitive and insensitive directions given by the eigenvector matrix W . That is, we can define a new transformed variable $\mathbf{v} := W^\top \mathbf{x}$ and continue the investigation of f w.r.t. this variable.

However, we still need to decide for a suitable reduction of the dimensionality. If the eigenvalues of C decay quickly enough, then we can split the n -dimensional input space correspondingly into a low-dimensional and a remaining higher-dimensional space. This is formally done by splitting the eigenvector matrix

$$W =: (W_1 \quad W_2), \quad (3.10)$$

where $W_1 \in \mathbf{R}^{n \times k}$ and $W_2 \in \mathbf{R}^{n \times (n-k)}$ for some $k \in \{1, \dots, n-1\}$. We call the column space of W_1 , i. e., $\text{ran}(W_1) := \{W_1 \mathbf{y} \mid \mathbf{y} \in \mathbf{R}^k\}$, the *active subspace (of f)* and the column space of W_2 the *inactive subspace (of f)*. Note that both spaces are subspaces of \mathbf{R}^n , but their dimensions are k and $n-k$, respectively. The new variable \mathbf{v} is also split correspondingly, i. e.,

$$\mathbf{v} =: \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix}, \quad (3.11)$$

where

$$\mathbf{y} = W_1^\top \mathbf{x} \in \mathbf{R}^k \quad \text{and} \quad \mathbf{z} = W_2^\top \mathbf{x} \in \mathbf{R}^{n-k}. \quad (3.12)$$

Analogously, we define

$$\mathbf{Y} := W_1^\top \mathbf{X} \quad \text{and} \quad \mathbf{Z} := W_2^\top \mathbf{X}. \quad (3.13)$$

The variables \mathbf{y} and \mathbf{Y} are called *active variables*; variables \mathbf{z} and \mathbf{Z} are called *inactive variables*.

Another way to (equivalently) derive the new coordinates \mathbf{y} and \mathbf{z} is through

$$\mathbf{x} = WW^\top \mathbf{x} = W_1 W_1^\top \mathbf{x} + W_2 W_2^\top \mathbf{x} = W_1 \mathbf{y} + W_2 \mathbf{z} \quad (3.14)$$

for arbitrary $\mathbf{x} \in \mathbf{R}^n$. This also gives a direct relation between \mathbf{x} and (\mathbf{y}, \mathbf{z}) and motivates the following notation,

$$\mathbf{x} = \llbracket \mathbf{y}, \mathbf{z} \rrbracket := \llbracket \mathbf{y}, \mathbf{z} \rrbracket_W := W_1 \mathbf{y} + W_2 \mathbf{z}. \quad (3.15)$$

We can note already at this point that $\mathbf{y} = W_1^\top \mathbf{x}$, as a low-dimensional linear combination of \mathbf{x} , suggests to choose $A = W_1$ in (3.5).

For convenience, we also define the sets

$$\mathcal{Y} := W_1^\top \mathcal{X} \quad \text{and} \quad \mathcal{Z} := W_2^\top \mathcal{X}, \quad (3.16)$$

where $M\mathcal{X} := \{M\mathbf{x} \mid \mathbf{x} \in \mathcal{X}\}$ for a matrix $M \in \mathbf{R}^{p \times n}$, $p \in \mathbf{N}$. The sets contain values for active and inactive variables originating from \mathcal{X} , the support of \mathbf{P}_X .

In addition, the transformation induces joint, marginal, and conditional densities, and distributions, for random variables \mathbf{Y} and \mathbf{Z} ; see, e. g., [30] for a formal probabilistic setup or compare with the construction in Sec. 2.2. The joint density of (\mathbf{Y}, \mathbf{Z}) is

$$\rho_{\mathbf{Y}, \mathbf{Z}}(\mathbf{y}, \mathbf{z}) = \rho_X(\llbracket \mathbf{y}, \mathbf{z} \rrbracket) \quad (3.17)$$

and the marginal densities are given by

$$\rho_{\mathbf{Y}}(\mathbf{y}) = \int_{\mathcal{Z}} \rho_{\mathbf{Y},\mathbf{Z}}(\mathbf{y}, \mathbf{z}) \, d\mathbf{z} \quad \text{and} \quad \rho_{\mathbf{Z}}(\mathbf{z}) = \int_{\mathcal{Y}} \rho_{\mathbf{Y},\mathbf{Z}}(\mathbf{y}, \mathbf{z}) \, d\mathbf{y}. \quad (3.18)$$

For a conditional density $\rho_{\mathbf{Z}|\mathbf{Y}}(\cdot|\mathbf{y})$, which becomes important below, we need to define

$$\mathcal{Y}^+ := \{\mathbf{y} \in \mathbf{R}^k \mid \rho_{\mathbf{Y}}(\mathbf{y}) > 0\} \quad (3.19)$$

as the set of all values for the active variable \mathbf{y} with a positive $\rho_{\mathbf{Y}}$ -density value. Then, we have that

$$\rho_{\mathbf{Z}|\mathbf{Y}}(\mathbf{z}|\mathbf{y}) = \frac{\rho_{\mathbf{Y},\mathbf{Z}}(\mathbf{y}, \mathbf{z})}{\rho_{\mathbf{Y}}(\mathbf{y})} \quad (3.20)$$

for all $\mathbf{y} \in \mathcal{Y}^+$.

In the remainder, we make use of the well-known tower property of conditional expectations, i. e., for any function $h : \mathbf{R}^n \rightarrow \mathbf{R}$ which is integrable w.r.t. $\mathbf{P}_{\mathbf{X}}$, it holds that

$$\mathbf{E}[h(\mathbf{X})] = \mathbf{E}[h(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket)] = \mathbf{E}[\mathbf{E}[h(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket) \mid \mathbf{Y}]]. \quad (3.21)$$

As the formal setup is now complete, we can continue with the actual goal of finding a low-dimensional representation of f in the form of a ridge function. As already mentioned, we decide to choose $A = W_1$ for the ridge approximation in (3.5). Hence, it remains to find a ridge profile g which is defined for $\mathbf{y} = W_1^\top \mathbf{x}$ on \mathbf{R}^k (or \mathcal{Y}).

Informally, given a value $\mathbf{y} = W_1^\top \mathbf{x} \in \mathcal{Y}$ for some $\mathbf{x} \in \mathcal{X}$, we need to find an approximation of $f(\mathbf{x})$. It is a well-known fact that, if $\mathbf{E}[f(\mathbf{X})^2] < \infty$, the conditional expectation of $f(\mathbf{X})$, given $\mathbf{Y} = \mathbf{y}$, minimizes the mean square error to f , i. e.,

$$\mathbf{E}[(f(\mathbf{X}) - \mathbf{E}[f(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket) \mid \mathbf{Y}])^2] \leq \mathbf{E}[(f(\mathbf{X}) - \mathbf{R})^2] \quad (3.22)$$

for any square-integrable random variable \mathbf{R} which is measurable w.r.t. the σ -algebra generated by \mathbf{Y} [156, Corollary 8.17]. Hence, we define

$$\begin{aligned} g(\mathbf{y}) &:= \mathbf{E}[f(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket) \mid \mathbf{Y} = \mathbf{y}] \\ &= \int_{\mathbf{R}^{n-k}} f(\llbracket \mathbf{y}, \mathbf{z} \rrbracket) \rho_{\mathbf{Z}|\mathbf{Y}}(\mathbf{z}|\mathbf{y}) \, d\mathbf{z} \end{aligned} \quad (3.23)$$

for $\mathbf{y} \in \mathcal{Y}^+$. Finally, we can define the desired ridge function as

$$f_g(\mathbf{x}) := g(W_1^\top \mathbf{x}) \quad (3.24)$$

for $\mathbf{x} \in \mathcal{X}^\circ$, where \mathcal{X}° denotes the interior of \mathcal{X} . The function f_g is well-defined since $W_1^\top \mathcal{X}^\circ \subseteq \mathcal{Y}^\circ \subseteq \mathcal{Y}^+$. Also, note that $\mathbf{P}_{\mathbf{X}}(\mathcal{X}^\circ) = 1$ for non-degenerate sets \mathcal{X} .

Remark (Orthogonal invariance). We want to mention an important and interesting concept that we call *orthogonal invariance*. It refers to the fact that we are actually not interested in a particular basis W_1 for the active subspace, but rather in the subspace itself that is spanned by the columns of W_1 . Indeed, the function g from (3.23) is invariant under an orthogonal transformation of the active subspace. Let $Q \in \mathbf{R}^{n \times n}$ be orthogonal and define

$$\bar{\mathbf{Y}} := Q^\top W_1^\top \mathbf{X} = Q^\top \mathbf{Y} \quad (3.25)$$

as the corresponding active variable in the transformed subspace; the inactive variable \mathbf{Z} is not transformed. We compute

$$\rho_{\mathbf{Z}|\bar{\mathbf{Y}}}(z|\bar{\mathbf{y}}) = \frac{\rho_{\mathbf{X}}(W_1 Q \bar{\mathbf{y}} + W_2 \mathbf{z})}{\int_{\mathbf{R}^{n-k}} \rho_{\mathbf{X}}(W_1 Q \bar{\mathbf{y}} + W_2 \mathbf{z}') d\mathbf{z}'} \quad (3.26)$$

$$= \frac{\rho_{\mathbf{X}}(W_1 \mathbf{y} + W_2 \mathbf{z})}{\int_{\mathbf{R}^{n-k}} \rho_{\mathbf{X}}(W_1 \mathbf{y} + W_2 \mathbf{z}') d\mathbf{z}'} \quad (3.27)$$

$$= \rho_{\mathbf{Z}|\mathbf{Y}}(z|\mathbf{y}) \quad (3.28)$$

and hence get that

$$g_Q(\mathbf{y}) := \mathbf{E}[f(W_1 Q \bar{\mathbf{Y}} + W_2 \mathbf{Z}) | \bar{\mathbf{Y}} = \mathbf{y}] \quad (3.29)$$

$$= \mathbf{E}[f(W_1 \mathbf{Y} + W_2 \mathbf{Z}) | \mathbf{Y} = \mathbf{y}] \quad (3.30)$$

$$= g(\mathbf{y}). \quad (3.31)$$

In [55], this property is used to justify the investigation of the optimization problem

$$\min_{W_1} \mathbf{E}[(f(\mathbf{X}) - f_g(\mathbf{X}))^2] \quad \text{such that } W_1 \in \mathbf{G}(k, n), \quad (3.32)$$

where $\mathbf{G}(k, n)$, called the *Grassmann manifold*, consists of all k -dimensional subspaces of \mathbf{R}^n and $f_g = f_g(W_1)$ is viewed as a function of W_1 . Note that the minimizing function in (3.32) is not convex in general; however, we can look for stationary points. For a Gaussian distribution $\mathbf{P}_{\mathbf{X}}$, it is shown in [55, Thm. 3] that the particular W_1 computed by ASM above is a *near-stationary* point in the sense that the Frobenius norm of the gradient on the Grassmann manifold (of the minimizing function) is bounded. For more details, see [55].

At this point, we have a ridge approximation for f , but do not know the approximation quality yet. As mentioned above, we determine the quality behind “ \approx ” in (3.5) by the mean square error between f and its approximation f_g , i. e., $\mathbf{E}[(f(\mathbf{X}) - f_g(\mathbf{X}))^2]$. Anticipating, we are interested in proving an upper bound on the mean square error involving eigenvalues corresponding to the *neglected* inactive subspace, i. e., we aim to show that

$$\mathbf{E}[(f(\mathbf{X}) - f_g(\mathbf{X}))^2] \leq C_{P,W}(\lambda_{k+1} + \cdots + \lambda_n) \quad (3.33)$$

for a Poincaré constant $C_{P,W} = C_{P,W}(W, \mathbf{P}_{\mathbf{X}}) > 0$, the origin of which gets clear below and which is indeed a central aspect in Sec. 3.4.

Note that $C_{P,W}$ particularly depends on the orthogonal transformation $W = W(f)$ from (3.8) and hence also indirectly on f . If necessary, we could get rid of this dependence by considering the supremum of $C_{P,W}$ over all orthogonal matrices, i. e.,

$$C_P := \sup_{W \text{ orth.}} C_{P,W}, \quad (3.34)$$

and get

$$\mathbf{E}[(f(\mathbf{X}) - f_g(\mathbf{X}))^2] \leq C_P(\lambda_{k+1} + \cdots + \lambda_n), \quad (3.35)$$

provided the constant $C_P = C_P(\mathbf{P}_{\mathbf{X}})$ exists. Deriving such an upper bound for a certain class of distributions would allow to choose $\mathbf{P}_{\mathbf{X}}$ independently of f . Note that [292, 293], which regard a related dimension reduction technique in the context of Bayesian inverse problems, based

on *logarithmic Sobolev inequalities* [108], also control the Poincaré constant for any orthogonal matrix W , but for a particular class of distributions; see Sec. 3.4.

The derivation of (3.33) starts with

$$\mathbf{E}[(f(\mathbf{X}) - f_g(\mathbf{X}))^2] = \mathbf{E}[\mathbf{E}[(f(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket) - g(\mathbf{Y}))^2 | \mathbf{Y}]] \quad (3.36)$$

$$\leq \mathbf{E}[C_{\mathbf{Y}} \cdot \mathbf{E}[\|\nabla^z f(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket)\|_2^2 | \mathbf{Y}]], \quad (3.37)$$

where we applied a probabilistic Poincaré inequality (see Def. A.2.14) in (3.37) w.r.t. the conditional distribution $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}(\cdot | \mathbf{Y}(\omega))$, $\omega \in \Omega$, with (random) Poincaré constant $C_{\mathbf{Y}} = C_{\mathbf{Y}}(W, \mathbf{P}_{\mathbf{X}}) > 0$.

Remark. Above, we assume continuous differentiability on f which is common in the literature. This assumption is, however, much too strong since it suffices to assume that $\nabla f \in L^2(\mathcal{X}, \mathbf{P}_{\mathbf{X}})$. Indeed, it implies that

$$\nabla^z f(\llbracket \mathbf{y}, \cdot \rrbracket) \in L^2(\mathbf{R}^{n-k}, \mathbf{P}_{\mathbf{Z}|\mathbf{Y}}(\cdot | \mathbf{y})) \quad (3.38)$$

for $\mathbf{P}_{\mathbf{Y}}$ -a.e. \mathbf{y} which also permits the step in (3.37); see [14, Def. 4.2.1]. This weaker assumption makes ASM applicable to much more scenarios of practical interest since it allows for functions being only differentiable $\mathbf{P}_{\mathbf{X}}$ -a.e.

If the random Poincaré constant $C_{\mathbf{Y}}$ is bounded, more exactly, if $\text{ess sup } C_{\mathbf{Y}} < \infty$, then we get that

$$\begin{aligned} \mathbf{E}[(f(\mathbf{X}) - f_g(\mathbf{X}))^2] &\leq C_{\mathbf{P},W} \mathbf{E}[\mathbf{E}[\|\nabla^z f(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket)\|_2^2 | \mathbf{Y}]] \\ &= C_{\mathbf{P},W} \mathbf{E}[\|\nabla^z f(\mathbf{X})\|_2^2], \end{aligned} \quad (3.39)$$

where $C_{\mathbf{P},W} = C_{\mathbf{P},W}(W, \mathbf{P}_{\mathbf{X}}) = \text{ess sup } C_{\mathbf{Y}} < \infty$. The following result finally yields the upper bound in (3.33).

Lemma 3.2.1 ([54, Lem. 2.2]). *The mean-squared L_2 norm of gradients of f w.r.t. \mathbf{y} and \mathbf{z} is*

$$\mathbf{E}[\|\nabla^{\mathbf{y}} f(\mathbf{X})\|_2^2] = \lambda_1 + \cdots + \lambda_k \quad (3.40)$$

and, respectively,

$$\mathbf{E}[\|\nabla^{\mathbf{z}} f(\mathbf{X})\|_2^2] = \lambda_{k+1} + \cdots + \lambda_n. \quad (3.41)$$

Proof. We only prove (3.40) since the proof for (3.41) is similar.

Note that for $\mathbf{x} \in \mathcal{X}$ such that $\mathbf{x} = \llbracket \mathbf{y}, \mathbf{z} \rrbracket = W_1 \mathbf{y} + W_2 \mathbf{z}$ for suitable $\mathbf{y} \in \mathcal{Y}$ and $\mathbf{z} \in \mathcal{Z}$, it holds that

$$\nabla^{\mathbf{y}} f(\mathbf{x}) = W_1^\top \nabla^{\mathbf{x}} f(\mathbf{x}). \quad (3.42)$$

It follows that

$$\begin{aligned} \mathbf{E}[\|\nabla^{\mathbf{y}} f(\mathbf{X})\|_2^2] &= \mathbf{E} \left[\text{trace} \left(\nabla^{\mathbf{y}} f(\mathbf{X}) \nabla^{\mathbf{y}} f(\mathbf{X})^\top \right) \right] \\ &= \text{trace} \left(\mathbf{E}[\nabla^{\mathbf{y}} f(\mathbf{X}) \nabla^{\mathbf{y}} f(\mathbf{X})^\top] \right) \\ &= \text{trace} \left(W_1^\top \mathbf{E}[\nabla^{\mathbf{x}} f(\mathbf{X}) \nabla^{\mathbf{x}} f(\mathbf{X})^\top] W_1 \right) \\ &= \text{trace} \left(W_1^\top C W_1 \right) \\ &= \text{trace} \left(W_1^\top W^\top \Lambda W W_1 \right) \\ &= \text{trace} (\Lambda_1) \\ &= \lambda_1 + \cdots + \lambda_k. \end{aligned} \quad (3.43)$$

■

Note that we make two assumptions in the derivation of (3.39) above:

1. We are allowed to apply a probabilistic Poincaré inequality w.r.t. $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}$ in (3.37).
2. The random Poincaré constant $C_{\mathbf{Y}}$ is bounded (has compact support), i.e.,

$$\text{ess sup } C_{\mathbf{Y}} < \infty. \quad (3.44)$$

The explicit indication of these assumptions is a main difference compared to the original exposition in [52, 54]. Whether the assumptions are satisfied indirectly depends on $\mathbf{P}_{\mathbf{X}}$ and directly, non-trivially on the related conditional distribution $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}$ induced by the orthogonal transformation from \mathbf{x} to (\mathbf{y}, \mathbf{z}) explained above. We emphasize that it is not sufficient that $\mathbf{P}_{\mathbf{X}}$ satisfies a probabilistic Poincaré inequality; it is the conditional distribution $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}(\cdot|\mathbf{y})$, $\mathbf{y} \in \mathcal{Y}^+$, and, additionally, it has to allow for a bounded Poincaré constant $C_{\mathbf{y}}$ in the sense described above.

The next two sections investigate the upper bound for distributions $\mathbf{P}_{\mathbf{X}}$ (and $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}$) that are known to satisfy the assumptions above (Sec. 3.3) and for distributions that do not follow the derivation above (Sec. 3.4) and thus only allow for *generalized* bounds [257].

3.3 Bounds for compactly supported and normal densities

Reference. This section follows [257, Sec. 3].

We look at two kinds of distribution classes for $\mathbf{P}_{\mathbf{X}}$ that are well-known to satisfy a probabilistic Poincaré inequality (see Def. A.2.14) and pass this property to a corresponding conditional distribution $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}(\cdot|\mathbf{y})$, $\mathbf{y} \in \mathcal{Y}^+$, which, most importantly, allows for a Poincaré constant $C_{\mathbf{y}}$ that is *uniform* in \mathbf{y} implying that $C_{\mathbf{P},W} = \text{ess sup } C_{\mathbf{Y}} < \infty$ as desired.

The *uniform distribution* $\mathbf{P}_{\mathbf{X}} = \mathcal{U}(\mathcal{X})$ on a Borel set \mathcal{X} is the canonical case for distributions with compact support. It has a Lebesgue density $\rho_{\mathbf{X}}(\mathbf{x}) = \mathbf{1}_{\mathcal{X}}(\mathbf{x})/\lambda(\mathcal{X})$, $\mathbf{x} \in \mathbf{R}^n$. We know that it satisfies a probabilistic Poincaré inequality on its own but also implies the same for conditional distributions $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}$ which are also uniform. Note that, in this case, the probabilistic Poincaré inequality coincides with a regular Poincaré inequality; see Thm. A.1.3 for functions in H^1 .

The following theorem from [257], which is actually a more general statement as it includes more compactly supported distributions than just the uniform one, proves that the two assumptions from above are indeed fulfilled. In order to allow for explicit Poincaré constants, we additionally assume convexity of \mathcal{X}° , the interior of \mathcal{X} . Recall that the Poincaré constant for a convex domain with diameter $d > 0$ is d/π [21].

Theorem 3.3.1 ([257, Thm. 3.1]). *Assume that \mathcal{X}° is a bounded and convex domain. If $0 < \delta \leq \rho_{\mathbf{X}}(\mathbf{x}) \leq D < \infty$ for all $\mathbf{x} \in \mathcal{X}^\circ$, then*

$$\mathbf{E}[(f(\mathbf{X}) - f_g(\mathbf{X}))^2] \leq C_{\mathbf{P}}(\lambda_{k+1} + \dots + \lambda_n) \quad (3.45)$$

for a constant

$$C_{\mathbf{P}} = C_{\mathbf{P}}(\delta, D, \mathcal{X}) = \frac{\text{diam}(\mathcal{X})}{\pi} \cdot \frac{D}{\delta} > 0. \quad (3.46)$$

Proof. Define

$$\mathcal{Z}_{\mathbf{y}}^{\circ} = \{\mathbf{z} \in \mathbf{R}^{n-k} \mid \llbracket \mathbf{y}, \mathbf{z} \rrbracket \in \mathcal{X}^{\circ}\} \subseteq \mathcal{Z} \quad (3.47)$$

and note that it is convex for $\mathbf{y} \in \mathcal{Y}^+$. It holds that

$$\text{diam}(\mathcal{Z}_{\mathbf{y}}^{\circ}) \leq \text{diam}(\mathcal{Z}) \leq \text{diam}(\mathcal{X}). \quad (3.48)$$

Note that

$$\frac{\delta}{\rho_{\mathbf{Y}}(\mathbf{y})} \leq \rho_{\mathbf{Z}|\mathbf{Y}}(\mathbf{z}|\mathbf{y}) \leq \frac{D}{\rho_{\mathbf{Y}}(\mathbf{y})} \quad (3.49)$$

for $\mathbf{y} \in \mathcal{Y}^+$ and $\mathbf{z} \in \mathcal{Z}_{\mathbf{y}}^{\circ}$. This justifies the following lines of computation for $\mathbf{y} \in \mathcal{Y}^+$,

$$\begin{aligned} & \mathbf{E}[(f(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket) - g(\mathbf{Y}))^2 \mid \mathbf{Y} = \mathbf{y}] \\ &= \int_{\mathcal{Z}_{\mathbf{y}}^{\circ}} (f(\llbracket \mathbf{y}, \mathbf{z} \rrbracket) - g(\mathbf{y}))^2 \rho_{\mathbf{Z}|\mathbf{Y}}(\mathbf{z}|\mathbf{y}) \, d\mathbf{z} \\ &\leq \frac{D}{\rho_{\mathbf{Y}}(\mathbf{y})} \int_{\mathcal{Z}_{\mathbf{y}}^{\circ}} (f(\llbracket \mathbf{y}, \mathbf{z} \rrbracket) - g(\mathbf{y}))^2 \, d\mathbf{z} \\ &\leq \frac{\text{diam}(\mathcal{Z}_{\mathbf{y}}^{\circ})}{\pi} \frac{D}{\rho_{\mathbf{Y}}(\mathbf{y})} \int_{\mathcal{Z}_{\mathbf{y}}^{\circ}} \|\nabla^{\mathbf{z}} f(\llbracket \mathbf{y}, \mathbf{z} \rrbracket)\|_2^2 \, d\mathbf{z} \\ &\leq \frac{\text{diam}(\mathcal{X})}{\pi} \frac{D}{\delta} \int_{\mathcal{Z}_{\mathbf{y}}^{\circ}} \|\nabla^{\mathbf{z}} f(\llbracket \mathbf{y}, \mathbf{z} \rrbracket)\|_2^2 \rho_{\mathbf{Z}|\mathbf{Y}}(\mathbf{z}|\mathbf{y}) \, d\mathbf{z} \\ &= \frac{\text{diam}(\mathcal{X})}{\pi} \frac{D}{\delta} \mathbf{E}[\|\nabla^{\mathbf{z}} f(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket)\|_2^2 \mid \mathbf{Y} = \mathbf{y}]. \end{aligned} \quad (3.50)$$

Then, combining (3.43) with (3.50) yields the result in (3.45). \blacksquare

Remark. Note that the Poincaré constant in (3.45) does not depend on W .

The *(multivariate) normal distribution* also passes the property of satisfying a probabilistic Poincaré inequality to the conditional distribution $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}$. It was shown in [47, Corollary 3.1] that the multivariate *standard* normal distribution has Poincaré constant 1. Shifting and scaling arguments used in [47, Corollary 3.2] give that, if $\mathbf{P}_{\mathbf{X}} = \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ for a mean vector $\boldsymbol{\mu} \in \mathbf{R}^n$ and a covariance matrix $\Sigma \in \mathbf{R}^{n \times n}$,

$$\mathbf{Var}(h(\mathbf{X})) \leq \mathbf{E}[\nabla h(\mathbf{X})^{\top} \Sigma \nabla h(\mathbf{X})] \quad (3.51)$$

for any continuously differentiable function h . If we continue the computation with

$$\begin{aligned} \mathbf{E}[\nabla h(\mathbf{X})^{\top} \Sigma \nabla h(\mathbf{X})] &= \mathbf{E}[\|\Sigma^{1/2} \nabla h(\mathbf{X})\|_2^2] \\ &\leq \|\Sigma^{1/2}\|_2^2 \mathbf{E}[\|\nabla h(\mathbf{X})\|_2^2] \\ &\leq \lambda_{\max}(\Sigma) \mathbf{E}[\|\nabla h(\mathbf{X})\|_2^2], \end{aligned} \quad (3.52)$$

we get that the Poincaré constant is $\lambda_{\max}(\Sigma)$ in the case of general multivariate normal distributions.

However, the derivation is only for $\mathbf{P}_{\mathbf{X}}$, but we need the conditional distribution $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}$ to satisfy the probabilistic Poincaré inequality. Fortunately, normal distributions are invariant

under orthogonal transformations meaning that $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}$ is also normally distributed. This property is exploited to show that also $C_P = C_{P,W} = \lambda_{\max}(\Sigma)$. The details for this derivation are given in Appendix B.

In fact, $\lambda_{\max}(\Sigma)$ as the Poincaré constant for multivariate normal distributions can also be deduced from a much more general framework involving so-called α -uniformly log-concave distributions covered in Sec. 3.4 which handles distributions with heavier (exponential) tails than discussed in this section.

3.4 Generalized bounds

Reference. This section strongly follows [257, Sec. 4].

So far, we only studied the two types of distributions from Sec. 3.3 (compactly supported and normal distributions) w.r.t. their applicability to ASM. Here, we investigate the large class of log-concave probability measures which have Lebesgue densities of the form

$$\rho_{\mathbf{X}}(\mathbf{x}) = \exp(-V(\mathbf{x})) \quad (3.53)$$

for a convex function $V : \mathbf{R}^n \rightarrow (-\infty, +\infty]$ and $\mathbf{x} \in \mathbf{R}^n$. Note that $+\infty$ is included in the codomain of V . The conditional density $\rho_{\mathbf{Z}|\mathbf{Y}}(\cdot|\mathbf{y})$ for a given $\mathbf{y} \in \mathcal{Y}^+$ is then given by

$$\rho_{\mathbf{Z}|\mathbf{Y}}(\mathbf{z}|\mathbf{y}) = \frac{\exp(-V(\llbracket \mathbf{y}, \mathbf{z} \rrbracket))}{\rho_{\mathbf{Y}}(\mathbf{y})} = \exp(-\tilde{V}_{\mathbf{y}}(\mathbf{z})), \quad (3.54)$$

where $\tilde{V}_{\mathbf{y}}(\mathbf{z}) := V(\llbracket \mathbf{y}, \mathbf{z} \rrbracket) + \log(\rho_{\mathbf{Y}}(\mathbf{y}))$. Note that $\tilde{V}_{\mathbf{y}}$ inherits convexity (in \mathbf{z}) from V . Bobkov [34] shows that log-concave distributions satisfy a Poincaré inequality and gives lower and upper bounds on the corresponding Poincaré constant.

First, we discuss the special case of α -uniformly convex functions V (Subsec. 3.4.1) for which the corresponding distribution $\mathbf{P}_{\mathbf{X}}$ is known to satisfy a Poincaré inequality with universal Poincaré constant $1/\alpha$. However, the assumption on $\mathbf{P}_{\mathbf{X}}$ being of uniformly log-concave type is somewhat restrictive since it excludes distributions with heavier tails as, for example, exponential or Laplace distributions. For this reason, we secondly investigate general log-concave distributions (Subsec. 3.4.2) and show that there might arise problems with this class due to arbitrary large Poincaré constants $C_{\mathbf{Y}}$. In particular, the problems and their proposed solution are exemplified on independent exponential distributions in $n \geq 2$ dimensions.

3.4.1 α -uniformly convex functions V

Definition 3.4.1 (α -uniformly convex function). A function $V \in C^2(\mathbf{R}^n)$ is said to be α -uniformly convex, if there is an $\alpha > 0$ such that for all $\mathbf{x} \in \mathbf{R}^n$ it holds that

$$\mathbf{u}^\top V''(\mathbf{x})\mathbf{u} \geq \alpha \|\mathbf{u}\|_2^2 \quad (3.55)$$

for all $\mathbf{u} \in \mathbf{R}^n$, where V'' denotes the Hessian matrix of V .

In [269, p. 43–44], it was shown that there is a dimension-free Poincaré constant $1/\alpha$ for α -uniformly log-concave $\mathbf{P}_{\mathbf{X}}$. Note that this says nothing about the special case $\alpha = 0$. The existence of a dimension-free Poincaré constant for this special case is actually a consequence of

the famous Kannan-Lovász-Simonovits conjecture; see, e. g., [11, 173]. However, since we need a Poincaré inequality for $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}(\cdot|\mathbf{y})$, $\mathbf{y} \in \mathcal{Y}^+$, we have to prove the following lemma similar to [293, Subsec. 7.2]

Lemma 3.4.2. *If $\mathbf{P}_{\mathbf{X}}$ is α -uniformly log-concave, then $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}(\cdot|\mathbf{y})$ is α -uniformly log-concave for each $\mathbf{y} \in \mathcal{Y}^+$.*

Proof. Let $\mathbf{y} \in \mathcal{Y}^+$. Recall that $\rho_{\mathbf{Z}|\mathbf{Y}}(\mathbf{z}|\mathbf{y}) = \exp(-\tilde{V}_{\mathbf{y}}(\mathbf{z}))$ for a convex function $\tilde{V}_{\mathbf{y}}(\mathbf{z}) := V(\llbracket \mathbf{y}, \mathbf{z} \rrbracket) + \log(\rho_{\mathbf{y}}(\mathbf{y}))$. The Hessian matrix $\tilde{V}_{\mathbf{y}}''(\mathbf{z})$ (w.r.t. \mathbf{z}) computes to

$$\tilde{V}_{\mathbf{y}}''(\mathbf{z}) = W_2^\top V''(\llbracket \mathbf{y}, \mathbf{z} \rrbracket) W_2. \quad (3.56)$$

Choose $\mathbf{w} \in \mathbf{R}^{n-k}$ arbitrarily. Then, for every $\mathbf{z} \in \mathbf{R}^{n-k}$, it holds that

$$\mathbf{w}^\top \tilde{V}_{\mathbf{y}}''(\mathbf{z}) \mathbf{w} = (W_2 \mathbf{w})^\top V''(\llbracket \mathbf{y}, \mathbf{z} \rrbracket) (W_2 \mathbf{w}) \quad (3.57)$$

$$\geq \alpha \|W_2 \mathbf{w}\|_2^2 = \alpha \|\mathbf{w}\|_2^2. \quad (3.58)$$

■

Since $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}(\cdot|\mathbf{y})$ inherits the universal Poincaré constant $1/\alpha$ from $\mathbf{P}_{\mathbf{X}}$, the result in (3.35) also holds for α -uniformly log-concave distributions with $C_P = 1/\alpha$ (independent of W) which is similar to [293, Corollary 2].

For example, α -uniformly log-concave distributions comprise multivariate normal distributions $\mathcal{N}(\mathbf{m}, \Sigma)$ with mean \mathbf{m} and covariance matrix Σ ($\alpha = 1/\lambda_{\max}(\Sigma)$). However, distributions that satisfy the assumption only for $\alpha = 0$ as, e. g., Weibull distributions with the exponential distribution as a special case or Gamma distributions with shape parameter $\beta \geq 1$, only belong to the class of general log-concave distributions.

3.4.2 General convex functions V

Since we cannot make use of a universal dimension-free Poincaré constant involving general convex functions $V : \mathbf{R}^n \rightarrow (-\infty, +\infty]$, we look at them more closely in this subsection. Recall that $\rho_{\mathbf{Z}|\mathbf{Y}}(\mathbf{z}|\mathbf{y}) = \exp(-\tilde{V}_{\mathbf{y}}(\mathbf{z}))$, $\mathbf{y} \in \mathcal{Y}^+$, for a convex function $\tilde{V}_{\mathbf{y}}$. We have to deal with the fact that the essential supremum of the random Poincaré constant $C_{\mathbf{Y}} = C_{\mathbf{Y}}(\omega)$ of $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}(\cdot|\mathbf{Y}(\omega))$ does possibly not exist. A corresponding example is given in Subsec. 3.4.3.

In the step from (3.37) to (3.39), we have applied Hölder's inequality with Hölder conjugates $(p, q) = (+\infty, 1)$. Since this is not possible for unbounded random variables $C_{\mathbf{Y}}$, we can only show a weaker result.

Lemma 3.4.3. *Let $\varepsilon > 0$. If $\|\nabla f(\mathbf{X})\|_2^2 \leq L$ \mathbf{P} -a.s. for some constant $L > 0$, then*

$$\mathbf{E}[(f(\mathbf{X}) - f_g(\mathbf{X}))^2] \leq C_{P,\varepsilon,W} (\lambda_{k+1} + \dots + \lambda_n)^{1/(1+\varepsilon)}, \quad (3.59)$$

where

$$C_{P,\varepsilon,W} = C_{P,\varepsilon,W}(\varepsilon, n, k, L, W, \mathbf{P}_{\mathbf{X}}) := L^{\varepsilon/(1+\varepsilon)} \mathbf{E}[C_{\mathbf{Y}}^{(1+\varepsilon)/\varepsilon}]^{\varepsilon/(1+\varepsilon)}. \quad (3.60)$$

Proof. The assumed boundedness of ∇f implies that also

$$\mathbf{E}[\|\nabla^z f(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket)\|_2^2 | \mathbf{Y}] \leq L \quad \mathbf{P}\text{-a.s.}$$

Choosing a weaker pair of conjugates $(p, q) = ((1 + \varepsilon)/\varepsilon, 1 + \varepsilon)$, $\varepsilon > 0$, we compute

$$\mathbf{E}[C_{\mathbf{Y}} \mathbf{E}[\|\nabla^z f(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket)\|_2^2 | \mathbf{Y}]] \quad (3.61)$$

$$\leq \mathbf{E}[C_{\mathbf{Y}}^p]^{1/p} \mathbf{E}[\mathbf{E}[\|\nabla^z f(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket)\|_2^2 | \mathbf{Y}]^q]^{1/q} \quad (3.62)$$

$$= \mathbf{E}[C_{\mathbf{Y}}^{(1+\varepsilon)/\varepsilon}]^{\varepsilon/(1+\varepsilon)} \mathbf{E}[\mathbf{E}[\|\nabla^z f(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket)\|_2^2 | \mathbf{Y}]^{1+\varepsilon}]^{1/(1+\varepsilon)} \quad (3.63)$$

$$\leq L^{\varepsilon/(1+\varepsilon)} \mathbf{E}[C_{\mathbf{Y}}^{(1+\varepsilon)/\varepsilon}]^{\varepsilon/(1+\varepsilon)} \mathbf{E}[\mathbf{E}[\|\nabla^z f(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket)\|_2^2 | \mathbf{Y}]]^{1/(1+\varepsilon)} \quad (3.64)$$

$$\leq L^{\varepsilon/(1+\varepsilon)} \mathbf{E}[C_{\mathbf{Y}}^{(1+\varepsilon)/\varepsilon}]^{\varepsilon/(1+\varepsilon)} (\lambda_{k+1} + \dots + \lambda_n)^{1/(1+\varepsilon)} \quad (3.65)$$

$$= C_{P,\varepsilon,W} (\lambda_{k+1} + \dots + \lambda_n)^{1/(1+\varepsilon)}. \quad (3.66)$$

The step in (3.65) uses Lem. 3.2.1. The result follows by (3.36) and (3.37). \blacksquare

Remark. The previous lemma requires the gradient of f to be uniformly bounded $\mathbf{P}_{\mathbf{X}}$ -a.e., an assumption that is not needed in [54] and [293].

However, first, applying ASM, in the sense that the matrix C from (3.6) is estimated by a finite Monte Carlo sum, requires an even slightly stronger assumption to prove results on corresponding approximations of eigenvalues λ_i and eigenvectors \mathbf{w}_i ; see Assumption 3.5.1 in Sec. 3.5.

Secondly, this assumption can be weakened by applying another Hölder's inequality analogous to (3.62). Indeed, for $\varepsilon \in (0, 1)$, we would get

$$\mathbf{E}[\mathbf{E}[\|\nabla^z f(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket)\|_2^2 | \mathbf{Y}]^{1+\varepsilon}]^{1/(1+\varepsilon)} \quad (3.67)$$

$$\leq \mathbf{E}[\mathbf{E}[\|\nabla^z f(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket)\|_2^2 | \mathbf{Y}]^{1/(1-\varepsilon)}]^{(1-\varepsilon)/(1+\varepsilon)} \quad (3.68)$$

$$\cdot \mathbf{E}[\mathbf{E}[\|\nabla^z f(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket)\|_2^2 | \mathbf{Y}]]^{\varepsilon/(1+\varepsilon)}. \quad (3.69)$$

Since

$$\mathbf{E}[\mathbf{E}[\|\nabla^z f(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket)\|_2^2 | \mathbf{Y}]^{1/(1-\varepsilon)}] \quad (3.70)$$

$$\leq \mathbf{E}[\mathbf{E}[\|\nabla^z f(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket)\|_2^{2/(1-\varepsilon)} | \mathbf{Y}]] \quad (3.71)$$

$$\leq \mathbf{E}[\|\nabla^x f(\mathbf{X})\|_2^{2/(1-\varepsilon)}], \quad (3.72)$$

we would only require $\|\nabla^x f(\mathbf{X})\|_2^{2/(1-\varepsilon)}$ to be integrable. What we, however, would have to accept in this case, is the resulting weaker order $\varepsilon/(1 + \varepsilon)$ in the eigenvalues belonging to the inactive subspace.

The L - and $\mathbf{P}_{\mathbf{X}}$ -dependence of $C_{P,\varepsilon,W}$ is notationally neglected in the following. If possible, we can choose a suitable $\varepsilon > 0$ to get $\mathbf{E}[C_{\mathbf{Y}}^{(1+\varepsilon)/\varepsilon}] < \infty$ and thus a finite constant $C_{P,\varepsilon,W}$. Note that we lose first order in the eigenvalues from the inactive subspace, but have instead order $1/(1 + \varepsilon) < 1$. Of course, the constant $C_{P,\varepsilon,W}$ could get arbitrarily large as $\varepsilon \rightarrow 0$, but this strongly depends on W and the moments of $C_{\mathbf{Y}}$; see the example given in Subsec. 3.4.3.

It is known by Bobkov [34, Eqs. (1.3), (1.8) and p. 1906] that there exists a (dimensionally dependent) Poincaré constant $C_{\mathbf{y}}$ for a general log-concave distribution $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}(\cdot|\mathbf{y})$ that is bounded from below and above by

$$\begin{aligned} \mathbf{E}[(\|\mathbf{Z} - \mathbf{z}_0\|_2 - \mathbf{E}[\|\mathbf{Z} - \mathbf{z}_0\|_2 | \mathbf{Y} = \mathbf{y}])^2 | \mathbf{Y} = \mathbf{y}] &\leq C_{\mathbf{y}} \\ &\leq K \mathbf{E}[\|\mathbf{Z} - \mathbf{z}_0\|_2^2 | \mathbf{Y} = \mathbf{y}] \\ &= K \sum_{i=1}^{n-k} \text{Var}(Z_i | \mathbf{Y} = \mathbf{y}), \end{aligned} \quad (3.73)$$

where $\mathbf{z}_0 := \mathbf{E}[\mathbf{Z} | \mathbf{Y} = \mathbf{y}]$ and $K = 432$ [34, Eqs. (1.8) and (3.4)] is a universal constant. To the authors' knowledge, the constant $C_{\mathbf{y}}$ is the best available. We provide a scenario in Subsec. 3.4.3 ("Rotation by $\theta = \pi/4$ ") in which the lower bound viewed as a random variable has no finite essential supremum implying the same for $C_{\mathbf{Y}}$.

However, to make use of Lem. 3.4.3, we need to investigate the involved constant $C_{\mathbf{P},\varepsilon,W}(\varepsilon, n, k)$.

Lemma 3.4.4. *It holds that*

$$\mathbf{E}[C_{\mathbf{Y}}^{(1+\varepsilon)/\varepsilon}]^{\varepsilon/(1+\varepsilon)} \leq K(n-k)^{1/(1+\varepsilon)} C_{\mathbf{V}_{\text{ar}},W}, \quad (3.74)$$

where

$$C_{\mathbf{V}_{\text{ar}},W} = C_{\mathbf{V}_{\text{ar}},W}(\varepsilon, n, k, W) := \left(\sum_{i=1}^{n-k} \mathbf{E}[\mathbf{V}_{\text{ar}}(Z_i | \mathbf{Y})^{(1+\varepsilon)/\varepsilon}] \right)^{\varepsilon/(1+\varepsilon)}. \quad (3.75)$$

Proof. Using Jensen's inequality for weighted sums, it follows that

$$\mathbf{E}[C_{\mathbf{Y}}^{(1+\varepsilon)/\varepsilon}] \leq K^{(1+\varepsilon)/\varepsilon} (n-k)^{(1+\varepsilon)/\varepsilon} \frac{1}{n-k} \sum_{i=1}^{n-k} \mathbf{E}[\mathbf{V}_{\text{ar}}(Z_i | \mathbf{Y})^{(1+\varepsilon)/\varepsilon}] \quad (3.76)$$

$$= K^{(1+\varepsilon)/\varepsilon} (n-k)^{1/\varepsilon} \sum_{i=1}^{n-k} \mathbf{E}[\mathbf{V}_{\text{ar}}(Z_i | \mathbf{Y})^{(1+\varepsilon)/\varepsilon}]. \quad (3.77)$$

The result follows. ■

Eventually, we get

$$C_{\mathbf{P},\varepsilon,W}(\varepsilon, n, k) \leq L^{\varepsilon/(1+\varepsilon)} K(n-k)^{1/(1+\varepsilon)} C_{\mathbf{V}_{\text{ar}},W}(\varepsilon, n, k). \quad (3.78)$$

As before, we can remove the dependence of $C_{\mathbf{V}_{\text{ar}},W}$ on $W = W(f)$ by considering the supremum over all orthogonal matrices. That is, we define

$$C_{\mathbf{P},\varepsilon} := \sup_{W \text{ orth.}} C_{\mathbf{P},\varepsilon,W} \quad (3.79)$$

and

$$C_{\mathbf{V}_{\text{ar}}} := \sup_{W \text{ orth.}} C_{\mathbf{V}_{\text{ar}},W}, \quad (3.80)$$

and get

$$\mathbf{E}[(f(\mathbf{X}) - f_g(\mathbf{X}))^2] \leq C_{\mathbf{P},\varepsilon}(\lambda_{k+1} + \dots + \lambda_n)^{1/(1+\varepsilon)}, \quad (3.81)$$

provided the constant $C_{\mathbf{P},\varepsilon} = C_{\mathbf{P},\varepsilon}(\varepsilon, n, k, L, \mathbf{P}_{\mathbf{X}})$ exists.

For $C_{\mathbf{V}_{\text{ar}}}$, we argue that it is actually enough to take the supremum only over the set of rotation matrices. Indeed, any orthogonal matrix W is either a *proper* ($\det W = 1$) or an *improper* ($\det W = -1$) rotation which is the combination of a proper rotation and an inversion of the axes; see, e.g., [153, 198]. However, since the constant $C_{\mathbf{V}_{\text{ar}},W}$ from (3.75) is invariant to inversions of the axes, it holds that

$$\sup_{W \text{ orth.}} C_{\mathbf{V}_{\text{ar}},W} = \sup_{R \text{ rot.}} C_{\mathbf{V}_{\text{ar}},R}. \quad (3.82)$$

This equality is exploited in the next subsection.

3.4.3 Independent exponential distributions

In this subsection, we use the lower bound of Bobkov from (3.73) to show that there exists a scenario, involving independent exponential distributions in $n \geq 2$ dimensions, in which the random Poincaré constant $C_{\mathbf{Y}}$ does not have an essential supremum implying that $C_{\mathbf{P}}$ from (3.34) does not exist. Therefore, the quantity $C_{\mathbf{V}_{\text{ar}}}$ from (3.80) is studied to derive a (finite) upper bound for $C_{\mathbf{P},\varepsilon}$ from (3.79) in this case.

We regard a random vector $\mathbf{X} = (X_1, \dots, X_n)^\top$ whose components are independently exponentially distributed with unit rates $\nu_i = 1$, $i = 1, \dots, n$, and will see that investigations with unit rates are sufficient to derive statements also involving other rates. The distribution of \mathbf{X} has the density

$$\rho_{\mathbf{X}}(\mathbf{x}) = \begin{cases} \exp(-x_1 - \dots - x_n) & \text{if } \mathbf{x} = (x_1, \dots, x_n)^\top \in \mathbf{R}_{\geq 0}^n, \\ 0 & \text{otherwise.} \end{cases} \quad (3.83)$$

That is, in this case $\mathcal{X} = \mathbf{R}_{\geq 0}^n$ and

$$V(\mathbf{x}) = \begin{cases} x_1 + \dots + x_n & \text{if } \mathbf{x} = (x_1, \dots, x_n)^\top \in \mathbf{R}_{\geq 0}^n, \\ +\infty & \text{otherwise.} \end{cases} \quad (3.84)$$

Note that V is convex.

Since we are interested in $C_{\mathbf{V}_{\text{ar}}}$ as a supremum over all orthogonal matrices, we assume that, in this subsection, $W = \begin{pmatrix} W_1 & W_2 \end{pmatrix}$ is an *arbitrary* orthogonal matrix not depending on f and $\mathbf{P}_{\mathbf{X}}$. Indeed, as the equality in (3.82) motivates, we can further assume that W is a rotation matrix.

2 dimensions

The joint density of two independently exponentially distributed random variables X_1 and X_2 both with unit rate is

$$\rho_{\mathbf{X}}(x_1, x_2) = \begin{cases} \exp(-x_1 - x_2) & \text{if } x_1, x_2 \geq 0, \\ 0 & \text{otherwise.} \end{cases} \quad (3.85)$$

First, let us regard a rotation of the two-dimensional Cartesian coordinate system by a general angle $\theta \in [-\pi, \pi)$ to a coordinate system for (y, z) , i.e.,

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = R_\theta \begin{pmatrix} y \\ z \end{pmatrix} \quad (3.86)$$

for a rotation matrix

$$W = R_\theta := \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. \quad (3.87)$$

That is, in two dimensions, it holds that

$$C_{\mathbf{V}_{\text{ar}}} = \sup_{\theta \in [-\pi, \pi)} C_{\mathbf{V}_{\text{ar}}, R_\theta}. \quad (3.88)$$

Subsequently, we look at the special case $\theta = \pi/4$ as an example for an unbounded Poincaré constant C_y of $\mathbf{P}_{z|y}(\cdot|y)$. Variables are written in thin letters in this subsection since they denote real values and not multidimensional vectors.

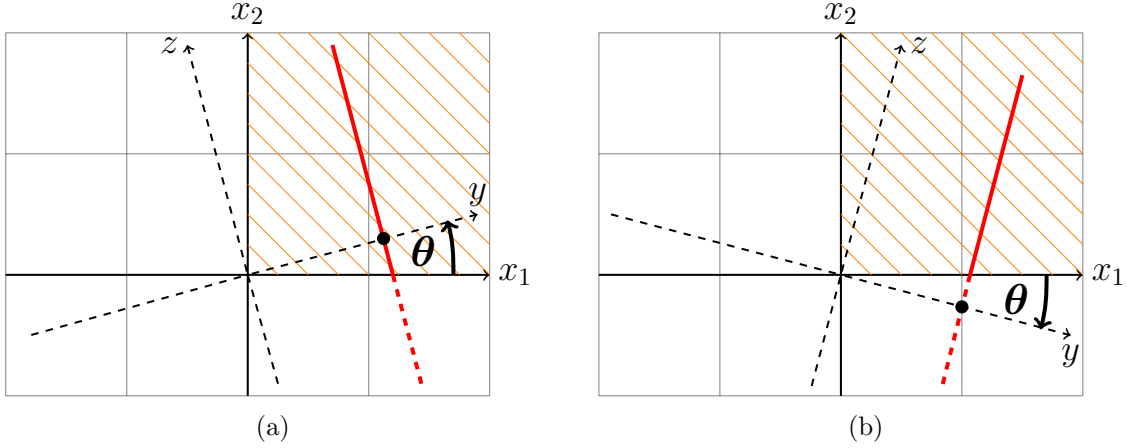


Figure 3.3: Rotations of the coordinate system with a positive (a) and a negative (b) angle. The orange lines depict contour lines in the support of $\rho_{\mathbf{X}}$. The red lines show the values of (y, z) for a given y . Their solid parts mark regions within the support of $\rho_{\mathbf{X}}$, whereas the dashed parts identify values with density zero.

Note that the bound from (3.74) in this two-dimensional setting becomes

$$\mathbf{E}[C_Y^{(1+\varepsilon)/\varepsilon}]^{\varepsilon/(1+\varepsilon)} \leq K C_{\mathbf{V}_{\text{ar}}, W}(\varepsilon, 2, 1) \quad (3.89)$$

with

$$C_{\mathbf{V}_{\text{ar}}, W}(\varepsilon, 2, 1) = \mathbf{E}[\mathbf{V}_{\text{ar}}(Z|Y)^{(1+\varepsilon)/\varepsilon}]^{\varepsilon/(1+\varepsilon)}. \quad (3.90)$$

Rotation by general θ

Let $\theta \in [-\pi, \pi)$. Then, the joint density of (Y, Z) is

$$\rho_{Y,Z}(y, z) = \exp(-(y \cos \theta - z \sin \theta) - (y \sin \theta + z \cos \theta)) \quad (3.91)$$

$$= \exp(-(\cos \theta + \sin \theta)y - (\cos \theta - \sin \theta)z). \quad (3.92)$$

for (y, z) with $\llbracket y, z \rrbracket \in \mathbf{R}_{\geq 0}^2$ and zero otherwise. If we define $a_{\theta}^+ := \cos \theta + \sin \theta$ and $a_{\theta}^- := \cos \theta - \sin \theta$, we have

$$\rho_{Y,Z}(y, z) = \begin{cases} \exp(-a_{\theta}^+ y - a_{\theta}^- z) & \text{if } \llbracket y, z \rrbracket \in \mathbf{R}_{\geq 0}^2, \\ 0 & \text{otherwise.} \end{cases} \quad (3.93)$$

Fig. 3.3 illustrates the situation for a positive (Fig. 3.3a) and a negative (Fig. 3.3b) angle θ .

The interval of investigation for $\theta \in [-\pi, \pi)$ can be reduced by reasons of periodicity and symmetry. First, note that the map

$$Q_{\varepsilon}(\theta) := C_{\mathbf{V}_{\text{ar}}, R_{\theta}}(\varepsilon, 2, 1), \quad (3.94)$$

is π -periodic in θ since an additional rotation by π corresponds to changing signs of y and z which is not important for integrals in Q_{ε} . Hence, it suffices to consider $\theta \in [-\pi/2, \pi/2)$. Secondly, from Fig. 3.3 it can be deduced that Q_{ε} , as a map of θ , is symmetric around $-\pi/4$

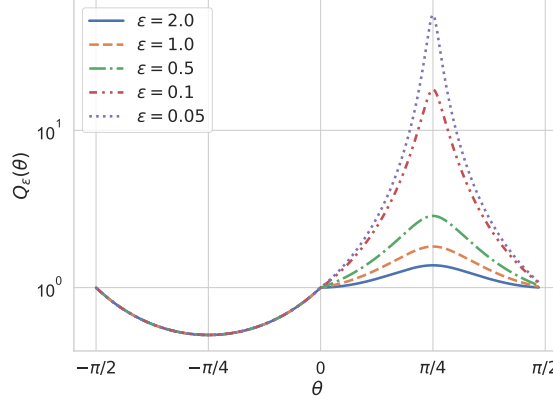


Figure 3.4: Illustration of symmetries in θ of the map $Q_\varepsilon(\theta)$ for several $\varepsilon > 0$.

in $[-\pi/2, 0]$ and symmetric around $\pi/4$ in $[0, \pi/2]$. This fact is also shown in Fig. 3.4. That is, it is enough to investigate angles $\theta \in [-\pi/4, \pi/4]$.

For the computation of integrals in $Q_\varepsilon(\theta)$, $\theta \in [-\pi/4, \pi/4]$, it is necessary, for a given y , to determine boundaries $\ell_0(y)$ and $\ell_1(y)$ of intervals for z that lie in the support of the joint density $\rho_{Y,Z}(y, z)$ (see the thick solid lines in Fig. 3.3). The integrals in $Q_\varepsilon(\theta)$ are computed using the computer algebra system *Wolfram Mathematica* [285]. The computation requires to treat the cases $\theta \in [-\pi/4, 0]$ and $\theta \in [0, \pi/4]$ differently (see Fig. 3.3).

For negative $\theta \in [-\pi/4, 0]$ and arbitrary $y \in \mathbf{R}$, we have that

$$\ell_0(y) = \begin{cases} |y| \cot(|\theta|) & \text{if } y < 0 \\ y \tan(|\theta|) & \text{if } y \geq 0 \end{cases} = |y| \tan(|\theta|)^{\text{sgn}(y)} \quad (3.95)$$

and $\ell_1(y) = \infty$, i. e.,

$$\rho_{Y,Z}(y, z) = \exp(-a_\theta^+ y - a_\theta^- z) \cdot \mathbf{1}_{[\ell_0(y), \ell_1(y)]}(z). \quad (3.96)$$

We compute that

$$\text{Var}(Z | Y = y) = (\cos(|\theta|) + \sin(|\theta|))^{-2} \quad (3.97)$$

which is constant in y and yields

$$Q_\varepsilon(\theta) = C_{\text{Var}, R_\theta}(\varepsilon, 2, 1) = (\cos(|\theta|) + \sin(|\theta|))^{-2}. \quad (3.98)$$

Note that this explains the left part of the graph of $Q_\varepsilon(\theta)$ in Fig. 3.4 which shows that $Q_\varepsilon(\theta)$ does not depend on ε for $\theta \in [-\pi/2, 0]$.

For non-negative $\theta \in [0, \pi/4]$ and a given $y \geq 0$, the boundaries are computed to $\ell_0(y) = -y \tan(\theta)$ and $\ell_1(y) = y \cot(\theta)$, i. e.,

$$\rho_{Y,Z}(y, z) = \exp(-a_\theta^+ y - a_\theta^- z) \cdot \mathbf{1}_{[0, \infty)}(y) \cdot \mathbf{1}_{[\ell_0(y), \ell_1(y)]}(z). \quad (3.99)$$

We compute that

$$\text{Var}(Z | Y = y)$$

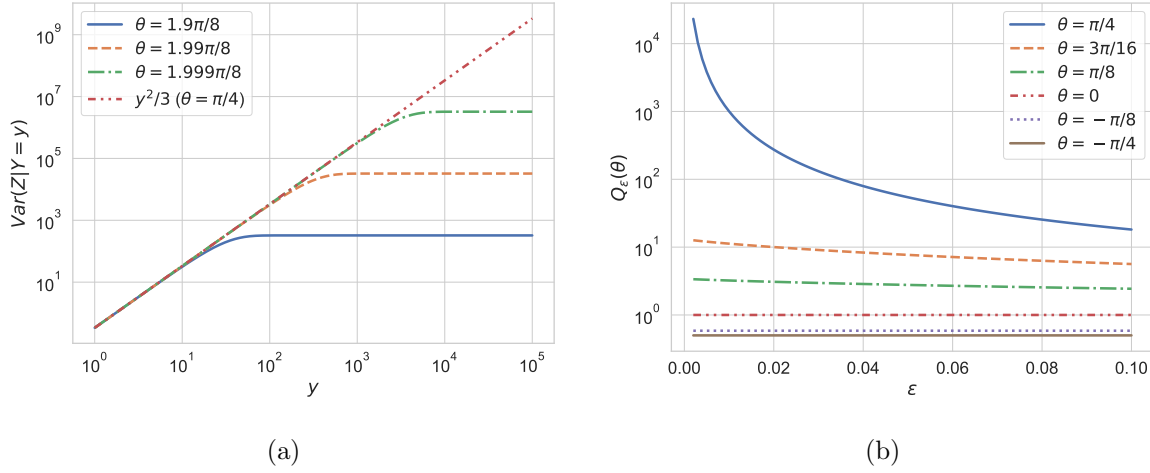


Figure 3.5: (a) The log-log plot of the map $y \mapsto \mathbf{Var}(Z|Y=y)$ shows that it is bounded for angles $\theta \in [0, \pi/4)$, but approaching the unbounded function $y^2/3$, which corresponds to $\theta = \pi/4$, as $\theta \rightarrow \pi/4$.

(b) The plot shows the map $\varepsilon \mapsto Q_\varepsilon(\theta)$ for several angles θ . Also, it illustrates the fact that $\theta = \pi/4$ is a special case for which $Q_\varepsilon(\theta)$ can get arbitrarily large.

$$= \frac{a_\theta}{8b_\theta^2} \left(\frac{1 - 2\exp(b_\theta y) + \exp(2b_\theta y) - 8\exp(b_\theta y)y^2(1 - d_\theta)}{(\exp(b_\theta y) - 1)^2} - c_\theta \right) \quad (3.100)$$

for $a_\theta := \csc(\theta)^4 \sec(\theta)^4$, $b_\theta := \sec(\theta) - \csc(\theta)$, $c_\theta := \cos(4\theta)$, and $d_\theta := \sin(2\theta)$. $\mathbf{Var}(Z|Y=y)$ can actually be bounded in y for $\theta \in [0, \pi/4)$. Indeed, since $d_\theta \in [0, 1)$, it holds that $1 - d_\theta \in (0, 1]$ implying that $8\exp(b_\theta y)y^2(1 - d_\theta) > 0$. It follows that

$$\mathbf{Var}(Z|Y=y) \leq \frac{a_\theta}{8b_\theta^2} \left(\frac{1 - 2\exp(b_\theta y) + \exp(2b_\theta y)}{(\exp(b_\theta y) - 1)^2} - c_\theta \right) \quad (3.101)$$

$$= \frac{a_\theta}{8b_\theta^2} \left(\frac{(\exp(b_\theta y) - 1)^2}{(\exp(b_\theta y) - 1)^2} - c_\theta \right) \quad (3.102)$$

$$= \frac{a_\theta(1 - c_\theta)}{8b_\theta^2}. \quad (3.103)$$

Fig. 3.5a illustrates the boundedness of $\mathbf{Var}(Z|Y=y)$ and additionally shows that it approaches the unbounded function $y \mapsto y^2/3$ as $\theta \rightarrow \pi/4$. Hence, for $\theta \in [0, \pi/4)$, it holds that

$$Q_\varepsilon(\theta) = C_{\mathbf{Var}, R_\theta}(\varepsilon, 2, 1) \leq \frac{a_\theta(1 - c_\theta)}{8b_\theta^2}. \quad (3.104)$$

This bound is itself unbounded in θ since $b_\theta \rightarrow 0$ and $a_\theta(1 - c_\theta) \rightarrow 32$ as $\theta \rightarrow \pi/4$ implying that we can see $\theta = \pi/4$ as a special case. This assessment is also supported by Fig. 3.5b. In particular, note that

$$C_{\mathbf{Var}} = C_{\mathbf{Var}, R_{\pi/4}}. \quad (3.105)$$

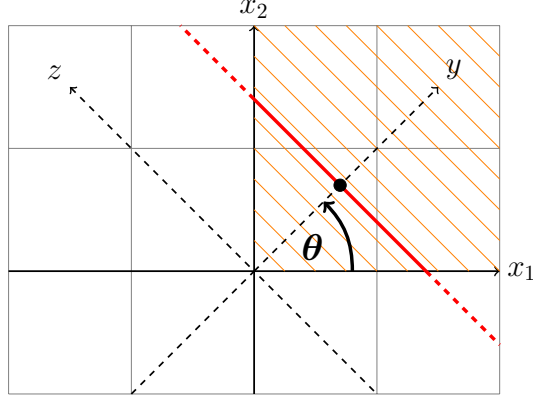


Figure 3.6: Exponential distribution in 2 dimensions with a coordinate system rotated by 45° . The orange lines depict the contour levels of the distribution in the support of $\rho_{\mathbf{X}}$. The solid red line marks the interval of the uniform distribution of $Z | Y = y$ for $y > 0$.

Rotation by $\theta = \pi/4$

A rotation of 45° , i.e., $\theta = \pi/4$ and $W = R_{\pi/4}$, is a limit case since $a_{\pi/4}^-$ from (3.93) becomes zero. The joint density for Y and Z is then

$$\rho_{Y,Z}(y, z) = \exp(-\sqrt{2}y) \cdot \mathbf{1}_{[0, \infty)}(y) \cdot \mathbf{1}_{[-y, y]}(z). \quad (3.106)$$

A graphical illustration of this case is given in Fig. 3.6. Consequently, the marginal distribution of Y is

$$\rho_Y(y) = \int_{-\infty}^{\infty} \rho_{Y,Z}(y, z) dz = 2y \exp(-\sqrt{2}y) \cdot \mathbf{1}_{[0, \infty)}(y) \quad (3.107)$$

and the conditional density $\rho_{Z|Y}(\cdot|y)$ computes to

$$\rho_{Z|Y}(z|y) = \frac{\mathbf{1}_{[-y, y]}(z)}{2y} \quad (3.108)$$

for $y > 0$. Note that $\rho_{Z|Y}(\cdot|y)$ is the density of a uniform distribution on the interval $[-y, y]$.

For $y > 0$, it follows that

$$\mathbf{Var}(Z | Y = y) = (2y)^{-1} \int_{-y}^y z^2 dz = y^2/3, \quad (3.109)$$

which is the expression that variances of $Z | Y = y$ for other angles θ^* approach to as $\theta^* \rightarrow \pi/4$ (see Fig. 3.5a).

Note that, since $|Z| | Y = y \sim \mathcal{U}([0, y])$, the lower bound from (3.73) for C_y in this case becomes

$$\mathbf{E}[(|Z| - \mathbf{E}[|Z| | Y = y])^2 | Y = y] = \mathbf{Var}(|Z| | Y = y) = y^2/12, \quad (3.110)$$

which is unbounded in y implying that the distribution of C_Y is not compactly supported. Therefore, we found a scenario in which the constants $C_{P,W}$ and C_P indeed do not exist.

However, there is still a chance that the constants $C_{P,\varepsilon,W}$ and $C_{P,\varepsilon}$ from (3.60) and, respectively, (3.79) exist. It holds that

$$C_{\mathbf{Var}}(\varepsilon, 2, 1) = C_{\mathbf{Var}, R_{\pi/4}}(\varepsilon, 2, 1) = \frac{1}{3} \mathbf{E}[Y^{2+2/\varepsilon}]^{\varepsilon/(1+\varepsilon)} \quad (3.111)$$

implying that the constant $C_{P,\varepsilon}(\varepsilon, 2, 1)$ can be bounded from above by

$$C_{P,\varepsilon}(\varepsilon, 2, 1) \leq L^{\varepsilon/(1+\varepsilon)} \frac{K}{3} \mathbf{E}[Y^{2+2/\varepsilon}]^{\varepsilon/(1+\varepsilon)}. \quad (3.112)$$

For example, choosing $\varepsilon = 2$ would give

$$C_{P,\varepsilon}(2, 2, 1) \leq 2K \left(\frac{L^2}{3} \right)^{1/3}. \quad (3.113)$$

n dimensions

This subsection aims to generalize the results of the previous subsection, i.e., we investigate the constant $C_{P,\varepsilon}$ from (3.79) for n independent exponential distributions.

Motivated by the two-dimensional case, we regard the rotation of the coordinate system by a matrix $W = R^*$ that rotates the vector $(1, 0, \dots, 0)^\top \in \mathbf{R}^n$ to $(1/\sqrt{n}, \dots, 1/\sqrt{n})^\top \in \mathbf{R}^n$. Note that in the two-dimensional case, a rotation by $\theta = \pi/4$ corresponds to a matrix rotating $(1, 0)^\top$ to $(1/\sqrt{2}, 1/\sqrt{2})^\top$. This is the worst case in the sense that $Z_i | \mathbf{Y} = \mathbf{y}$ is uniformly distributed for each component Z_i in $\mathbf{Z} = (Z_1, \dots, Z_{n-k})^\top$ and hence, similar to the two-dimensional case, the conditional variance of $Z_i | \mathbf{Y} = \mathbf{y}$ has no finite essential supremum. In the context from above, it holds that

$$C_{\mathbf{V}\text{ar}}(\varepsilon, n, k) = C_{\mathbf{V}\text{ar}, R^*}(\varepsilon, n, k). \quad (3.114)$$

The following theorem studies this case and investigates the dimensional dependence of the involved constant.

Theorem 3.4.5. *For $\rho_{\mathbf{X}}$ as in (3.83), it holds that*

$$\mathbf{E}[(f(\mathbf{X}) - f_g(\mathbf{X}))^2] \leq C_{\exp^n} (\lambda_{k+1} + \dots + \lambda_n)^{1/(1+\varepsilon)} \quad (3.115)$$

for a constant

$$C_{\exp^n} = C_{\exp^n}(\varepsilon, n, k, L, \mathbf{P}_{\mathbf{X}}) \geq C_{P,\varepsilon} \quad (3.116)$$

Proof. In the support of $\rho_{\mathbf{X}}$, i.e., in $\mathcal{X} = \mathbf{R}_{\geq 0}^n$, $\rho_{\mathbf{X}}$ is greater than zero and constant on the intersection of $\mathbf{R}_{\geq 0}^n$ and planes

$$P_a := \{\mathbf{x} \mid x_1 + \dots + x_n = a\} = \{\mathbf{x} \mid (1, \dots, 1)^\top \mathbf{x} = a\} \subset \mathbf{R}^n, \quad a > 0, \quad (3.117)$$

i.e., on hypersurfaces $T_a := P_a \cap \mathbf{R}_{\geq 0}^n$. The situation is illustrated by Fig. 3.7 for $n = 3$ dimensions.

For $\mathbf{x} = \llbracket \mathbf{y}, \mathbf{z} \rrbracket \in \mathbf{R}_{\geq 0}^n$, the value of $\rho_{\mathbf{Y}, \mathbf{Z}}(\mathbf{y}, \mathbf{z})$ is only determined by $y_1 \geq 0$. Reversely, if $y_1 < 0$, then $\rho_{\mathbf{Y}, \mathbf{Z}}(\mathbf{y}, \mathbf{z}) = 0$. We know that the point at $\mathbf{x}_0 := \beta(1, \dots, 1)^\top \in \mathbf{R}^n$ with $\|\mathbf{x}_0\|_2 = y_1$ is supposed to lie on P_a for some $\beta > 0$. It follows immediately that $\beta = y_1/\sqrt{n}$. Also, we determine a with

$$a = (1, \dots, 1)^\top \mathbf{x}_0 = \frac{y_1}{\sqrt{n}} n = \sqrt{n} y_1. \quad (3.118)$$

Let us define $T(y_1) := T_{\sqrt{n}y_1}$. That is,

$$\rho_{\mathbf{Y}, \mathbf{Z}}(\mathbf{y}, \mathbf{z}) = \exp(-\sqrt{n}y_1) \cdot \mathbf{1}_{[0, \infty)}(y_1) \cdot \mathbf{1}_{T(y_1)}(\mathbf{y}, \mathbf{z}). \quad (3.119)$$

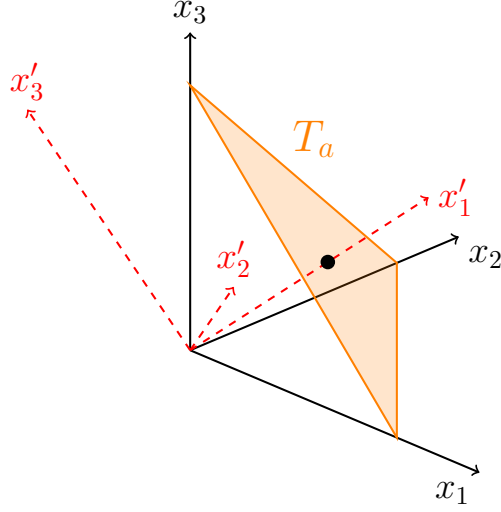


Figure 3.7: Exponential distribution in 3D with a rotated coordinate system.

$T(y_1)$, as a geometric figure, is a *regular* $(n-1)$ -simplex in n dimensions. $T(y_1)$ is intrinsically $(n-1)$ -dimensional and has n corners which are

$$(\sqrt{n}y_1, 0, \dots, 0), \dots, (0, \dots, 0, \sqrt{n}y_1) \in \mathbf{R}^n. \quad (3.120)$$

It follows that the side length of $T(y_1)$ is $\sqrt{2n}y_1$. Note that the coordinates $\tilde{\mathbf{y}} = (y_2, \dots, y_k)^\top$ and $\mathbf{z} = (z_1, \dots, z_{n-k})^\top$ all move on $T(y_1)$.

We can rewrite $T(y_1)$ as

$$T(y_1) = \{\mathbf{x} \in \mathbf{R}_{\geq 0}^n \mid (W^\top \mathbf{x})_1 = y_1\} \quad (3.121)$$

$$= \{\llbracket \tilde{\mathbf{y}}, \tilde{\mathbf{z}} \rrbracket \mid \llbracket \tilde{\mathbf{y}}, \tilde{\mathbf{z}} \rrbracket \in \mathbf{R}_{\geq 0}^n, \tilde{y}_1 = y_1\}. \quad (3.122)$$

This motivates to view $T(y_1)$ as an $(n-1)$ -dimensional set in the rotated coordinate system, i.e., we define

$$\tilde{T}(y_1) := \{(\tilde{\mathbf{y}}, \mathbf{z}) \in \mathbf{R}^{k-1} \times \mathbf{R}^{n-k} \mid \llbracket (y_1, \tilde{\mathbf{y}}), \mathbf{z} \rrbracket \in T(y_1)\} \subset \mathbf{R}^{n-1}. \quad (3.123)$$

We observe that the conditioned random variable $(\tilde{\mathbf{Y}}, \mathbf{Z}) \mid Y_1 = y_1$ is uniformly distributed on the regular $(n-1)$ -simplex $\tilde{T}(y_1)$.

The basic idea to get a bound for $\mathbf{E}[\mathbf{Var}(Z_i | \mathbf{Y})^{(1+\varepsilon)/\varepsilon}]$ is based on the fact that z_i , moving as the $(k+i-1)$ -th coordinate inside $\tilde{T}(y_1)$, takes values in $[0, h_i(y_1)]$, where $h_i(y_1)$ is the height of a regular $(k+i-1)$ -simplex with side length $\sqrt{2n}y_1$ and is thus bounded. In general, the height of a regular n -simplex is the distance of a vertex to the circumcentre of its opposite regular $(n-1)$ -simplex. By [41, p. 367], it holds that

$$h_i(y_1) = \sqrt{\frac{n(k+i)}{k+i-1}} y_1. \quad (3.124)$$

We start the computation by noting that

$$\rho_{\mathbf{Y}, \mathbf{Z}}(\mathbf{y}, \mathbf{z}) = \exp(-\sqrt{n}y_1) \cdot \mathbf{1}_{[0, \infty)}(y_1) \cdot \mathbf{1}_{\tilde{T}(y_1)}(\tilde{\mathbf{y}}, \mathbf{z}). \quad (3.125)$$

The marginal distribution of $Z_i | \mathbf{Y} = \mathbf{y}$ is given by

$$\rho_{Z_i|\mathbf{Y}}(z_i|\mathbf{y}) = \int \cdots \int \rho_{\mathbf{Z}|\mathbf{Y}}(\mathbf{z}|\mathbf{y}) dz_1 \dots dz_{i-1} dz_{i+1} \dots dz_{n-k}. \quad (3.126)$$

and so we get

$$\rho_{\mathbf{Y}}(\mathbf{y}) \rho_{Z_i|\mathbf{Y}}(z_i|\mathbf{y}) \quad (3.127)$$

$$= \int \cdots \int \rho_{\mathbf{Y},\mathbf{Z}}(\mathbf{y}, \mathbf{z}) dz_1 \dots dz_{i-1} dz_{i+1} \dots dz_{n-k} \quad (3.128)$$

$$= \exp(-\sqrt{n}y_1) \cdot \mathbf{1}_{[0,\infty)}(y_1) \quad (3.129)$$

$$\cdot \int \cdots \int \mathbf{1}_{T(y_1)}(\check{\mathbf{y}}, \mathbf{z}) dz_1 \dots dz_{i-1} dz_{i+1} \dots dz_{n-k}. \quad (3.130)$$

Using Jensen's inequality in a first step, we can continue with

$$\mathbf{E}[\mathbf{Var}(Z_i|\mathbf{Y})^{(1+\varepsilon)/\varepsilon}] \leq \mathbf{E}[\mathbf{E}[Z_i^{2(1+\varepsilon)/\varepsilon} | \mathbf{Y}]] \quad (3.131)$$

$$= \int \left(\int z_i^{2(1+\varepsilon)/\varepsilon} \rho_{Z_i|\mathbf{Y}}(z_i|\mathbf{y}) dz_i \right) \rho_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} \quad (3.132)$$

$$= \int_0^\infty \exp(-\sqrt{n}y_1) \left(\int \int z_i^{2(1+\varepsilon)/\varepsilon} \cdot \mathbf{1}_{\check{T}(y_1)}(\check{\mathbf{y}}, \mathbf{z}) d\mathbf{z} d\check{\mathbf{y}} \right) dy_1 \quad (3.133)$$

$$\leq \int_0^\infty \exp(-\sqrt{n}y_1) h_i(y_1)^{2(1+\varepsilon)/\varepsilon} \left(\int \int \mathbf{1}_{\check{T}(y_1)}(\check{\mathbf{y}}, \mathbf{z}) d\mathbf{z} d\check{\mathbf{y}} \right) dy_1 \quad (3.134)$$

$$= \int_0^\infty \exp(-\sqrt{n}y_1) \left(\sqrt{\frac{n(k+i)}{k+i-1}} y_1 \right)^{2(1+\varepsilon)/\varepsilon} \frac{\sqrt{n^n}}{(n-1)!} y_1^{n-1} dy_1 \quad (3.135)$$

$$= \left(\frac{n(k+i)}{k+i-1} \right)^{(1+\varepsilon)/\varepsilon} \frac{\sqrt{n^n}}{(n-1)!} \int_0^\infty y_1^{n+1+2/\varepsilon} \exp(-\sqrt{n}y_1) dy_1 \quad (3.136)$$

$$= \left(\frac{n(k+i)}{k+i-1} \right)^{(1+\varepsilon)/\varepsilon} \frac{\sqrt{n^n}}{(n-1)!} \frac{\Gamma(n+2+2/\varepsilon)}{n^{(1+\varepsilon)/\varepsilon} \sqrt{n^n}} \quad (3.137)$$

$$= \left(\frac{k+i}{k+i-1} \right)^{(1+\varepsilon)/\varepsilon} \frac{\Gamma(n+2+2/\varepsilon)}{(n-1)!}. \quad (3.138)$$

Note that an intermediate step of the previous calculation uses the fact that the volume of the regular $(n-1)$ -simplex $\check{T}(y_1)$ with side length $\sqrt{2n}y_1$ is (see [41, p. 367])

$$\int \int \mathbf{1}_{\check{T}(y_1)}(\check{\mathbf{y}}, \mathbf{z}) d\mathbf{z} d\check{\mathbf{y}} = \frac{\sqrt{n^n}}{(n-1)!} y_1^{n-1}. \quad (3.139)$$

Remember from (3.74) and (3.75) that

$$\mathbf{E}[C_{\mathbf{Y}}^{(1+\varepsilon)/\varepsilon}]^{\varepsilon/(1+\varepsilon)} \leq K(n-k)^{1/(1+\varepsilon)} C_{\mathbf{Var}}(\varepsilon, n, k) \quad (3.140)$$

with

$$C_{\mathbf{Var}}(\varepsilon, n, k) = \left(\sum_{i=1}^{n-k} \mathbf{E}[\mathbf{Var}(Z_i|\mathbf{Y})^{(1+\varepsilon)/\varepsilon}] \right)^{\varepsilon/(1+\varepsilon)} \quad (3.141)$$

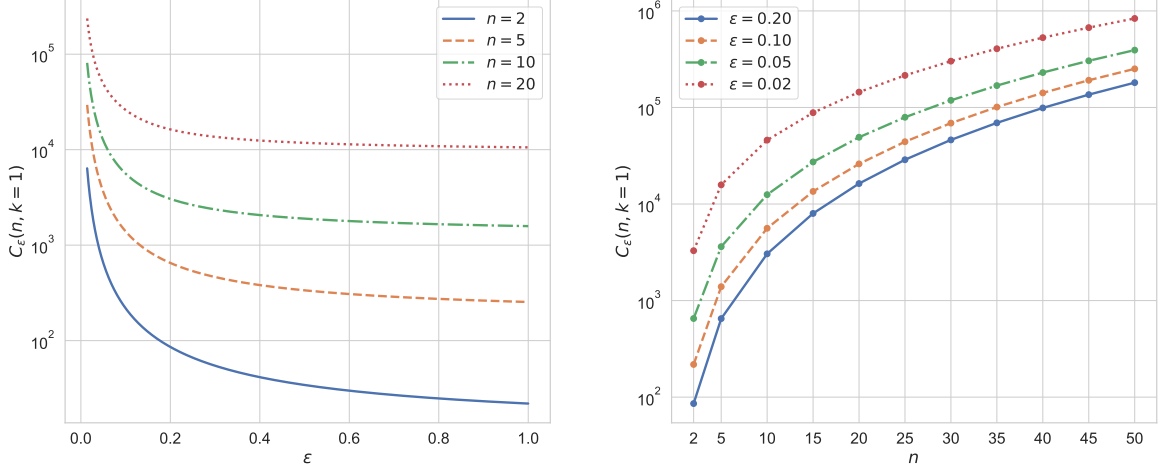


Figure 3.8: The left plot shows curves of the map $\varepsilon \mapsto C_\varepsilon(n, k = 1)$ for $n \in \{2, 5, 10, 20\}$. Curves of $n \mapsto C_\varepsilon(n, k = 1)$ for $\varepsilon \in \{0.02, 0.05, 0.1, 0.2\}$ are displayed on the right.

$$\leq \left(\frac{\Gamma(n+2+2/\varepsilon)}{(n-1)!} \sum_{i=1}^{n-k} \left(\frac{k+i}{k+i-1} \right)^{(1+\varepsilon)/\varepsilon} \right)^{\varepsilon/(1+\varepsilon)}. \quad (3.142)$$

Defining

$$C_\varepsilon(n, k) := (n-k)^{1/(1+\varepsilon)} \left(\frac{\Gamma(n+2+2/\varepsilon)}{(n-1)!} \sum_{i=1}^{n-k} \left(\frac{k+i}{k+i-1} \right)^{(1+\varepsilon)/\varepsilon} \right)^{\varepsilon/(1+\varepsilon)} \quad (3.143)$$

then yields

$$\mathbf{E}[C_{\mathbf{Y}}^{(1+\varepsilon)/\varepsilon}]^{\varepsilon/(1+\varepsilon)} \leq K C_\varepsilon(n, k). \quad (3.144)$$

Combining all bounds, we get that

$$C_{P,\varepsilon}(\varepsilon, n, k) \leq K \cdot L^{\varepsilon/(1+\varepsilon)} \cdot C_\varepsilon(n, k) =: C_{\exp^n}(\varepsilon, n, k, L), \quad (3.145)$$

where $C_{P,\varepsilon}(\varepsilon, n, k)$ was defined in (3.79). We recall that n denotes the dimension of the problem, k the dimension of the active subspace, L is the upper bound on $\|\nabla f\|_2^2$, and K the universal constant from (3.74).

The result follows by Lem. 3.4.3. ■

Fig. 3.8 depicts the quantity $C_\varepsilon(n, k = 1)$ from (3.143) as a function of $\varepsilon > 0$ for some $n \in \mathbf{N}$ (left plot) and as a function of $n \geq 2$ for several $\varepsilon > 0$ (right plot). We set $k = 1$ since this gives the maximum value for C_ε over all $k \geq 1$. As expected, the curves increase quickly as ε approaches zero or, respectively, n becomes large.

Remark. In the previous theorem, the exponential distributions are assumed to have unit rates. The computations can also be made for arbitrary rates ν_i , $i = 1, \dots, n$. However, some modifications are necessary.

Let $\boldsymbol{\nu} = (\nu_1, \dots, \nu_n)^\top$ denote the vector of rates. To get again the worst case scenario as in the previous subsection (uniform distribution on a simplex structure), the coordinate system

has to be rotated in such a way that the vector $(1, 0, \dots, 0)^\top$ rotates to $\boldsymbol{\nu}/\|\boldsymbol{\nu}\|_2$. The structure of a regular simplex that is used in the estimates above is not present in this more general case. Instead, we get a general simplex whose heights are not as easy to compute as in the regular case. However, rough estimates can be achieved by enclosing the general simplex with a larger regular one.

3.4.4 Future work with MGH distributions

The generalized bound from Lem. 3.4.3 and the study of corresponding Poincaré type constants $C_{P,\varepsilon,W}$ and $C_{P,\varepsilon}$ for independent exponential distributions in Subsec. 3.4.3 motivate further similar investigations of more general distributions.

From a statistical perspective, a study of the class of *multivariate generalized hyperbolic distributions* (MGH) (see, e.g., [17]) can be considered as a next step since it allows for distributions with both non-zero skewness and heavier tails.

An MGH is a distribution of the random vector

$$\mathbf{X} = \boldsymbol{\mu} + \boldsymbol{\beta}A + \sqrt{A}\mathbf{M}\mathbf{V} \quad (3.146)$$

with location parameter $\boldsymbol{\mu} \in \mathbf{R}^n$, skewness parameter $\boldsymbol{\beta} \in \mathbf{R}^n$, and a symmetric positive definite matrix $\mathbf{M} \in \mathbf{R}^{n \times n}$. The scalar random variable A , called the mixing variable, follows a generalized inverse Gaussian distribution (GIG) [140], and $\mathbf{V} \sim \mathcal{N}(0, I)$ is independent of A . As a particular example, for \mathbf{X} to be Laplace distributed, we set $\boldsymbol{\beta} = 0$ and let A be exponentially distributed [157]. Note that, however, the example from Subsec. 3.4.3, assuming independent exponential distributions, is not an MGH. In order to include this case, we would have to extend A to a random positive definite matrix.

Nevertheless, MGH is a large class containing classical distributions like the normal-inverse Gaussian, generalized Laplace, and Student's t-distribution. In particular, these distributions are interesting since they have been used in areas like, for instance, economics and financial markets [18, 19, 85], spatial and Geostatistics [35, 36, 278], and linear mixed-effects [13, 182, 294] which are used, e.g., for linear non-Gaussian time series models in medical longitudinal studies [13].

We mention that, under an assumption on a parameter, MGH distributions are log-concave [290], i.e., we can use the bounds on Poincaré constants $C_{\mathbf{Y}}$ of Bobkov from (3.73).

In our opinion, it is preferable to start the investigation with the subclass of symmetric MGH distributions, i.e., $\boldsymbol{\beta} = 0$ in (3.146). The following lines demonstrate particular difficulties that we already encounter in this smaller subclass. Let us choose $\boldsymbol{\mu} = \boldsymbol{\beta} = 0$ and $\mathbf{M} = I$ in (3.146) such that

$$\mathbf{X} = \sqrt{A}\mathbf{V} \quad (3.147)$$

with $\mathbf{V} \sim \mathcal{N}(0, I)$. A common first step is to study \mathbf{X} conditioned on $A = a$, $a > 0$, i.e., $\mathbf{X} | A = a \sim \mathcal{N}(0, aI)$, and to use the tower property of conditional expectations. That is, analogously to (3.6), we define

$$C := \mathbf{E}[C_A] = W\Lambda W^\top \quad (3.148)$$

with

$$C_a := \mathbf{E}[\nabla f(\mathbf{X})\nabla f(\mathbf{X})^\top | A = a] = W_a\Lambda_a W_a^\top. \quad (3.149)$$

Choosing $k \leq n - 1$ independent of A , we further set

$$\mathbf{Y}_a := W_{a,1}^\top \mathbf{X} \quad \text{and} \quad \mathbf{Z}_a := W_{a,2}^\top \mathbf{X}. \quad (3.150)$$

The computation starts, similar to (3.36), with

$$\mathbf{E}[(f(\mathbf{X}) - f_g(\mathbf{X}))^2 \mid A = a] \quad (3.151)$$

$$= \mathbf{E}[\mathbf{E}[(f(\llbracket \mathbf{Y}_a, \mathbf{Z}_a \rrbracket_{W_a}) - g(\mathbf{Y}_a))^2 \mid \mathbf{Y}_a] \mid A = a] \quad (3.152)$$

$$\leq \mathbf{E}[C_{\mathbf{Y}_a} \mathbf{E}[\|\nabla^{z_a} f(\llbracket \mathbf{Y}_a, \mathbf{Z}_a \rrbracket_{W_a})\|_2^2 \mid \mathbf{Y}_a] \mid A = a] \quad (3.153)$$

$$= a \mathbf{E}[\mathbf{E}[\|\nabla^{z_a} f(\llbracket \mathbf{Y}_a, \mathbf{Z}_a \rrbracket_{W_a})\|_2^2 \mid \mathbf{Y}_a] \mid A = a] \quad (3.154)$$

$$= a \operatorname{trace}(\Lambda_{a,2}). \quad (3.155)$$

In (3.154), we use the fact that the Poincaré constant of a normal distribution $\mathcal{N}(0, aI)$ is $\lambda_{\max}(aI) = a$; see Sec. 3.3. The last step to (3.155) follows from Lem. 3.2.1. This yields

$$\mathbf{E}[(f(\mathbf{X}) - f_g(\mathbf{X}))^2] = \mathbf{E}[\mathbf{E}[(f(\mathbf{X}) - f_g(\mathbf{X}))^2 \mid A]] \quad (3.156)$$

$$\leq \mathbf{E}[A \cdot \operatorname{trace}(\Lambda_{A,2})], \quad (3.157)$$

where the random variable $A \cdot \operatorname{trace}(\Lambda_{A,2})$ is assumed to have finite first moment.

At this point, as long as A is not compactly supported, we can only continue by applying another Hölder's inequality similar to the proof of Lem. 3.4.3. However, in any case, we have to face the problem that $\mathbf{E}[\operatorname{trace}(\Lambda_{A,2})]$ is, in general, not equal to $\operatorname{trace}(\Lambda_2)$ which is the sum of eigenvalues belonging to the inactive subspace that we actually aim for. Nevertheless, we know that

$$\mathbf{E}[\operatorname{trace}(\Lambda_A)] = \operatorname{trace}(\Lambda), \quad (3.158)$$

but it is unclear whether, and how, this equality can be exploited for our purposes.

3.5 Practical considerations

In contrast to the last two sections, which treat rather theoretical aspects of ASM, this section considers ASM from a practical perspective. We need the results described here for the following chapters, where we apply ASM in the context of Bayesian inverse problems involving high-dimensional complex physical models. The general procedure of applying ASM in Bayesian inversion is described in Ch. 4.

Here, we consider two kinds of issues that were neglected so far but frequently arise when using ASM in practice. It is the approximation/estimation of

1. the auto-correlation matrix C from (3.6) determining the active and inactive subspace,
2. the ridge profile g from (3.23) for the approximation of f on a low-dimensional space.

Our exposition is based on the contents in [52].

Estimation of C

The matrix C from (3.6) is estimated by a finite Monte Carlo sum, i. e.,

$$C \approx \tilde{C} := \frac{1}{N_{\tilde{C}}} \sum_{j=1}^{N_{\tilde{C}}} \nabla f(\mathbf{X}_j) \nabla f(\mathbf{X}_j)^\top, \quad (3.159)$$

where $N_{\tilde{C}} > 0$ and $\mathbf{X}_j \stackrel{\text{i.i.d.}}{\sim} \mathbf{P}_{\mathbf{X}}$, $j = 1, \dots, N_{\tilde{C}}$. Analogously to the steps from above, we get the eigendecomposition

$$\tilde{C} = \tilde{W} \tilde{\Lambda} \tilde{W}^\top \quad (3.160)$$

with eigenvectors $\tilde{\mathbf{w}}_i \in \mathbf{R}^n$ and eigenvalues $\tilde{\lambda}_i \geq 0$, $i = 1, \dots, n$.

A natural immediate question from practitioners is on the size of $N_{\tilde{C}}$, i. e., how many gradient samples do we need to approximate C sufficiently accurate? And what does “sufficiently accurate” mean?

First, we should note that we actually do *not* want to approximate C itself but rather its eigenvalues and eigenvectors. The following paragraphs first present corresponding results for the approximation of eigenvalues and subsequently results for eigenvectors, both under the following assumption.

Assumption 3.5.1. The spectral norm of the gradient of f is uniformly bounded from above, i. e.,

$$\|\nabla^{\mathbf{x}} f(\mathbf{x})\|_2 \leq L \quad (3.161)$$

for some $L > 0$ and all $\mathbf{x} \in \mathcal{X}$.

The following theorem gives a probabilistic bound for the approximation quality of the perturbed eigenvalues $\tilde{\lambda}_i$.

Theorem 3.5.2 ([52, Thm. 3.3]). *For $\varepsilon \in (0, 1]$, it holds that*

$$\mathbf{P} \left(\tilde{\lambda}_\ell \geq (1 + \varepsilon) \lambda_\ell \right) \leq (n - \ell + 1) \exp \left(- \frac{N_{\tilde{C}} \lambda_\ell \varepsilon^2}{4L^2} \right) \quad (3.162)$$

and

$$\mathbf{P} \left(\tilde{\lambda}_\ell \leq (1 - \varepsilon) \lambda_\ell \right) \leq \ell \exp \left(- \frac{N_{\tilde{C}} \lambda_\ell^2 \varepsilon^2}{4\lambda_1 L^2} \right) \quad (3.163)$$

for $\ell = 1, \dots, n$.

The proof for this theorem uses an *eigenvalue Bernstein inequality for subexponential matrices* which is a concentration inequality for eigenvalues of a finite sum of random, independent, and symmetric matrices satisfying a subexponential growth condition; see Thm. A.2.15.

The appearance of $N_{\tilde{C}}$ in the negative exponential expressions in (3.162) and (3.163) is the key to provide evidence for the practical usability of ASM. Gradient samples are generally considered to be computationally expensive, but, actually, we do not need to compute many of them as the following corollary shows.

Corollary 3.5.3 ([52, Corollary 3.5]). *Let $\chi_\ell := \lambda_1 / \lambda_\ell$. Then, for $\varepsilon \in (0, 1]$, using*

$$N_{\tilde{C}} = \Omega \left(\frac{L^2 \chi_\ell^2}{\lambda_1 \varepsilon^2} \log(n) \right) \quad (3.164)$$

samples in (3.159) implies that

$$|\tilde{\lambda}_\ell - \lambda_\ell| \leq \varepsilon \lambda_\ell \quad (3.165)$$

with high probability.

Note that the Ω notation means that $N_{\tilde{C}}$ can be bounded from below by the expression in the parentheses multiplied by a constant; see Def. A.1.1. The logarithm in the problem dimension n opens the possibility to approximate the eigenvalues “sufficiently accurate” without an enormous amount of computational costs.

As a practical recipe, the author of [52] suggests to choose

$$N_{\tilde{C}} = \beta_{\tilde{C}} m \log(n) \quad (3.166)$$

gradient samples for the accurate estimation of the first m eigenvalues of C , where $\beta_{\tilde{C}} > 0$ represents the unknown constants in the bounds from above and is called a *sampling factor*. It is recommended to choose $\beta_{\tilde{C}}$ between 2 and 10.

Of course, the quality of the estimated eigenvalues $\tilde{\lambda}_i$ needs to be explicitly examined in practice, i.e., we have to check if the number of samples $N_{\tilde{C}}$ from (3.159) is large enough to allow for an accurate estimate. This can be done by a *bootstrap* approach [86, 167], where an ensemble of gradient samples itself is used to quantify the variability in the corresponding estimated eigenvalues $\tilde{\lambda}_i$, i.e., no further gradients need to be computed.

In the ℓ -th iteration of the bootstrap algorithm, a predefined number of samples from the ensemble is randomly picked, the matrix $\tilde{C}^{(\ell)}$ is built with the picked samples according to (3.159), and the corresponding eigenvalues $\tilde{\lambda}_i^{(\ell)}$ are computed. This process is done many times such that we end up with an ensemble of eigenvalues $\tilde{\lambda}_i^{(\ell)}$, $\ell = 1, \dots, n$, for every $i = 1, \dots, n$. Plotting the minimum and maximum value for each i , i.e.,

$$\tilde{\lambda}_{i,\min} := \min_{\ell} \tilde{\lambda}_i^{(\ell)} \quad \text{and} \quad \tilde{\lambda}_{i,\max} := \max_{\ell} \tilde{\lambda}_i^{(\ell)} \quad (3.167)$$

on a logarithmic scale then provides a notion of the variability in each eigenvalue.

Fig. 3.9 shows the resulting plot for a synthetic toy example with $f(\mathbf{x}) := \frac{1}{2} \mathbf{x}^\top A \mathbf{x}$ for a given matrix $A \in \mathbf{R}^{20 \times 20}$ with predefined eigenvalues and $\mathbf{P}_X := \mathcal{N}(0, I)$. It illustrates two cases. The left plot depicts a scenario in which we take a small number of gradient samples $N_{\tilde{C}} = 10$ resulting in a larger variability in the estimated eigenvalues. However, a larger number of samples $N_{\tilde{C}} = 100$ allows for a more accurate estimation, i.e., the variability is clearly reduced; see the plot on the right.

The quality of the estimated eigenvector matrix \tilde{W} is measured by the *distance* between the exact and perturbed subspace which is defined as

$$\text{dist}(\text{ran}(W_1), \text{ran}(\tilde{W}_1)) := \|W_1 W_1^\top - \tilde{W}_1 \tilde{W}_1^\top\|_2; \quad (3.168)$$

see [106, Sec. 2.6.3]. Note that [106, Thm. 2.6.1] gives that

$$\text{dist}(\text{ran}(W_1), \text{ran}(\tilde{W}_1)) = \|W_1^\top \tilde{W}_2\|_2 = \|\tilde{W}_1^\top W_2\|_2 \quad (3.169)$$

and recall that $W_1, \tilde{W}_1 \in \mathbf{R}^k$.

It is known that this distance has an informative upper bound. More concretely, for each $\varepsilon > 0$ small enough denoting the relative error $\|\tilde{C} - C\|_2 / \|C\|_2$, there exists $N_{\tilde{C}} = N_{\tilde{C}}(\varepsilon) > 0$ such that

$$\text{dist}(\text{ran}(W_1), \text{ran}(\tilde{W}_1)) \leq \frac{4\lambda_1 \varepsilon}{\lambda_k - \lambda_{k+1}} \quad (3.170)$$

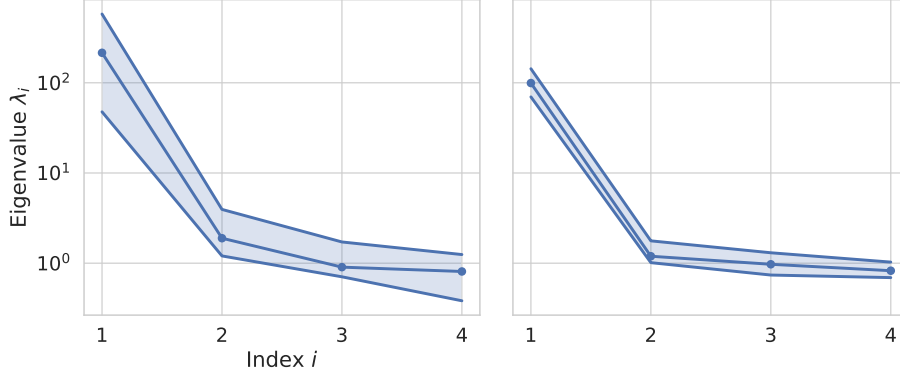


Figure 3.9: For $f(\mathbf{x}) := \frac{1}{2}\mathbf{x}^\top A\mathbf{x}$ with $A \in \mathbf{R}^{20 \times 20}$ and $\mathbf{P}_{\mathbf{X}} := \mathcal{N}(0, I)$, the variability in the estimated eigenvalues λ_i is approximated by a bootstrap approach. The left plot shows a scenario with a small number of samples $N_{\tilde{C}} = 10$ causing a larger variability as the estimation on the right where we take $N_{\tilde{C}} = 100$ gradient samples.

with high probability; see [52, Corollary 3.10] (especially for more concrete expressions for ε and $N_{\tilde{C}}$). Hence, for a large *spectral gap* between the k -th and $(k+1)$ -th eigenvalue, we get a good estimate of the (corresponding) active subspace. On the other hand, since the result is only an upper bound, a small or even zero spectral gap gives only a hint for an inaccurate estimation; further analysis is necessary to confirm that the estimated active subspace is indeed of bad quality.

In practice, a more direct assertion for the approximation of the subspace, that we follow in the next chapters, can be done by a sufficient summary plot (see Sec. 3.1) and an evaluation of the *coefficient of determination* (see Def. A.2.4) [81, 102] if the profile function g is approximated with a polynomial surrogate; see Sec. 4.4 discussing this topic in the context of Bayesian inverse problems.

Remark. In the context of Bayesian inverse problems, it is not clear whether a large spectral gap is actually necessary for an accurate approximation of the posterior distribution. For example, in [293], which has a similar context as ours but where the posterior approximations are evaluated with the Kullback–Leibler (KL) divergence [159, 160] (see Def. A.2.10), a large spectral gap is not discussed at all since the results are independent of it.

Remark. The authors of [127] also derive upper bounds similar to (3.170) for the subspace estimation. They bound the sine of the *principal angle* [33] between $\text{ran}(W_1)$ and $\text{ran}(\tilde{W}_1)$, denoted by $\angle(\text{ran}(W_1), \text{ran}(\tilde{W}_1))$, which is, however, equivalent to our notion of distance since

$$\text{dist}(\text{ran}(W_1), \text{ran}(\tilde{W}_1)) = \angle(\text{ran}(W_1), \text{ran}(\tilde{W}_1)) \quad (3.171)$$

due to [106, Sec. 12.4.3]. The main difference of their results is that, instead of having a $\log(2n)$ term in the number of samples $N_{\tilde{C}}$ (as is the case in [52, Corollary 3.10]), it is enough to only have the *numerical rank* of C , also called *intrinsic dimension* and defined as

$$\text{intdim}(C) := \frac{\text{trace}(C)}{\|C\|_2} \geq 1, \quad (3.172)$$

inside the logarithm. That is, we actually overestimate the number of samples $N_{\tilde{C}}$ by the practical recipe given in (3.166). Nevertheless, we follow this recommendation since it is still acceptable in our context.

The estimated matrix \tilde{W} , and hence \tilde{W}_1 , was a random quantity so far. In the following, we fix $\tilde{W} = \tilde{W}(\omega)$, $\omega \in \Omega$, and interpret its deviation from the exact matrix W as *perturbation*. We consequently agree on the following assumption related to this perturbation for an investigation of its impact on the approximation of f .

Assumption 3.5.4. There exists an $\varepsilon > 0$ such that

$$\text{dist}(\text{ran}(W_1), \text{ran}(\tilde{W}_1)) \leq \varepsilon. \quad (3.173)$$

The perturbation also affects the active and inactive variables, i. e., we define

$$\tilde{\mathbf{Y}} := \tilde{W}_1^\top \mathbf{X} \quad \text{and} \quad \tilde{\mathbf{Z}} := \tilde{W}_2^\top \mathbf{X}. \quad (3.174)$$

Consequently, we have

$$\rho_{\tilde{\mathbf{Y}}, \tilde{\mathbf{Z}}}(\tilde{\mathbf{y}}, \tilde{\mathbf{z}}) = \rho_{\mathbf{X}}(\llbracket \tilde{\mathbf{y}}, \tilde{\mathbf{z}} \rrbracket_{\tilde{W}}) \quad (3.175)$$

for $\tilde{\mathbf{y}} \in \tilde{\mathcal{Y}} := \tilde{W}_1^\top \mathcal{X}$ and $\tilde{\mathbf{z}} \in \tilde{\mathcal{Z}} := \tilde{W}_2^\top \mathcal{X}$. The marginal densities $\rho_{\tilde{\mathbf{Y}}}$ and $\rho_{\tilde{\mathbf{Z}}}$ are given analogously to (3.18). For the conditional density $\rho_{\tilde{\mathbf{Z}}|\tilde{\mathbf{Y}}}$, which is defined analogously to (3.20), we define

$$\tilde{\mathcal{Y}}^+ := \{\tilde{\mathbf{y}} \in \mathbf{R}^k \mid \rho_{\tilde{\mathbf{Y}}}(\tilde{\mathbf{y}}) > 0\} \quad (3.176)$$

analogously to (3.19).

Finally, we are able to define the corresponding approximation to f on the perturbed active subspace, i. e., for $\tilde{\mathbf{y}} \in \tilde{\mathcal{Y}}^+$,

$$\tilde{g}(\tilde{\mathbf{y}}) := \mathbf{E}[f(\llbracket \tilde{\mathbf{Y}}, \tilde{\mathbf{Z}} \rrbracket_{\tilde{W}}) \mid \tilde{\mathbf{Y}} = \tilde{\mathbf{y}}] \quad (3.177)$$

and

$$f_{\tilde{g}}(\mathbf{x}) := \tilde{g}(\tilde{W}_1^\top \mathbf{x}) \quad (3.178)$$

for $\mathbf{x} \in \mathcal{X}$ with $\tilde{W}_1^\top \mathbf{x} \in \tilde{\mathcal{Y}}^+$.

For proving an upper bound on the mean square error of $f_{\tilde{g}}$ approximating f , analogously to (3.33), we have to additionally assume that the perturbed conditional distribution $\mathbf{P}_{\tilde{\mathbf{Z}}|\tilde{\mathbf{Y}}}$ satisfies a probabilistic Poincaré inequality with bounded Poincaré constant $C_{\tilde{\mathbf{Y}}}$ in the sense from above, i. e., $\text{ess sup } C_{\tilde{\mathbf{Y}}} < \infty$. For example, the distributions $\mathbf{P}_{\mathbf{X}}$ discussed in Sec. 3.3 fulfill this assumption, independent of the perturbation \tilde{W} of W . The following result is a slight variation of [52, Thm. 4.6], differing only in the additional assumption.

Theorem 3.5.5. *Under Assumption 3.5.4, suppose the perturbed conditional distribution $\mathbf{P}_{\tilde{\mathbf{Z}}|\tilde{\mathbf{Y}}}$ satisfies a probabilistic Poincaré inequality with bounded Poincaré constant $C_{\tilde{\mathbf{Y}}}$ in the sense from above. Then, the mean square error of $f_{\tilde{g}}$ approximating f is bounded from above by*

$$\mathbf{E}[(f(\mathbf{X}) - f_{\tilde{g}}(\mathbf{X}))^2] \leq C_{\tilde{\mathbf{P}}, \tilde{W}} \left(\varepsilon(\lambda_1 + \dots + \lambda_k)^{1/2} + (\lambda_{k+1} + \dots + \lambda_n)^{1/2} \right)^2, \quad (3.179)$$

where $C_{\tilde{\mathbf{P}}, \tilde{W}} = C_{\tilde{\mathbf{P}}, \tilde{W}}(\tilde{W}, \mathbf{P}_{\mathbf{X}}) = \text{ess sup } C_{\tilde{\mathbf{Y}}} < \infty$.

Proof. The proof follows the same arguments as the proof of [52, Thm. 4.6]. ■

Approximation of g

The second type of perturbation is caused by the general unavailability of the ridge profile g in closed form. At first glance, it appears to be computationally costly to approximate the integral expression in (3.23). But since, by construction, the original function f is supposed to vary only mildly when changing the inactive variable \mathbf{z} , it is natural to estimate g also by a Monte Carlo sum, i. e., we define

$$g(\mathbf{y}) \approx g_N(\mathbf{y}) := \frac{1}{N} \sum_{j=1}^N f(\llbracket \mathbf{y}, \mathbf{Z}_j^{\mathbf{y}} \rrbracket) \quad (3.180)$$

for $\mathbf{y} \in \mathcal{Y}^+$, where $N > 0$ is the number of samples $\mathbf{Z}_j^{\mathbf{y}} \stackrel{\text{i.i.d.}}{\sim} \mathbf{P}_{\mathbf{Z}|\mathbf{Y}}(\cdot|\mathbf{y})$ for $j = 1, \dots, N$. Note that g_N is a random variable. We explicitly denote the samples $\mathbf{Z}_j^{\mathbf{y}}$ with a superscript \mathbf{y} to emphasize that they are drawn from the *conditional* distribution $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}$ depending on \mathbf{y} . Furthermore, if f is a ridge function, then $N = 1$ sample is enough for an exact approximation. But also in many more general cases, it is enough to have only $N = 1$ sample; see, e. g., the two complex scenarios described in Ch. 5 and 6.

In fact, it is not trivial to sample from $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}$ in general. For example, if $\mathbf{P}_{\mathbf{X}} = \mathcal{N}(0, I)$, then, by the rotational symmetry of the standard normal distribution, we have that

$$\mathbf{Y} \sim \mathcal{N}(0, I) \quad \text{and} \quad (\mathbf{Z} | \mathbf{Y} = \mathbf{y}) = \mathbf{Z} \sim \mathcal{N}(0, I), \quad (3.181)$$

which makes sampling from $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}$ rather simple. However, for such a common case as the uniform distribution on a hypercube, i. e., if $\mathbf{P}_{\mathbf{X}} = \mathcal{U}([0, 1]^n)$, the random variables \mathbf{Y} and \mathbf{Z} are not independent in general and hence sampling from the conditional distribution becomes difficult as briefly discussed in [52, Sec. 4.1]. Nevertheless, we see in Ch. 5 and 6 that the corresponding scenarios are not affected by this issue since we have to take only $N = 1$ sample. Ch. 4 explains why we can get this single sample “for free.”

It is natural to derive a bound on the mean square error also for this type of perturbation. Defining

$$f_{g_N}(\mathbf{x}) := g_N(W_1^\top \mathbf{x}) \quad (3.182)$$

for $\mathbf{x} \in \mathcal{X}$ with $W_1^\top \mathbf{x} \in \mathcal{Y}^+$, we can prove a lemma that is central to prove an upper bound.

Strictly speaking, the corresponding derivation in [52] is not fully correct since it does not consider g_N as a random variable. Also, a similar exposition in [54], which is however formally correct, does not provide much insight. As already mentioned, g_N , and hence also f_{g_N} , is a random variable. That is the reason why expressions like

$$\int_{\mathcal{X}} (f(\mathbf{x}) - f_{g_N}(\mathbf{x}))^2 \rho_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \quad (3.183)$$

are also random quantities, but the similar expression

$$\mathbf{E}[(f(\mathbf{X}) - f_{g_N}(\mathbf{X}))^2] \quad (3.184)$$

is a deterministic real value.

The following lemma, which is the main step for deriving the upper bound and a standard result from Monte Carlo theory [207], is based on the observation that

$$g(\mathbf{Y}) = \mathbf{E}[g_N(\mathbf{Y}) | \mathbf{Y}] \quad \mathbf{P}\text{-a.s.} \quad (3.185)$$

Lemma 3.5.6. *It holds that*

$$\mathbf{E}[(f_g(\mathbf{X}) - f_{g_N}(\mathbf{X}))^2] = \frac{1}{N} \mathbf{E}[(f(\mathbf{X}) - f_g(\mathbf{X}))^2]. \quad (3.186)$$

Proof. We compute that

$$\begin{aligned} \mathbf{E}[(g(\mathbf{Y}) - g_N(\mathbf{Y}))^2 | \mathbf{Y}] &= \mathbf{E}[(\mathbf{E}[g_N(\mathbf{Y}) | \mathbf{Y}] - g_N(\mathbf{Y}))^2 | \mathbf{Y}] \\ &= \mathbf{Var}(g_N(\mathbf{Y}) | \mathbf{Y}) \\ &= \frac{1}{N^2} \sum_{j=1}^N \mathbf{Var}(f(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket) | \mathbf{Y}) \\ &= \frac{1}{N} \mathbf{E}[(g(\mathbf{Y}) - f(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket))^2 | \mathbf{Y}], \end{aligned} \quad (3.187)$$

where in each line we only have equality \mathbf{P} -a.s. Applying the tower property of conditional expectations, this yields

$$\begin{aligned} \mathbf{E}[(f_g(\mathbf{X}) - f_{g_N}(\mathbf{X}))^2] &= \mathbf{E}[(g(\mathbf{Y}) - g_N(\mathbf{Y}))^2] \\ &= \mathbf{E}[\mathbf{E}[(g(\mathbf{Y}) - g_N(\mathbf{Y}))^2 | \mathbf{Y}]] \\ &= \frac{1}{N} \mathbf{E}[(g(\mathbf{Y}) - f(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket))^2] \\ &= \frac{1}{N} \mathbf{E}[(f(\mathbf{X}) - f_g(\mathbf{X}))^2]. \end{aligned} \quad (3.188)$$

■

This immediately gives the following result on the upper bound of the mean square error of f_{g_N} approximating f .

Theorem 3.5.7. *Suppose that $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}$ satisfies a probabilistic Poincaré inequality with bounded Poincaré constant $C_{\mathbf{Y}}$ in the sense from above. Then, the mean square error of f_{g_N} approximating f is bounded from above by*

$$\mathbf{E}[(f(\mathbf{X}) - f_{g_N}(\mathbf{X}))^2] \leq C_{\mathbf{P},W} (1 + N^{-1/2})^2 (\lambda_{k+1} + \cdots + \lambda_n), \quad (3.189)$$

where $C_{\mathbf{P},W} = C_{\mathbf{P},W}(W, \mathbf{P}_{\mathbf{X}}) = \text{ess sup } C_{\mathbf{Y}} < \infty$.

Proof. The triangle inequality in $L^2(\Omega, \mathbf{P})$ and Lem. 3.5.6 give that

$$\begin{aligned} \mathbf{E}[(f(\mathbf{X}) - f_{g_N}(\mathbf{X}))^2]^{1/2} &\leq \mathbf{E}[(f(\mathbf{X}) - f_g(\mathbf{X}))^2]^{1/2} + \mathbf{E}[(f_g(\mathbf{X}) - f_{g_N}(\mathbf{X}))^2]^{1/2} \\ &= (1 + N^{-1/2}) \mathbf{E}[(f(\mathbf{X}) - f_g(\mathbf{X}))^2]^{1/2}. \end{aligned} \quad (3.190)$$

After squaring both sides, the assumption allows to use (3.39) and Lem. 3.2.1 yielding the result in (3.189). ■

The final result in this section concerns the approximation of f taking both types of perturbations into account. That is, we define

$$f_{\tilde{g}_N}(\mathbf{x}) := \tilde{g}_N(\tilde{W}_1^\top \mathbf{x}) \quad (3.191)$$

for $\mathbf{x} \in \mathcal{X}$ with $\tilde{W}_1^\top \mathbf{x} \in \tilde{\mathcal{Y}}^+$, where

$$\tilde{g}_N(\tilde{\mathbf{y}}) := \frac{1}{N} \sum_{j=1}^N f(\|\tilde{\mathbf{y}}, \tilde{\mathbf{Z}}_j^{\tilde{\mathbf{y}}}\|_{\tilde{W}}) \quad (3.192)$$

for $\tilde{\mathbf{y}} \in \tilde{\mathcal{Y}}^+$ and samples $\tilde{\mathbf{Z}}_j^{\tilde{\mathbf{y}}} \stackrel{\text{i.i.d.}}{\sim} \mathbf{P}_{\tilde{\mathbf{Z}}|\tilde{\mathbf{Y}}}(\cdot|\tilde{\mathbf{y}})$, $j = 1, \dots, N$. With the same reasoning for $f_{\tilde{g}_N}$ as for f_{g_N} , we apply the triangle equality to get the following final upper bound for the mean square error.

Theorem 3.5.8. *Under the same assumptions as in Thm. 3.5.5, the mean square error of $f_{\tilde{g}_N}$ approximating f is bounded from above by*

$$\begin{aligned} & \mathbf{E}[(f(\mathbf{X}) - f_{\tilde{g}_N}(\mathbf{X}))^2] \\ & \leq C_{\tilde{\mathbf{P}}, \tilde{W}} (1 + N^{-1/2})^2 \left(\varepsilon(\lambda_1 + \dots + \lambda_k)^{1/2} + (\lambda_{k+1} + \dots + \lambda_n)^{1/2} \right)^2, \end{aligned} \quad (3.193)$$

where $C_{\tilde{\mathbf{P}}, \tilde{W}} = C_{\tilde{\mathbf{P}}, \tilde{W}}(\tilde{W}, \mathbf{P}_{\mathbf{X}}) = \text{ess sup } C_{\tilde{\mathbf{Y}}} < \infty$.

As already mentioned, the two types of perturbations that we discussed in this section frequently arise in practical applications of ASM. A particular kind of application is the combination of ASM with Bayesian inverse problems (Ch. 2) which is the central topic of the next chapter. Moreover, it is the foundation for Ch. 5 and 6.

Active subspaces and Bayesian inverse problems

Since both concepts, Bayesian inverse problems and ASM, were introduced separately in Ch. 2 and 3, respectively, we are looking at their combination in this chapter.

In Sec. 2.5, it is explained that rather primitive sampling strategies as rejection and importance sampling suffer from the curse of dimensionality and thus must be excluded from consideration as potential approaches to solve high-dimensional Bayesian inverse problems. MCMC methods, in particular the MH algorithm (Alg. 4), provide a way to alleviate the impact of the number of dimensions on the computational costs by proposing samples that reflect the shape of the posterior distribution sufficiently well in order to decrease the auto-correlation time between succeeding samples.

But still, out-of-the-box MCMC methods cannot completely eliminate deteriorating performance that comes with increasing dimensions since they still operate in the full-dimensional parameter space. Hence, it is natural to ask if there exists a certain structure of low dimensionality for a given Bayesian inverse problem on which the update from the prior to the posterior distribution can be restricted without losing too much information that would actually be necessary for a sufficiently accurate inference. For checking the existence of such a low-dimensional structure and its exploitation, we can apply ASM introduced in the previous chapter, i. e., we look for a low-dimensional *subspace* that can be used to perform the main parts of the overall inference process.

Related literature

Before we provide the setup for performing ASM with Bayesian inverse problems, let us have a look on related literature.

We mainly follow the approach from [56], where the authors describe how ASM can be exploited for accelerating MCMC. Our surrogate construction, discussed in Sec. 4.4, is similar to the study in [62].

ASM is, however, not the only approach to speed up MCMC by finding low-dimensional structure. In [69], a *likelihood-informed subspace* (LIS) is studied for Bayesian inverse problems with a Gaussian prior. In this case, it is the Hessian matrix of the data misfit function, i. e., a second order instead of a first order quantity, that is used to construct a subspace in which the

change from prior to posterior is dominant. A follow-up investigation of how a LIS-adjusted MCMC algorithm can improve the performance of producing posterior samples is provided in [68]. Also, LIS was lately used in the multilevel context (for Multilevel Monte Carlo; see Sec. 1.2) [26, 65]. Large-scale Bayesian inverse problems, solved using likelihood-informed parameter and state reduction, were studied in [70].

Recently, the *certified* approach from [293] finds bounds on the Kullback–Leibler divergence (see Def. A.2.10) between a subspace-based posterior approximation and the exact posterior. The constructed subspace is based on (probabilistic) logarithmic Sobolev inequalities [108] (in contrast to Poincaré inequalities) involving gradients of the data misfit and likelihood function. Assuming α -uniformly log-concave distributions (see Subsec. 3.4.1), the authors provide a simple way to choose the dimension of the subspace, here denoted by k , to get a particular guarantee on the posterior approximation error. The authors also provide a comprehensive comparison between several dimension reduction approaches for Bayesian inversion including ASM and LIS.

We also refer to [25] which combines geometric MCMC methods (like Riemannian manifold MALA [100]) on a low-dimensional subspace with mesh-independent infinite-dimensional approaches. ASM and LIS might be brought into the geometric setting as well, but this requires further research.

Finally, we mention the approaches in [89, 247] as these developments can also be seen as dimension reduction in Bayesian inverse problems. They exploit low-dimensional structure in *linear* Bayesian inverse problems by a low-rank Hessian approximation of the data misfit function.

Setup

Recall from Ch. 2 that the data \mathbf{D} and parameters \mathbf{X} of a Bayesian inverse problem are regarded as random variables and we assume that

$$\mathbf{D} = \mathcal{G}(\mathbf{X}) + \boldsymbol{\eta} \quad (4.1)$$

for a parameter-to-observation map $\mathcal{G} : \mathcal{X} \rightarrow \mathcal{D}$ and a term $\boldsymbol{\eta} \sim \mathcal{N}(0, \Gamma)$ modeling observational noise with zero mean and covariance matrix Γ .

Starting from a prior distribution μ_0 on the space of parameters, the ultimate goal is to compute samples from the posterior distribution $\mu^{\mathbf{d}} := \mathbf{P}_{\mathbf{X}|\mathbf{D}}(\cdot|\mathbf{d})$ incorporating the information given by a particular realization $\mathbf{d} \in \mathcal{D}$ of the data. The mathematical setup as well as arguments for the well-posedness of this problem are addressed in Sec. 2.2 and 2.4, respectively.

By Bayes' Theorem (Thm. 2.2.1), the density of the posterior distribution, denoted by $\rho^{\mathbf{d}} := \rho_{\mathbf{X}|\mathbf{D}}(\cdot|\mathbf{d})$, is given (up to a normalizing constant) by

$$\rho^{\mathbf{d}}(\mathbf{x}) \propto \exp(-f^{\mathbf{d}}(\mathbf{x})) \rho_0(\mathbf{x}) \quad (4.2)$$

for $\mathbf{x} \in \mathcal{X}$, where

$$f^{\mathbf{d}}(\mathbf{x}) := \frac{1}{2} \|\mathbf{d} - \mathcal{G}(\mathbf{x})\|_{\Gamma}^2 \quad (4.3)$$

is the data misfit function measuring the (Γ -weighted) deviation between a forward run with parameter \mathbf{x} and the given data \mathbf{d} .

As mentioned above, the idea is to perform Bayesian inversion on a low-dimensional subspace, provided it is present. That is, we apply ASM by setting

$$f(\mathbf{x}) = f^{\mathbf{d}}(\mathbf{x}) \quad \text{and} \quad \mathbf{P}_{\mathbf{X}} = \mu_0 \quad (4.4)$$

in the context of Ch. 3.

The following section, Sec. 4.1, presents a simple toy example that motivates the low-dimensional approach to Bayesian inverse problems. Subsequently, Sec. 4.2 and 4.3 provide details for the general application of ASM to the solution of Bayesian inverse problems by an MCMC method in the low-dimensional subspace. The construction of a cheap surrogate for saving computational costs while performing MCMC is addressed in Sec. 4.4.

4.1 An example

To illustrate that there might be cases in which we do not necessarily need to perform Bayesian inversion, i. e., in our case Markov chain Monte Carlo, in the full parameter space, we present a rather simple example that, however, provides the basic idea exploited in more complex scenarios later.

We regard again a linear forward model $\mathcal{G}(\mathbf{x}) := A\mathbf{x}$ for a square matrix $A \in \mathbf{R}^{4 \times 4}$ and random parameters $\mathbf{X} \sim \mathcal{N}(0, I) = \mu_0$ (implying that $\mathcal{X} = \mathbf{R}^4$). For the sake of simplicity, let us assume that the true underlying parameter is $\mathbf{x}_0 = 0$ and the data are $\mathbf{d} = 0$. The covariance matrix of the observational noise is set to be diagonal with $\Gamma_{ii} = \sigma^2 := 1/2^2$, $i = 1, \dots, 4$.

To really have an example of low-dimensionality, we manually choose an orthogonal eigenvector matrix $U \in \mathbf{R}^{4 \times 4}$ and a diagonal eigenvalue matrix $S \in \mathbf{R}^{4 \times 4}$ to determine the dominant directions of the linear forward model which is then set to

$$A = USU^\top. \quad (4.5)$$

The matrix U is randomly sampled but ensured to be orthogonal and the eigenvalues are set to

$$S = \text{diag}((3 \cdot 10^1, 10^1, 0, 0)^\top) \quad (4.6)$$

to have only two dominant directions (the first two columns in U).

In this case, the data misfit function becomes

$$f^{\mathbf{d}}(\mathbf{x}) = \frac{1}{2\sigma^2} \|A\mathbf{x}\|_2^2 = \frac{1}{2\sigma^2} \mathbf{x}^\top A^\top A \mathbf{x} \quad (4.7)$$

for $\mathbf{x} \in \mathcal{X}$. The corresponding matrix C from (3.6) can then be analytically computed to

$$C = \frac{1}{\sigma^4} US^4U^\top, \quad (4.8)$$

which means that

$$W = U \quad \text{and} \quad \Lambda = \frac{1}{\sigma^4} S^4 \quad (4.9)$$

referring to (3.7).

We run the RWM algorithm (Alg. 3) for the full model in the full (untransformed) four-dimensional space to compute a sequence of approximate posterior samples $\mathbf{X}_i^{\mathbf{d}} \sim \mu^{\mathbf{d}}$, $i = 1, \dots, 10^5$. Their realizations are denoted by $\mathbf{x}_i^{\mathbf{d}}$ in the following. The first $2 \cdot 10^4$ samples are

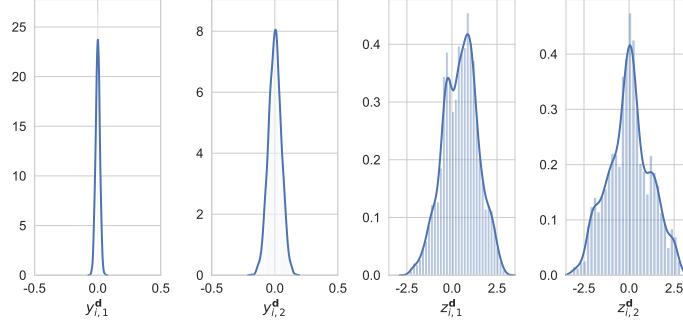


Figure 4.1: Marginal posterior distributions of the variables transformed according to (4.10). We observe that only the linear combinations which are dominant for the action of the map $\mathcal{G}(\mathbf{x}) := A\mathbf{x}$, $\mathbf{x} \in \mathcal{X}$, are informed during the inference process. Note that the x -axes of the four subfigures are different. The remaining components in the inactive subspace stay uninformed as the huge variance of their marginal posterior distribution indicates.

regarded as part of the burn-in period and removed from the final sequence. Subsequently, we transform the final sequence to the coordinate system based on the columns in $W = U$, i. e., we set

$$\begin{pmatrix} \mathbf{y}_i^d \\ \mathbf{z}_i^d \end{pmatrix} = U^\top \mathbf{x}_i^d, \quad (4.10)$$

where $\mathbf{y}_i^d, \mathbf{z}_i^d \in \mathbf{R}^2$.

The marginal distributions of the two components of $\mathbf{y}_i^d = (y_{i,1}^d, y_{i,2}^d)^\top$ and $\mathbf{z}_i^d = (z_{i,1}^d, z_{i,2}^d)^\top$ are depicted in Fig. 4.1. Noting that the x -axes are different, we can see clearly that only the components of \mathbf{y}_i^d are informed whereas the components of \mathbf{z}_i^d are not. That means, in this admittedly simple case it is actually enough to run the inference in a two- rather than in the full four-dimensional space. In more complex scenarios, the information about such a hidden low-dimensional structure can lead to a substantial reduction in computational expenses.

The description of a general procedure for performing MCMC in the low-dimensional subspace is the subject of the next section.

4.2 MCMC in the active subspace

As described in Ch. 3, we formally need to compute

$$C = \mathbf{E}[\nabla f^d(\mathbf{X}) \nabla f^d(\mathbf{X})^\top] \quad (4.11)$$

and its eigendecomposition $C = W\Lambda W^\top$ according to (3.7) and (3.8). The gradient of the misfit function f^d is

$$\nabla f^d(\mathbf{x}) = \nabla \mathcal{G}(\mathbf{x})^\top \Gamma^{-1}(\mathcal{G}(\mathbf{x}) - \mathbf{d}) \quad (4.12)$$

for arbitrary $\mathbf{x} \in \mathcal{X}$.

Once we decided for a particular subspace spanned by the columns of the submatrix W_1 from (3.10), the low-dimensional approximation to the misfit function, also called the *low-dimensional misfit function*, is given by

$$g^d(\mathbf{y}) := \mathbf{E}[f^d(\llbracket \mathbf{Y}, \mathbf{Z} \rrbracket) \mid \mathbf{Y} = \mathbf{y}] \quad (4.13)$$

for each $\mathbf{y} \in \mathcal{Y}^+$.

We can now be clearer about what we mean by performing Bayesian inversion in the low-dimensional subspace. Since the misfit function $f^{\mathbf{d}}$ is only mildly varying on the inactive subspace (spanned by the columns of the submatrix W_2) by construction, it is justified to interpret that the parameters are only negligibly informed along the inactive subspace which suggests that we focus the Bayesian update on the complementary directions. That is, we want to create samples of the posterior distribution $\mu^{\mathbf{d}}$ only in the active subspace and the remaining part from the inactive subspace is conditionally sampled from the prior distribution μ_0 .

Formally, this means that, for $\mathbf{x} \in \mathcal{X}$, $\mathbf{y} \in \mathcal{Y}$, and $\mathbf{z} \in \mathcal{Z}$ with $\mathbf{x} = \llbracket \mathbf{y}, \mathbf{z} \rrbracket$, we decompose the density of the posterior distribution by

$$\begin{aligned} \rho^{\mathbf{d}}(\mathbf{x}) &\propto \exp(-f^{\mathbf{d}}(\mathbf{x})) \rho_0(\mathbf{x}) \\ &\approx \exp(-g^{\mathbf{d}}(W_1^\top \mathbf{x})) \rho_{\mathbf{Y}}(W_1^\top \mathbf{x}) \rho_{\mathbf{Z}|\mathbf{Y}}(W_2^\top \mathbf{x} | W_1^\top \mathbf{x}) \\ &= \exp(-g^{\mathbf{d}}(\mathbf{y})) \rho_{\mathbf{Y}}(\mathbf{y}) \rho_{\mathbf{Z}|\mathbf{Y}}(\mathbf{z}|\mathbf{y}). \end{aligned} \quad (4.14)$$

Recall that $\rho_{\mathbf{Y}}$ and $\rho_{\mathbf{Z}|\mathbf{Y}}$ are marginal and, respectively, conditional densities of the prior distribution μ_0 in this chapter. Defining new densities (up to a normalizing constant)

$$\rho_{g,\mathbf{Y}}^{\mathbf{d}}(\mathbf{y}) := \exp(-g^{\mathbf{d}}(\mathbf{y})) \rho_{\mathbf{Y}}(\mathbf{y}) \quad (4.15)$$

and

$$\rho^{\mathbf{d}}(\mathbf{x}) := \rho_{g,\mathbf{Y}}^{\mathbf{d}}(W_1^\top \mathbf{x}) \rho_{\mathbf{Z}|\mathbf{Y}}(W_2^\top \mathbf{x} | W_1^\top \mathbf{x}), \quad (4.16)$$

we get a *posterior distribution on the active subspace*, denoted by $\mu_{g,\mathbf{Y}}^{\mathbf{d}}$, induced by $\rho_{g,\mathbf{Y}}^{\mathbf{d}}$ and a corresponding approximation $\mu_g^{\mathbf{d}}$ to the full posterior induced by $\rho_g^{\mathbf{d}}$, i.e., we have that

$$\frac{d\mu_{g,\mathbf{Y}}^{\mathbf{d}}}{d\lambda} := \rho_{g,\mathbf{Y}}^{\mathbf{d}} \quad \text{and} \quad \frac{d\mu_g^{\mathbf{d}}}{d\lambda} := \rho_g^{\mathbf{d}}. \quad (4.17)$$

Similar to the bounds derived in Ch. 3, we can make a statement on the approximation quality of $\mu_g^{\mathbf{d}}$ w.r.t. the Hellinger distance involving the eigenvalues of C . The following result is a slight variation of [56, Thm. 3.1], differing only by an assumption on the conditional distribution $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}$ shown to be necessary for the application of the theorems in Ch. 3.

Theorem 4.2.1. *Suppose that $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}$ satisfies a probabilistic Poincaré inequality with bounded Poincaré constant $C_{\mathbf{Y}}$ in the sense from Sec. 3.2. Then, the Hellinger distance between $\mu^{\mathbf{d}}$ and $\mu_g^{\mathbf{d}}$ is bounded from above by*

$$d_{\text{Hell}}(\mu^{\mathbf{d}}, \mu_g^{\mathbf{d}})^2 \leq M C_{\mathbf{P},W} (\lambda_{k+1} + \dots + \lambda_n), \quad (4.18)$$

where

$$M = \frac{1}{8} \left(\mathbf{E}[\exp(-f^{\mathbf{d}}(\mathbf{X}))] \cdot \exp(-\mathbf{E}[f^{\mathbf{d}}(\mathbf{X})]) \right)^{-1/2} \quad (4.19)$$

and $C_{\mathbf{P},W} = \text{ess sup } C_{\mathbf{Y}} < \infty$.

Proof. The proof follows the same steps as the proof of [56, Thm. 3.1] with a correct incorporation of the additional assumption on $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}$. ■

Remark. Recall that we actually only have perturbations of the exact quantities available, i. e.,

$$\tilde{C} := \frac{1}{N_{\tilde{C}}} \sum_{j=1}^{N_{\tilde{C}}} \nabla f^{\mathbf{d}}(\mathbf{X}_j) \nabla f^{\mathbf{d}}(\mathbf{X}_j)^{\top} = \tilde{W} \tilde{\Lambda} \tilde{W}^{\top}, \quad (4.20)$$

where $N_{\tilde{C}} > 0$ and $\mathbf{X}_j \sim \mathbf{P}_{\mathbf{X}}$, and

$$\tilde{g}_N^{\mathbf{d}}(\tilde{\mathbf{y}}) := \frac{1}{N} \sum_{j=1}^N f^{\mathbf{d}}(\llbracket \tilde{\mathbf{y}}, \tilde{\mathbf{Z}}_j \rrbracket_{\tilde{W}}) \quad (4.21)$$

for $\tilde{\mathbf{y}} \in \tilde{\mathcal{Y}}^+$, $N > 0$, and $\tilde{\mathbf{Z}}_j \sim \mathbf{P}_{\tilde{\mathbf{Z}}|\tilde{\mathbf{Y}}}(\cdot|\tilde{\mathbf{y}})$. However, similar to the bounds shown in Sec. 3.5, we can derive a bound on the Hellinger distance between the exact posterior distribution and corresponding approximations. Indeed, under the same assumptions as in Thm. 3.5.5, we get that

$$\mathbf{E}[d_{\text{Hell}}(\mu^{\mathbf{d}}, \mu_{\tilde{g}_N}^{\mathbf{d}})] \leq \sqrt{MC_{\tilde{\mathbf{P}}, \tilde{W}}} (1 + N^{-1/2}) \left(\varepsilon(\lambda_1 + \dots + \lambda_k)^{1/2} + (\lambda_{k+1} + \dots + \lambda_n)^{1/2} \right), \quad (4.22)$$

where M is from (4.19) and $C_{\tilde{\mathbf{P}}, \tilde{W}} = \text{ess sup } C_{\tilde{\mathbf{Y}}} < \infty$.

This result is a variation of [56, Thm. 3.1, (3.13)], differing by the expectation value on the left side of the inequality in (4.22) which is missing in the reference but necessary since $\mu_{\tilde{g}_N}^{\mathbf{d}}$ is a *random measure* [147].

For the remaining presentation in this section, however, we use exact quantities for the sake of a clear notation and readability.

The former results suggest a two-step approach for the construction of samples following the posterior approximation $\mu_g^{\mathbf{d}}$:

1. Run a RWM algorithm (see Sec. 2.5) in the active subspace to create samples distributed according to $\mu_{g, \mathbf{Y}}^{\mathbf{d}}$.
2. Take the low-dimensional samples from step 1 and (conditionally) complete them to samples in the full-dimensional space.

The adjusted RWM algorithm in step 1 is given in Alg. 5. It inherits the convenient fact from the original RWM algorithm that it is enough to know the target density $\rho_{g, \mathbf{Y}}^{\mathbf{d}}$ only up to a constant.

In Subsec. 2.5.2, we claim that running a RWM algorithm instead of the more advanced MH algorithm allowing for more general proposal distributions is enough in our case. At this point, we are finally able to provide an argument for this statement.

What makes the MH algorithm more powerful than the RWM algorithm is the possibility to have a proposal distribution that incorporates parts of the structure of the target distribution as, for example, sensitivities, low-dimensional behavior, or its local and global geometry, which can lead to more effective proposals [40, 99]. In contrast, the approach we are following makes the construction of a more advanced proposal distribution unnecessary since the coordinate system the Markov chain is moving in, i. e., the coordinate system transformed by the eigenvector matrix W from (3.8), is itself already aligned with the directions dominant for the

Algorithm 5: Random Walk Metropolis algorithm in the active subspace

Input : Starting point $\mathbf{Y}_0 := \mathbf{y}_0 \in \mathcal{Y}$, target distribution $\mu_{g,\mathbf{Y}}^d$ with density $\rho_{g,\mathbf{Y}}^d$, symmetric distribution \mathbf{Q} to create proposals.

Output: Samples approximately following $\mu_{g,\mathbf{Y}}^d$.

Iterating for $j = 1, 2, \dots, N_{\text{RWM-AS}}$, the j -th step is:

1. Draw $\varepsilon \sim \mathbf{Q}$ independently of \mathbf{Y}_{j-1} and define the proposal

$$\mathbf{Y}' = \mathbf{Y}_{j-1} + \varepsilon. \quad (4.23)$$

2. Compute acceptance probability

$$\begin{aligned} \alpha(\mathbf{Y}'|\mathbf{Y}_{j-1}) &:= \min \left\{ 1, \frac{\rho_{g,\mathbf{Y}}^d(\mathbf{Y}')}{\rho_{g,\mathbf{Y}}^d(\mathbf{Y}_{j-1})} \right\} \\ &= \min \left\{ 1, \frac{\exp(-g^d(\mathbf{Y}')) \rho_{\mathbf{Y}}(\mathbf{Y}')}{\exp(-g^d(\mathbf{Y}_{j-1})) \rho_{\mathbf{Y}}(\mathbf{Y}_{j-1})} \right\}. \end{aligned} \quad (4.24)$$

3. Set $\mathbf{Y}_j = \mathbf{Y}'$ with probability $\alpha(\mathbf{Y}'|\mathbf{Y}_{j-1})$, otherwise set $\mathbf{Y}_j = \mathbf{Y}_{j-1}$.
-

behavior of the target distribution, i.e., the posterior distribution. For this reason, it is then enough to propose symmetric samples which are simple to compute.

Details for the computation of suitable samples in the full parameter space in step 2 are provided in the next section.

4.3 Construction of samples in the full space

We assume that step 1 from above is executed successfully such that we have samples \mathbf{Y}_j , $j = 1, \dots, N_{\text{RWM-AS}}$, that approximately follow the distribution $\mu_{g,\mathbf{Y}}^d$, available.

The fact that these samples originate from a Markov chain implies that they are correlated. However, we can hope for a small correlation since the Markov chain is low-dimensional and hence mixes faster provided the proposal variance is chosen in a way to get a suitable acceptance rate.

Sec. 2.5 mentions two additional important parameters for the evaluation of the samples \mathbf{Y}_j besides the acceptance rate. First, we have to remove the first $N_{\text{burn-in}}$ samples in the Markov chain since they might belong to the burn-in period and thus can introduce bias in corresponding estimations. The concrete number of samples regarded as part of the burn-in period strongly depends on the application; we comment on it in Ch. 5 and 6 at the right places. Secondly, we need to regard the number of effective samples N_{eff} . Instead of “all” $N_{\text{RWM-AS}} - N_{\text{burn-in}}$ samples \mathbf{Y}_j , we might want to keep only the subset of “effective” samples, for example, if the post-processing of these samples is computationally expensive.

Technically, as already indicated in Sec. 2.5, we can keep only every K -th sample to reduce the number of samples and their correlations, where K is the (estimated) auto-correlation time of the Markov chain. The samples in the resulting reduced set are then supposed to be nearly uncorrelated. This technique is also called *thinning* in the literature [183]. However, in general

thinning is not necessary since the effects of correlations can average out if the chain is long enough. The important point actually is to check the chain's effective sample size N_{eff} .

In the following, we keep the notation \mathbf{Y}_j for samples according to $\mu_{g,\mathbf{Y}}^d$ even if they were reduced by thinning as described in the previous paragraph.

The main task we need to accomplish at this point is the construction of samples in the original full-dimensional space with the samples \mathbf{Y}_j from the active subspace. Motivated by the lines in (4.14) and (4.16), we “complete” a sample \mathbf{Y}_j by sampling \mathbf{Z}_j from the *conditional* prior distribution, i. e.,

$$\mathbf{Z}_j \sim \mathbf{P}_{\mathbf{Z}|\mathbf{Y}}(\cdot|\mathbf{Y}_j), \quad (4.25)$$

and a subsequent re-transformation of the coordinate system to get a sample $\mathbf{X}_j \sim \mu_g^d$, i. e.,

$$\mathbf{X}_j := \llbracket \mathbf{Y}_j, \mathbf{Z}_j \rrbracket. \quad (4.26)$$

Note that in the case of a standard normal prior distribution μ_0 , it is easy to sample from $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}$ since $\mathbf{Y}, \mathbf{Z} \sim \mathcal{N}(0, I)$ are independent. In general, however, it is not straightforward and obvious how to sample from the conditional distribution $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}$. For this general case, we explain two alternatives.

A first option is to run another MCMC algorithm. Fortunately, the density of $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}$, i. e., $\rho_{\mathbf{Z}|\mathbf{Y}}$, has only to be known up to a constant. Thus, we can use that

$$\rho_{\mathbf{Z}|\mathbf{Y}}(\mathbf{z}|\mathbf{y}) \propto \rho_0(\llbracket \mathbf{y}, \mathbf{z} \rrbracket) \quad (4.27)$$

for $\mathbf{y} \in \mathcal{Y}^+$ and $\mathbf{z} \in \mathcal{Z}$ such that $\llbracket \mathbf{y}, \mathbf{z} \rrbracket \in \mathcal{X}$. This expression is computationally cheap to evaluate if we choose one of the common well-known continuous distributions for the prior, as, e. g., a uniform distribution.

A second alternative which is, however, only feasible for small dimensions of the active subspace, i. e., for $k \leq 3$, is to use a pre-computed set $\hat{\mathcal{X}}$ of prior samples $\hat{\mathbf{X}}_j \sim \mu_0$, $j = 1, \dots, J$, such that, for a given $\mathbf{y} \in \mathcal{Y}$ and tolerance $\varepsilon \geq 0$ with $\varepsilon \ll 1$, we uniformly sample $\mathbf{X}_{\mathbf{y},\varepsilon}$ from the discrete finite set

$$\mathcal{X}_{\mathbf{y},\varepsilon} := \{\hat{\mathbf{X}}_j \in \hat{\mathcal{X}} \mid \|\mathbf{y} - W_1^\top \hat{\mathbf{X}}_j\|_2 \leq \varepsilon\}. \quad (4.28)$$

and set

$$\mathbf{Z}^{\mathbf{y}} := W_2^\top \mathbf{X}_{\mathbf{y},\varepsilon}. \quad (4.29)$$

Consequently, the sample $\mathbf{Z}^{\mathbf{y}}$ only approximately follows the conditional distribution $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}(\cdot|\mathbf{y})$. Assuming that ε is chosen sufficiently small, we argue that this does not significantly affect further analyses since the inactive part of the posterior samples does not dominate the data misfit function by construction. Of course, the number of samples J in the set $\hat{\mathcal{X}}$ has to get larger with an increasing dimension of the active subspace and a decreasing tolerance to obtain a non-empty set $\mathcal{X}_{\mathbf{y},\varepsilon}$. Also, in many cases J has to be large if \mathbf{y} is located near the boundary of \mathcal{Y} . Admittedly, this approach is somewhat ad hoc but can give reasonable samples that approximately follow $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}$.

4.4 Surrogate for the low-dimensional misfit

The previous sections describe that we perform the Bayesian update via an adjusted MCMC algorithm (Alg. 5) only in the low-dimensional active subspace. This algorithm needs to

evaluate the low-dimensional misfit function g^d from (4.13) for which we often only have a perturbation \tilde{g}_N^d available.

Recall that \tilde{g}_N^d is a random variable in general and thus the MCMC algorithm would have to produce samples from a random distribution which is not theoretically covered by our explanation in Sec. 2.5. Although there exist investigations for this case [15], we rather decide to construct a (deterministic) surrogate model for \tilde{g}_N^d .

If the number of samples N in (4.21) is large enough, the random fluctuations around the exact function g^d are negligible such that we can assume that a surrogate is not only a reasonable approximation but also increases the performance of the MCMC algorithm since we do not need to evaluate the full forward map for the misfit function f^d at every single step.

Note that each computation of a gradient $\nabla f^d(\mathbf{X}_j)$, $j = 1, \dots, N_{\tilde{C}}$, from (4.12) for the approximation \tilde{C} of C in (4.20) also computes a forward evaluation $\mathcal{G}(\mathbf{X}_j)$ that we can reuse “for free” in the construction of a surrogate for \tilde{g}_N^d .

There are two standard approaches for such a surrogate in the context of ASM. In Ch. 5 and 6, we make use of a polynomial regression which is explained in Alg. 6 and was already described in [52, Alg. 4.1]. Basically, we project the samples \mathbf{X}_j from (4.20) on the active

Algorithm 6: Construction of polynomial surrogate for g^d on the active subspace

Input : Samples $\mathbf{X}_j \sim \mu_0$ and corresponding misfit values $f_j := f^d(\mathbf{X}_j)$, $j = 1, \dots, N_{\tilde{C}}$, approximation to active directions stored in \tilde{W}_1 .

Output: Polynomial surrogate g_{poly} .

- 1 Project each sample on the active subspace by

$$\tilde{\mathbf{Y}}_j = \tilde{W}_1^\top \mathbf{X}_j. \quad (4.30)$$

- 2 Fit a polynomial g_{poly} with degree p to the pairs $(\tilde{\mathbf{Y}}_j, f_j)$.
- 3 For arbitrary $\mathbf{x} \in \mathcal{X}$, evaluate the surrogate with

$$g_{\text{poly}}(\tilde{W}_1^\top \mathbf{x}) \approx f(\mathbf{x}). \quad (4.31)$$

coordinate and use corresponding function evaluations

$$f_j := f^d(\mathbf{X}_j) = \frac{1}{2} \|\mathbf{d} - \mathcal{G}(\mathbf{X}_j)\|_\Gamma^2, \quad (4.32)$$

where we reuse the forward evaluations $\mathcal{G}(\mathbf{X}_j)$ as mentioned above, to construct a polynomial surrogate denoted by g_{poly} .

We emphasize that g_{poly} is actually a surrogate for $\tilde{g}_{N=1}^d$, i. e., for $N = 1$ sample of the conditional distribution $\mathbf{P}_{\tilde{\mathbf{Z}}|\tilde{\mathbf{Y}}}$ in (4.21). Note that we also get this single conditional sample “for free” since it is inherited from the full samples \mathbf{X}_j . This approach is suitable for scenarios in which it is enough to have only $N = 1$ sample of $\mathbf{P}_{\tilde{\mathbf{Z}}|\tilde{\mathbf{Y}}}$. The appropriateness (in statistics referred to as *goodness of fit*) of the polynomial surrogate model can be checked qualitatively with a sufficient summary plot (see Sec. 3.1) and quantitatively by the corresponding *coefficient of determination* (see Def. A.2.4) denoted by r^2 , also known as *r² score* [81, 102]; we use both opportunities later.

A second alternative with *Gaussian Process Regression* (GPR) [283], which we employ in Ch. 7, is regularly used in the literature and briefly discussed in this paragraph. Here, a prior Gaussian process with a predefined mean function $m(\tilde{\mathbf{y}})$ and covariance operator $K(\tilde{\mathbf{y}}, \tilde{\mathbf{y}}')$, $\tilde{\mathbf{y}}, \tilde{\mathbf{y}}' \in \tilde{\mathcal{Y}}$, for the low-dimensional misfit function is *updated* with the (noisy) value pairs $(\tilde{\mathbf{Y}}_j, f_j)$ that are also used above for the polynomial fit. We refer to [249, 283] for details on GPR and to [54] for their first application in the context of ASM. We also mention that we make use of a GPR implementation from the `scikit-learn` library [210] called `GaussianProcessRegressor`. On the one hand, GPR provides a lot of flexibility, i. e., the free choice of the covariance structure allows to approximate a large class of functions. On the other hand, however, this freedom also comes with some responsibility meaning that the user has to decide for a certain covariance structure a priori which creates assumptions that are rather hard to verify in practice. More recent developments investigate so-called *Deep Gaussian Processes* [72, 82] which are constructed to *learn* the covariance structure by given data. But these approaches were not applied for ASM so far.

Parameter estimation for a methane hydrate model

Reference. The content of this chapter is based on our article in [256].

Methane gas hydrates are crystalline solids formed when water molecules enclatharate methane molecules [242]. Gas hydrates are stable at low temperatures and high pressures and occur naturally in permafrost regions and marine off-shores [74]. If warmed or depressurized, gas hydrates destabilize and dissociate into water and gas. It is estimated that the energy content of methane occurring as hydrates exceeds the *combined* energy content of all other conventional fossil fuels [212]. Natural gas hydrates are, therefore, deemed a promising future energy resource.

Several methods have been proposed for gas extraction from hydrate reservoirs, such as thermal stimulation, depressurization, and chemical activation [199, 200]. Application of these methods at large scales is, however, very challenging due to the inherent geotechnical risks associated with gas hydrate destabilization such as rapid consolidation, seafloor subsidence, well collapse, uncontrolled sand migration, and local and regional slope instability [254, 255].

In order to quantify these risks under various production scenarios and to make realistic assessments regarding the viability of these production methods, a number of multiphysics models (e.g., [110, 133, 152, 154, 155, 231]) have been developed in the recent years. It is known that the gas hydrate-bearing sediments (GHBS) are very complex geomaterials which show a wide range of geomechanical behaviors depending on their distribution, saturation, morphology, formation, and consolidation history, etc.

The predictive capability of these models, therefore, depends heavily on the accuracy of the input constitutive model and parameters. A number of constitutive models have been proposed so far to describe the geomechanical behavior of GHBS [94, 213, 215, 268]. One common feature of these models is their large number of empirical parameters, often exceeding ten. The models themselves are highly complex, and traditional techniques of estimating the model parameters not only require large experimentation effort but also very large (often even prohibitive) computational efforts from solving multi-dimensional nonlinear optimization problems which may be ill-posed.

In this chapter, we formulate a Bayesian inverse problem that involves a constitutive model for methane hydrates and approximately solve it by the use of ASM described in Ch. 3 and 4.

That is, we look for a low-dimensional subspace on which the data misfit function $f^{\mathbf{d}}$ is dominant, i. e., an active subspace that is mostly informed by given experimental data \mathbf{d} . In this manner, ASM speeds up the computations necessary for parameter inference which would otherwise be computationally expensive or even intractable when applying standard techniques.

In Sec. 5.1, we explain the experimental setting and give details on the constitutive model. The setup for the corresponding Bayesian inverse problem, i. e., a definition of calibration parameters, quantities of interest, etc., is provided in Sec. 5.2. Sec. 5.3 presents the computational results of applying ASM to the given Bayesian inverse problem. Finally, a discussion in Sec. 5.4 relating the results to the model complements this chapter.

5.1 Case study

Reference. Subsec. 5.1.1 follows [256, Sec. 2] and is included for the sake of completeness. The experiments were organized, carried out, and kindly provided by Christian Deusner (GEOMAR Helmholtz Centre for Ocean Research Kiel) and his colleagues; see [77].

The description of the model in Subsec. 5.1.2 is taken from [256, Sec. 3]. It was developed by Shubhangi Gupta and her colleagues at TUM and GEOMAR; see [109, 110].

5.1.1 Experimental study

Experimental data were obtained in a controlled triaxial compression test with gas hydrate-bearing sand (GHBS). GHBS was formed under controlled isotropic effective stress using the *excess-gas-method* [49, 218]. In the excess-gas-method, gas hydrates are formed in partially water saturated porous or granular media by supplying gas within gas hydrate stability boundaries, i. e., at high pressure and low temperature. The use of the excess-gas-method enables the formation of homogeneously distributed gas hydrates in the porous matrix and adjustment of well-defined gas hydrate saturations (S_h) as a consequence of the limited availability of water.

Further, due to initial phase distributions and wetting behavior, gas hydrates are preferentially formed on grain surfaces and in pore throats. This microscale phase distribution is recognized to result in mechanical strengthening of the bulk sediment [134, 189]. After completion of methane hydrate formation, drained triaxial compression tests were performed at controlled axial strain rates under quasi-static loading and constant confining effective stress.

Experimental setup and measurements

Experiments were carried out in the custom-made high pressure apparatus NESSI (Natural Environment Simulator for Sub-seafloor Interactions) [77] (see Fig. 5.1), which is equipped with a triaxial cell mounted in a 40 l stainless steel vessel (APS GmbH Wille Geotechnik, Rosdorf, Germany). The sample sleeve is made from FKM. Other wetted parts of the setup are made of stainless steel. Axial and confining stresses and sample volume changes were monitored throughout the overall experimental period using high-precision hydraulic pumps. Pore pressure was measured in the influent and the effluent fluid streams close to the sample top and bottom. Pressure control was achieved using automated high-pressure piston pumps (Tedyne ISCO, Lincoln, USA). The experiment was carried out under constant temperature conditions, temperature control was achieved with a thermostat system (T1200, Lauda, Lauda-Königshofen, Germany).

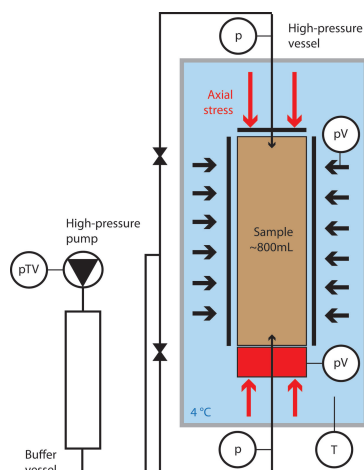


Figure 5.1: Experimental setup

Experimental control and process monitoring was carried out using high-precision piston pumps which individually control pressure and volumes of hydraulic (axial and confining) and pore fluids. During triaxial compression, pressure and fluid volumes were measured and recorded at 1 s intervals. The accuracy of the individual pressure measurements is $\pm 0.5\%$ at constant temperature. Random errors resulting from temperature changes or leakage of fluids can be neglected due to the short duration of the compression tests and large thermal buffer capacity of the high-pressure systems. The accuracy of volume and strain measurements is related to pressure measurements since system volume changes are calibrated depending on the system pressure. Thus, erroneous pressure measurements can result in an overall error of volume measurement of 4 ml, which converts to 0.4% of volumetric strain.

The sediment sample was prepared from quartz sand (initial sample porosity: 0.35, grain size: 0.1 – 0.6 mm, G20TEAS, Schlingmeier, Schwülper, Germany), which was mixed with de-ionized water to achieve a final water saturation of 0.2 relative to the initial sample porosity. The partially water-saturated and thoroughly homogenized sediment was filled into the triaxial sample cell to obtain final sample dimensions of 160 mm in height and 80 mm in diameter. The sample geometry was assured using a sample forming device. The sample was cooled to 4° C after the triaxial cell was mounted inside the pressure vessel.

Experimental procedure

Prior to the gas hydrate formation, the partially water-saturated sediment sample was isotropically consolidated to 1 MPa effective stress under drained conditions. The sample was flushed with CH_4 gas and, subsequently, pressurized with CH_4 gas to obtain a pore pressure of approximately 10 MPa. During pressurization with CH_4 gas, and throughout the overall gas hydrate formation period, isotropic effective stress was controlled to remain constant at 1 MPa using an automated control algorithm. The formation process was continuously monitored by logging the CH_4 gas pressures. Mass balances and volume saturations were calculated based on CH_4 gas pressure to confirm that available pore water was fully converted into gas hydrates.

After completion of gas hydrate formation, the triaxial compression test was conducted at a controlled axial strain rate of 0.1 mm/min. During axial loading and compression, the confining effective stress was controlled to remain constant by adjusting the confining hydraulic

fluid volume in the pressure vessel. Accumulated volumetric strain was calculated based on changes of axial and confining volumes, which are monitored by the hydraulic pumps.

5.1.2 Model

From a geomechanics point of view, the gas hydrate-bearing sands (GHBS) are cohesive-frictional granular materials. The most important features of the mechanical behavior of GHBS include plastic deformations and the onset of critical state (i. e., isochoric deformations), pressure-dependence, and shear-volumetric strain coupling (or *dilatancy*). The strength and the stiffness of GHBS are strongly influenced by gas hydrate saturation and hydrate morphology (i. e., pore-filling, load-bearing, cementing, etc.), as well as the hydrate formation method [277, 291]. In general, the higher the gas hydrate saturation, the higher is the bulk compressive strength of the GHBS. The gas hydrate saturation also enhances the cohesive strength, frictional resistance, and dilatancy of GHBS.

We model the geomechanical behavior of GHBS within an incremental elasto-plasticity framework, and for simplicity we assume infinitesimal strains. This subsection presents the main elements of our elasto-plastic material model for GHBS, including the yield function, plastic flow directions, and hardening and softening evolution laws.

Notation. For any second order tensors \mathbf{a} and \mathbf{b} , the inner product is given by $\mathbf{a} : \mathbf{b} = a_{ij}b_{ij}$, and the dyadic product is given by $(\mathbf{a} \otimes \mathbf{b}) = a_{ij}b_{kl}$. The tensor product between a fourth order tensor \mathbf{A} and a second order tensor \mathbf{b} is given by $\mathbf{A} : \mathbf{b} = A_{ijkl}b_{kl}$. The Euclidean norm of \mathbf{a} is given by $\|\mathbf{a}\| = (\mathbf{a} : \mathbf{a})^{1/2}$. The second order unit tensor is given by $\mathbf{I} = \delta_{ij}$, where δ_{ij} denotes the Kronecker delta function. The fourth order unit tensor is given by $\mathbf{I} = \frac{1}{2}(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})$. The trace of \mathbf{a} is given by $\text{Tr}\mathbf{a} = \mathbf{I} : \mathbf{a}$. Any second order tensor \mathbf{a} can be decomposed into a dilational (or volumetric) part, $\text{dila}\mathbf{a}$, and a deviatoric part, $\text{dev}\mathbf{a}$. In 3D, $\text{dila}\mathbf{a} = \frac{1}{3}\text{Tr}\mathbf{a}$ and $\text{dev}\mathbf{a} = \mathbf{a} - \frac{1}{3}\text{Tr}\mathbf{a}$.

Preliminaries

Let $\boldsymbol{\sigma}$ be the Cauchy stress tensor and $\boldsymbol{\epsilon} = \frac{1}{2}(\nabla\mathbf{u} + \nabla\mathbf{u}^\top)$ the infinitesimal strain tensor. The vector \mathbf{u} denotes the displacement field. Both $\boldsymbol{\sigma}$ and $\boldsymbol{\epsilon}$ are symmetric second order tensors. The total infinitesimal strain $\boldsymbol{\epsilon}$ is decomposed *additively* into the elastic strain $\boldsymbol{\epsilon}^e$ and the plastic strain $\boldsymbol{\epsilon}^p$, i. e., $\boldsymbol{\epsilon} = \boldsymbol{\epsilon}^e + \boldsymbol{\epsilon}^p$.

In classical plasticity [138], the state of stress depends on the loading-unloading history and is calculated incrementally. The stress and the strain rate tensors (i. e., $\dot{\boldsymbol{\sigma}}$, $\dot{\boldsymbol{\epsilon}}$, $\dot{\boldsymbol{\epsilon}}^e$, and $\dot{\boldsymbol{\epsilon}}^p$) are approximated using an implicit Euler finite difference method, i. e., for any time interval $[t_i, t_{i+1}]$, the stress or strain rates are approximated as

$$\dot{\boldsymbol{\cdot}} = \frac{[\boldsymbol{\cdot}] - [\boldsymbol{\cdot}]}{t_{i+1} - t_i}, \quad (5.1)$$

where $[\boldsymbol{\cdot}]$ denotes the state at time t_i , and $[\boldsymbol{\cdot}]$ denotes the state at time t_{i+1} .

We define the plasticity relationships in terms of the following stress invariants,

$$p = \frac{1}{3}\text{Tr}\boldsymbol{\sigma} \quad \text{and} \quad q = \sqrt{\frac{3}{2}}\|\text{dev}\boldsymbol{\sigma}\|, \quad (5.2)$$

where p denotes hydrostatic or mean stress, and q denotes shear stress. Corresponding invariants of strain rate are

$$\dot{\epsilon}_v = \text{Tr} \dot{\boldsymbol{\epsilon}} \quad \text{and} \quad \dot{\epsilon}_s = \sqrt{\frac{2}{3}} \|\text{dev } \dot{\boldsymbol{\epsilon}}\|, \quad (5.3)$$

where $\dot{\epsilon}_v$ denotes the volumetric strain rate, and $\dot{\epsilon}_s$ denotes the shear strain rate. The invariants of elastic and plastic strain rates are defined similarly.

Elasticity

In the elastic range, we assume linear isotropic material behavior, i. e., the stress $\boldsymbol{\sigma}$ is related to the elastic strain $\boldsymbol{\epsilon}^e$ through a linear Hooke's law,

$$\boldsymbol{\sigma} = \mathbb{C}^e : \boldsymbol{\epsilon}^e, \quad (5.4)$$

where \mathbb{C}^e is the elastic stiffness of the material which is a symmetric positive definite fourth order tensor,

$$\mathbb{C}^e = L_1 \mathbf{I} \otimes \mathbf{I} + 2L_2 \mathbf{I}. \quad (5.5)$$

L_1 and L_2 are the Lamé coefficients,

$$L_1 = \frac{\nu E(S_h)}{(1 + \nu)(1 - 2\nu)} \quad \text{and} \quad L_2 = \frac{E(S_h)}{2(1 + \nu)}, \quad (5.6)$$

where E is the elastic Young's modulus and ν is the Poisson's ratio of the GHBS, respectively. S_h denotes the gas hydrate saturation. It is observed that the Young's modulus of GHBS increases with increasing S_h , while the Poisson's ratio does not vary much over a wide range of S_h and can be assumed constant [169, 195]. In [233], the authors have proposed an empirical relationship for E of the form

$$E = E_s(\sigma_c) + S_h^m E_h, \quad (5.7)$$

where E_s and E_h denote the Young's modulus of the sand and gas hydrates, respectively, and σ_c is the confining stress. The exponent m varies over a wide range. In their experiment and modeling study in [109], the authors found that the effect of S_h on E was linear ($m = 1$) during hydrate formation, while during hydrate dissociation the effect of S_h on E was stronger ($m = 3$).

Yield function

There exists a yield surface F in the stress space that encompasses the elastic region. The stress states lying *inside* the yield surface produce *elastic* deformations, while the stress states lying *on* the surface produce *plastic* deformations. The stress states outside the yield surface are inadmissible.

We consider a Drucker-Prager yield criterion where the yield function is given as

$$F(\boldsymbol{\sigma}, \boldsymbol{\chi}) := q + \alpha(\boldsymbol{\chi})p - c(\boldsymbol{\chi}) = 0. \quad (5.8)$$

The function F describes a conical surface in the principal stress space (see Fig. 5.2a). The parameter α indicates the mobilized frictional resistance at any given stress state. The parameter c indicates the cohesive strength of the granular material. $\boldsymbol{\chi}$ denotes the vector of internal plastic variables which affect the hardening-softening behavior of GHBS due to changes in internal structure or grain contacts, packing density of the sand grains, hydrate saturation, hydrate pore habit, etc.

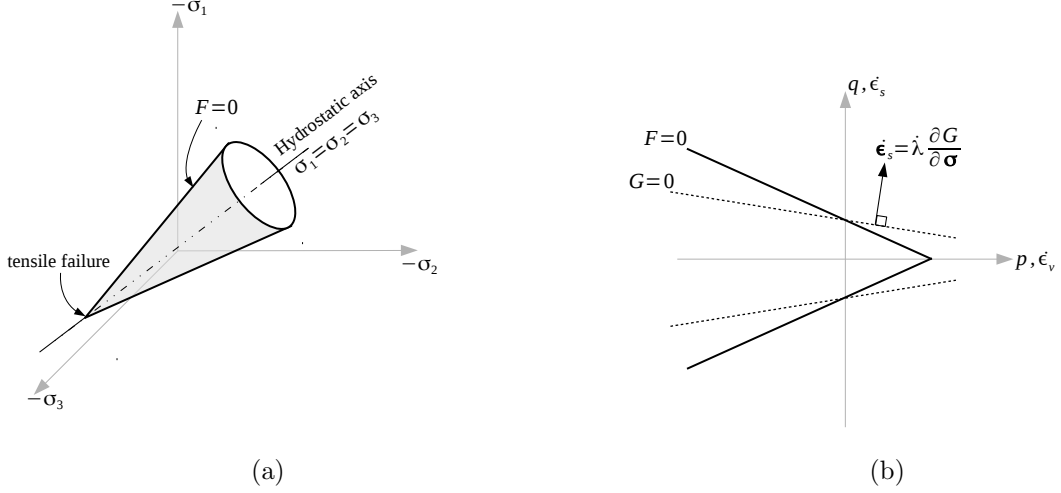


Figure 5.2: (a) Drucker-Prager yield surface in principal stress space. (b) Potential surface and plastic strains in $p - q$ space.

Plastic strains

Similar to the yield surface F , there exists a plastic potential surface G in the stress space such that the plastic flow occurs in a direction normal to this surface (see Fig. 5.2b). The incremental plastic strains (i.e., the plastic strain rate) can be derived from the plastic potential G as

$$\dot{\epsilon}^p = \dot{\lambda} \frac{\partial G}{\partial \boldsymbol{\sigma}}, \quad (5.9)$$

where $\partial G / \partial \boldsymbol{\sigma}$ describes the normal to the surface G , and $\dot{\lambda}$ is a proportionality constant indicative of the magnitude of the plastic strain increment. It can further be shown that the invariants of the plastic strain rate can be written as

$$\dot{\epsilon}_v^p = \dot{\lambda} \frac{\partial G}{\partial p} \quad \text{and} \quad \dot{\epsilon}_s^p = \dot{\lambda} \frac{\partial G}{\partial q}. \quad (5.10)$$

We consider a non-associative flow rule, i.e., $G \neq F$,

$$G(\boldsymbol{\sigma}, \boldsymbol{\chi}) := q + \beta(\boldsymbol{\chi})p = 0, \quad (5.11)$$

where $\beta < \alpha$. The parameter β denotes the *dilatancy* of the material. Dilatancy is a characteristic property of frictional granular materials. It contributes to the strength of the material and effectively couples the deviatoric and volumetric components of plastic deformation. From (5.10) and (5.11), dilatancy can be written as $\beta = \dot{\epsilon}_v^p / \dot{\epsilon}_s^p$. Depending on the relative packing density of the grains, it can allow for macroscopic contraction or dilation of the material under external loads.

Loading-unloading conditions

Along any process of loading-unloading, if $F < 0$, the stress state is elastic and $\dot{\lambda} = 0$, while, if $F = 0$, the stress state is plastic and $\dot{\lambda} > 0$. These nonlinear inequality constraints can be

reformulated as the following Karush-Kuhn-Tucker [158] optimality conditions,

$$F(\boldsymbol{\sigma}, \boldsymbol{\chi}) \leq 0, \quad \dot{\lambda} \geq 0, \quad \text{and} \quad \dot{\lambda} F = 0. \quad (5.12)$$

To confine the stress trajectory to the yield surface during plastic loading, an additional plastic consistency condition is considered [75],

$$\dot{\lambda} \dot{F} = \dot{\lambda} \left(\frac{\partial F}{\partial \boldsymbol{\sigma}} : \dot{\boldsymbol{\sigma}} + \frac{\partial F}{\partial \boldsymbol{\chi}} : \dot{\boldsymbol{\chi}} \right) = 0. \quad (5.13)$$

Evolution laws

Experiments have shown that an increase in gas hydrate saturation tends to increase the initial frictional resistance, apparent cohesive resistance, peak strength, and peak dilatancy of GHBS [134, 170, 189]. It is also observed that GHBS show a distinct strain hardening-softening behavior [195, 196]. In [268], this behavior is explained in detail, and strain dependent evolution laws are presented to capture the observed strain hardening-softening. Other strain dependent evolution laws have also been presented in [154, 214, 215], among others. In our experiments, in addition to strain hardening-softening, we additionally observe a distinct *secondary* hardening phase in the stress-strain response of the GHBS samples.

Assuming that the frictional resistance of any geomaterial can be expressed as a sum of dilatancy and some *residual* frictional resistance [286], i. e.,

$$\alpha = \beta + \alpha_{\text{res}}, \quad (5.14)$$

where, at critical state, $\beta = 0$ and $\alpha = \alpha_{\text{res}}$, we hypothesize that the primary hardening occurs due to the dilatancy of the sample, while the secondary hardening occurs due to an increase in residual frictional resistance under plastic loading. This is likely the result of internal damage and hydrate redistribution in the pore spaces causing higher particle density, increased interlocking of grains, and higher friction at the grain contacts. This effect becomes dominant when the material has spent all its dilatancy and has achieved a critical state. We ignore the contribution of the cohesive strength on the observed strain dependent hardening-softening-hardening behavior.

Let us consider the plastic internal variables $\boldsymbol{\chi} = (\dot{\epsilon}_s^p, \epsilon_s^p, S_h)^\top$. From (5.10) and (5.11), we get $\dot{\epsilon}_s^p = \dot{\lambda}$ and $\epsilon_s^p = \int_{t_i}^{t_{i+1}} \dot{\lambda} dt = \lambda$. To capture the macroscopic stress-dilatancy behavior of GHBS observed during our triaxial compression experiments, we describe smooth empirical evolution laws for the properties α , β , and c in (5.8) and (5.11) as

$$c = c(S_h), \quad (5.15)$$

$$\beta = \beta^*(S_h) \cdot \bar{\lambda} \cdot \exp(1 - \bar{\lambda}^{m_\beta}), \quad (5.16)$$

$$\alpha_{\text{res}} = \alpha_{\text{res}}^1(S_h) + \Delta\alpha_{\text{res}}(S_h) \cdot \left(1 + 1/\dot{\bar{\lambda}}\right)^{-1} \cdot \bar{\lambda}^{m_\alpha}, \quad (5.17)$$

where $\bar{\lambda} = \lambda/\lambda^*(S_h)$ and $\dot{\bar{\lambda}} = \dot{\lambda}/\dot{\lambda}^*(S_h)$. The functional dependence of the parameters c , β^* , λ^* , α_{res}^1 , $\Delta\alpha_{\text{res}}$, and $\dot{\lambda}^*$ on S_h can be derived through empirical correlation by repeating these experiments over a range of hydrate saturations. In this work, we consider only a single GHBS sample with a constant hydrate saturation. So, the exact functional dependence of the plasticity parameters on S_h is not of direct relevance for the purpose of presenting our use of ASM and will not be discussed further.

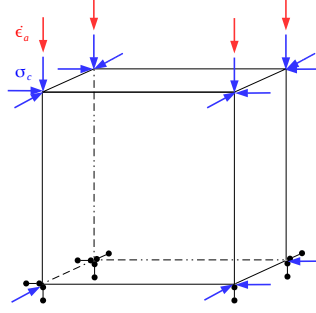


Figure 5.3: One element triaxial setup.

(5.16) and (5.17) are extensions of the evolution functions proposed in [12]. The parameter β^* denotes the peak dilatancy and λ^* denotes the corresponding accumulated plastic shear strain. The parameter α_{res}^1 denotes the minimum frictional resistance of the intact material before loading. It is interesting to note that in (5.17), when $\dot{\epsilon}_s \rightarrow 0$, we get $\alpha_{\text{res}} = \alpha_{\text{res}}^1$. Physically, this implies that, under quasi-static loading conditions, the material does not undergo microscopic damage and the residual frictional resistance of the material remains constant. At higher loading rates ($\dot{\epsilon}_s > 0$), however, the effects of microscopic damage, sand and hydrate grain rearrangement, friction at grain contacts, etc. become progressively larger, resulting in an overall increase in the macroscopic residual frictional resistance.

Finite element implementation

We solve the *global* nonlinear equilibrium equation using a Galerkin finite element formulation defined on Q_1 elements. The nonlinearities are resolved iteratively using a full Newton-Raphson method with a continuum tangent matrix [296]. Within each global Newton iteration step, a *local* problem is solved at each Gauss point to determine the new stress state. The local problem involves the integration of the material model (see previous paragraphs) over the load increment of the current global step. We use an implicit return mapping algorithm [130, 240] to solve the local problem. The implicit algorithm uses the *final* point in the stress space to evaluate the relevant derivatives and variables. Since this point is not known in advance, a Newton-Raphson method is used to advance the solution iteratively toward the final solution. In a more generalized solution method, the nonlinear equilibrium equation as well as the inequality constraints can be treated within a single Newton iteration, which can be implemented as a primal-dual active set strategy (e. g., [116, 117]). We have implemented our numerical scheme in C++ based on the DUNE PDELab framework [20, 76].

Numerical simulation of the triaxial compression experiments

We consider a *one element* triaxial setup with unit dimensions as shown in Fig. 5.3. Load is applied in two stages. In the first stage, an isotropic load equal to the confining stress of $\sigma_c = 1$ MPa is applied. This corresponds to the initial stress state of the GHBS sample in the triaxial compression experiment. In the second stage, the strain-controlled triaxial compression of the GHBS sample is simulated by specifying an axial strain rate of $\dot{\epsilon}_a = -1.04167 \times 10^{-5} \text{ s}^{-1}$,

Table 5.1: Prior intervals of the eight model parameters.

No.	Parameter	Min	Max	Unit
1	c	1.8×10^6	2.4×10^6	Pa
2	α_{res}^l	0.5	0.6	—
3	$\Delta\alpha_{\text{res}}$	0.2	0.3	—
4	$\dot{\lambda}^*$	1.6×10^{-3}	1.9×10^{-3}	—
5	m_α	0.75	1.05	—
6	β^*	0.3	0.45	—
7	λ^*	0.01	0.011	—
8	m_β	0.67	0.74	—

which corresponds to a vertical displacement rate of 0.1 mm/minute in the $-z$ direction. The strain-controlled load is applied incrementally in 1350 steps with a step size of 10 s.

5.2 Setup

For the formulation of a Bayesian inverse problem, as described in Sec. 2.2, we need to specify some necessary quantities as the set of parameters, quantities of interest, the data \mathbf{d} , the covariance matrix Γ of the noise distribution, and the prior distribution.

Model parameters

The parameters for the model described in Subsec. 5.1.2 are

$$(c \ \alpha_{\text{res}}^l \ \Delta\alpha_{\text{res}} \ \dot{\lambda}^* \ m_\alpha \ \beta^* \ \lambda^* \ m_\beta)^\top \in \mathbf{R}^8. \quad (5.18)$$

Note that $\Delta\alpha_{\text{res}} := \alpha_{\text{res}}^h - \alpha_{\text{res}}^l$ and that the inference is done in an 8-dimensional space, i. e., $n = 8$. The parameters in (5.18) are called *model parameters*; prior intervals for them are given in Tab. 5.1. These intervals were pre-determined by engineering knowledge.

Calibration parameters

Since ASM prefers centered inputs with normalized variances, we actually do not infer the model parameters above but corresponding so-called *calibration parameters*, i. e.,

$$\mathbf{x} = (\bar{c} \ \overline{\alpha_{\text{res}}^l} \ \overline{\Delta\alpha_{\text{res}}} \ \overline{\dot{\lambda}^*} \ \overline{m_\alpha} \ \overline{\beta^*} \ \overline{\lambda^*} \ \overline{m_\beta})^\top \in \mathbf{R}^8. \quad (5.19)$$

The prior distribution μ_0 for Bayesian inversion is given on the space of these calibration parameters which, in this case, is chosen to be

$$\mu_0 = \mathbf{P}_\mathbf{X} = \mathcal{U}([-1, 1]^8) \quad (5.20)$$

implying that $\mathcal{X} = [-1, 1]^8$. Particular prior samples of calibration parameters are then mapped to their corresponding counterpart in the space of model parameters and vice versa which is just a linear transformation.

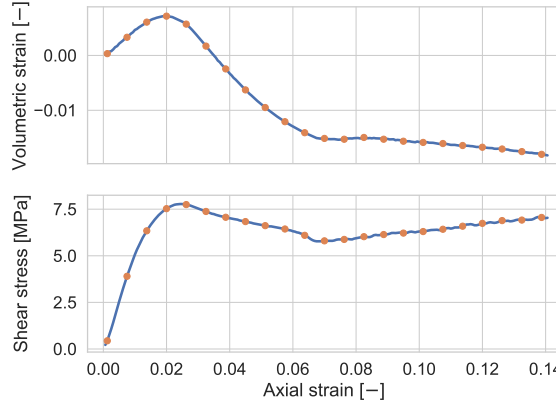


Figure 5.4: Given volumetric strain and shear stress data (orange dots) for 23 given axial strain values.

Quantities of interest

In the context of this chapter, the quantities of interest are values of volumetric strain ϵ and shear stress σ each for 23 given axial strain values which yields the corresponding forward operator

$$\mathcal{G}(\mathbf{x}) := (\mathcal{G}_\epsilon(\mathbf{x}), \mathcal{G}_\sigma(\mathbf{x}))^\top \in \mathbf{R}^{n_d}, \quad (5.21)$$

where $\mathbf{x} \in \mathcal{X}$, $n_d = n_\epsilon + n_\sigma = 46$; the values of $\mathcal{G}_\epsilon : \mathbf{R}^8 \rightarrow \mathbf{R}^{n_\epsilon}$ and $\mathcal{G}_\sigma : \mathbf{R}^8 \rightarrow \mathbf{R}^{n_\sigma}$ are ordered according to the given axial strain values. The measured volumetric strain and shear stress data

$$\mathbf{d} := (\mathbf{d}_\epsilon, \mathbf{d}_\sigma)^\top \in \mathbf{R}^{n_\epsilon + n_\sigma}, \quad (5.22)$$

that is needed as a conditioning argument for the posterior distribution $\mu^{\mathbf{d}}$, are depicted in Fig. 5.4. Following Ch. 2, the experimental noise $\boldsymbol{\eta}$ is Gaussian distributed with mean zero and covariance matrix $\Gamma \in \mathbf{R}^{n_d \times n_d}$. We assume that

$$\Gamma = \begin{pmatrix} \Gamma_\epsilon & \\ & \Gamma_\sigma \end{pmatrix}, \quad (5.23)$$

where $\Gamma_\epsilon \in \mathbf{R}^{n_\epsilon \times n_\epsilon}$ and $\Gamma_\sigma \in \mathbf{R}^{n_\sigma \times n_\sigma}$ are diagonal matrices corresponding to a 2% noise level in the measurements, i. e.,

$$\Gamma_{ii} = (0.02 \times d_i)^2 \quad (5.24)$$

for $i = 1, \dots, n_d$, where d_i denotes the i -th component of \mathbf{d} . The data misfit function can thus be written as

$$f^{\mathbf{d}}(\mathbf{x}) = \frac{1}{2} \|\mathbf{d} - \mathcal{G}(\mathbf{x})\|_\Gamma^2 \quad (5.25)$$

$$= \frac{1}{2} (\|\mathbf{d}_\epsilon - \mathcal{G}_\epsilon(\mathbf{x})\|_{\Gamma_\epsilon}^2 + \|\mathbf{d}_\sigma - \mathcal{G}_\sigma(\mathbf{x})\|_{\Gamma_\sigma}^2) \quad (5.26)$$

for $\mathbf{x} \in \mathcal{X}$.

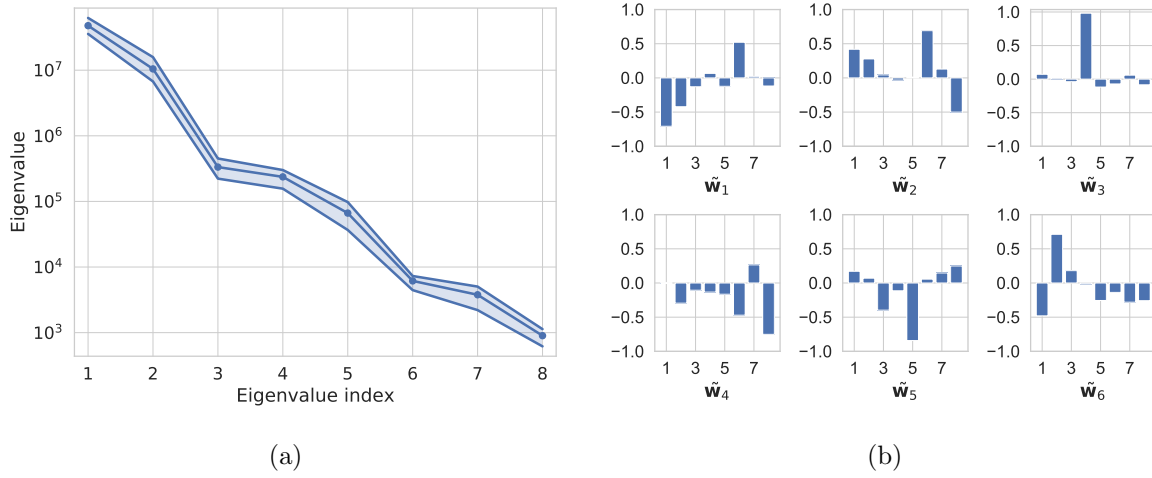


Figure 5.5: (a) Estimated eigenvalues of C and their bootstrap intervals. (b) First six estimated eigenvectors of C .

5.3 Computational results

Computational costs

Using ASM, we need to approximate the matrix C from (3.6) with a finite Monte Carlo sum as described in Sec. 3.5, i. e., we need to compute gradients of the misfit function; see (4.12).

Since the model is not readily capable of an adjoint formulation which would make gradient computations easier, central finite differences are used to approximate partial derivatives of the forward operator, i. e.,

$$\frac{\partial \mathcal{G}_i(\mathbf{x})}{\partial x_j} \approx \frac{\mathcal{G}_i(\mathbf{x} + h\mathbf{e}_j) - \mathcal{G}_i(\mathbf{x} - h\mathbf{e}_j)}{2h} \quad (5.27)$$

for $i = 1, \dots, n_d$, $j = 1, \dots, n$, $h > 0$ small enough, and $\mathbf{x} \in \mathcal{X}^\circ$. Hence, a single sample of ∇f^d requires $2 \times n + 1 = 17$ forward evaluations. Following the heuristic from (3.166) with choosing pessimistic parameters, we need at least $N_{\tilde{C}} = \lceil 10 \times 8 \times \log(8) \rceil = 167$ gradient samples. However, we decide to use $N_{\tilde{C}} = 250$ samples to be sure.

Since the gradient samples used for the Monte Carlo sum are statistically independent, it is possible to parallelize their computation. The $17 \times 250 = 4,250$ forward evaluations were computed on 35 cores within 15.12 h, i. e., in 529.2 core hours. That is, a single forward computation took 7.47 min on average. The additional calculations required for the eigendecomposition and corresponding bootstrap intervals are negligibly cheap.

Active subspace for the inference

We take the $N_{\tilde{C}} = 250$ gradient samples and approximate the matrix C with them according to (3.159). Fig. 5.5a shows the estimated eigenvalues of C and corresponding bootstrap intervals which indicate that the number of samples are enough since they are reasonably small. To find a suitable active subspace, we look for spectral gaps due to (3.170) which can indeed be observed after the second and fifth eigenvalue. The first gap is about 1.5 orders of magnitude in size whereas the second is about one order of magnitude. This suggests to choose a 2D

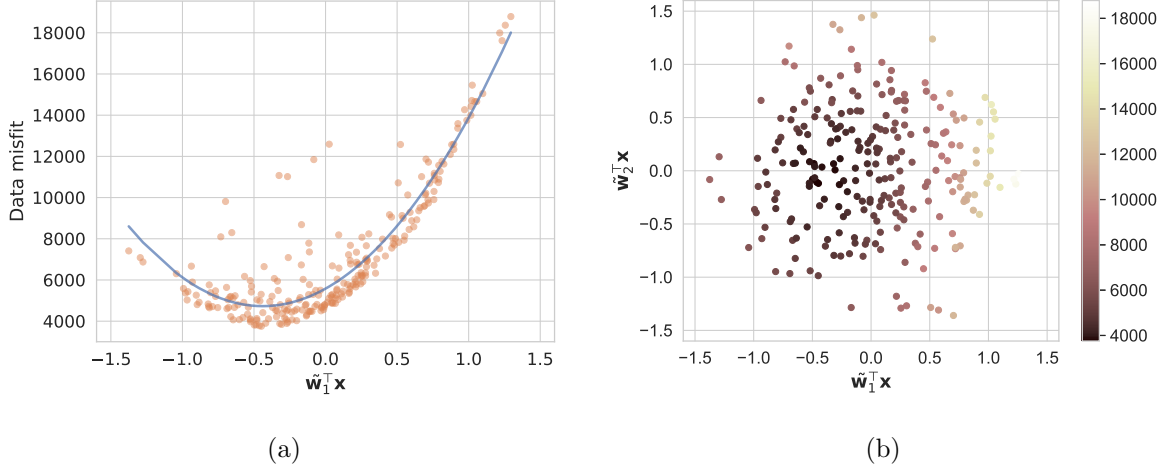


Figure 5.6: (a) 1D sufficient summary plot on the axis $\tilde{\mathbf{w}}_1^\top \mathbf{x}$ and the corresponding quadratic surrogate giving $r_{1D}^2 = 0.8434$. (b) Sufficient summary plot in two dimensions on the axes $\tilde{\mathbf{w}}_1^\top \mathbf{x}$ and $\tilde{\mathbf{w}}_2^\top \mathbf{x}$. Here, the quadratic surrogate yields $r_{2D}^2 = 0.9776$.

or a 5D subspace. In the following, we investigate both options for comparison. The first six eigenvectors (associated with the largest six eigenvalues) that are taken to span the active subspace are depicted in Fig. 5.5b. Recall that the active subspace, in this context, consists of directions in the original space of calibration parameters that are dominantly informed by the given data \mathbf{d} , on average.

For a given subspace, it remains to find a surrogate for the low-dimensional misfit function $g^{\mathbf{d}}$, or, more accurately, $\tilde{g}_N^{\mathbf{d}}$ (with $N = 1$); see Sec. 4.4. We decide to fit a regression-based polynomial

$$g_{\text{poly}} \approx \tilde{g}_N^{\mathbf{d}} \quad (5.28)$$

as described by Alg. 6 by taking samples \mathbf{X}_j and $f^{\mathbf{d}}(\mathbf{X}_j)$, $j = 1, \dots, N_{\tilde{C}}$, that were already computed for the approximation of C . We emphasize that there is no need for further forward evaluations, we instead reuse samples $f^{\mathbf{d}}(\mathbf{X}_j)$ that came out of the computation of gradients; for details see Sec. 4.4.

The 1D sufficient summary plot in Fig. 5.6a shows a quadratic surrogate (a parabola) in one dimension. Recall from Sec. 3.1 that a sufficient summary plot projects the samples \mathbf{X}_j on the axes of the active subspace and then plots function values corresponding to these axes. Since the number of coefficients in a general parabola is only 3, we can be sure to avoid the issue of overfitting. The coefficient of determination, as an indicator for the quality of an active subspace, is $r_{1D}^2 = 0.8434$.

However, we see that there are some outliers that are not fitted sufficiently well by the surrogate which means, and confirms, that a 1D subspace is not enough. Correspondingly, Fig. 5.6b depicts a sufficient summary plot in two dimensions; the color of the dots stand for the function value (see the colorbar on the right). The color gradient in this figure looks reasonably smooth and does not contain huge jumps, i.e., there is a good chance to approximate $f^{\mathbf{d}}$ by a two-dimensional surrogate. In this case, a quadratic surrogate can be chosen to obtain a coefficient of determination of $r_{2D}^2 = 0.9776$ which is quite good. The number of coefficients here is 6, so we also avoid overfitting here.

Finally, if we choose a 5D subspace and a corresponding quadratic surrogate, resulting in only 21 degrees of freedom, then $r_{5D}^2 = 0.9824$.

Such polynomial regression-based surrogates are extremely cheap to evaluate, also in higher dimensions. While a larger dimension of the active subspace brings less error in the final posterior approximation, a smaller dimension accelerates mixing in the MCMC algorithm and thus needs less time to produce a particular amount of effective samples that are necessary for the inference of calibration parameters.

Inference with MCMC in the active subspace

After the surrogate g_{poly} to the low-dimensional misfit function is computed, we can exploit it for constructing samples that approximately follow the posterior distribution on the active subspace

$$\mu_{g_{\text{poly}}, \tilde{\mathbf{Y}}}^d \approx \mu_{g, \mathbf{Y}}^d \quad (5.29)$$

from (4.17).

Remark. We neither have the exact variables \mathbf{Y} , \mathbf{Z} nor the exact function g^d available due to the finite approximation of C explained above. That is, we instead need to use $\tilde{\mathbf{Y}}$, $\tilde{\mathbf{Z}}$, and g_{poly} , respectively.

As mentioned, we compare results for 2D and 5D subspaces. However, we expect that the posterior approximation using a 5D subspace is more accurate since more calibration parameters (through linear combinations) are taken into account. Also, the bound on the Hellinger distance between the exact posterior distribution and its low-dimensional approximation in Thm. 4.2.1, which is proportional to the sum of eigenvalues belonging to the inactive subspace, supports our expectation.

Recall that the density $\rho_{g_{\text{poly}}, \tilde{\mathbf{Y}}}^d$ of $\mu_{g_{\text{poly}}, \tilde{\mathbf{Y}}}^d$ from (4.15) depends on the prior density $\rho_{\tilde{\mathbf{Y}}}$ which is the prior density ρ_0 marginalized on the active variable $\tilde{\mathbf{Y}}$. That is, in order to run the MCMC algorithm in the active subspace (Alg. 5), we need to estimate $\rho_{\tilde{\mathbf{Y}}}$. In the case of our uniform prior distribution, it is generally difficult to find analytical expressions for $\rho_{\tilde{\mathbf{Y}}}$ which means that we have to approximate it numerically. We produce a sufficiently large amount of prior samples, project them on the active subspace, and run a kernel density estimation using the Python library `scikit-learn` [210].

In order to actually start the MCMC algorithm, we still have to decide for a particular (symmetric) proposal distribution \mathbf{Q} . We choose $\mathbf{Q} = \mathcal{N}(0, \gamma^2 I)$ with a so-called *proposal variance* $\gamma^2 > 0$. The actual proposal variances for the 2D and 5D case have to be chosen such that the corresponding algorithm produces reasonable acceptance rates between 0.2 and 0.4; see Sec. 2.5. For this, we run a few trial runs with different proposal variances to observe whether they result in the mentioned acceptance rates. We emphasize that this trial and error approach is cheap in our case because of the polynomial surrogate. However, recall from Sec. 2.5 that, in general, there exist more complex techniques to find suitable proposal distributions.

For the 2D subspace, choosing $\gamma_{2D}^2 = 0.02$ results in an acceptance rate of approximately 35%. The corresponding Markov chain consists of 10^6 steps with a chosen burn-in period of 10^5 samples. Using the formula from (2.87) to compute the effective sample size, we get that $N_{\text{eff}, 2D} \approx 75,000$ samples.

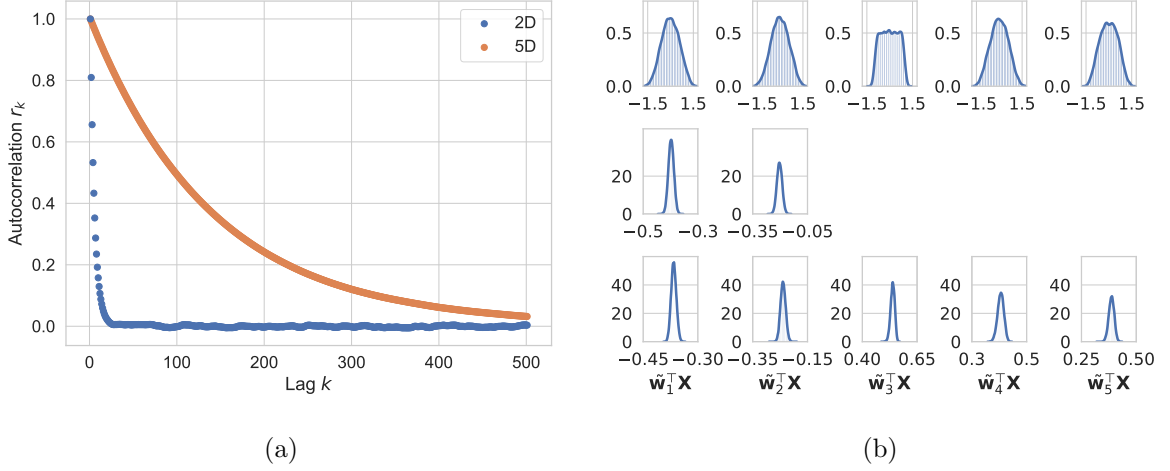


Figure 5.7: (a) Auto-correlations r_k with lag k of the second component \tilde{Y}_2 in the resulting 2D and 5D chain. (b) The top row shows histograms of the marginal prior distributions of the active variable $\tilde{\mathbf{Y}}$. The middle and bottom rows depict the approximate marginal posterior distributions of $\mu_{g_{\text{poly}}, \tilde{\mathbf{Y}}}^d$ computed in the 2D and 5D subspace, respectively.

For comparison, a proposal variance of $\gamma_{5D}^2 = 0.0017$ for the 5D case also gives an acceptance rate of about 35% in a run with 10^7 steps and a burn-in period of 10^6 samples resulting in $N_{\text{eff}, 5D} \approx 21,000$ effective samples.

Fig. 5.7a plots the auto-correlation r_k with lag k of the second component in both the 2D and 5D chain. We observe that the autocorrelation of the 5D chain decreases much slower than in the 2D case which is a common behavior. However, we argue that choosing the 5D subspace is nevertheless preferable since it yields a better posterior approximation with acceptable computational costs due to the cheap surrogate. According to the calculated auto-correlations, we finally apply thinning (see Sec. 4.3) to reduce the set of all samples to only (approximately uncorrelated) “effective” samples.

Histograms for the marginal prior distributions of the active variable $\tilde{\mathbf{Y}}$ (top row) and the approximate marginal posterior distributions of $\mu_{g_{\text{poly}}, \tilde{\mathbf{Y}}}^d$ for the 2D and 5D case (middle and bottom row) are displayed in Fig. 5.7b. It is worthwhile to first take a look at the marginal prior distributions in the top row. The distribution of $\tilde{Y}_3 = \tilde{\mathbf{w}}_3^\top \mathbf{X}$ is still quite uniform which is reasonable. Indeed, regarding the third eigenvector $\tilde{\mathbf{w}}_3$ in Fig. 5.5b, we see that almost only the fourth component contributes, i. e., $\tilde{Y}_3 \approx X_4 \sim \mathcal{U}([-1, 1])$. All the other histograms in the top row show distributions that are more concentrated around the mean zero which is typical for linear combinations (with sufficiently large weights) of uniform random variables. For the marginal posterior distributions, note that the x - and y -axes in each row have different scales. This difference confirms that the active variables $\tilde{\mathbf{Y}}$ are dominantly informed by the data as their posterior variances are greatly reduced compared to the prior variances. Moreover, note that the variables $\tilde{Y}_1 = \tilde{\mathbf{w}}_1^\top \mathbf{X}$ and $\tilde{Y}_2 = \tilde{\mathbf{w}}_2^\top \mathbf{X}$ are further informed from the 2D to the 5D case.

Eventually, we need to construct full samples in \mathbf{R}^n with the “effective” samples $\tilde{\mathbf{Y}}_j \in \mathbf{R}^k$ which approximately follow $\mu_{g_{\text{poly}}, \tilde{\mathbf{Y}}}^d$ and result from the RWM algorithm explained above. Recall from Sec. 4.3 that this means to construct samples $\tilde{\mathbf{Z}}_j \sim \mathbf{P}_{\tilde{\mathbf{Z}}|\tilde{\mathbf{Y}}}(\cdot|\tilde{\mathbf{Y}}_j)$. Since $\tilde{\mathbf{Y}}$ and $\tilde{\mathbf{Z}}$

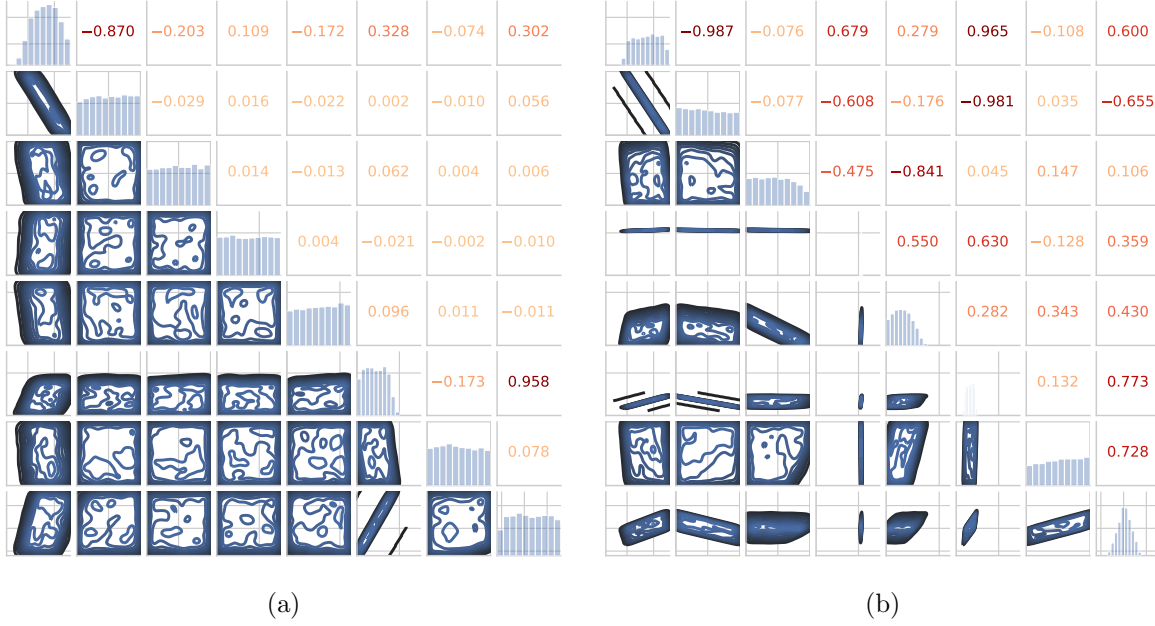


Figure 5.8: Univariate posterior histograms are shown on the diagonal. Bivariate posterior correlations are depicted on the lower triangular with corresponding correlation coefficients (see Def. A.2.3b)) reflected on the upper triangular. Higher coefficients (in absolute value) result in darker red color. (a) 2D case, (b) 5D case.

depend in a quite nontrivial manner due to the uniform prior distribution, we run another MCMC algorithm for each $\tilde{\mathbf{Y}}_j$ to approximately sample from

$$\rho_{\tilde{\mathbf{Z}}|\tilde{\mathbf{Y}}}(\tilde{\mathbf{z}}|\tilde{\mathbf{Y}}_j) \propto \rho_0(\|\tilde{\mathbf{Y}}_j, \tilde{\mathbf{z}}\|_{\tilde{W}}), \quad (5.30)$$

viewed as a density in $\tilde{\mathbf{z}}$, as described in Sec. 4.3. In fact, we compute a whole set of samples

$$\{\tilde{\mathbf{Z}}_{j,\ell} \mid \tilde{\mathbf{Z}}_{j,\ell} \stackrel{\text{approx.}}{\sim} \mathbf{P}_{\tilde{\mathbf{Z}}|\tilde{\mathbf{Y}}}(\cdot|\tilde{\mathbf{Y}}_j), \ell = 1, \dots, N_{\tilde{\mathbf{Z}}}\} \quad (5.31)$$

for every $\tilde{\mathbf{Y}}_j$ while ensuring that $N_{\tilde{\mathbf{Z}}}$ does not depend on $\tilde{\mathbf{Y}}_j$, i.e., the amount of samples $\tilde{\mathbf{Z}}_{j,\ell}$ is the same for each $\tilde{\mathbf{Y}}_j$. Then, the final set of samples approximately following the posterior on the full parameter space is

$$\{\|\tilde{\mathbf{Y}}_j, \tilde{\mathbf{Z}}_{j,\ell}\|_{\tilde{W}} \mid j = 1, \dots, N_{\text{eff}}, \ell = 1, \dots, N_{\tilde{\mathbf{Z}}}\}. \quad (5.32)$$

The set in (5.32) is computed for both the 2D and 5D active subspace. The resulting pairwise bivariate posterior relationships are plotted in Fig. 5.8. The left plot (Fig. 5.8a) shows the 2D and the right plot (Fig. 5.8b) the 5D case.

So far, the results shown are all related to the calibration parameters from (5.19). To actually derive quantitative statements for the model parameters, which are particularly interesting for the practitioner, we need to re-transform the samples from the set in (5.32) to the space of model parameters. Recall that relationship between model and calibration parameters is linear. The resulting approximate posterior means and standard deviations, in both the 2D and 5D case, are provided in Tab. 5.2.

Table 5.2: Approximate posterior means ($\tilde{\mu}_{\text{post}}$) and standard deviations ($\tilde{\sigma}_{\text{post}}$) for each model parameter using both the 2D and 5D subspace.

		c	α_{res}^1	$\Delta\alpha_{\text{res}}$	$\dot{\lambda}^*$
2D	$\tilde{\mu}_{\text{post}}$	2.169×10^6	0.551	0.251	1.749×10^{-3}
	$\tilde{\sigma}_{\text{post}}$	1.165×10^5	2.850×10^{-2}	2.882×10^{-2}	8.668×10^{-5}
5D	$\tilde{\mu}_{\text{post}}$	2.183×10^6	0.548	0.246	1.812×10^{-3}
	$\tilde{\sigma}_{\text{post}}$	1.148×10^5	2.874×10^{-2}	2.724×10^{-2}	2.626×10^{-6}
		m_α	β^*	λ^*	m_β
2D	$\tilde{\mu}_{\text{post}}$	0.903	0.343	1.049×10^{-2}	0.706
	$\tilde{\sigma}_{\text{post}}$	8.639×10^{-2}	2.427×10^{-2}	2.864×10^{-4}	1.946×10^{-2}
5D	$\tilde{\mu}_{\text{post}}$	0.827	0.332	1.053×10^{-2}	0.701
	$\tilde{\sigma}_{\text{post}}$	4.361×10^{-2}	7.745×10^{-3}	2.830×10^{-4}	7.248×10^{-3}

5.4 Discussion

Fig. 5.8a shows that the 2D case greatly informed the correlation between $X_1 = \bar{c}$ and $X_2 = \alpha_{\text{res}}^1$, and between $X_6 = \beta^*$ and $X_8 = \bar{m}_\beta$ which is reasonable since these components have larger weights/contributions in the first two eigenvectors $\tilde{\mathbf{w}}_1$ and $\tilde{\mathbf{w}}_2$. This implies that the cohesion parameter c and the initial residual friction parameter α_{res}^1 are dominant for the data misfit and thus constitute a central role for the methane hydrate model. Actually, this matches very well with the expectation that the initial yield surface, i. e.,

$$q + \alpha_{\text{res}}^1 - c = 0, \quad (5.33)$$

controls the onset of plasticity in any Coulomb-type plasticity model. Since the parameters c and α_{res}^1 can even technically be estimated with relatively high confidence through triaxial testing, we could perform an additional analysis of low-dimensional structures for other parameters for which we do not have straightforward estimation procedures.

Furthermore, looking at the 5D case in Fig. 5.8b, we see that additional uni- and bivariate relationships are informed as expected. In particular, this can be seen by the darker red colored upper triangular denoting the actual correlation coefficients. Also, the smaller standard deviations demonstrate that running the inversion in the 5D subspace leads to more informed posterior distributions compared to the 2D subspace which can also be observed in Tab. 5.2. We mention once more that the inference in the 5D subspace indeed requires more steps in the MCMC algorithm; but this computational overhead is acceptable due to the cheap polynomial surrogate, especially if the result is a more accurate posterior approximation.

Finally, we look at forward runs corresponding to the posterior means of both the 2D and the 5D case; see Tab. 5.2. That is, in both cases we compute

$$\mathcal{G}(\tilde{\boldsymbol{\mu}}_{\text{post}}) = \begin{pmatrix} \mathcal{G}_\epsilon(\tilde{\boldsymbol{\mu}}_{\text{post}}) \\ \mathcal{G}_\sigma(\tilde{\boldsymbol{\mu}}_{\text{post}}) \end{pmatrix}, \quad (5.34)$$

where $\tilde{\boldsymbol{\mu}}_{\text{post}}$ denotes the corresponding posterior mean in the space of calibration parameters. The outcomes are plotted in Fig. 5.9. We observe that all curves match the experimental

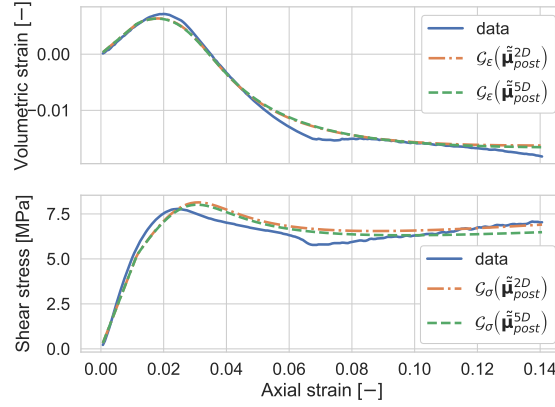


Figure 5.9: Volumetric strain (top) and shear stress (bottom) computed for posterior means of the 2D ($\tilde{\mu}_{\text{post}}^{2D}$) and 5D ($\tilde{\mu}_{\text{post}}^{5D}$) case.

data relatively well which particularly demonstrates that our inference procedure does not produce unreasonable posterior estimates. Also, while the 2D and 5D case give different posterior approximations, the deviation of the curves for both cases is rather small which can be explained by the fact that the data misfit function does not change significantly when we enlarge the subspace as is shown above (the coefficient of variation of the polynomial surrogate is almost equal for both cases). We conclude that, hence, both cases yield reasonable posterior approximations for the given quantity of interest.

In future work, we might consider alternative Bayesian inverse problems with different quantities of interest, noise levels, and parameters yielding results that better reflect other important physical characteristics of the volumetric strain and shear stress as, for example, the second hardening phase and peak locations and heights.

Calibration and sensitivity analysis for a karst hydrological model

Reference. The content of this chapter is based on our article in [258]. It is an outcome of the interdisciplinary project *UNcertainties due to boundary conditions in predicting MIXing in groundwater* (UNMIX) supported by *Deutsche Forschungsgemeinschaft* (DFG) through *TUM International Graduate School for Science and Engineering* (IGSSE), GSC 81.

Models are commonly used in karst systems to investigate the dominant hydrological processes and the quantity and quality of water resources in well-defined surface or subsurface catchments. Karst systems represent heterogeneous groundwater aquifers that are formed over longer periods of time due to an extensive dissolution of their dominantly carbonate bedrock, e. g., limestone, dolomite, or gypsum [91]. The formed dissolution channels constitute preferential flow paths in the underground which interact with the rock matrix and may outcrop as springs on the ground surface.

Various karst modeling approaches exist, ranging from black-box models [142, 161, 162], i. e., transferring an input signal to a desired output signal, over lumped parameter models (gray-box) [90, 190, 241] to distributed process-based models [97, 125, 220, 235]. Given their ability to represent the physical characteristics of a catchment in detail, distributed process-based models are usually the first choice in water resources research. In the particular case of karst aquifers, however, acquiring the relevant data for these models is challenging due to the heterogeneous nature of karstic systems and their mostly unknown subsurface drainage systems [288]. Also, past studies have shown that even if physical parameters may be obtained from field observations, the fact that they mostly represent point measurements can lead to a severe mismatch when using these parameters in distributed hydrological models [126, 227].

For these reasons, lumped process-based models are commonly accepted modeling approaches in karst water resources research [122, 141, 143]. The parameters of such lumped modeling approaches are typically not directly measurable in the field and need to be estimated in the framework of model calibration [119]. This leads to a decisive trade-off: on the one hand, lumped models based on a low number of calibration parameters, e. g., 4 to 6, are less prone to non-uniqueness in parameter identification [28, 135], i. e., different parameter combinations lead to the same result. However, the representation of the dominant hydrolog-

ical processes in karst systems may be too simple and not sufficiently represented by this low number of parameters [121]. In contrast, by including more calibration parameters to better represent relevant processes in the model structure, such as the effect of land use changes on spring discharges, the parameters may become unidentifiable, which can reduce the prediction accuracy of the model [120]. To tackle the challenge of applying lumped parameter models with a high-dimensional parameter space for karst hydrological research studies, there is a need to perform comprehensive parameter studies to avoid model overparametrization and to reduce model parameter and output uncertainties.

Approaches for inverse problems in hydro(geo)logy [45, 295] increasingly adopted the Bayesian viewpoint (Ch. 2) during the last two decades [87, 150, 259], especially due to the rise of available computational power. A particular popular MCMC approach in the hydrology community is the *DiffeRential Evolution Adaptive Metropolis* (DREAM) algorithm; see [163, 272, 273, 274, 275] for its definition, variants, and extensions; for a related discussion regarding questions on equifinality [28], see [29, 276]. The DREAM algorithm runs multiple chains in parallel and adaptively adjusts the scale and orientation of the proposal distribution according to *differential evolution* [250, 251] as genetic algorithm for population evolution. The algorithm was proven to maintain the detailed balance condition from (2.80) and demonstrated to be valuable for parameter inferences of high-dimensional hydrological models [163]. However, DREAM still performs inferences in the full space, although a variant based on “self-adaptive randomized subspace sampling” reduces the influence of dimensionality [273].

Similar to the previous chapter, the following sections present our approach to solve a Bayesian inverse problem involving a particular karst hydrological model by techniques described in Ch. 3 and 4. In addition, we provide a global sensitivity analysis giving insight in parameter-to-output relationships for almost no further computational costs. We again study dominant directions in the parameter space that drive the change from the prior to the posterior distribution and hence target the mentioned challenge of unidentifiable parameters in high-dimensional spaces. We mention that ASM was already applied to another hydrological model [137] which is, however, not suitable for karst systems.

The structure of the chapter is as follows. Sec. 6.1 gives details on the investigated spring recharge area and the hydrological model with its parameters. The setup for the corresponding Bayesian inverse problem is provided in Sec. 6.2. In Sec. 6.3, we describe how to derive global sensitivity metrics from the eigenvalues and eigenvectors computed for ASM. Finally, the computational results are presented in Sec. 6.4.

6.1 Case study

Reference. The description of the spring recharge area in Subsec. 6.1.1 is taken from [258, Subsec. 2.1] and included for the sake of completeness.

Subsec. 6.1.2 follows [258, Subsec. 2.2] for details on the model which was developed by Daniel Bittner (TUM) and his colleagues in [31].

6.1.1 Kerschbaum Spring Recharge Area

The karst spring that we investigate in the present study is the Kerschbaum spring located about 10 km south of the city of Waidhofen a.d. Ybbs (Austria); see Fig. 6.1a,b. Its recharge area was delimited in a former study by [114] and comprises about 2.5 km². This pre-alpine

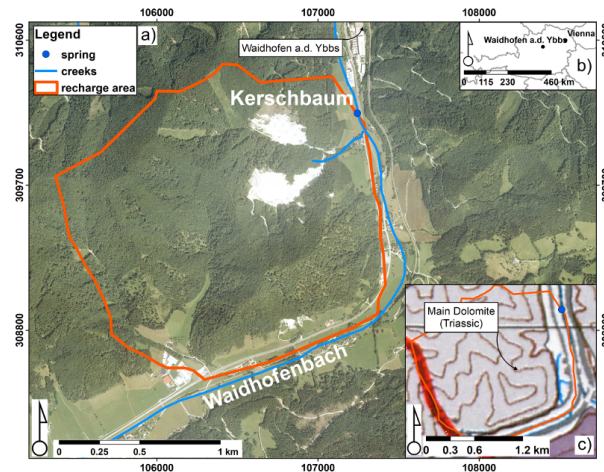


Figure 6.1: Overview of the characteristics of the Kerschbaum spring recharge area and its geographical localization. a) An orthophoto and the boundary of the recharge area with the location of the Kerschbaum spring. b) The geographical position of Waidhofen a.d. Ybbs in Austria. c) The dominant presence of dolomitic basement rocks in the catchment [95]. The isolines represent different elevation levels.

catchment is part of the eastern-most foothills of the Northern Calcareous Alps with the lowest elevation of 435 m at the Kerschbaum spring and a maximum elevation of 868 m on the summit of the mountain Glashüttenberg. The climate of the study area can be described as warm-moderate, with an annual mean temperature of 8° C and an annual mean precipitation of 1379 mm, both determined from daily measuring data recorded at the Hinterlug weather station between 1981 and 2014. Forests represent the dominant land cover in the study area with beeches as primary tree species. Moreover, parts of the recharge area are used for dolomite mining.

From a geological point of view, the entire recharge area of the Kerschbaum spring is dominated by a lithologic sequence of Triassic dolostones; see Fig. 6.1c). Apart from the absence of significant sinkholes in the regarded recharge area, leading to the fact that diffuse infiltration plays a key role for groundwater recharge, [114] also provided evidence for a deep karstified aquifer system with a well-connected drainage system through fractures and conduits in the Kerschbaum spring aquifer. It is important to note that the Kerschbaum spring represents the most important source for the freshwater supply of the city and the surroundings of Waidhofen and is thus of particular interest for water resources research studies [31].

6.1.2 Model

The LuKARS model was recently proposed by [31] with the aim to investigate the hydrological effects of land use changes in karst systems. LuKARS therefore considers the dominant hydrotopes in a defined recharge area, i. e., areas characterized by homogeneous soil and land cover properties, as distinct spatial units. The sum of the individual hydrotope responses to a given input signal (e. g., precipitation) plus the contribution of a shared linear baseflow storage is then the total spring discharge that should be modeled at a catchment's outlet.

As input data, the model itself needs a precipitation time series as well as the hydrotope soil information to run. If further processes affecting the effective precipitation are considered,

such as interception and evapotranspiration, further input data is required. In our case, we also take into account snow melt and accumulation, interception, and evapotranspiration, for which we further need a temperature time series with a daily resolution.

Moreover, a measured discharge time series is needed from the spring of interest to calibrate and validate the model. In the particular case of the Kerschbaum spring, the discharge is measured with a flowmeter directly in the spring. The discharge, precipitation, and temperature time series with a daily resolution for our model period from 2006 to 2008 were kindly provided by the water works Waidhofen a.d. Ybbs.

The LuKARS model for the Kerschbaum spring in Waidhofen a.d. Ybbs was set up in [31] and includes four spatially lumped dominant hydrotopes in the considered recharge area. Hydrotopes 1–3 have beeches as dominant tree species; however, they differ in terms of their individual soil characteristics and spatial shares. While the first hydrotope (denoted by Hyd 1) covers 13% of the recharge area and is characterized by shallow soils with mostly coarse-grained soil particles, hydrotope 3 (denoted by Hyd 3), in contrast, covers 27% of the catchment and is defined by deeper and fine textured soils. Hydrotope 2 (denoted by Hyd 2) has the largest spatial share in the Kerschbaum spring recharge area (56%) and represents a transition hydrotope between Hyd 1 and Hyd 3 with moderate soil thicknesses and coarse to fine-textured soils. Hydrotope Q (denoted by Hyd Q) characterizes the dolomite quarries, which covered about 4% of space in the recharge area during the model period (2006–2008) in this study.

From a hydrological point of view, the areas of the dolomite quarries are drained by surface runoff and do not contribute to the Kerschbaum spring discharge. As an obligation to avoid a possible contamination of the aquifer from the quarry areas, a protective layer consisting of fine material prevents infiltration into the groundwater system. Thus, Hyd Q is excluded from model calibration and will not be mentioned hereafter. Also, [31] derived the baseflow coefficient k_b to match the relatively constant baseflow discharge of the Kerschbaum spring with its low temporal variability. For this reason, as well as to put the focus on calibrating the hydrotope parameters, k_b was chosen as calibrated by [31]. More details about the LuKARS model, i. e., a description of the equations used in LuKARS and the relevant parameters, are provided in Appendix C. In the following, we use an index $i \in \{1, 2, 3\}$ to denote specifications for Hyd i .

Each hydrotope is modeled as an independent bucket that has three different discharge components. The first, representing quickflow ($Q_{\text{hyd},i}$) occurring via preferential flow paths (e. g., conduits), is described by a non-linear hysteresis function that is activated once a defined storage threshold ($e_{\text{max},i}$) is reached and stops after the storage value falling below a predefined minimum storage value ($e_{\text{min},i}$). The second and third discharge components are both implemented by a linear discharge function and represent the discharge to a shared baseflow storage ($Q_{\text{is},i}$) as well as secondary spring discharge ($Q_{\text{sec},i}$), i. e., a discharge component that transfers water out of the catchment and does not contribute to the spring discharge.

All together, seven parameters need to be calibrated for the implementation of each single hydrotope. These are the discharge parameter $k_{\text{hyd},i}$ and the dimensionless exponent α_i for $Q_{\text{hyd},i}$, the storage thresholds for the quickflow activation $e_{\text{min},i}$ and $e_{\text{max},i}$, parameter $k_{\text{is},i}$ as the discharge coefficient of $Q_{\text{is},i}$ and, finally, $k_{\text{sec},i}$ and $e_{\text{sec},i}$ as the discharge coefficient and the activation level for $Q_{\text{sec},i}$, respectively.

Given the different physical characteristics of all defined hydrotopes, the parameters of one hydrotope need to follow some constraints with respect to the parameters used for the implementation of other hydrotopes. From a practical point of view, this means that a hydrotope

with shallow and coarse-grained soils (e.g., Hyd 1) needs to have a lower storage capacity and higher discharge coefficient as compared to a hydrotope with deep and fine-textured soils (e.g., Hyd 3). For the particular case of the three hydrotopes in the Kerschbaum spring recharge area, the parameter constraints are given as

$$\begin{aligned}
k_{\text{hyd},1} &\geq k_{\text{hyd},2} \geq k_{\text{hyd},3}, \\
e_{\text{min},1} &\leq e_{\text{min},2} \leq e_{\text{min},3}, \\
e_{\text{max},1} &\leq e_{\text{max},2} \leq e_{\text{max},3}, \\
\alpha_1 &\geq \alpha_2 \geq \alpha_3, \\
k_{\text{is},1} &\geq k_{\text{is},2} \geq k_{\text{is},3}, \\
k_{\text{sec},1} &\geq k_{\text{sec},2} \geq k_{\text{sec},3}, \\
e_{\text{sec},1} &\leq e_{\text{sec},2} \leq e_{\text{sec},3}.
\end{aligned} \tag{6.1}$$

Also, it has to hold that

$$e_{\text{min},i} \leq e_{\text{max},i} \tag{6.2}$$

for each hydrotope $i = 1, 2, 3$. Although the introduced condition for the α_i values is not strictly necessary, we implemented it to further enhance the quick response of hydrotopes with a low difference between $e_{\text{min},i}$ and $e_{\text{max},i}$ and a generally low value of $e_{\text{max},i}$ during high precipitation events.

The LuKARS model for the Kerschbaum spring recharge area was manually calibrated in [31]. Based on this trial-and-error calibration, it was possible to reliably determine possible ranges of all model parameters. These are shown in Tab. 6.1 and will be used as prior parameter intervals for the presented study in a Bayesian setting.

6.2 Setup

The next subsections provide the setting for the Bayesian inference. In particular, we specify the set of model and calibration parameters, the assumed prior distribution, the quantities of interest, and the noise level that is assumed on the experimental data.

Model parameters

The model parameters are all the quantities appearing in (6.1), i.e.,

$$k_{\text{hyd},i}, e_{\text{min},i}, e_{\text{max},i}, \alpha_i, k_{\text{is},i}, k_{\text{sec},i}, \text{ and } e_{\text{sec},i}$$

for each hydrotope $i = 1, 2, 3$. Intervals for the prior distribution that we assume on the space of model parameters are provided in Tab. 6.1. We should mention that the bounds for $e_{\text{max},i}$ were derived by considering a given difference value between $e_{\text{min},i}$ and $e_{\text{max},i}$, i.e., for each hydrotope $i = 1, 2, 3$, we set

$$e_{\text{max},i,\text{lb}} = e_{\text{min},i,\text{lb}} + \Delta e_{i,\text{lb}} \quad \text{and} \quad e_{\text{max},i,\text{ub}} = e_{\text{min},i,\text{ub}} + \Delta e_{i,\text{ub}}, \tag{6.3}$$

where the subscripts lb and ub denote lower and, respectively, upper bounds of $e_{\text{min},i}$ and $e_{\text{max},i}$ values. The difference values are denoted by $\Delta e_{i,\text{lb}}$ and $\Delta e_{i,\text{ub}}$.

Table 6.1: Prior intervals for model parameters.

No.	Model par.	Lower bound	Upper bound	Unit	Description
1	$k_{\text{hyd},1}$	9	900	m^2d^{-1}	discharge parameter for $Q_{\text{hyd},1}$
2	$e_{\text{min},1}$	10	50	mm	min. storage capacity Hyd 1
3	$e_{\text{max},1}$	15	75	mm	max. storage capacity Hyd 1
4	α_1	0.7	1.6	–	quickflow exponent of Hyd 1
5	$k_{\text{is},1}$	0.002	0.2	$\text{m mm}^{-1}\text{d}^{-1}$	discharge parameter for $Q_{\text{is},1}$
6	$k_{\text{sec},1}$	0.0095	0.95	$\text{m mm}^{-1}\text{d}^{-1}$	discharge parameter for $Q_{\text{sec},1}$
7	$e_{\text{sec},1}$	25	70	mm	activation level for $Q_{\text{sec},1}$
8	$k_{\text{hyd},2}$	8.5	850	m^2d^{-1}	discharge parameter for $Q_{\text{hyd},2}$
9	$e_{\text{min},2}$	40	80	mm	min. storage capacity Hyd 2
10	$e_{\text{max},2}$	80	160	mm	max. storage capacity Hyd 2
11	α_2	0.5	1.3	–	quickflow exponent of Hyd 2
12	$k_{\text{is},2}$	0.00055	0.055	$\text{m mm}^{-1}\text{d}^{-1}$	discharge parameter for $Q_{\text{is},2}$
13	$k_{\text{sec},2}$	0.0023	0.23	$\text{m mm}^{-1}\text{d}^{-1}$	discharge parameter for $Q_{\text{sec},2}$
14	$e_{\text{sec},2}$	130	220	mm	activation level for $Q_{\text{sec},2}$
15	$k_{\text{hyd},3}$	7.7	770	m^2d^{-1}	discharge parameter for $Q_{\text{hyd},3}$
16	$e_{\text{min},3}$	75	120	mm	min. storage capacity Hyd 3
17	$e_{\text{max},3}$	160	255	mm	max. storage capacity Hyd 3
18	α_3	0.2	0.7	–	quickflow exponent of Hyd 3
19	$k_{\text{is},3}$	0.00025	0.025	$\text{m mm}^{-1}\text{d}^{-1}$	discharge parameter for $Q_{\text{is},3}$
20	$k_{\text{sec},3}$	0.0015	0.15	$\text{m mm}^{-1}\text{d}^{-1}$	discharge parameter for $Q_{\text{sec},3}$
21	$e_{\text{sec},3}$	320	450	mm	activation level for $Q_{\text{sec},3}$

Calibration parameters

Since we distinguish between model and calibration parameters, we need to specify the transformation rules which are not as trivial as in the previous chapter.

First, all the k_* parameters are calibrated on a log scale, i. e., we define

$$k_*^{\log} := \log(k_*) \quad (6.4)$$

for each $k_* \in \{k_{\text{hyd},i}, k_{\text{is},i}, k_{\text{sec},i}\}$, $i = 1, 2, 3$.

Secondly, we need to ensure that the constraints in (6.1) and (6.2) are fulfilled. In other words, the dependence between the model parameters is removed by the definition of new parameters that reflect the *differences* between two dependent parameters. We have two types of dependencies in the present case:

1. cross-hydrotope dependencies according to (6.1),
2. dependency between $e_{\text{min},i}$ and $e_{\text{max},i}$ due to (6.2).

The first point concerns only parameters in hydrotope 2 and 3. We set

$$p_i := p_{i,\text{lb}} + \Delta p_{(i-1,i)}(\min\{p_{i,\text{ub}}, p_{i-1}\} - p_{i,\text{lb}}) \quad (6.5)$$

or

$$p_i = \max\{p_{i-1}, p_{i,\text{lb}}\} + \Delta p_{(i-1,i)}(p_{i,\text{ub}} - \max\{p_{i-1}, p_{i,\text{lb}}\}), \quad i = 2, 3, \quad (6.6)$$

depending on whether the model parameter

$$p_i \in \left\{ k_{\text{hyd},i}^{\log}, e_{\text{min},i}, \alpha_i, k_{\text{is},i}^{\log}, k_{\text{sec},i}^{\log}, e_{\text{sec},i}^{\log} \right\} \quad (6.7)$$

follows a decreasing or, respectively, increasing behavior; see (6.1). The fixed values for $p_{i,\text{lb}}$ and $p_{i,\text{ub}}$ denote the lower and upper bound of the prior intervals for model parameters from

Tab. 6.1. The quantities $\Delta p_i \in [0, 1]$, which are all independent, are included in the set of calibration parameters. To meet the second dependency, we bring in parameters Δe_i , $i = 1, 2, 3$, that reflect the *difference* between model parameters $e_{\min,i}$ and $e_{\max,i}$, i. e., we have the relationship

$$e_{\max,i} = e_{\min,i} + \Delta e_i \quad (6.8)$$

for each $i = 1, 2, 3$. Note that Δe_i is independent of $e_{\min,i}$ in contrast to $e_{\max,i}$. The minimum and maximum values for Δe_i are determined by respective lower and upper bounds from Tab. 6.1, i. e.,

$$\Delta e_i \in [e_{\max,i,\text{lb}} - e_{\min,i,\text{lb}}, e_{\max,i,\text{ub}} - e_{\min,i,\text{ub}}]. \quad (6.9)$$

As a last step, all calibration parameters have to be normalized, i. e., their original prior interval is translated and scaled to the interval $[-1, 1]$. That is, the final (random) vector of all normalized calibration parameters, which are denoted again with a bar, is

$$\begin{aligned} \mathbf{x} = & (\overline{k_{\text{hyd},1}^{\log}}, \overline{e_{\min,1}}, \overline{\Delta e_1}, \overline{\alpha_1}, \overline{k_{\text{is},1}^{\log}}, \overline{k_{\text{sec},1}^{\log}}, \overline{e_{\text{sec},1}}, \\ & \overline{\Delta k_{\text{hyd},(1,2)}^{\log}}, \overline{\Delta e_{\min,(1,2)}}, \overline{\Delta e_2}, \overline{\Delta \alpha_{(1,2)}}, \overline{\Delta k_{\text{is},(1,2)}^{\log}}, \overline{\Delta k_{\text{sec},(1,2)}^{\log}}, \overline{\Delta e_{\text{sec},(1,2)}}), \\ & \overline{\Delta k_{\text{hyd},(2,3)}^{\log}}, \overline{\Delta e_{\min,(2,3)}}, \overline{\Delta e_3}, \overline{\Delta \alpha_{(2,3)}}, \overline{\Delta k_{\text{is},(2,3)}^{\log}}, \overline{\Delta k_{\text{sec},(2,3)}^{\log}}, \overline{\Delta e_{\text{sec},(2,3)}})^{\top} \in \mathbf{R}^{21}. \end{aligned} \quad (6.10)$$

The prior distribution is assumed on the space of these calibration parameters and we again choose

$$\mu_0 = \mathbf{P}_{\mathbf{X}} = \mathcal{U}([-1, 1]^{21}) \quad (6.11)$$

implying that $\mathcal{X} = [-1, 1]^{21}$. A particular sample $\mathbf{X} \sim \mu_0$ can be transformed to a model parameter by the rules above to serve as input for a model run. Note that the (transformed) prior distribution on the model parameters is *not* uniform since the transformations are non-linear and “indirect,” in the sense that there is no direct analogue to, for example, Δe_i in the set of model parameters.

Quantities of interest

The hydrological model in this chapter maps a certain instance of model parameters to the corresponding discharge values which constitute our quantities of interest. That is, all in all, we have a forward operator $\mathcal{G} : \mathcal{X} \rightarrow \mathbf{R}^{n_d}$, where $n_d = 1096$. The actual measured discharge values $\mathbf{d} \in \mathbf{R}^{n_d}$, that we need as a condition argument for the posterior $\mu^{\mathbf{d}}$, are plotted in Fig. 6.2.

Also here, we assume a zero-centered Gaussian distributed experimental noise at a level of 5% for the used flowmeter as was suggested by the water works owner in Waidhofen a.d. Ybbs. More concretely, the noise covariance matrix Γ becomes

$$\Gamma_{ij} = (0.05 \times d_i)^2 \delta_{ij} \quad (6.12)$$

for $i, j = 1, \dots, n_d$, where d_i denotes the i -th component of the data \mathbf{d} . That is, Γ is diagonal.

6.3 Global sensitivity analysis

Approaches for global sensitivity analysis (GSA) are a standard tool in the field of UQ as is already mentioned in the introductory chapter; see Sec. 1.2. Recall that we briefly introduced

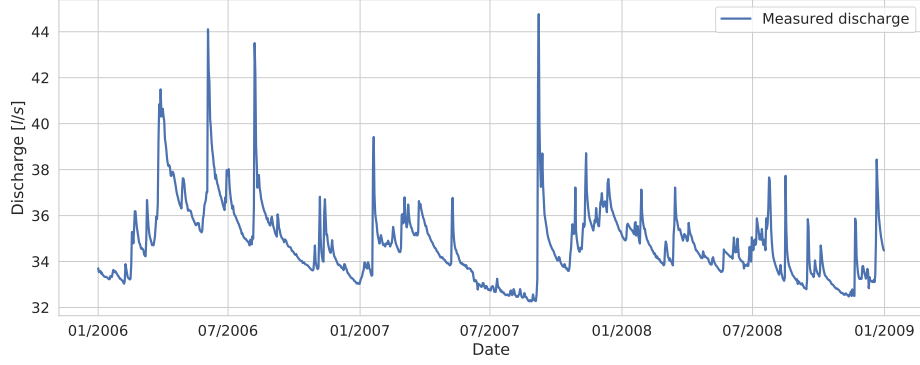


Figure 6.2: Measured discharge data for the calibration period 2006–2008.

total-effect Sobol indices as a technique to quantify the expected variance of a given function f caused by a certain single input variable; see (1.10). We repeat and emphasize that this type of sensitivity index is coordinate-aligned, i. e., it refers to a single original input variable X_i . This is in contrast to sensitivities computed by ASM which are linear combinations of input variables and thus yield sensitive directions in the input space. However, since sensitive directions contain more information than just coordinate-aligned sensitivities, there is hope that the former can be constructed by the latter. Indeed, the authors of [53] demonstrate that the eigendecomposition of the matrix C in (3.6) can be exploited to compute sensitivity metrics that are comparable to total-effect Sobol indices in most practical situations. But they also mention that it is possible to construct scenarios which lead to different, non-comparable results.

Particularly, in [53], the eigenvectors in W are weighted with corresponding eigenvalues from Λ to derive global sensitivity indices on coordinates. These indices are stored in a vector $\mathbf{s} \in \mathbf{R}^n$ in which the value of the i -th component reflects the global sensitivity of f w.r.t. X_i , i. e., we define

$$s_i := s_i(n_{\text{GS}}) := \sum_{j=1}^{n_{\text{GS}}} \lambda_j(\mathbf{w}_j)_i^2 \quad (6.13)$$

for $i = 1, \dots, n_{\text{GS}}$, where $n_{\text{GS}} \leq n$. Since we set $n_{\text{GS}} = n$ in the following, we can write more compactly that

$$\mathbf{s}(n) = (W \circ W)\boldsymbol{\lambda}, \quad (6.14)$$

where $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_n)^\top \in \mathbf{R}^n$ and \circ denotes elementwise multiplication.

We should also mention at this point (again) that the exact quantities W and Λ are often not available in practice meaning that we can only use approximations \tilde{W} and $\tilde{\Lambda}$ to compute an approximate vector of global sensitivities $\tilde{\mathbf{s}} \approx \mathbf{s}$. It is difficult to specify strict lower bounds on the number of samples $N_{\tilde{C}}$ in (3.159) in general to compute sufficiently accurate sensitivity values $\tilde{\mathbf{s}}$. For this reason, we follow the results in [53] and use as many samples as were shown to be sufficient therein.

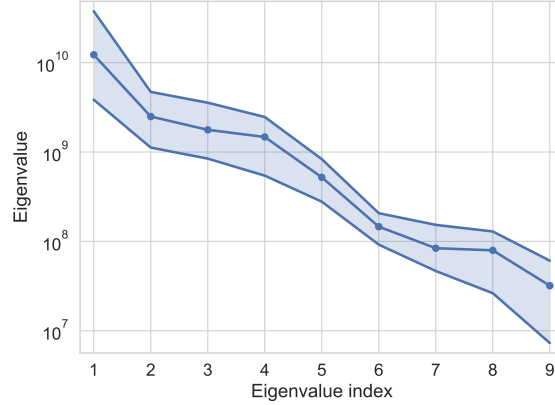


Figure 6.3: Spectrum of the matrix \tilde{C} for the karst hydrological model LuKARS.

6.4 Computational results

Computational costs

With a pessimistic sampling factor $\beta_{\tilde{C}} = 10$, we would need to compute about $N_{\tilde{C}} = 250$ gradient samples according to (3.166) to approximate the first $m = 8$ eigenvalues of \tilde{C} from (3.159) sufficiently accurately. Nevertheless, we decide to choose $N_{\tilde{C}} = 1,000$ gradient samples since we want to ensure that the global sensitivities, mentioned in Subsec. 6.3, are also approximated well enough.

Recall from (4.12) that we need the Jacobian matrix $\nabla \mathcal{G}$ of the forward operator \mathcal{G} to compute gradients of the data misfit function f^d . Since the LuKARS model is not capable of exact gradients, we need to approximate them by central finite differences analogous to the previous chapter; see (5.27). Using seven CPU cores, each at 3 GHz, the required

$$N_{\tilde{C}} \cdot (2n + 1) = 1,000 \cdot (2 \cdot 21 + 1) = 43,000 \quad (6.15)$$

forward runs for the computation of \tilde{C} require about 4.3 hours with an average of about 2.5 seconds for a single model run.

Active subspace for the inference and global sensitivities

The resulting eigenvalues of \tilde{C} are plotted in Fig. 6.3. The light blue band around the eigenvalues depicts the estimated variability due to the random nature of \tilde{C} as described in Sec. 3.5. We observe that there is a first spectral gap between the first and second eigenvalue, and a second gap between the fourth and fifth eigenvalue which might be exploited for our purpose of dimension reduction. However, these gaps are not as large as in the previous chapter, i.e., we expect that the respective surrogates do not have such a remarkably high goodness of fit as the polynomial response surface for the methane hydrate model.

Fig. 6.4 shows the corresponding first four eigenvectors $\tilde{w}_1, \dots, \tilde{w}_4$ of \tilde{C} and the global sensitivities computed according to (6.13). The colors of the bars visualize the different hydrotopes. Note that we included all eigenpairs to compute the approximate vector of sensitivities \tilde{s} , i.e., we have that $n_{\text{GS}} = n$. The first three eigenvectors (Fig. 6.4a–c) are all dominated by

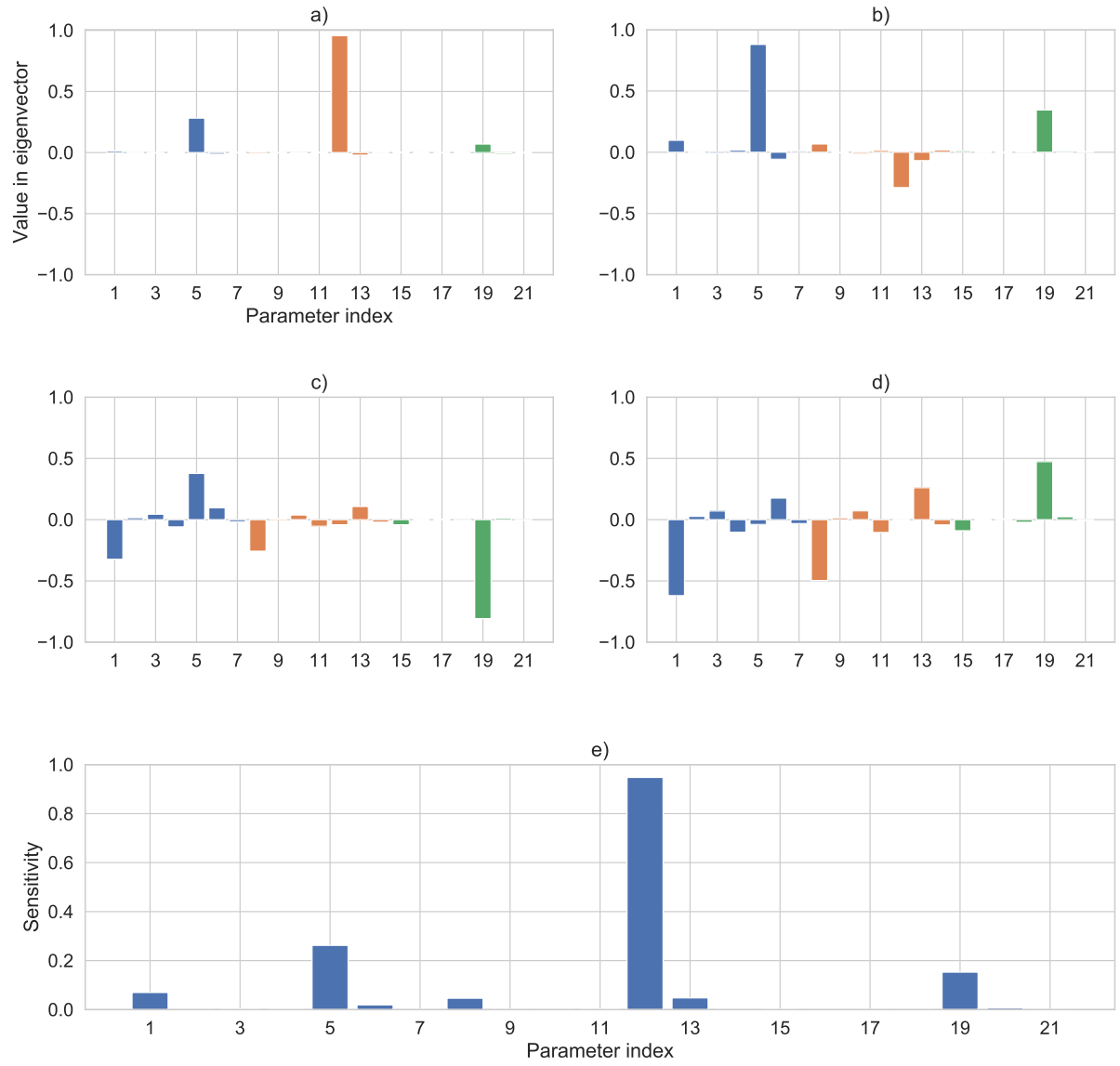


Figure 6.4: **a)–d)**: First four eigenvectors $\tilde{w}_1, \dots, \tilde{w}_4$ of \tilde{C} . The colors of the bars visualize the different hydrotopes. (Blue: Hyd 1; Orange: Hyd 2; Green: Hyd 3) **e)**: Approximate global sensitivity values \tilde{s} computed according to (6.13).

the parameters X_5 , X_{12} , and X_{19} which interestingly all involve the $k_{\text{is},i}$ values for each of the three hydrotopes. Recall that the eigendecomposition is performed for the calibration, and not the model parameters. That is, in descending dominance, we have

$$1. \overline{\Delta k_{\text{is},(1,2)}^{\log}}, 2. \overline{k_{\text{is},1}^{\log}}, 3. \overline{\Delta k_{\text{is},(2,3)}^{\log}},$$

which is, of course, also reflected in the global sensitivities; see Fig. 6.4e. The fourth eigenvector (Fig. 6.4d) displays that also the parameters

$$X_1 = \overline{k_{\text{hyd},1}^{\log}} \text{ and } X_8 = \overline{\Delta k_{\text{hyd},(1,2)}^{\log}}$$

involving k_{hyd} values from hydrotopes 1 and 2 have some importance.

To actually start the inference using the RWM algorithm in the active subspace (Alg. 5), we still need to construct a polynomial surrogate

$$g_{\text{poly}} \approx \tilde{g}_N^d \quad (6.16)$$

for the low-dimensional data misfit function with $N = 1$ by reusing the samples $f^d(\mathbf{X}_j)$, that were already computed for the approximation \tilde{C} of C , for Alg. 6. This procedure is already explained in Sec. 5.3 for the methane hydrate model.

We decide for a 4D subspace since we regard that

$$\frac{\tilde{\lambda}_5 + \dots + \tilde{\lambda}_n}{\tilde{\lambda}_1 + \dots + \tilde{\lambda}_n} \approx 0.0475 = 4.75\%, \quad (6.17)$$

as the percentage share of the eigenvalues belonging to the inactive subspace to the sum of all eigenvalues, is sufficiently small. The corresponding polynomial surrogate, having fourth order, results in a coefficient of determination of $r_{4D}^2 \approx 0.77$ which is assessed as sufficient but, as expected, not as good as for the methane hydrate model. Since a fourth order polynomial with four inputs already has 70 degrees of freedom, we compute another 20,000 samples $f^d(\mathbf{X}_\ell)$, where $\mathbf{X}_\ell \sim \mu_0$, to ensure to avoid overfitting. The corresponding r^2 score results in a very similar value ≈ 0.77 confirming that we do not run into the mentioned issue.

Remark. For completeness, a fourth order polynomial on a 1D subspace yields a coefficient of determination of $r_{1D}^2 \approx 0.23$ which is certainly not good enough.

Inference with MCMC in the active subspace

The polynomial response surface g_{poly} is consequently exploited for running the adjusted RWM algorithm in the active subspace (Alg. 5). Recall that we want to construct samples approximately following the posterior distribution on the active subspace denoted by $\mu_{g_{\text{poly}}, \tilde{\mathbf{Y}}}^d$.

Note again that we only have perturbations $\tilde{\mathbf{Y}}$, $\tilde{\mathbf{Z}}$, and g_{poly} available which implies that the resulting posterior approximation is computed on the perturbed subspace with the polynomial surrogate.

Since the general procedure in this subsection coincides with the methodology from the respective subsection of the previous chapter, we only show the results that are related to the present hydrological study case and do not repeat the explanations.

The variance of the proposal distribution $\mathbf{Q} \sim \mathcal{N}(0, \gamma^2 I)$ is adjusted to $\gamma^2 = 0.005$ to result in an acceptance rate of 35% which is reasonably good according to Sec. 2.5. We compute

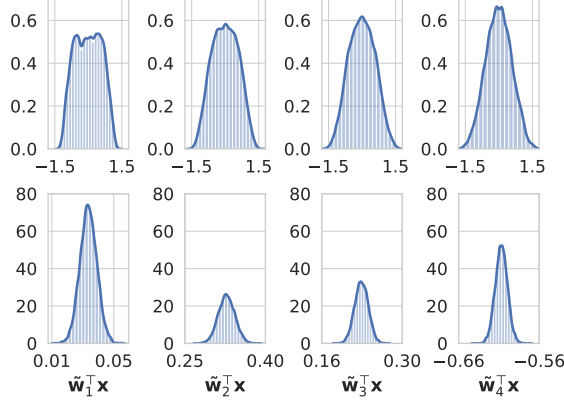


Figure 6.5: The top row shows histograms of the marginal prior distributions of the active variable $\tilde{\mathbf{Y}}$. The bottom row depicts the approximate marginal posterior distributions of $\mu_{g_{poly}, \tilde{\mathbf{Y}}}^d$.

10^6 (correlated) samples and regard the first 10^5 as part of the burn-in period yielding an effective sample size $N_{\text{eff}} \approx 12,000$.

The marginal distributions of $\mu_{g_{poly}, \tilde{\mathbf{Y}}}^d$ are shown in Fig. 6.5. The x - and y -axes, substantially changing from the prior (top row) to the posterior (bottom row), display that the posterior distribution in the active subspace is considerably informed as expected and desired. We adjusted the x -axes in the bottom row to make the variance visible; otherwise, the distributions would have been very thin lines which display only little information.

In order to have samples in the full space following $\mu_{g_{poly}}^d \approx \mu_g^d$, we follow the same approach as in Sec. 5.3 to compute conditional samples

$$\tilde{\mathbf{Z}}_{j,\ell} \sim \mathbf{P}_{\tilde{\mathbf{Z}}|\tilde{\mathbf{Y}}}(\cdot|\tilde{\mathbf{Y}}_j) \quad (6.18)$$

for each $\tilde{\mathbf{Y}}_j \stackrel{\text{approx.}}{\sim} \mu_{g_{poly}, \tilde{\mathbf{Y}}}^d$ from the reduced set of “effective” samples and subsequently put them together as in (5.32).

These samples in the full space are then translated back to the space of model parameters and give results shown in Tab. 6.2. On the left, we see the approximate posterior means and standard deviations for each model parameter from Tab. 6.1. The means and standard deviations of model parameters that are informed substantially are highlighted in bold. As expected by the global sensitivity values from Fig. 6.4e, the $k_{\text{is},i}$ parameters are mostly informed by the discharge data \mathbf{d} . For bivariate relationships, the right subtable displays the highest (in absolute value) *sample Pearson correlation coefficients* (see Def. A.2.3b)) of the model parameters. Note that the smallest negative correlations are smaller, in absolute value, than the largest positive ones. The correlation of $e_{\text{min},i}$ and $e_{\text{max},i}$ values, as well as the correlation of $k_{\text{is},1}$ and $k_{\text{is},2}$, is due to the constraints in (6.1) and thus already present in the prior on the model parameters. Most of the other correlations, as e.g., of $k_{\text{hyd},2}$ and $k_{\text{sec},2}$ or of $k_{\text{hyd},2}$ and $k_{\text{is},2}$, are found by the actual inference. It is worth mentioning that the highest correlations found by the inference only involve k_* parameters.

As a kind of validation, we show a plot of the approximate prior and posterior push-forward distributions, i. e., of

$$\mu_0(\{\mathbf{x} \in \mathcal{X} | \mathcal{G}(\mathbf{x}) \in \cdot\}) \text{ and } \mu_{g_{poly}}^d(\{\mathbf{x} \in \mathcal{X} | \mathcal{G}(\mathbf{x}) \in \cdot\}), \quad (6.19)$$

Table 6.2: **Left:** Posterior means and standard deviations of model parameters. (The informed parameters are highlighted in bold.) **Right:** Highest 2D correlations for model parameters. The smallest negative correlations are smaller, in absolute value, than the largest positive ones.

No.	Model par.	Mean	Std.
1	$k_{\text{hyd},1}$	3.07×10^2	2.34×10^2
2	$e_{\text{min},1}$	29.86	11.57
3	$e_{\text{max},1}$	44.49	12.90
4	α_1	1.17	0.26
5	$k_{\text{is},1}$	5.18×10^{-2}	3.98×10^{-3}
6	$k_{\text{sec},1}$	0.17	0.22
7	$e_{\text{sec},1}$	47.78	12.95
8	$k_{\text{hyd},2}$	70.62	55.81
9	$e_{\text{min},2}$	60.46	11.27
10	$e_{\text{max},2}$	1.20×10^2	16.14
11	α_2	0.82	0.21
12	$k_{\text{is},2}$	4.52×10^{-3}	1.61×10^{-4}
13	$k_{\text{sec},2}$	2.03×10^{-2}	3.23×10^{-2}
14	$e_{\text{sec},2}$	1.76×10^2	25.99
15	$k_{\text{hyd},3}$	25.94	21.75
16	$e_{\text{min},3}$	95.71	14.18
17	$e_{\text{max},3}$	2.06×10^2	20.23
18	α_3	0.43	0.14
19	$k_{\text{is},3}$	6.35×10^{-4}	1.69×10^{-5}
20	$k_{\text{sec},3}$	6.21×10^{-3}	1.07×10^{-2}
21	$e_{\text{sec},3}$	3.85×10^2	37.48

Model par.		Corr. coef.
$e_{\text{min},1}$	$e_{\text{max},1}$	0.89
$k_{\text{is},1}$	$k_{\text{is},2}$	0.77
$e_{\text{min},3}$	$e_{\text{max},3}$	0.70
$e_{\text{min},2}$	$e_{\text{max},2}$	0.70
$k_{\text{hyd},2}$	$k_{\text{sec},2}$	0.66
$k_{\text{is},1}$	$k_{\text{hyd},2}$	0.64
$k_{\text{is},1}$	$k_{\text{sec},2}$	0.63
$k_{\text{hyd},2}$	$k_{\text{is},2}$	0.59
$k_{\text{sec},1}$	$k_{\text{is},3}$	0.57
$k_{\text{sec},2}$	$k_{\text{sec},3}$	0.56
$k_{\text{is},2}$	$k_{\text{sec},2}$	0.52

which are measures on the data space. That is, we randomly take 1,000 samples of both the prior μ_0 and the posterior $\mu_{g_{\text{poly}}}^d$ constructed by the RWM algorithm explained above and run the forward operator \mathcal{G} for each of them. Note that these computations are only possible since the model is cheap to evaluate which is not the case for the methane hydrate model from Ch. 5.

Fig. 6.6 depicts the measured discharge values (blue lines) and the sample medians (orange lines) and means (green lines) of both the prior (Fig. 6.6a) and the approximate posterior (Fig. 6.6b) push-forward distribution for comparison over the whole study period. Also, Fig. 6.6 shows 95% quantile bands, i. e., the interval between the 2.5% and 97.5% quantile, of the data (noise) distribution (light blue areas) and the two push-forward distributions (light orange areas). It is apparent that the inference in the 4D subspace not only yields more suitable means and medians but also substantially reduces the uncertainty on the data space as can be seen by the smaller quantile band on the right in Fig. 6.6b.

6.5 Discussion

The following discussion on the most sensitive parameters is taken from [258, Sec. 5] to which we also refer for a general hydrological interpretation of our results.

Parameters X_5 , X_{12} , and X_{19} show the largest contributions in the first four eigenvectors $\tilde{w}_1, \dots, \tilde{w}_4$ in Fig. 6.4a–d. These parameters correspond to $k_{\text{is},i}$ model parameters, which delimit the flow contributions from the hydrotopes to the linear baseflow storage. As derived in [31], the baseflow storage exhibits a relatively constant discharge behavior with a small temporal variability and its discharge coefficient k_b was not changed within the presented

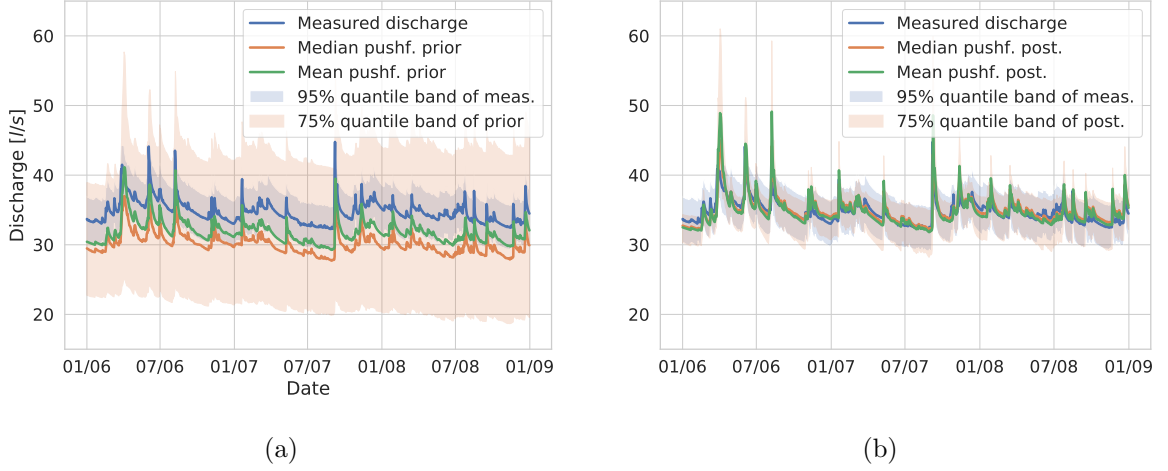


Figure 6.6: Prior (a) and approximate posterior (b) push-forward distributions (defined in (6.19)).

research study. Since the outflow from the baseflow storage is controlled by its variable storage (e_b) and its constant discharge coefficient (k_b), the hydrotope discharge coefficients for the groundwater recharge ($k_{is,i}$) also affect the baseflow discharge and its temporal dynamics since they control e_b . Given that k_b was not included as a calibration parameter, the $k_{is,i}$ parameters are responsible to maintain the baseflow contribution as derived by [31] and are most informed as displayed in the first eigenvector $\tilde{\mathbf{w}}_1$.

Although parameters X_5 , X_{12} , and X_{19} have the same physical interpretation, we can observe that they display different sensitivities for the different hydrotopes. This is due to the fact that different hydrotopes cover areas which are different in extension (Hyd 1 — 13%, Hyd 2 — 56%, and Hyd 3 — 27%). Therefore, the interpretation of the most important parameters occurring in an eigenvector, should both consider the physical meaning of the parameter and the relative contribution of each single hydrotope to the total spring discharge, which is highly affected by the relative area covered by the hydrotope. In this specific case, parameter X_{12} , associated with Hyd 2, displays the largest value since it covers the largest area in the Kerschbaum spring catchment and thus has a significant contribution to the total spring discharge. Parameter X_5 has the second largest value although Hyd 1 ranks as third in terms of coverage area. This is explained by the fact that Hyd 1 provides the most dynamic and variable discharge behavior of all hydrotopes. Hence, the discharge contribution from Hyd 1 is essential to reproduce the discharge dynamics observed in the Kerschbaum spring. Hyd 3 has the smallest contribution in $\tilde{\mathbf{w}}_1$, which can be explained by its more constant and less variable discharge behavior as compared to Hyd 1 and its smaller spatial share as compared to Hyd 2. Hence, although Hyd 3 has a larger area covered than Hyd 1, parameter X_{19} is less dominant than parameter X_5 .

We also comment on Fig. 6.6b. The sample medians and means of the approximate posterior push-forward distribution computed with the 4D subspace match the low-flow conditions reasonably well; also, the quantile band representing the uncertainty is reduced considerably. However, the high-flow conditions stay uncertain and are not well matched by the median and mean which means that they are not sufficiently reflected by the approximated posterior distribution. To achieve this, we would have to adjust the data misfit function in the sense

that deviations in high-flow conditions are weighted more than in low-flow-conditions, but this would entail a re-formulation of the Bayesian inverse problem.

Finally, we refer to our follow-up investigation in [32] which contains an in-depth validation of the model with further sensitivity analyses that are more interesting from a hydrological point of view. In particular, we study the hydrological features that actually lead to the dimension reduction presented in this chapter.

Iterative active subspaces

This final chapter discusses an *iterative* approach for the construction of active subspaces in Bayesian inverse problems. We start with a motivation in Sec. 7.1 and formulate two goals that we aim to achieve by an iterative algorithm defined in Sec. 7.2. A corresponding analysis with an illustrative case is provided in Sec. 7.3. Sec. 7.4 discusses an implementation of the algorithm on a real scenario involving a model for the 2014 Ebola outbreak in West Africa. Finally, conclusions are given in Sec. 7.5.

7.1 Motivation

A critical point of applying ASM for solving Bayesian inverse problems is the use of the *prior* distribution as weighting for sensitivities (or gradients), i. e., so far we compute

$$C_0 := \mathbf{E}[\nabla f^d(\mathbf{X}) \nabla f^d(\mathbf{X})^\top] = \int_{\mathcal{X}} \nabla f^d(\mathbf{x}) \nabla f^d(\mathbf{x})^\top \rho_0(\mathbf{x}) \, d\mathbf{x} \quad (7.1)$$

as it was done in (4.11) and also for the applications in Ch. 5 and 6.

Indeed, the active subspace, i. e., the column space of W_1 from (3.10), computed w.r.t. the prior can be crucially different from the one computed w.r.t. the posterior, especially if the prior substantially differs from the posterior distribution in the sense that the support of the prior contains relatively large regions with high misfit values. That is, we actually would like to approximate

$$C^d := \mathbf{E}[\nabla f^d(\mathbf{X}) \nabla f^d(\mathbf{X})^\top \mid \mathbf{D} = \mathbf{d}] = \int_{\mathcal{X}} \nabla f^d(\mathbf{x}) \nabla f^d(\mathbf{x})^\top \rho^d(\mathbf{x}) \, d\mathbf{x} \quad (7.2)$$

for given data $\mathbf{d} \in \mathcal{D}$ instead of C_0 .

As an illustration, regard the sufficient summary plot in Fig. 7.1 referring to a scenario that is discussed in Sec. 7.4. The orange dots depict the misfit values of samples $\mathbf{X}_j \sim \mu_0$ from (4.20) (projected on the axis of the first active variable by $\tilde{\mathbf{w}}_1^\top \mathbf{X}_j$) that are used to approximate C_0 . Note that the y axis is set to a log scale. Although the orange dots follow a one-dimensional behavior on the left part, we can see that they do not so near the high posterior region, i. e., the region with lower data misfit values (given a reasonably flat prior), which is around $\tilde{\mathbf{w}}_1^\top \mathbf{x} = 0.5$ in this example. This is confirmed by manually added samples \mathbf{X}'_j

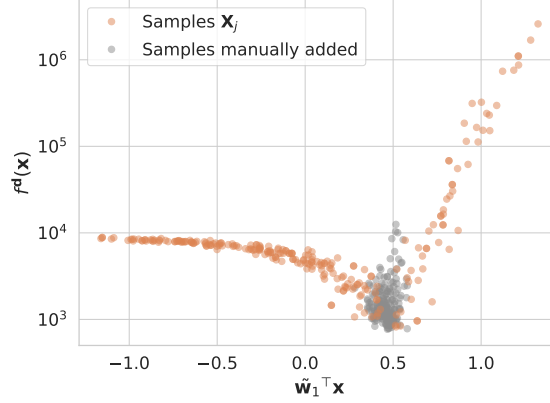


Figure 7.1: A one-dimensional subspace (spanned by $\tilde{\mathbf{w}}_1$) that is not suitable for samples in regions with high posterior probability, i. e., lower misfit values (given a reasonably flat prior).

(gray dots), which are randomly constructed such that $\tilde{\mathbf{w}}_1^\top \mathbf{X}'_j$ is near 0.5, as they do not show a one-dimensional structure at all. The reason for this problem is the fact that the dominant (active) directions are computed w.r.t. the prior distribution, i. e., all orange dots, and not only the dots with a relatively low misfit value.

So, on the one hand, we want to approximate the posterior via ASM, as described in Ch. 4, but on the other hand we need the posterior in order to approximate C^d . This constitutes a circular task and suggests an *iterative* scheme. In the following, we formulate and discuss an iterative algorithm that aims for constructing an active subspace that

1. is more aligned to regions with higher posterior probability and, hence,
2. can serve as a better starting point for further inference procedures as the subspace given by C_0 .

Related literature

A related iterative approach is presented in [293, Subsec. 3.3.3 & Alg. 2] which employs a *sequential importance sampling* framework. Remember that this work is already mentioned in the related literature section of Ch. 3 and 4. It suggests an algorithm that, in each step, chooses the dimension k of an active subspace such that a so-called *reconstruction error* falls under a predefined threshold. This dimension can generally take on every value $k \leq n$ which, on the one hand, might imply higher computational costs since cheap surrogates are not easily available in higher dimensions but is, on the other hand, necessary in order to make importance sampling for C^d function and useful.

In contrast, our approach does not aim to fully approximate C^d in each single step but rather proposes a *step-wise* scheme which allows the computation of cheap surrogates and targets the two goals from above. In particular, it is Goal 2 that constitutes a main difference to the attempt undertaken in [293].

7.2 Iterative algorithm

In the remainder, we set $f := f^d$.

The iterative algorithm that we suggest as a solution to the problem mentioned above is given by the steps in Alg. 7. Note that all quantities used in the definition of the algorithm are *exact*. Of course, in practice (i.e., in Sec. 7.4) we have to fall back on estimations and approximations as discussed in Sec. 3.5.

Most of the steps in the algorithm are straightforward to understand; however, some of them need emphasis and clarification. At every iteration ℓ , we compute an eigenvector matrix $W_1^{(\ell)}$ containing directions that dominate (w.r.t. $\mu^{(\ell)}$) the behavior of the data misfit function. In general, each active subspace (spanned by the columns of $W_1^{(\ell)}$) can have a *different* dimension $k^{(\ell)} \in \{1, \dots, n-1\}$; see Step 3.

But the most important step to understand is Step 7. In (7.8), we decompose the density of the next posterior approximation $\mu^{(\ell+1)}$, denoted by $\rho^{(\ell+1)}$, into

1. $\rho_g^{(\ell)} \propto \exp(-g^{(\ell)}(\cdot)) \rho_{0, \mathbf{Y}, W^{(\ell)}}(\cdot)$, the posterior density on the *active* subspace spanned by the columns of $W_1^{(\ell)}$, and
2. $\rho_{\mathbf{Z}^{(\ell)} | \mathbf{Y}^{(\ell)}}$, the density corresponding to the subspace which is *inactive* w.r.t. $\mu^{(\ell)}$.

It is important to note that the densities $\rho_{0, \mathbf{Y}, W^{(\ell)}}$ and $\rho_{\mathbf{Z}^{(\ell)} | \mathbf{Y}^{(\ell)}}$ are *not* derived from the same “parent” density as it was the case so far in previous chapters. Actually, both share the same parent density only in the first iteration when $\ell = 0$ (which then is the prior density ρ_0). The density $\rho_{0, \mathbf{Y}, W^{(\ell)}}$, which can be named the prior density on the active subspace spanned by the columns of $W_1^{(\ell)}$, is *always* referring to the prior density ρ_0 . In contrast, the conditional density $\rho_{\mathbf{Z}^{(\ell)} | \mathbf{Y}^{(\ell)}}$ relates to $\rho^{(\ell)}$, the density of $\mu^{(\ell)}$, in every iteration ℓ . The reason for this decomposition is that we aim for a *gradual* approximation of the *posterior*, which is proportional to the product of the likelihood and the *prior*. Using $\rho_{\mathbf{Z}^{(\ell)} | \mathbf{Y}^{(\ell)}}$ (instead of the conditional prior density $\rho_{\mathbf{Z}^{(0)} | \mathbf{Y}^{(0)}}$) ensures that we do not lose information (in the form of informed directions) that we already computed in previous iterations.

Referring back to the goals formulated at the end of Sec. 7.1 (Goal 1 and 2), we see that the directions stored in $W^{(\ell)}$ are more aligned to regions with higher posterior probability already from iteration $\ell = 1$ on. Indeed, we see in Sec. 7.4 that it is enough there to perform only two iterations (i.e., for $\ell = 0$ and $\ell = 1$) to meet the two goals mentioned above.

But before we investigate this more complex scenario, we continue with an analysis of Alg. 7 by looking at a rather simple case which however provides insight into the evolution of posterior approximations $\mu^{(\ell)}$ driven by the iterations.

7.3 Analysis with an illustrative case

The analysis for Alg. 7 resumes the example from Sec. 2.3 that involves a Bayesian inverse problem with a Gaussian prior $\mu_0 := \mathcal{N}(\mathbf{m}_0, \Sigma_0)$, Gaussian observational noise $\boldsymbol{\eta} \sim \mathcal{N}(0, \Gamma)$, and a linear forward map $\mathcal{G}(\mathbf{x}) := A\mathbf{x}$ for some matrix $A \in \mathbf{R}^{n_d \times n}$. For the sake of a clear illustration, we further specialize this example by assuming that $\Sigma_0 = I$, $\Gamma = \gamma^2 I$ for $\gamma > 0$, and $\mathbf{d} = \mathcal{G}(\mathbf{m}_0) = A\mathbf{m}_0$.

According to (2.35), we get a posterior $\mu^d \sim \mathcal{N}(\mathbf{m}, \Sigma)$ with $\mathbf{m} = \mathbf{m}_0$ and

$$\Sigma = I - A^\top (\gamma^2 I + AA^\top)^{-1} A \quad (7.10)$$

Algorithm 7: Iterative active subspace construction

Input : Starting distribution $\mu^{(0)} := \mu_0$.

Output: A sequence of posterior approximations $\mu^{(\ell)}$ and eigenvector matrices $W^{(\ell)}$ containing (non-)dominant directions w.r.t. $\mu^{(\ell)}$.

Let $\mathbf{X}^{(0)} \sim \mu^{(0)}$.

Iterating for $\ell = 0, 1, 2, \dots$, the ℓ -th step is:

1. Compute

$$\begin{aligned} C^{(\ell)} &:= \mathbf{E}[\nabla f(\mathbf{X}^{(\ell)}) \nabla f(\mathbf{X}^{(\ell)})^\top] \\ &= \int_{\mathcal{X}} \nabla f(\mathbf{x}) \nabla f(\mathbf{x})^\top \rho^{(\ell)}(\mathbf{x}) d\mathbf{x}. \end{aligned} \quad (7.3)$$

2. Eigendecompose $C^{(\ell)} =: W^{(\ell)} \Lambda^{(\ell)} W^{(\ell)\top}$.

3. Decide for a dimension of the active subspace $k^{(\ell)} \in \{1, \dots, n-1\}$ and split the active directions accordingly, i. e.,

$$W^{(\ell)} =: \begin{pmatrix} W_1^{(\ell)} & W_2^{(\ell)} \end{pmatrix}, \quad (7.4)$$

where $W_1^{(\ell)} \in \mathbf{R}^{n \times k^{(\ell)}}$ and $W_2^{(\ell)} \in \mathbf{R}^{n \times (n-k^{(\ell)})}$.

4. Define active and inactive (random) variables,

$$\mathbf{Y}^{(\ell)} := W_1^{(\ell)\top} \mathbf{X}^{(\ell)} \in \mathbf{R}^{k^{(\ell)}}, \quad \mathbf{Z}^{(\ell)} := W_2^{(\ell)\top} \mathbf{X}^{(\ell)} \in \mathbf{R}^{n-k^{(\ell)}}. \quad (7.5)$$

5. Compute

$$\begin{aligned} g^{(\ell)}(\mathbf{y}) &:= \mathbf{E}[f(\llbracket \mathbf{Y}^{(\ell)}, \mathbf{Z}^{(\ell)} \rrbracket_{W^{(\ell)}}) \mid \mathbf{Y}^{(\ell)} = \mathbf{y}] \\ &= \int f(\llbracket \mathbf{y}, \mathbf{z} \rrbracket_{W^{(\ell)}}) \rho_{\mathbf{Z}^{(\ell)} \mid \mathbf{Y}^{(\ell)}}(\mathbf{z} \mid \mathbf{y}) d\mathbf{z}. \end{aligned} \quad (7.6)$$

6. Compute the prior on the current active subspace,

$$\rho_{0, \mathbf{Y}, W^{(\ell)}}(\mathbf{y}) := \int \rho_0(\llbracket \mathbf{y}, \mathbf{z} \rrbracket_{W^{(\ell)}}) d\mathbf{z}. \quad (7.7)$$

7. Define the next posterior approximation, i. e., for $\mathbf{x} \in \mathcal{X}$ with $\mathbf{x} = \llbracket \mathbf{y}, \mathbf{z} \rrbracket_{W^{(\ell)}}$, where $\mathbf{y} = W_1^{(\ell)\top} \mathbf{x}$ and $\mathbf{z} = W_2^{(\ell)\top} \mathbf{x}$, set

$$\begin{aligned} \rho^{(\ell+1)}(\mathbf{x}) &: \propto \exp(-g^{(\ell)}(\mathbf{y})) \rho_{0, \mathbf{Y}, W^{(\ell)}}(\mathbf{y}) \rho_{\mathbf{Z}^{(\ell)} \mid \mathbf{Y}^{(\ell)}}(\mathbf{z} \mid \mathbf{y}) \\ &\propto: \rho_g^{(\ell)}(\mathbf{y}) \rho_{\mathbf{Z}^{(\ell)} \mid \mathbf{Y}^{(\ell)}}(\mathbf{z} \mid \mathbf{y}) \end{aligned} \quad (7.8)$$

and let $\mu_g^{(\ell)}$ be the distribution induced by $\rho_g^{(\ell)}$. For $\mathring{\mathbf{Y}}^{(\ell+1)} \sim \mu_g^{(\ell)}$ and $\mathring{\mathbf{Z}}^{(\ell+1)} := (\mathbf{Z}^{(\ell)} \mid \mathbf{Y}^{(\ell)} = \mathring{\mathbf{Y}}^{(\ell+1)})$, we define

$$\mathbf{X}^{(\ell+1)} := W^{(\ell)} \begin{pmatrix} \mathring{\mathbf{Y}}^{(\ell+1)} \\ \mathring{\mathbf{Z}}^{(\ell+1)} \end{pmatrix} \sim \mu^{(\ell+1)}. \quad (7.9)$$

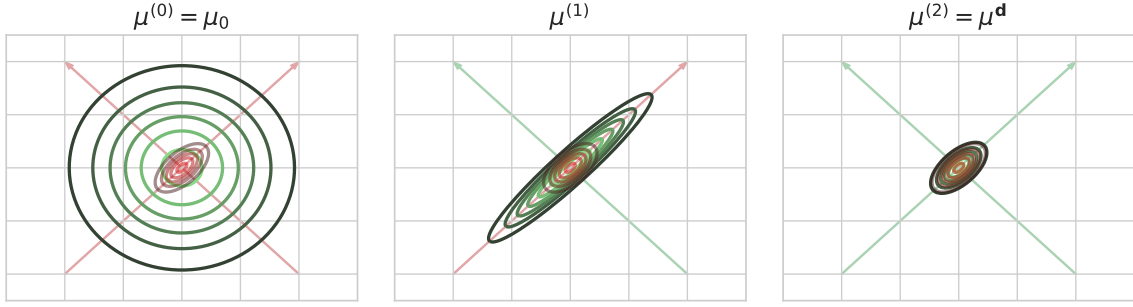


Figure 7.2: Evolution of posterior approximations for the described linear Gaussian Bayesian inverse problem in $n = 2$ dimensions. The green/black contour lines indicate $\mu^{(\ell)}$, whereas the transparent red contour lines display the posterior. The color of an arrow shows if a corresponding direction is already informed.

$$= \left(I + \frac{1}{\gamma^2} A^\top A \right)^{-1}. \quad (7.11)$$

We do not intend to provide a formal analysis demonstrating that the two goals mentioned above get achieved. The analysis (for the given simplified setting) rather aims for showing desirable mathematical properties of the algorithm as the approximation of the true posterior and *consistency* in the sense that $\mu^{(\ell)} = \mu^{\mathbf{d}}$ for some $\ell \in \mathbf{N}_0$ implies that also $\mu^{(\ell+1)} = \mu^{\mathbf{d}}$.

The evolution of the gradual posterior approximations $\mu^{(\ell)}$ (described by Alg. 7) in the given setting is illustrated for $n = 2$ dimensions in Fig. 7.2. In all three subfigures, the green/black circles display the contour lines of $\mu^{(\ell)}$, whereas the transparent red contour lines indicate the posterior distribution. The arrows show the directions of active/inactive directions which actually coincide with the *principal components* [139] of the posterior and stay constant throughout the iterations in this example which is demonstrated later. The color of an arrow indicates whether the corresponding direction is already informed. The iterations start on the left with $\mu^{(0)} = \mu_0$. In the first iteration ($\ell = 0$) from the left to the middle subfigure, we assume to have only one active direction (spanned by $\mathbf{w}_1^{(0)} = (-1/\sqrt{2}, 1/\sqrt{2})^\top$) that gets informed (red colored on the left but green in the middle). The second iteration ($\ell = 1$) from the middle to the right subfigure then also informs the second direction such that $\mu^{(2)} = \mu^{\mathbf{d}}$. The second direction becomes active for $\ell = 1$ since the misfit function changes more in this direction (w.r.t. $\mu^{(1)}$) than in the already informed one.

The visualization of the iterations above is formalized in the following statement.

Proposition 7.3.1. *Let $\mu_0 = \mathcal{N}(\mathbf{m}_0, I)$ with $\mathbf{m}_0 \in \mathbf{R}^n$. Suppose that $\mathcal{G}(\mathbf{x}) := A\mathbf{x}$ for $\mathbf{x} \in \mathcal{X}$ with $A \in \mathbf{R}^{n_{\mathbf{d}} \times n}$, and $\boldsymbol{\eta} \sim \mathcal{N}(0, \gamma^2 I)$ for $\gamma > 0$. Furthermore, assume that $\mathbf{d} = \mathcal{G}(\mathbf{m}_0)$.*

Considering Alg. 7, set $\Lambda := \Lambda^{(0)}$ and $W := W^{(0)}$. Then, for every iteration $\ell \in \mathbf{N}_0$, it holds that

$$\mathbf{X}^{(\ell)} \sim \mathcal{N}(\mathbf{m}_0, \Sigma^{(\ell)}) \quad (7.12)$$

with

$$\Sigma^{(\ell)} = W \begin{pmatrix} \left(I + \Lambda_{1:K^{(\ell)}}^{1/2} \right)^{-1} & 0 \\ 0 & I \end{pmatrix} W^\top \quad (7.13)$$

for some natural number $0 \leq K^{(\ell)} \leq n$.

Remark. (i) For a diagonal matrix $D \in \mathbf{R}^{n \times n}$, $D_{1:K}$ denotes the upper left submatrix of dimension $K \times K$ for a natural number $0 \leq K \leq n$. (ii) For real diagonal matrices D with non-negative diagonal elements d_{jj} , the square root $D^{1/2}$ is chosen to be the diagonal matrix with *positive* square roots $+\sqrt{d_{jj}}$ on the diagonal to get uniqueness. (iii) As mentioned above, the posterior distribution in the given setting is $\mu^{\mathbf{d}} = \mathcal{N}(\mathbf{m}_0, \Sigma)$ with

$$\Sigma = \left(I + \frac{1}{\gamma^2} A^\top A \right)^{-1}. \quad (7.14)$$

Note that Σ can also be rewritten as

$$\Sigma = W(I + \Lambda^{1/2})^{-1} W^\top \quad (7.15)$$

by using the Woodbury matrix identity; see Lem. A.1.2. Indeed, we get that

$$W \Lambda W^\top = W^{(0)} \Lambda^{(0)} W^{(0)\top} = C^{(0)} = \mathbf{E}[\nabla f^{\mathbf{d}}(\mathbf{X}^{(0)}) \nabla f^{\mathbf{d}}(\mathbf{X}^{(0)})^\top] = \left(\frac{1}{\gamma^2} A^\top A \right)^2 \quad (7.16)$$

implying that

$$\frac{1}{\gamma^2} A^\top A = W \Lambda^{1/2} W^\top. \quad (7.17)$$

The proof of Prop. 7.3.1 can be found in Sec. D.1. Comparing $\Sigma^{(\ell)}$ from (7.13) and Σ from (7.15), we see that the directions (stored as columns in W) are gradually informed which was illustrated in Fig. 7.2 for $n = 2$ dimensions. Depending on the sequence $(K^{(\ell)})_{\ell \in \mathbf{N}_0}$ approaching the full dimension n , the sequence of covariance matrices $(\Sigma^{(\ell)})_{\ell \in \mathbf{N}_0}$ can gradually approach the posterior covariance matrix Σ .

7.4 Experiment with an Ebola spread model

The problem described in Sec. 7.1 is experimentally investigated in this section with a model for the dynamics of the 2014 Ebola outbreak in West Africa which also accounts for intervention strategies [16]. We employ Alg. 7 to meet the two goals from above (Goal 1 and 2) and will see that one or two iterations are enough to improve the initial situation for a Bayesian inference in this case.

Note that ASM was already applied for computing global sensitivity metrics in the context of a similar model for the Ebola spread in West Africa to guide resource allocation [78].

Subsec. 7.4.1 describes the Ebola spread model that we use for the study in this section. We do not provide full detail on the model and related existing work but rather give references where appropriate. The setup for a corresponding Bayesian inverse problem is given in Subsec. 7.4.2. Finally, the actual outcomes of our study, the computational results, are discussed in Subsec. 7.4.3.

7.4.1 Model description

The 2014 Ebola Virus Disease (EVD) outbreak in Guinea, Liberia, and Sierra Leone is in fact the largest and longest epidemic ever reported since the first identification of this disease

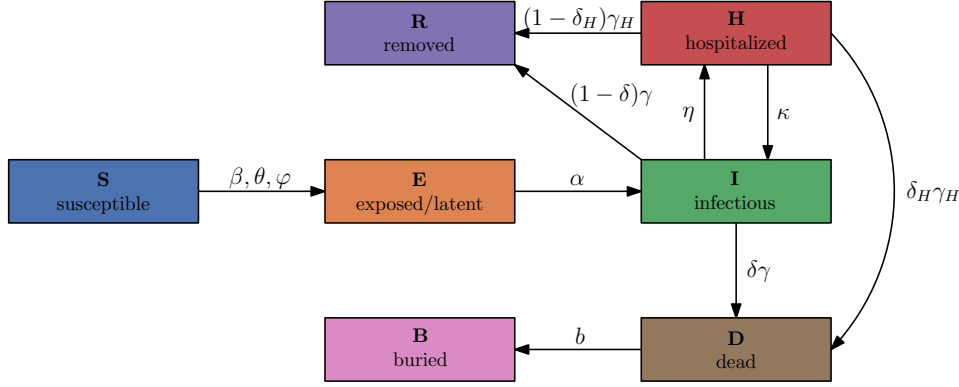


Figure 7.3: Model compartments on the basis of [16, Fig. 1] including parameters for the transitions between them.

from 1976 in Zaire (now the Democratic Republic of Congo). Beginning with the first case reported on December 26, 2013 in Guinea, the virus rapidly spread to Liberia and Sierra Leone causing a total of 28,616 reported cases and 11,310 reported deaths as of June 10, 2016 [7].

Mathematical modeling supported efforts of international organizations as the World Health Organization (WHO) by, for example, estimating crucial quantities like the *basic reproduction number* \mathcal{R}_0 , an epidemiological metric that indicates the average number of secondary infections caused by one infected host in a fully susceptible population over the whole time of the host's infection.

The model, which we employ for our experiment and call “the” model in the following, was developed by the authors of [16]. It is a *compartmental* model, i.e., it divides a population into so-called *compartments*, and is based on simpler models like the SIR or SEIR model that involve compartments for susceptible (S), exposed/latent (E), infectious (I), and removed/recovered (R) subpopulations and describe suitable transitions among them. Susceptible persons get infected (with a certain average rate) and move to the compartment of exposed ones. However, the exposed people are not yet infectious but do become after the *incubation period*. As soon as they recovered from the corresponding disease, they are removed from the chain of transmission.

In addition, the model extends the work of [174] which adds three more compartments to account for hospitalized (H), dead (D), and buried (B) individuals. These additional compartments play a central role in the model since the Ebola virus is transmitted not only by contact with infectious hosts in the community but also often in hospitals and at funerals; see [9]. The difference of the model to the one from [174] lies in an inclusion of patients that abandon healthcare facilities and return back to the community at their own (and others’) risk, as it was reported several times by WHO [9]. For a more detailed review of related existing models and work, we refer to [16, Introduction].

Parameters for a standard version of the model, as the transmission rates in the community, in hospitals, and at funerals, are listed in the upper and middle part of Tab. 7.1. A transmission diagram including the mentioned parameters is depicted in Fig. 7.3 for better visualization. Without consideration of intervention strategies to prevent further spread of the virus, these (pre-intervention) parameters would stay constant/stationary over the whole simulation period.

However, the model does account for the fact that control measures were initiated at a certain point in time. These efforts encompass, for example, surveillance, isolation of suspected

Table 7.1: Parameters for the Ebola spread model; see [16, Tab. 1 and 2].

Par.	Description	Unit	x_{fit} [Range]
β	Transmission rate in the community before intervention	1/week	0.532 [0.350–0.575]
θ	Transmission rate in hospitals before intervention	1/week	0.328 [0.100–0.480]
φ	Transmission rate at funerals before intervention	1/week	2.104 [1.402–2.475]
$1/\eta$	Mean time from symptoms onset to hospitalization before intervention	days	4.8 [4.8–5.3]
$1/b$	Mean time from death to burial before intervention	days	5.4 [4–6.6]
$1/\gamma$	Mean duration of the infection	days	10.4 [9.5–10.5]
$1/\alpha$	Mean duration of the incubation period	days	10 [9.5–10.5]
$1/\gamma_H$	Average permanence in the hospital	days	4.6 [4.4–4.9]
δ	Case fatality ratio in the community	–	0.73 [0.69–0.73]
δ_H	Case fatality ratio in hospitals	–	0.61 [0.52–0.64]
κ	Hospital leaving rate	1/week	0.0025 [0.0022–0.0028]
$\tilde{\beta}$	Transmission rate in the community after intervention	1/week	0.505 [0.425–0.532]
t_{β}	Time to achieve $(\beta + \tilde{\beta})/2$	days	12.5 [9.70–48.52]
$\tilde{\theta}$	Transmission rate in hospitals after intervention	1/week	0.095 [0.033–0.328]
t_{θ}	Time to achieve $(\theta + \tilde{\theta})/2$	days	11.1 [9.70–48.52]
$\tilde{\varphi}$	Transmission rate at funerals after intervention	1/week	1.115 [0.210–2.104]
t_{φ}	Time to achieve $(\varphi + \tilde{\varphi})/2$	days	10.3 [9.70–48.52]
$1/\tilde{\eta}$	Mean time from symptoms onset to hospitalization after intervention	days	4.1 [2–5]
t_{η}	Time to achieve $(\eta + \tilde{\eta})/2$	days	27.1 [16.17–48.52]
$1/\tilde{b}$	Mean time from death to burial after intervention	days	4.9 [1–5.4]
t_b	Time to achieve $(b + \tilde{b})/2$	days	21.2 [6.93–48.52]

cases, and a reduction of the duration between deaths and corresponding burials, and consequently imply changes of some parameters. Following [80], the model identifies five parameters that need to be adjusted (see upper part of Tab. 7.1),

- transmission rates in the community (β), in hospitals (θ), and at funerals (φ),
- the hospitalization rate (η), and
- the burial rate (b).

It is assumed that these control parameters change *gradually* in time once the intervention strategies start. Abusing notation, we introduce corresponding *time-varying* parameters, i. e., for $a \in \{\beta, \theta, \varphi, \eta, b\}$, we define

$$a(t) := \begin{cases} a & \text{for } t < T_{\text{interv}} \\ \tilde{a} + (a - \tilde{a}) \exp(-q_a(t - T_{\text{interv}})) & \text{for } t \geq T_{\text{interv}}, \end{cases} \quad (7.18)$$

where $q_a > 0$ specifies how fast the transition from a to a new value \tilde{a} occurs and $T_{\text{interv}} \geq 0$ denotes the starting time of interventions. Note that the mean value $(a + \tilde{a})/2$ is reached at time $T_{\text{interv}} + t_a$ with $t_a = \log(2)/q_a$. The (intervention) parameters that are added due to (7.18) are specified in the lower part of Tab. 7.1.

With this setup, we can finally provide the system of ordinary differential equations (ODEs) that describe the dynamics of the mentioned compartments [16, Appendix A.1],

$$\begin{cases} S' &= -\lambda(t)S \\ E' &= \lambda(t)S - \alpha E \\ I' &= \alpha E - (\gamma + \eta(t))I + \kappa H \\ R' &= (1 - \delta)\gamma I + (1 - \delta_H)\gamma_H H \\ H' &= \eta(t)I - (\gamma_H + \kappa)H \\ D' &= \delta\gamma I + \delta_H\gamma_H H - b(t)D. \end{cases} \quad (7.19)$$

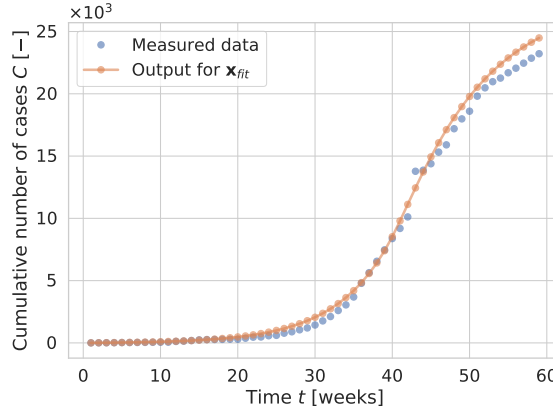


Figure 7.4: Measured data and model output using parameters \mathbf{x}_{fit} for the cumulative number of cases.

The expression

$$\lambda(t) := \beta(t) \frac{I}{N} + \theta(t) \frac{H}{N} + \varphi(t) \frac{D}{N} \quad (7.20)$$

is called the *force of infection*, where $N = N(t) := S(t) + E(t) + I(t) + R(t)$ denotes the (current) number of individuals in the community. The auxiliary variables for the cumulative (total) number of burials $B = B(t)$ and the cumulative number of infected cases $C = C(t)$ are given by

$$\begin{cases} B' &= b(t)D \\ C' &= \alpha E. \end{cases} \quad (7.21)$$

Initial values for the system in (7.19), adopted from the authors of [16], are set to

$$(S(0), E(0), I(0), R(0), H(0), D(0)) = (2.2 \cdot 10^7, 3.42, 2.31, 0, 0.1, 4.25) \quad (7.22)$$

and

$$(B(0), C(0)) = (0, 0). \quad (7.23)$$

The information, that is used in [16] to fit the model parameters from Tab. 7.1, is based on weekly incidences reported through WHO situation reports [8] from December 28, 2013 to February 13, 2015 (59 weeks). First, pre-intervention parameters and parameters independent of interventions (middle part of Tab. 7.1) were fitted to data from December 28, 2013 to October 3, 2014 (40 weeks) and, subsequently, intervention parameters to data from October 3, 2014 to February 13, 2015 (20 weeks). Hence, we set the starting time of interventions to

$$T_{\text{interv}} = 40. \quad (7.24)$$

For concrete values of the fitted parameters, which we denote by \mathbf{x}_{fit} , see the right column in Tab. 7.1. The measured data and the model output using parameters \mathbf{x}_{fit} for the cumulative number of cases C are plotted in Fig. 7.4.

7.4.2 Setup

Also here, we distinguish between model and calibration parameters. Their relationship is, however, simple in this case. According to the model description in Subsec. 7.4.1, we have the following set of 21 model parameters to be inferred,

$$\{\beta, \tilde{\beta}, t_\beta, \theta, \tilde{\theta}, t_\theta, \varphi, \tilde{\varphi}, t_\varphi, \eta, \tilde{\eta}, t_\eta, b, \tilde{b}, t_b, \gamma, \alpha, \gamma_H, \delta, \delta_H, \kappa\}; \quad (7.25)$$

see also Tab. 7.1.

Following [16], we assume a uniform prior distribution on these model parameters. The corresponding ranges provided in Tab. 7.1 are taken from [16, Tab. 1 and 2]. We map these model parameters to calibration parameters \mathbf{X} by a linear transformation such that

$$\mathbf{X} \sim \mu_0 = \mathcal{U}([-1, 1]^{21}), \quad (7.26)$$

which implies that $\mathcal{X} = [-1, 1]^{21}$ and $n = 21$.

The quantity of interest is chosen to be weekly snapshots of the cumulative number of cases that we denoted by $C = C(t)$. Since corresponding data is available for 59 weeks, the forward operator is

$$\mathcal{G} : \mathcal{X} \rightarrow \mathbf{R}^{n_d}, \quad \mathbf{x} \mapsto (C(t_i))_{i=1, \dots, n_d}, \quad (7.27)$$

where $n_d = 59$ and t_i denotes the point in time corresponding to i weeks after the beginning of the outbreak ($t = 0$). The vector of *measured* cumulative cases $\mathbf{d} \in \mathbf{R}^{n_d}$ contains the data provided by WHO as mentioned above. We assume an observational noise level of 5% which translates to a diagonal covariance matrix Γ for the noise term $\boldsymbol{\eta}$ from (2.5), i.e.,

$$\Gamma_{ij} = (0.05 \times d_i)^2 \delta_{ij}, \quad (7.28)$$

where d_i denotes the i -th component of \mathbf{d} .

7.4.3 Computational results

First, we emphasize that the application of Alg. 7 to iteratively approximate the posterior distribution $\mu^d = \mathbf{P}_{\mathbf{X}|\mathcal{D}}(\cdot|\mathbf{d})$ is not directly possible for the described experiment since the definition of the algorithm uses *exact* quantities which are not available here. Therefore, we partially have to fall back on estimations and approximations of these quantities as described in Sec. 3.5. This has consequences for the execution of the algorithm which we discuss and comment on later.

Following (3.159) and (3.166), we use $N_{\tilde{C}} = 250$ gradient samples $\nabla f(\mathbf{X}_j^{(\ell)})$, $\mathbf{X}_j^{(\ell)} \sim \mu^{(\ell)}$, to estimate $C^{(\ell)}$ from (7.3) at each iteration ℓ . This corresponds to a pessimistic sampling factor of $\beta_{\tilde{C}} = 10$ for an accurate estimation of the first $m = 8$ eigenvalues.

In contrast to the experiments from Ch. 5 and 6, in which the Jacobians of the forward operator \mathcal{G} were approximated by central finite differences, here we can compute the “exact” Jacobians by using an adjoint formulation of the ODE system in (7.19). This is a well-known and common approach for computing derivatives of functionals [217]. To demonstrate the “exactness” of the computed Jacobians at a point $\mathbf{x}_0 \in \mathcal{X}^\circ$, we use Taylor’s theorem to write

$$\mathcal{G}_i(\mathbf{x}_0 + \Delta^h) = \mathcal{G}_i(\mathbf{x}_0) + \nabla \mathcal{G}_i(\mathbf{x}_0)^\top \cdot \Delta^h + O(h^2), \quad (7.29)$$

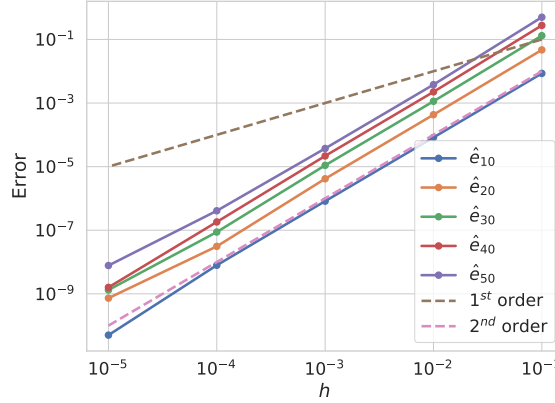


Figure 7.5: Verification for Jacobians of the forward operator \mathcal{G} . For $\mathbf{x}_0 = \mathbf{x}_{\text{fit}}$, the *relative* errors $\hat{e}_i = \hat{e}_i(h) := e_i(h)/|\mathcal{G}_i|$, $i = 10, 20, 30, 40, 50$, show a second order decay for $h \rightarrow 0$ which also holds for absolute errors e_i from (7.31) as it is expected for a correctly computed Jacobian.

where \mathcal{G}_i is the i -th component of \mathcal{G} , $\nabla \mathcal{G}_i$ denotes its gradient, and

$$\Delta^h := h \cdot (1, \dots, 1)^\top \quad (7.30)$$

with $h > 0$ sufficiently small. This corresponds to a first order approximation of \mathcal{G}_i , for which we denote the error by $e_i = e_i(h)$ with

$$e_i(h) := |\mathcal{G}_i(\mathbf{x}_0 + \Delta^h) - \mathcal{G}_i(\mathbf{x}_0) - \nabla \mathcal{G}_i(\mathbf{x}_0)^\top \cdot \Delta^h|. \quad (7.31)$$

Hence, a correctly computed Jacobian would yield a second order decay of $e_i(h)$ as $h \rightarrow 0$. For $\mathbf{x}_0 = \mathbf{x}_{\text{fit}}$, this is demonstrated in Fig. 7.5 for $i = 10, 20, 30, 40, 50$. Thus, we assume that the Jacobians $\nabla \mathcal{G}(\mathbf{X}_j^{(\ell)})$ and hence the gradient samples

$$\nabla f(\mathbf{X}_j^{(\ell)}) = \nabla \mathcal{G}(\mathbf{X}_j^{(\ell)})^\top \Gamma^{-1} (\mathcal{G}(\mathbf{X}_j^{(\ell)}) - \mathbf{d}) \quad (7.32)$$

are computed correctly using the adjoint formulation.

In the first iteration, i. e., for $\ell = 0$, we compute $C^{(0)}$ from (7.3) as before according to the prior $\mu^{(0)} = \mu_0$ (see Ch. 5 and 6). Estimations for the corresponding first $m = 8$ eigenvalues of $\Lambda^{(0)}$ and the first four eigenvectors $\mathbf{w}_i^{(0)}$, $i = 1, 2, 3, 4$, of $W^{(0)}$ are depicted in Fig. 7.6. The eigenvalues are quite large due to the high misfit values caused by the prior which includes regions in the parameter space that do not fit the data very well. However, we see that the first eigenvalue, which is three orders of magnitude larger than the second one, is highly dominating. Also, the variability in the estimated eigenvalues, caused by their random nature (see Sec. 3.5), is estimated quite high. But the minimum values seem to be outliers since the averages (solid line between the minimum and maximum values) lie near the maximum values. These rather large intervals are assessed as an additional indication of a “bad” prior.

We identify two dominating parameters in the first eigenvector $\tilde{\mathbf{w}}_1^{(0)}$, φ , the transmission rate at funerals before intervention, and b , the burial rate before intervention. Recall that $\tilde{\mathbf{w}}_i^{(0)}$ denotes the approximation of $\mathbf{w}_i^{(0)}$. It is intuitive that φ and b but also most of the other

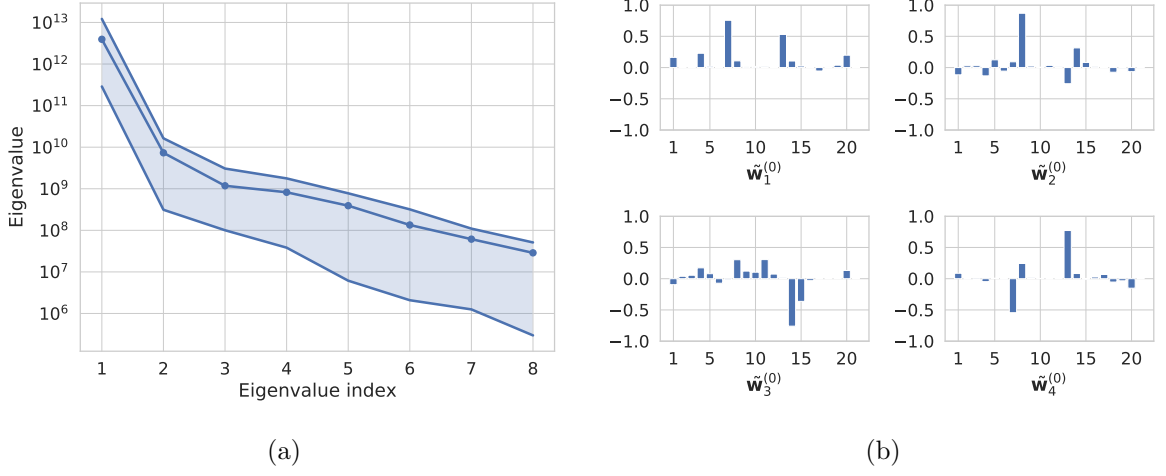


Figure 7.6: Estimations for the first eight eigenvalues in $\Lambda^{(0)}$ (a) and for the first four eigenvectors in $W^{(0)}$ (b).

dominating parameters (θ , δ_H , and β) are *pre-intervention* parameters since, first, the majority of data points lies in the pre-intervention period and, secondly, the intervention parameters are not able to compensate once the pre-intervention data points are not fitted sufficiently well.

Consequently, we decide for a one-dimensional active subspace in the first iteration, i. e., we set $k^{(0)} = 1$ in Step 3 of Alg. 7. The misfit values $f^d(\mathbf{X}_j^{(0)})$ that correspond to the samples $\mathbf{X}_j^{(0)}$, $j = 1, \dots, N_{\tilde{C}}$, are plotted on the axis $\tilde{\mathbf{w}}_1^{(0)\top} \mathbf{x}$ in the left part of Fig. 7.7 (orange dots). Note that they do not need to be additionally computed but reuse evaluations of the forward map $\mathcal{G}(\mathbf{X}_j^{(0)})$ which are already computed for the approximation of $C^{(0)}$; see (4.12).

Due to the non-polynomial shape of the misfit values on the first active axis, we prefer to perform GPR (instead of polynomial regression as in Ch. 5 and 6) which fits a Gaussian process to the given data points; see Sec. 4.4. We emphasize that, in this case of the first iteration, the regression is done on a log scale. The noise level corresponding to GPR is estimated with the data points. The solid blue line in Fig. 7.7 displays the mean function of the fitted Gaussian process; its exponential is then chosen as a surrogate for $g_N^{(0)}$ ($N = 1$) and thus also for $g^{(0)}$.

In the left part of Fig. 7.7, the misfit values (orange dots) follow the GPR surrogate, which we denote by $g_{\text{GPR}}^{(0)}$, quite well in the intervals, say, $[-1, 0]$ and $[1, 1.5]$. However, both regions do not contain the lowest misfit values which are located in, say, $[0.3, 0.7]$, where the surrogate does not fit the samples as well as in the other mentioned intervals. Hence, we guess that the misfit function does not have a one-dimensional structure w.r.t. the computed first active axis in the “relevant,” i. e., high posterior, interval. An ad hoc fix would increase the dimension of the active subspace $k^{(0)}$, but this is not promising in general since also the remaining directions were computed w.r.t. $\mu^{(0)} = \mu_0$.

To check our guess, we manually produce additional samples $\tilde{\mathbf{Y}}_{\text{add},j}$ (randomly) in the mentioned relevant interval in a first step. Since $\mu^{(0)} = \mu_0$ is not a standard normal distribution, it is not trivial to sample $\tilde{\mathbf{Z}}_{\text{add},j} \sim \mathbf{P}_{\tilde{\mathbf{Z}}|\tilde{\mathbf{Y}}}(\cdot|\tilde{\mathbf{Y}}_{\text{add},j})$ to complete the additional samples to full samples

$$\mathbf{X}_{\text{add},j} = [\tilde{\mathbf{Y}}_{\text{add},j}, \tilde{\mathbf{Z}}_{\text{add},j}]_{\tilde{W}^{(0)}}. \quad (7.33)$$

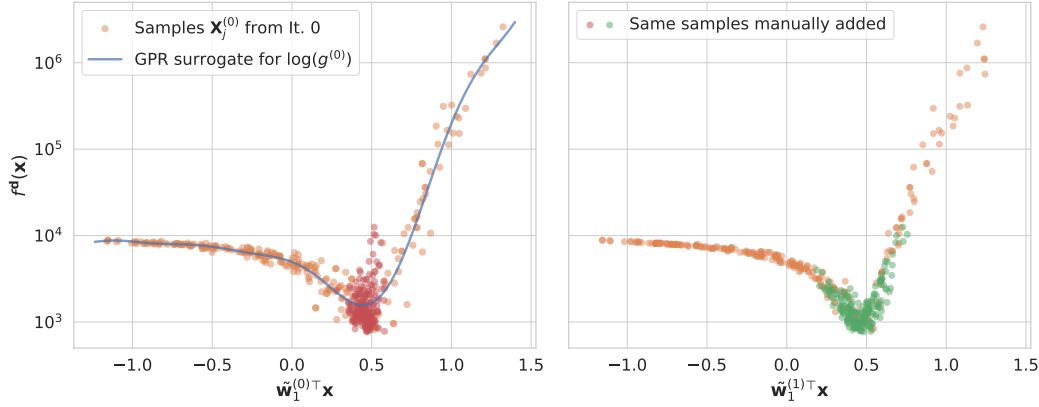


Figure 7.7: Both plots show the same set of misfit values $f^d(\mathbf{X}_j^{(0)})$ (orange dots) and the same set of samples that are manually (randomly) added such that they lie around $\tilde{\mathbf{w}}_1^{(0)\top} \mathbf{x} = 0.5$ (red dots on the left, green dots on the right). The left (resp. right) x axis is defined by the first active direction computed in the first (resp. second) iteration, i. e., for $\ell = 0$ (resp. $\ell = 1$).

Instead, we use the alternative explained in Sec. 4.3 with $\hat{\mathcal{X}} = \{\hat{\mathbf{X}}_j^{(0)} \stackrel{\text{i.i.d.}}{\sim} \mu^{(0)}\}_{j=1,\dots,J}$. The corresponding misfit values $f^d(\mathbf{X}_{\text{add},j})$, which do require to execute the model again, are subsequently computed and added to the left plot in Fig. 7.7 (red dots). We emphasize that the GPR surrogate $g_{\text{GPR}}^{(0)}$ is fitted to the orange dots, i. e., the original misfit values, only. The added misfit values do not follow a one-dimensional shape at all implying that $g_{\text{GPR}}^{(0)}$ is of bad quality in the region with higher posterior probability which would, however, be necessary in order to get a good posterior approximation.

Following Alg. 7, we aim for a second iteration ($\ell = 1$) which incorporates and reuses the information of the first iteration ($\ell = 0$). In order to approximate $C^{(1)}$, we first have to (approximately) produce samples $\mathbf{X}_j^{(1)} \stackrel{\text{i.i.d.}}{\sim} \mu^{(1)}$ by carrying out Steps 6 and 7, which we explain for general iterations ℓ .

So, let $\ell \in \mathbf{N}_0$ be arbitrary for the moment. The prior $\rho_{0,\tilde{\mathbf{Y}},\tilde{\mathbf{W}}^{(\ell)}}$ on the active subspace spanned by $\tilde{\mathbf{W}}_1^{(\ell)}$ is approximated (Step 6) as before in Ch. 5 and 6. The more difficult part is the approximation of $\mu_g^{(\ell)}$ ((7.8) in Step 7), or actually its density

$$\rho_g^{(\ell)}(\mathbf{y}) \propto \exp(-g^{(\ell)}(\mathbf{y})) \rho_{0,\mathbf{Y},W^{(\ell)}}(\mathbf{y}). \quad (7.34)$$

As mentioned, the low-dimensional misfit $g^{(\ell)}$ is not available, but we would like to use a surrogate for it, generally denoted by $g_{\text{surr}}^{(\ell)}$, which, however, might not fit well to the misfit values in high posterior regions (as demonstrated above for $\ell = 0$).

To reduce the impact of an inaccurate surrogate and to avoid biasing subsequent iterations while still utilizing (at least part of) the information that it contains, we “regularize” the surrogate by employing a technique called *tempering* [84]. This approach reduces concentration of probability mass and widens a posterior distribution. Regarding $\rho_{g_{\text{surr}}}^{(\ell)}$ adapted from (7.34), tempering means to look at a density

$$\rho_{g_{\text{surr}},T}^{(\ell)}(\tilde{\mathbf{y}}) \propto \exp(-g_{\text{surr}}^{(\ell)}(\tilde{\mathbf{y}}))^{\beta_T^{(\ell)}} \rho_{0,\tilde{\mathbf{Y}},\tilde{\mathbf{W}}^{(\ell)}}(\tilde{\mathbf{y}}), \quad (7.35)$$

where $\beta_T^{(\ell)} \in [0, 1]$ has to be chosen. Note that $\beta_T^{(\ell)} = 0$ corresponds to $\rho_{g_{\text{surr}}, T}^{(\ell)} = \rho_{0, \tilde{\mathbf{Y}}, \tilde{W}^{(\ell)}}$, i. e., just the prior, and $\beta_T^{(\ell)} = 1$ means $\rho_{g_{\text{surr}}, T}^{(\ell)} = \rho_{g_{\text{surr}}}^{(\ell)}$, i. e., the full posterior. Indeed, the factor $\beta_T^{(\ell)}$ controls the level of widening since

$$\exp(-g_{\text{surr}}^{(\ell)}(\tilde{\mathbf{y}}))^{\beta_T^{(\ell)}} = \exp\left(-\frac{1}{1/\beta_T^{(\ell)}} g_{\text{surr}}^{(\ell)}(\tilde{\mathbf{y}})\right). \quad (7.36)$$

The choice of a particular value for $\beta_T^{(\ell)}$ is difficult to justify in general. Actually, it is of greater importance to be rather pessimistic by choosing a value that is not too large. Recall that choosing $\beta_T^{(\ell)}$ rather small almost corresponds to sampling from the (wider) prior which does not “destroy” the current iteration, i. e., does not introduce a misleading bias caused by an inaccurate surrogate. That is, we recommend to decide for rather small values, particularly in the first iteration. We come back to this circumstance later when discussing concrete surrogates.

Eventually, we compute an approximation $\mu_{g_{\text{surr}}, T}^{(\ell)} \approx \mu_g^{(\ell)}$, where $\mu_{g_{\text{surr}}, T}^{(\ell)}$ denotes the distribution induced by $\rho_{g_{\text{surr}}, T}^{(\ell)}$. The samples

$$\tilde{\mathbf{Y}}_j^{(\ell+1)} \sim \mu_{g_{\text{surr}}, T}^{(\ell)} \quad (7.37)$$

are computed by running the RWM algorithm in the active subspace (Alg. 5 with $\mu_{g, \mathbf{Y}}^d = \mu_{g_{\text{surr}}, T}^{(\ell)}$). The remaining conditional samples

$$\tilde{\mathbf{Z}}_j^{(\ell+1)} = (\tilde{\mathbf{Z}}_j^{(\ell)} | \tilde{\mathbf{Y}}_j^{(\ell)} = \tilde{\mathbf{Y}}_j^{(\ell+1)}), \quad (7.38)$$

which are necessary to get samples $\mathbf{X}_j^{(\ell+1)} \stackrel{\text{approx.}}{\sim} \mu^{(\ell+1)}$ from (7.9), are again generated by the alternative procedure explained in Sec. 4.3 with

$$\hat{\mathcal{X}} = \{\hat{\mathbf{X}}_j^{(\ell)} \stackrel{\text{approx.}}{\sim} \mu^{(\ell)}\}_{j=1, \dots, J}. \quad (7.39)$$

We resume the description of the first iteration, i. e., $\ell = 0$. Above, we fit a GPR surrogate to the misfit values $f^d(\mathbf{X}_j^{(0)})$ which means that

$$g_{\text{surr}}^{(0)} = g_{\text{GPR}}^{(0)}. \quad (7.40)$$

As recommended, the factor for the tempering approach (widening the posterior distribution) is chosen to be rather small,

$$\beta_T^{(0)} = 5 \cdot 10^{-3}, \quad (7.41)$$

since the misfit values are quite different and thus the posterior is rather concentrated.

After the samples $\mathbf{X}_j^{(1)} \stackrel{\text{approx.}}{\sim} \mu^{(1)}$ are computed, we proceed to the second iteration $\ell = 1$ and approximate $C^{(1)}$. The corresponding first eight eigenvalues from $\tilde{\Lambda}^{(1)}$ and first four eigenvectors in $\tilde{W}^{(1)}$ are depicted in Fig. 7.8. We observe that, first, the magnitude of the estimated eigenvalues drop about five orders of magnitude compared to the first iteration, which indicates that they are computed in regions with much smaller misfit values. Secondly, the variability in the estimates (light blue band) becomes quite small, which we assess as a more “proper” behavior.

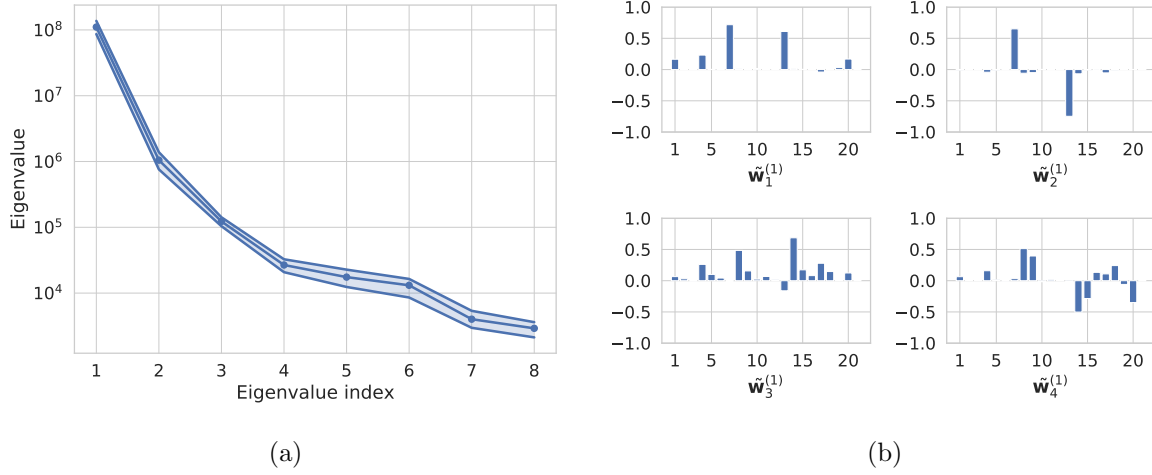


Figure 7.8: Estimations of the first eight eigenvalues (a) and the first four eigenvectors (b) of $C^{(1)}$ in the second iteration ($\ell = 1$).

Let us have a look again at Fig. 7.7. The right part plots the same misfit values (orange dots) from the first iteration as on the left part, and also plots the same samples that are manually added (red dots on the left, green dots on the right). The only difference is that the axis on the right is defined by the first active direction computed in the *second* iteration ($\ell = 1$), i. e., $\tilde{\mathbf{w}}_1^{(1)\top} \mathbf{x}$. Although the deviation of $\tilde{\mathbf{w}}_1^{(1)}$ from $\tilde{\mathbf{w}}_1^{(0)}$ is not substantial, we see that the green dots much better follow a one-dimensional shape which is exactly what we intend to occur by computing a second active subspace built with samples from regions with higher posterior probability.

The surrogate of $g^{(1)}$ for the second iteration on the first active axis $\tilde{\mathbf{w}}_1^{(1)\top} \mathbf{x}$ and the corresponding misfit values $f^d(\mathbf{X}_j^{(1)})$ are plotted in Fig. 7.9. We decide for a polynomial surrogate of fourth order, i. e.,

$$g_{\text{sur}}^{(1)} = g_{\text{poly}}^{(1)}. \quad (7.42)$$

The coefficient of determination in this case is $r_{\ell=1}^2 = 0.7071$. Note that, in contrast to the sufficient summary plot of the first iteration (left part of Fig. 7.7), the y axis in Fig. 7.9 is linear (instead of logarithmic) and the misfit values are much smaller, on average. Although there are quite some outliers that deviate from the surrogate in the interval, say, $[0.2, 0.35]$, they do not lie in the region of highest posterior probability which is, say, $[0.4, 0.5]$.

Anyway, as before we apply tempering to reduce the impact of the inaccurate surrogate in sampling from $\mu_{g_{\text{sur}}, T}^{(1)} = \mu_{g_{\text{poly}}, T}^{(1)}$ with a factor

$$\beta_T^{(1)} = 1 \cdot 10^{-2}. \quad (7.43)$$

As a final computational step in this demonstration, we compute another set of estimated eigenvalues and eigenvectors of $C^{(2)}$ after we generated samples $\mathbf{X}_j^{(2)} \stackrel{\text{approx.}}{\sim} \mu^{(2)}$ following the instructions above.

Let us have a look at the evolution of subspaces spanned by the columns of eigenvector matrices $\tilde{W}^{(\ell)}$, $\ell = 0, 1, 2$. In particular, we can ask how large we would have needed to choose, for instance, the dimension $k^{(0)}$ for $\tilde{W}_1^{(0)}$ in the first iteration in order to span the

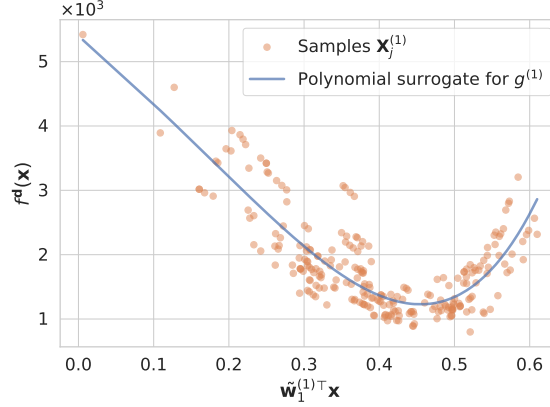


Figure 7.9: Misfit values $f^d(\mathbf{X}_j^{(1)})$ (orange dots) and a polynomial surrogate of fourth order for $g^{(1)}$ (blue solid line). The corresponding coefficient of determination is $r_{\ell=1}^2 = 0.7071$.

same (or a similar) subspace as we do for a particular $k^{(1)}$ and $\tilde{W}_1^{(1)}$. If this dimension is large (for a certain $k^{(1)}$), we know that the subspaces changed substantially. Of course, the same argumentation holds for the transition from iteration $\ell = 1 \rightarrow \ell = 2$.

The results are illustrated in Fig. 7.10. On the left, dimension $k^{(0)}$ varies on the x axis and the y axis plots the corresponding maximum principal angle [33] (in degree) between $\tilde{W}_1^{(0)} \in \mathbf{R}^{n \times k^{(0)}}$ and $\tilde{W}_1^{(1)} \in \mathbf{R}^{n \times k^{(1)}}$ for fixed $k^{(1)} = 1, 2, 3, 4$. The right subplot shows the same quantities for iterations $\ell = 1, 2$. We observe that the subspaces change more from iteration $\ell = 0 \rightarrow \ell = 1$ than from $\ell = 1 \rightarrow \ell = 2$ which might be caused by the substantial difference of the “bad” prior $\mu_0 = \mu^{(0)}$ and the approximation of $\mu^{(1)}$.

Finally, we additionally show the benefit of the iterative algorithm on another illustration. Of course, the deviation of the iterative approximations for posteriors $\mu^{(\ell)}$ from the true posterior μ^d (quantified with a suitable measure as, e. g., the Hellinger distance (see Def. A.2.8) or the KL divergence (see Def. A.2.10)) is difficult to estimate in practice. Nevertheless, to derive a statement on the quality of the posterior approximations, we compute corresponding push-forward distributions as in Ch. 6. That is, we choose 1,000 samples $\mathbf{X}_j^{(\ell)} \stackrel{\text{approx.}}{\sim} \mu^{(\ell)}$ for each $\ell = 0, 1, 2$ and evaluate $\mathcal{G}(\mathbf{X}_j^{(\ell)})$ for each sample to approximate the push-forward measure

$$\mu^{(\ell)}(\mathcal{G}^{-1}(\cdot)) : \mathcal{B}(\mathbf{R}^{n_d}) \rightarrow [0, 1]. \quad (7.44)$$

The results are provided in Fig. 7.11. The x axes again plot the time t in weeks and the y axes show the cumulative number of Ebola cases $C(t)$ on a logarithmic scale. Measured data (see also Fig. 7.4) is depicted by a dotted blue line; the mean and median of the push-forward approximations are drawn as dashed green and, respectively, solid orange line. The light orange area displays the intervals between the (empirical) 0.025- and 0.975-quantile denoted by $q_{0.025}$ and $q_{0.975}$, respectively, i. e., the intervals contain 95% of the computed outputs $\mathcal{G}(\mathbf{X}_j^{(\ell)})$. For the computation of (empirical) quantiles, the corresponding (empirical) *cumulative distribution function* (CDF) is *linearly* interpolated. Note that the top subplot shows the mentioned quantities for the push-forward *prior* $\mu_0 = \mu^{(0)}$.

We observe that the “forward uncertainty,” depicted by the light orange quantile intervals,

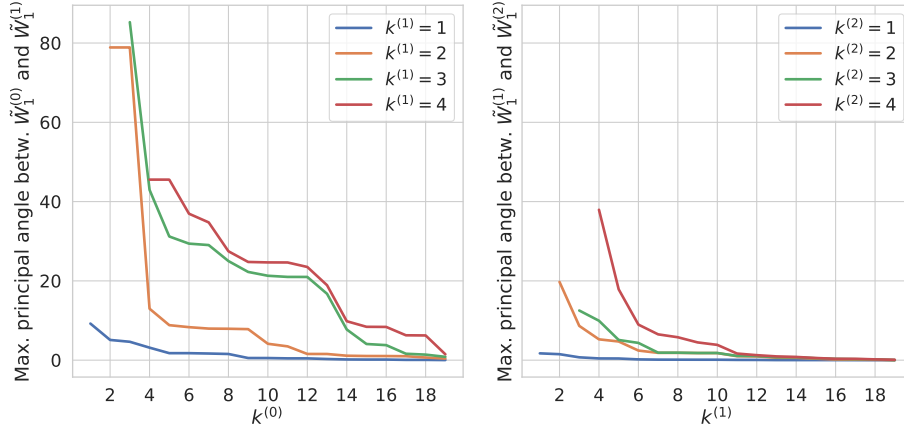


Figure 7.10: Both subplots display the distance of two successive subspaces spanned by the columns of $\tilde{W}_1^{(\ell)} \in \mathbf{R}^{n \times k^{(\ell)}}$ and $\tilde{W}_1^{(\ell+1)} \in \mathbf{R}^{n \times k^{(\ell+1)}}$ for $\ell = 0$ (left) and $\ell = 1$ (right). The x axis plots the dimension $k^{(\ell)}$ and the y axis shows, for fixed $k^{(\ell+1)} = 1, 2, 3, 4$, the distance which is measured by the maximum principal angle [33].

gradually reduces in each iteration. Also, the mean and median become reasonably compatible with the measured data.

7.5 Conclusions

Sec. 7.1 includes a motivational example demonstrating that it might be useful to consider an iterative approach for the construction of active subspaces in a Bayesian inverse problem. The computational experiment involving a model for the 2014 Ebola outbreak in West Africa (Sec. 7.4) shows that the two goals from above (Goal 1 and 2) are achieved in this case.

Indeed, we see that the first active direction changes from iteration $\ell = 0 \rightarrow \ell = 1$ such that it is more aligned to the behavior of the misfit function f^d in regions with higher posterior probability (Goal 1). Additionally, we confirm that subspaces containing more than one active direction can also vary substantially, especially regarding the first two iterations (see Fig. 7.10), which leads to the assumption that subspaces spanned by a certain number of columns from (iterated) eigenvector matrices $W^{(1)}$ or $W^{(2)}$ serve as a better starting point for further inference procedures (Goal 2) that intend to include more than, say, 2 or 3 directions.

Another supporting argument, which is partially included in the achievement of Goal 1, lies in the fact that the approach is robust w.r.t. “bad” prior distributions. Often, especially newly developed models have neither experience nor evidence on suitable values for involved model parameters which leads to the prior having a rather large support in which high posterior regions might occupy a relatively small part. We expect that our iterative scheme needs only one or a maximum of two iterations to detect regions containing more suitable parameter values (as is demonstrated in Sec. 7.4).

Although we do regard these outcomes as a first validation of the discussed iterative approach, further research is necessary to make it more “robust” in our opinion. For example, it is not clear so far how much the choice of the tempering factor $\beta_T^{(\ell)}$ from Sec. 7.4, which

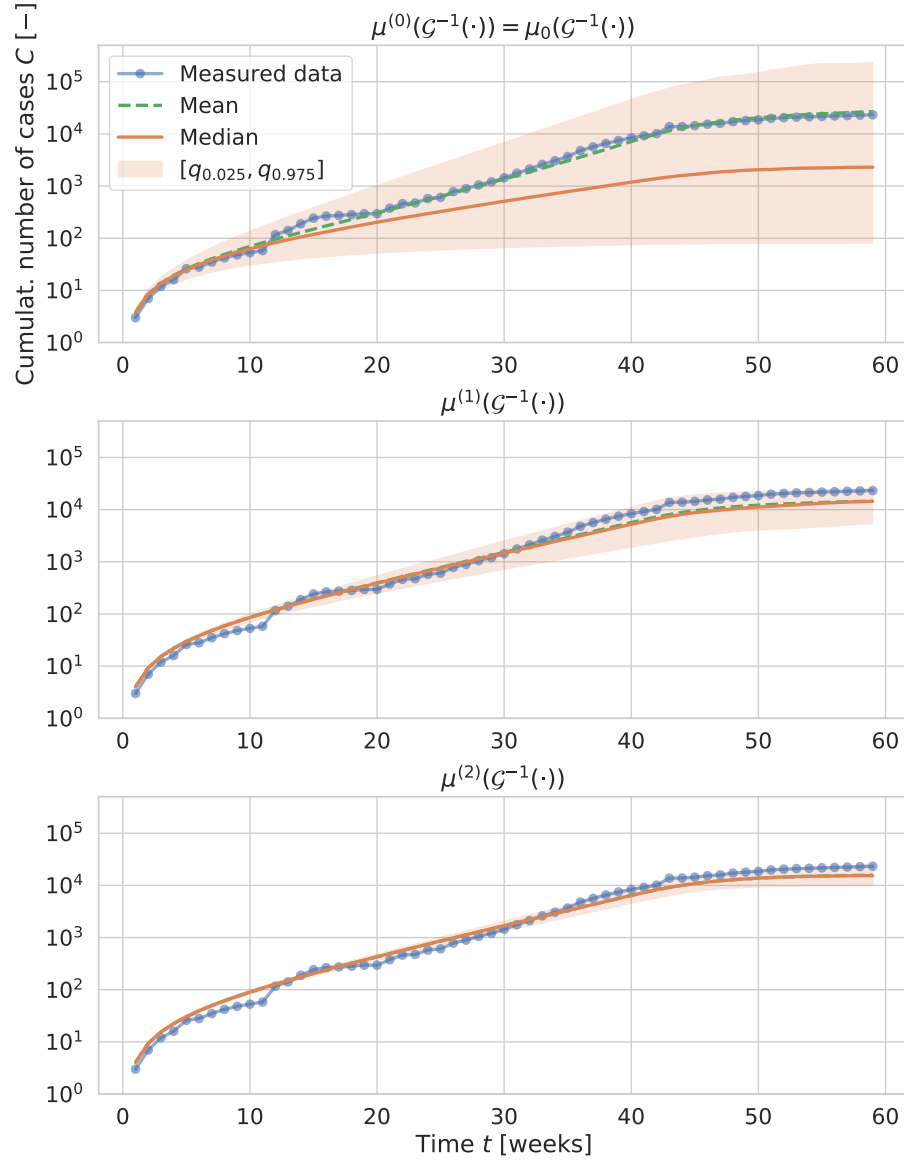


Figure 7.11: Gradual reduction of “forward uncertainty” demonstrated by approximations of push-forward posterior distributions $\mu^{(\ell)}(\mathcal{G}^{-1}(\cdot))$, $\ell = 0, 1, 2$. The light orange area displays the intervals between the (empirical) 0.025- and 0.975-quantile, i.e., it contains 95% of the computed outputs $\mathcal{G}(\mathbf{X}_j^{(\ell)})$, $j = 1, \dots, 10^3$.

aims to reduce the impact of an inaccurate surrogate while still exploiting information from it, influences the approximations of subspaces. Also, it would be interesting to see statistics on the maximum principal angles between subspaces (Fig. 7.10) as, e.g., their variability w.r.t. the random estimations of $C^{(\ell)}$.

Appendix A

Common results and definitions

A.1 General

Definition A.1.1. (Big Ω notation) For two compatible functions f and g , we define that

$$f \in \Omega(g) : \Longleftrightarrow g \in O(f). \quad (\text{A.1})$$

Lemma A.1.2 (Woodbury matrix identity). Let $A \in \mathbf{R}^{n \times n}$, $U \in \mathbf{R}^{n \times k}$, $C \in \mathbf{R}^{k \times k}$, and $V \in \mathbf{R}^{k \times n}$. Assume that A and C are invertible. Then, it holds that

$$(A + UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}. \quad (\text{A.2})$$

Theorem A.1.3 (Poincaré inequality for H^1). Assume that $D \subset \mathbf{R}^n$ is a bounded domain. Then, there exists a constant $C_P > 0$, depending only on D , such that for each $f \in H^1(D)$, it holds that

$$\|f - \bar{f}\|_{L^2(D)} \leq C_P \|\nabla f\|_{L^2(D)}, \quad (\text{A.3})$$

where $\bar{f} := \int_D f(\mathbf{x}) \, d\mathbf{x}$ denotes the mean of f on D . The constant C_P is called the Poincaré constant (of D).

If, in addition, D is convex, then $C_P = d/\pi$, where $d > 0$ is the diameter of D [21].

A.2 Stochastics

Definition A.2.1 ((Cross-)Covariance matrix). The *cross-covariance matrix* of two random vectors \mathbf{X}_1 and \mathbf{X}_2 is defined as

$$\mathbf{Cov}(\mathbf{X}_1, \mathbf{X}_2) := \mathbf{E}[(\mathbf{X}_1 - \mathbf{E}[\mathbf{X}_1])(\mathbf{X}_2 - \mathbf{E}[\mathbf{X}_2])^\top]. \quad (\text{A.4})$$

Similarly, we define the *covariance matrix* of a random vector \mathbf{X} as

$$\mathbf{Cov}(\mathbf{X}) := \mathbf{Cov}(\mathbf{X}, \mathbf{X}). \quad (\text{A.5})$$

Definition A.2.2 (Auto-correlation matrix). The *auto-correlation matrix* of a random vector \mathbf{X} is defined as

$$\mathbf{Cor}(\mathbf{X}) := \mathbf{E}[\mathbf{X}\mathbf{X}^\top]. \quad (\text{A.6})$$

Note that $\mathbf{Cov}(\mathbf{X}) = \mathbf{Cor}(\mathbf{X}) - \mathbf{E}[\mathbf{X}]\mathbf{E}[\mathbf{X}]^\top$.

Definition A.2.3 ((Sample) Pearson correlation coefficient). a) The *Pearson correlation coefficient* $r_{X,Y}$ of two scalar random variables X and Y is defined as

$$r_{X,Y} := \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}, \quad (\text{A.7})$$

where σ_X and σ_Y denote standard deviations of X and Y , respectively.

b) The *sample Pearson correlation coefficient* $r_{\mathbf{x}\mathbf{y}}$ for two given associated vectors $\mathbf{x} = (x_1, \dots, x_q)^\top \in \mathbf{R}^q$, $q \in \mathbf{N}$, and $\mathbf{y} = (y_1, \dots, y_q)^\top \in \mathbf{R}^q$ is defined as

$$r_{\mathbf{x}\mathbf{y}} = r_{\mathbf{x}\mathbf{y}}(\mathbf{x}, \mathbf{y}) := \frac{\|\mathbf{x} - \mathbf{y}\|_2^2}{\|\mathbf{x} - \bar{\mathbf{x}}\|_2 \|\mathbf{y} - \bar{\mathbf{y}}\|_2}, \quad (\text{A.8})$$

where $\bar{\mathbf{x}} = \bar{x} \cdot (1, \dots, 1)^\top \in \mathbf{R}^q$ and $\bar{\mathbf{y}} = \bar{y} \cdot (1, \dots, 1)^\top \in \mathbf{R}^q$ with sample means \bar{x} of \mathbf{x} and \bar{y} of \mathbf{y} .

Definition A.2.4 (Coefficient of determination, r^2 score [81, 102]). Let $\mathbf{y} = (y_1, \dots, y_q)^\top \in \mathbf{R}^q$, $q \in \mathbf{N}$, denote the vector of observed data with sample mean \bar{y} and $\mathbf{f} = (f_1, \dots, f_q)^\top \in \mathbf{R}^q$ associated predicted values. The *coefficient of determination*, denoted by r^2 (thus also called r^2 score), is defined as

$$r^2 = r^2(\mathbf{y}, \mathbf{f}) := 1 - \frac{\|\mathbf{f} - \bar{\mathbf{y}}\|_2^2}{\|\mathbf{y} - \bar{\mathbf{y}}\|_2^2}, \quad (\text{A.9})$$

where $\bar{\mathbf{y}} := \bar{y} \cdot (1, \dots, 1)^\top \in \mathbf{R}^q$.

It quantifies the proportion of variance in a dependent variable that is explained by independent variables. Normally, $r^2 \in [0, 1]$; in regression contexts, $r^2 = 0$ indicates that the sample mean of observed data \bar{y} is a better fit than the predicted values, whereas $r^2 = 1$ means that the predictions perfectly fit the data. In the context of *simple linear regression*, (i. e., fitting data by a linear function in one variable by using *ordinary least square estimators*) we have that

$$r^2 = r_{\mathbf{x}\mathbf{y}}^2$$

with $\mathbf{x} = \mathbf{f}$.

Lemma A.2.5 (Completing the squares in the multivariate case). Let $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2 \in \mathbf{R}^n$ and $\Sigma_1, \Sigma_2 \in \mathbf{R}^{n \times n}$ be symmetric matrices such that $M := \Sigma_1 + \Sigma_2$ is invertible. Then, we can write a sum of two corresponding quadratic forms as a sum of a single quadratic form and a constant term, i. e., for variable $\mathbf{x} \in \mathbf{R}^n$ we have

$$\begin{aligned} & (\mathbf{x} - \boldsymbol{\mu}_1)^\top \Sigma_1 (\mathbf{x} - \boldsymbol{\mu}_1) + (\mathbf{x} - \boldsymbol{\mu}_2)^\top \Sigma_2 (\mathbf{x} - \boldsymbol{\mu}_2) \\ &= \\ & (\mathbf{x} - M^{-1}\mathbf{v})^\top M (\mathbf{x} - M^{-1}\mathbf{v}) + R, \end{aligned} \quad (\text{A.10})$$

where

$$\mathbf{v} := \Sigma_1 \boldsymbol{\mu}_1 + \Sigma_2 \boldsymbol{\mu}_2 \quad \text{and} \quad R := -\mathbf{v}^\top M^{-1} \mathbf{v} + \boldsymbol{\mu}_1^\top \Sigma_1 \boldsymbol{\mu}_1 + \boldsymbol{\mu}_2^\top \Sigma_2 \boldsymbol{\mu}_2. \quad (\text{A.11})$$

Proof. First, note that for any vectors $\mathbf{x}, \mathbf{v} \in \mathbf{R}^n$ and a symmetric invertible matrix $M \in \mathbf{R}^{n \times n}$, it holds that

$$\begin{aligned} (\mathbf{x} - M^{-1}\mathbf{v})^\top M (\mathbf{x} - M^{-1}\mathbf{v}) &= \mathbf{x}^\top M \mathbf{x} - 2\mathbf{v}^\top M^{-1} M \mathbf{x} + \mathbf{v}^\top M^{-1} M M^{-1} \mathbf{v} \\ &= \mathbf{x}^\top M \mathbf{x} - 2\mathbf{v}^\top \mathbf{x} + \mathbf{v}^\top M^{-1} \mathbf{v}. \end{aligned} \quad (\text{A.12})$$

Expanding the left hand side of (A.10) gives

$$\begin{aligned} (\mathbf{x} - \boldsymbol{\mu}_1)^\top \Sigma_1 (\mathbf{x} - \boldsymbol{\mu}_1) + (\mathbf{x} - \boldsymbol{\mu}_2)^\top \Sigma_2 (\mathbf{x} - \boldsymbol{\mu}_2) = \\ = \mathbf{x}^\top (\Sigma_1 + \Sigma_2) \mathbf{x} - 2\mathbf{x}^\top (\Sigma_1 \boldsymbol{\mu}_1 + \Sigma_2 \boldsymbol{\mu}_2) + \boldsymbol{\mu}_1^\top \Sigma_1 \boldsymbol{\mu}_1 + \boldsymbol{\mu}_2^\top \Sigma_2 \boldsymbol{\mu}_2. \end{aligned} \quad (\text{A.13})$$

Now, use (A.12) with the definitions in (A.11) to get the result (A.10). \blacksquare

Theorem A.2.6 (Conditional distribution for jointly Gaussian random variables [149]). *Let $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2) \in \mathbf{R}^n$ be a random vector with $\mathbf{X}_1 \in \mathbf{R}^q$ and $\mathbf{X}_2 \in \mathbf{R}^{n-q}$ for $q \in \{1, \dots, n-1\}$. Assume that $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ with*

$$\boldsymbol{\mu} = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2) \quad \text{and} \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \quad (\text{A.14})$$

such that the dimensions of both subcomponents correspond to \mathbf{X}_1 and \mathbf{X}_2 . Then, for $\mathbf{x}_2 \in \mathbf{R}^{n-q}$,

$$\mathbf{X}_1 | \mathbf{X}_2 = \mathbf{x}_2 \sim \mathcal{N}(\bar{\boldsymbol{\mu}}, \bar{\Sigma}) \quad (\text{A.15})$$

with

$$\bar{\boldsymbol{\mu}} = \boldsymbol{\mu}_1 + \Sigma_{12} \Sigma_{22}^{-1} (\mathbf{x}_2 - \boldsymbol{\mu}_2) \quad \text{and} \quad \bar{\Sigma} = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}. \quad (\text{A.16})$$

Remark. (1) The expression for $\bar{\Sigma}$ is called the *Schur complement* of Σ_{22} in Σ . (2) Note that $\bar{\Sigma}$ does not depend on the condition variable $\mathbf{X}_2 = \mathbf{x}_2$ which is typical for conditional Gaussian distributions.

Theorem A.2.7 (Fernique's theorem). *If μ is a Gaussian measure on a Banach space \mathcal{X} , then there exists $\alpha > 0$ such that*

$$\int_{\mathcal{X}} \exp(\alpha \|\mathbf{x}\|_{\mathcal{X}}^2) \mu(d\mathbf{x}) < \infty. \quad (\text{A.17})$$

Definition A.2.8 (Hellinger distance). Let \mathbf{P} and \mathbf{Q} be two probability measures that are absolutely continuous w.r.t. a σ -finite reference measure λ . Then, the *Hellinger distance* of \mathbf{P} and \mathbf{Q} is defined as

$$d_{\text{Hell}}(\mathbf{P}, \mathbf{Q})^2 := \frac{1}{2} \int \left(\sqrt{\frac{d\mathbf{P}}{d\lambda}} - \sqrt{\frac{d\mathbf{Q}}{d\lambda}} \right)^2 d\lambda, \quad (\text{A.18})$$

where $d\mathbf{P}/d\lambda$ and $d\mathbf{Q}/d\lambda$ denote Radon-Nikodym derivatives of \mathbf{P} and \mathbf{Q} , respectively.

Definition A.2.9 (Total variation distance of probability measures). The *total variation distance* of two probability measures \mathbf{P} and \mathbf{Q} , defined on the same σ -algebra \mathcal{A} , is defined as

$$d_{\text{TV}}(\mathbf{P}, \mathbf{Q}) := \|\mathbf{P} - \mathbf{Q}\|_{\text{TV}} := \sup_{A \in \mathcal{A}} |\mathbf{P}(A) - \mathbf{Q}(A)|. \quad (\text{A.19})$$

Definition A.2.10 (Kullback-Leibler divergence). Let \mathbf{P} and \mathbf{Q} two probability measures such that \mathbf{P} is absolutely continuous w.r.t. \mathbf{Q} . The *Kullback-Leibler divergence* of \mathbf{Q} from \mathbf{P} is defined as

$$d_{\text{KL}}(\mathbf{P} \parallel \mathbf{Q}) := \int_{\mathcal{X}} \log \left(\frac{d\mathbf{P}}{d\mathbf{Q}} \right) d\mathbf{P}, \quad (\text{A.20})$$

where $d\mathbf{P}/d\mathbf{Q}$ is the Radon-Nikodym derivative of \mathbf{P} w.r.t. \mathbf{Q} .

Moreover, if λ is a measure on \mathcal{X} dominating \mathbf{P} and \mathbf{Q} , such that densities $p := d\mathbf{P}/d\lambda$ and $q := d\mathbf{Q}/d\lambda$ exist, then

$$d_{\text{KL}}(\mathbf{P} \parallel \mathbf{Q}) = \int_{\mathcal{X}} p \log \left(\frac{p}{q} \right) d\lambda. \quad (\text{A.21})$$

Theorem A.2.11 (Strong law of large numbers). *Let $(\mathbf{X}_j)_{j \in \mathbf{N}}$ be an infinite sequence of i.i.d. random variables with $\mathbf{E}[\mathbf{X}_1] = \mathbf{E}[\mathbf{X}_2] = \dots = \mathbf{m} \in \mathbf{R}^n$. Then, the sample average converges to \mathbf{m} almost surely, i. e.,*

$$\frac{1}{N} \sum_{j=1}^N \mathbf{X}_j \xrightarrow[N \rightarrow \infty]{a.s.} \mathbf{m}. \quad (\text{A.22})$$

Definition A.2.12 (Markov kernel). Let $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ and $(\mathcal{Y}, \mathcal{B}(\mathcal{Y}))$ be two measurable topological spaces (equipped with their Borel σ -algebra). A map $K : \mathcal{X} \times \mathcal{B}(\mathcal{Y}) \rightarrow [0, 1]$ is called a *Markov kernel* if

1. $K(\mathbf{x}, \cdot)$ is a probability measure on $(\mathcal{Y}, \mathcal{B}(\mathcal{Y}))$ for each $\mathbf{x} \in \mathcal{X}$,
2. $K(\cdot, A)$ is $\mathcal{B}(\mathcal{X})$ -measurable for each $A \in \mathcal{B}(\mathcal{Y})$.

Definition A.2.13 (Reversible Markov chain [193, Def. 6.44]). A stationary Markov chain $(\mathbf{X}_j)_{j \in \mathbf{N}_0}$ with values in \mathcal{X} is called *reversible* if the distribution of $\mathbf{X}_j \mid \mathbf{X}_{j-1} = \mathbf{x}$ is equal to the distribution of $\mathbf{X}_j \mid \mathbf{X}_{j+1} = \mathbf{x}$ for every $j \in \mathbf{N}$ and $\mathbf{x} \in \mathcal{X}$.

Definition A.2.14 (Probabilistic Poincaré inequality). Let \mathbf{X} be a random variable with values in \mathbf{R}^n . The distribution of \mathbf{X} , denoted by $\mathbf{P}_{\mathbf{X}}$, is said to *satisfy a probabilistic Poincaré inequality with Poincaré constant $C_P > 0$* if

$$\text{Var}(f(\mathbf{X})) \leq C_P \mathbf{E}[\|\nabla f(\mathbf{X})\|_2^2] \quad (\text{A.23})$$

for each real-valued function f defined on the support of $\mathbf{P}_{\mathbf{X}}$ with

$$\mathbf{E}[\|\nabla f(\mathbf{X})\|_2^2] < \infty. \quad (\text{A.24})$$

Theorem A.2.15 (Eigenvalue Bernstein inequality for subexponential matrices [101, Thm. 5.3]). *Consider a finite sequence $(M_j)_{j=1, \dots, N}$ of random, independent, and symmetric matrices from $\mathbf{R}^{n \times n}$, all of which satisfy the subexponential moment growth condition*

$$\mathbf{E}[M_j^p] \preceq \frac{p!}{2} B^{p-2} \Sigma_j^2 \quad (\text{A.25})$$

for each $p = 2, 3, \dots$, where B is a positive constant and Σ_j^2 are positive semi-definite matrices. Given an integer $\ell \leq n$, set

$$\mu_\ell = \lambda_\ell \left(\sum_{j=1}^N \mathbf{E}[M_j] \right). \quad (\text{A.26})$$

Choose V_+ as an orthogonal matrix of size $n \times (n - \ell + 1)$ that satisfies

$$\mu_\ell = \lambda_{\max} \left(\sum_{j=1}^N V_+^\top \mathbf{E}[M_j] V_+ \right) \quad (\text{A.27})$$

and define

$$\sigma_\ell^2 := \lambda_{\max} \left(\sum_{j=1}^N V_+^\top \Sigma_j^2 V_+ \right). \quad (\text{A.28})$$

Then, for any $t \geq 0$,

$$\mathbf{P} \left(\lambda_\ell \left(\sum_{j=1}^N M_j \right) \geq \mu_\ell + t \right) \leq \begin{cases} (n - \ell + 1) \exp(-t^2/(4\sigma_\ell^2)) & \text{if } t \leq \sigma_\ell^2/B, \\ (n - \ell + 1) \exp(-t/(4B)) & \text{if } t \geq \sigma_\ell^2/B. \end{cases} \quad (\text{A.29})$$

Invariance of Poincaré constants for multivariate normal distributions

Adopting the setup of Ch. 3, we show that the Poincaré constant of a multivariate normal distribution $\mathbf{P}_{\mathbf{X}} = \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ for a mean vector $\boldsymbol{\mu} \in \mathbf{R}^n$ and a covariance matrix $\Sigma \in \mathbf{R}^{n \times n}$ is the same for $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}(\cdot|\mathbf{y})$, $\mathbf{y} \in \mathbf{R}^k$, in general.

We already know that the Poincaré constant of $\mathbf{P}_{\mathbf{X}}$ is $\lambda_{\max}(\Sigma)$ by [47] and the computation in (3.52). The random variable \mathbf{X} , orthogonally transformed by the eigenvector matrix W from (3.8), is distributed according to

$$W^\top \mathbf{X} \sim \mathcal{N}(W^\top \boldsymbol{\mu}, W^\top \Sigma W). \quad (\text{B.1})$$

Taking active and inactive variables and also the split of W from (3.10) into account, we get that

$$\begin{pmatrix} \mathbf{Y} \\ \mathbf{Z} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} W_1^\top \boldsymbol{\mu} \\ W_2^\top \boldsymbol{\mu} \end{pmatrix}, \begin{pmatrix} W_1^\top \Sigma W_1 & W_1^\top \Sigma W_2 \\ W_2^\top \Sigma W_1 & W_2^\top \Sigma W_2 \end{pmatrix}\right). \quad (\text{B.2})$$

Thm. A.2.6 gives that $\mathbf{Z} | \mathbf{Y} = \mathbf{y} \sim \mathcal{N}(\boldsymbol{\mu}_{\mathbf{Z}|\mathbf{Y}}, \Sigma_{\mathbf{Z}|\mathbf{Y}})$ with

$$\boldsymbol{\mu}_{\mathbf{Z}|\mathbf{Y}} = W_2^\top \boldsymbol{\mu} + W_2^\top \Sigma W_1 (W_1^\top \Sigma W_1)^{-1} (\mathbf{y} - W_1^\top \boldsymbol{\mu}) \quad (\text{B.3})$$

and

$$\Sigma_{\mathbf{Z}|\mathbf{Y}} = W_2^\top \Sigma W_2 - W_2^\top \Sigma W_1 (W_1^\top \Sigma W_1)^{-1} W_1^\top \Sigma W_2. \quad (\text{B.4})$$

Note that $\Sigma_{\mathbf{Z}|\mathbf{Y}}$ is the *Schur complement* of the block $W_2^\top \Sigma W_2$ in $W^\top \Sigma W$ which does not depend on the condition variable \mathbf{y} .

The *interlacing property* of Schur complements of (either positive or negative) semi-definite Hermitian matrices [244, Thm. 5] implies that

$$\lambda_{\min}(W^\top \Sigma W) \leq \lambda_{\min}(\Sigma_{\mathbf{Z}|\mathbf{Y}}) \leq \lambda_{\max}(\Sigma_{\mathbf{Z}|\mathbf{Y}}) \leq \lambda_{\max}(W^\top \Sigma W). \quad (\text{B.5})$$

Since $W^\top = W^{-1}$, $W^\top \Sigma W$ is *similar* to Σ implying that the eigenvalues of $W^\top \Sigma W$ are the same as the eigenvalues of Σ . Therefore,

$$\lambda_{\min}(\Sigma) \leq \lambda_{\min}(\Sigma_{\mathbf{Z}|\mathbf{Y}}) \leq \lambda_{\max}(\Sigma_{\mathbf{Z}|\mathbf{Y}}) \leq \lambda_{\max}(\Sigma). \quad (\text{B.6})$$

Hence, we might get a smaller Poincaré constant for $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}$ than $\lambda_{\max}(\Sigma)$ for a *specific* orthogonal transformation W , but *in general* the Poincaré constant for $\mathbf{P}_{\mathbf{Z}|\mathbf{Y}}$ is equal to $\lambda_{\max}(\Sigma)$.

Model equations for LuKARS model

This appendix is taken from [258, Appendix A].

In LuKARS, the following balance equation is solved for each individual hydrotope:

$$\frac{de_i}{dt} = \begin{cases} S_i - \frac{Q_{\text{sec},i} + Q_{\text{is},i} + Q_{\text{hyd},i}}{a_i} & \text{if } e_i > 0, \\ 0 & \text{if } e_i = 0. \end{cases} \quad (\text{C.1})$$

e_i is the water level [L] in hydrotope i , t [T] indicates the time and S_i is a hydrotope-specific sink and source term in form of a mass balance of precipitation, snow melt, evapotranspiration, and interception. We used the temperature index approach from [187] to calculate snow melt. Interception was estimated based on indications for beech forests in [83]. Then, evapotranspiration was calculated based on the method of [260]. $Q_{\text{sec},i}$ [L^3T^{-1}] summarizes all flow terms that do not contribute to the discharge at an investigated karst spring, i. e., secondary spring discharge and overland flow. $Q_{\text{is},i}$ [L^3T^{-1}] represents the discharge from hydrotope i to a linear baseflow storage, considered as groundwater recharge. $Q_{\text{hyd},i}$ [L^3T^{-1}] is a hydrotope-specific quickflow component through preferential flow paths (e. g., subsurface conduits) with a direct connection to the spring outlet. a_i [L^2] is the space covered by a respective hydrotope.

The following balance equation is solved for the baseflow storage:

$$\frac{de_b}{dt} = \begin{cases} \frac{\sum_i Q_{\text{is},i} - Q_b}{A} & \text{if } e_b > 0, \\ 0 & \text{if } e_b = 0. \end{cases} \quad (\text{C.2})$$

e_b is the water level [L] in the baseflow storage and $\sum_i Q_{\text{is},i}$ [L^3T^{-1}] integrates the flows from all hydrotopes to the baseflow storage. Q_b [L^3T^{-1}] indicates water flow from the storage B to the spring and simulates the matrix contribution from the saturated zone to the spring discharge. The variable A [L^2] is the space of the entire recharge area. The discretized forms of (C.1) and (C.2), as shown in (C.3) and (C.4), are solved for each time step n :

$$e_{i,n+1} = \max \left[0, e_{i,n} + \left(S_{i,n} - \frac{Q_{\text{sec},i,n} + Q_{\text{is},i,n} + Q_{\text{hyd},i,n}}{a_i} \right) \cdot \Delta t \right], \quad (\text{C.3})$$

$$e_{b,n+1} = \max \left[0, e_{b,n} + \left(\frac{\sum_i (Q_{\text{is},i,n}) - Q_{b,n}}{A} \right) \cdot \Delta t \right]. \quad (\text{C.4})$$

The discharge terms are computed as:

$$Q_{\text{hyd},i,n} = \varepsilon \left[\frac{\max(0, e_{i,n} - e_{\min,i})}{e_{\max,i} - e_{\min,i}} \right]^{\alpha_i} \cdot \frac{k_{\text{hyd},i}}{l_{\text{hyd},i}} \cdot a_i, \quad (\text{C.5})$$

$$Q_{\text{is},i,n} = k_{\text{is},i} \cdot e_{i,n} \cdot a_i, \quad (\text{C.6})$$

$$Q_{\text{sec},i,n} = k_{\text{sec},i} \cdot \max(0, e_{i,n} - e_{\text{sec},i}) \cdot a_i, \quad (\text{C.7})$$

$$Q_{\text{b},n} = k_{\text{b}} \cdot e_{\text{b},n} \cdot A. \quad (\text{C.8})$$

$e_{\max,i}$ [L] and $e_{\min,i}$ [L] are the upper and lower storage thresholds of hydrotope i . The exponent α_i controls the magnitude of the quickflow component from each hydrotope. $e_{\text{sec},i}$ [L] represents a hydrotope-specific activation level for $Q_{\text{sec},i}$. $k_{\text{is},i}$ [LT^{-1}] and $k_{\text{sec},i}$ [LT^{-1}] are the specific discharge parameters for $Q_{\text{is},i}$ [L^3T^{-1}] and $Q_{\text{sec},i}$ [L^3T^{-1}]. $k_{\text{hyd},i}$ [L^2T^{-1}] indicates the specific discharge parameter for the quickflow and $l_{\text{hyd},i}$ [L] is the mean distance of hydrotope i to the adjacent spring, allowing to account for the relative location and distribution of hydrotope i in a specific recharge area. The ratio between $k_{\text{hyd},i}$ and $l_{\text{hyd},i}$ is the hydrotope discharge coefficient. Then, the dimensionless connectivity/activation indicator ε is defined as:

$$\varepsilon_{n+1} = 0 \quad \text{if} \quad \begin{cases} \varepsilon_n = 0 \ \& \ e_{i,n+1} < e_{\max,i} \ \text{or} \\ \varepsilon_n = 1 \ \& \ e_{i,n+1} \leq e_{\min,i}, \end{cases} \quad (\text{C.9})$$

$$\varepsilon_{n+1} = 1 \quad \text{if} \quad \begin{cases} \varepsilon_n = 0 \ \& \ e_{i,n+1} \geq e_{\max,i} \ \text{or} \\ \varepsilon_n = 1 \ \& \ e_{i,n+1} > e_{\min,i}. \end{cases} \quad (\text{C.10})$$

Appendix D

Pending proofs

D.1 Proof of Prop. 7.3.1

Proof of Prop. 7.3.1. Without loss of generality, assume that $\mathbf{m}_0 = 0$ implying that $\mathbf{d} = 0$ and

$$f^{\mathbf{d}}(\mathbf{x}) = \frac{1}{2\gamma^2} \|\mathbf{d} - A\mathbf{x}\|_2^2 = \frac{1}{2\gamma^2} \|A\mathbf{x}\|_2^2 = \frac{1}{2\gamma^2} \mathbf{x}^\top A^\top A \mathbf{x}. \quad (\text{D.1})$$

If $\mathbf{m}_0 \neq 0$, regard the translated variable $\hat{\mathbf{X}} := \mathbf{X} - \mathbf{m}_0$ with $\hat{\mathcal{G}}(\hat{\mathbf{x}}) := \mathcal{G}(\hat{\mathbf{x}} + \mathbf{m}_0)$, $\hat{\mathbf{d}} = \hat{\mathcal{G}}(\hat{\mathbf{m}}_0) = \mathbf{d}$ for $\hat{\mathbf{m}}_0 = 0$, and $\hat{\mu}_0 = \mathcal{N}(0, I)$. Then,

$$\hat{f}^{\hat{\mathbf{d}}}(\hat{\mathbf{x}}) = \frac{1}{2\gamma^2} \|\hat{\mathbf{d}} - \hat{\mathcal{G}}(\hat{\mathbf{x}})\|_2^2 \quad (\text{D.2})$$

$$= \frac{1}{2\gamma^2} \|\mathbf{d} - A\hat{\mathbf{x}} - A\mathbf{m}_0\|_2^2 \quad (\text{D.3})$$

$$= \frac{1}{2\gamma^2} \|A\hat{\mathbf{x}}\|_2^2 \quad (\text{D.4})$$

$$= f^{\mathbf{d}}(\hat{\mathbf{x}}). \quad (\text{D.5})$$

Of course, the translation to the posterior $\hat{\mu}^{\hat{\mathbf{d}}}$ has to be reversed to obtain the posterior $\mu^{\mathbf{d}}$.

With $\mathbf{m}_0 = 0$ and $\mathbf{d} = 0$, we have that

$$\nabla f(\mathbf{x}) = \frac{1}{\gamma^2} A^\top A \mathbf{x}. \quad (\text{D.6})$$

It follows that

$$C := C^{(0)} = \frac{1}{\gamma^2} A^\top A \left(\int \mathbf{x} \mathbf{x}^\top \rho^{(0)}(\mathbf{x}) \, \mathrm{d}\mathbf{x} \right) A^\top A \frac{1}{\gamma^2} \quad (\text{D.7})$$

$$= \frac{1}{\gamma^4} A^\top A I A^\top A = \left(\frac{1}{\gamma^2} A^\top A \right)^2 = W^{(0)} \Lambda^{(0)} W^{(0)\top}. \quad (\text{D.8})$$

Note that a singular value decomposition $A = U_A \Sigma_A V_A^\top$ yields that

$$W^{(0)} = V_A \quad \text{and} \quad \Lambda^{(0)} = \left(\frac{1}{\gamma^2} \Sigma_A^\top \Sigma_A \right)^2. \quad (\text{D.9})$$

Let $\ell \in \mathbf{N}_0$. Anticipating, we will see that $W^{(\ell)} = W^{(0)}P^{(\ell)} = WP^{(\ell)}$ for a permutation matrix $P^{(\ell)}$. The identities

$$W_1 = W \begin{pmatrix} I \\ 0 \end{pmatrix} \quad \text{and} \quad W_2 = W \begin{pmatrix} 0 \\ I \end{pmatrix} \quad (\text{D.10})$$

motivate the definition of

$$P_1^{(\ell)} := P^{(\ell)} \begin{pmatrix} I \\ 0 \end{pmatrix} \in \mathbf{R}^{n \times k^{(\ell)}} \quad \text{and} \quad P_2^{(\ell)} := P^{(\ell)} \begin{pmatrix} 0 \\ I \end{pmatrix} \in \mathbf{R}^{n \times (n-k^{(\ell)})}. \quad (\text{D.11})$$

Then, for $\mathbf{y} \in \mathcal{Y}^{(\ell)} := W_1^{(\ell)\top} \mathcal{X}$ and $\mathbf{z} \in \mathcal{Z}^{(\ell)} := W_2^{(\ell)\top} \mathcal{X}$, it holds that

$$f(\llbracket \mathbf{y}, \mathbf{z} \rrbracket_{W^{(\ell)}}) = \frac{1}{2\gamma^2} \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix}^\top W^{(\ell)\top} A^\top A W^{(\ell)} \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix} \quad (\text{D.12})$$

$$= \frac{1}{2} \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix}^\top W^{(\ell)\top} C^{1/2} W^{(\ell)} \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix} \quad (\text{D.13})$$

$$= \frac{1}{2} \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix}^\top P^{(\ell)\top} \Lambda^{1/2} P^{(\ell)} \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix} \quad (\text{D.14})$$

$$= \frac{1}{2} \mathbf{y}^\top P_1^{(\ell)\top} \Lambda^{1/2} P_1^{(\ell)} \mathbf{y} + \frac{1}{2} \mathbf{z}^\top P_2^{(\ell)\top} \Lambda^{1/2} P_2^{(\ell)} \mathbf{z} \quad (\text{D.15})$$

$$= \frac{1}{2} \mathbf{y}^\top E^{(\ell)1/2} \mathbf{y} + r(\mathbf{z}), \quad (\text{D.16})$$

where $E^{(\ell)} := P_1^{(\ell)\top} \Lambda P_1^{(\ell)}$ is a diagonal matrix with eigenvalues said to be *selected by* $P_1^{(\ell)}$ on its diagonal and r is some function depending only on \mathbf{z} . It follows that

$$g^{(\ell)}(\mathbf{y}) = \frac{1}{2} \mathbf{y}^\top E^{(\ell)1/2} \mathbf{y} + R \quad (\text{D.17})$$

for some $R \geq 0$.

Since the prior, as a standard normal distribution, is rotationally symmetric, the prior density on the active subspace at step ℓ is

$$\rho_{0,\mathbf{Y}}(\mathbf{y}) = \rho_{0,\mathbf{Y},W^{(\ell)}}(\mathbf{y}) \propto \exp\left(-\frac{1}{2}\|\mathbf{y}\|_2^2\right). \quad (\text{D.18})$$

Note that it does not depend on ℓ .

To compute the distribution of $\mathring{\mathbf{Y}}^{(\ell+1)}$, we continue with

$$\rho_g^{(\ell)}(\mathbf{y}) \propto \exp(-g^{(\ell)}(\mathbf{y})) \rho_{0,\mathbf{Y}}(\mathbf{y}) \quad (\text{D.19})$$

$$\propto \exp\left(-\frac{1}{2} \mathbf{y}^\top E^{(\ell)1/2} \mathbf{y} - \frac{1}{2} \|\mathbf{y}\|_2^2\right) \quad (\text{D.20})$$

$$= \exp\left(-\frac{1}{2} \mathbf{y}^\top \left(I + E^{(\ell)1/2}\right) \mathbf{y}\right) \quad (\text{D.21})$$

yielding that

$$\mathring{\mathbf{Y}}^{(\ell+1)} \sim \mathcal{N}\left(0, \left(I + E^{(\ell)1/2}\right)^{-1}\right). \quad (\text{D.22})$$

We are ready to prove the statement above for the initial step $\ell = 0$. Let $k^{(0)} \in \{1, \dots, n-1\}$ be the dimension of the active subspace according to Alg. 7. Since $P^{(0)} = I$, we get that

$$E^{(0)} = P_1^{(0)\top} \Lambda P_1^{(0)} = \Lambda_{1:k^{(0)}}. \quad (\text{D.23})$$

Combining this and the fact that

$$\dot{\mathbf{Z}}^{(1)} = (\mathbf{Z}^{(0)} \mid \mathbf{Y}^{(0)} = \mathbf{y}) = \mathbf{Z}^{(0)} \sim \mathcal{N}(0, I), \quad (\text{D.24})$$

it holds that

$$\mathbf{X}^{(1)} = W^{(0)} \begin{pmatrix} \dot{\mathbf{Y}}^{(1)} \\ \dot{\mathbf{Z}}^{(1)} \end{pmatrix} \sim \mathcal{N}(0, \Sigma^{(1)}), \quad (\text{D.25})$$

where

$$\Sigma^{(1)} = W^{(0)} \begin{pmatrix} \left(I + \Lambda_{1:k^{(0)}}^{1/2}\right)^{-1} & 0 \\ 0 & I \end{pmatrix} W^{(0)\top}. \quad (\text{D.26})$$

That is, we have that $K^{(0)} = k^{(0)}$.

Now, assume that the statement above holds for some $\ell \in \mathbf{N}_0$ and regard the step from ℓ to $\ell + 1$. We compute

$$C^{(\ell)} = \frac{1}{\gamma^2} A^\top A \Sigma^{(\ell)} A^\top A \frac{1}{\gamma^2} \quad (\text{D.27})$$

$$= \frac{1}{\gamma^2} A^\top A W \begin{pmatrix} \left(I + \Lambda_{1:K^{(\ell)}}^{1/2}\right)^{-1} & 0 \\ 0 & I \end{pmatrix} W^\top A^\top A \frac{1}{\gamma^2} \quad (\text{D.28})$$

$$= W \Lambda^{1/2} \begin{pmatrix} \left(I + \Lambda_{1:K^{(\ell)}}^{1/2}\right)^{-1} & 0 \\ 0 & I \end{pmatrix} \Lambda^{1/2} W^\top \quad (\text{D.29})$$

$$= W \begin{pmatrix} \Lambda_{1:K^{(\ell)}}^{1/2} \left(I + \Lambda_{1:K^{(\ell)}}^{1/2}\right)^{-1} \Lambda_{1:K^{(\ell)}}^{1/2} & 0 \\ 0 & \Lambda_{K^{(\ell)}+1:n} \end{pmatrix} W^\top \quad (\text{D.30})$$

$$= W P^{(\ell)} P^{(\ell)\top} \begin{pmatrix} \Lambda_{1:K^{(\ell)}}^{1/2} \left(I + \Lambda_{1:K^{(\ell)}}^{1/2}\right)^{-1} \Lambda_{1:K^{(\ell)}}^{1/2} & 0 \\ 0 & \Lambda_{K^{(\ell)}+1:n} \end{pmatrix} P^{(\ell)} P^{(\ell)\top} W^\top \quad (\text{D.31})$$

$$=: W^{(\ell)} \Lambda^{(\ell)} W^{(\ell)\top}, \quad (\text{D.32})$$

where $W^{(\ell)} = W P^{(\ell)}$ and

$$\Lambda^{(\ell)} = P^{(\ell)\top} \begin{pmatrix} \Lambda_{1:K^{(\ell)}}^{1/2} \left(I + \Lambda_{1:K^{(\ell)}}^{1/2}\right)^{-1} \Lambda_{1:K^{(\ell)}}^{1/2} & 0 \\ 0 & \Lambda_{K^{(\ell)}+1:n} \end{pmatrix} P^{(\ell)}. \quad (\text{D.33})$$

The permutation matrix $P^{(\ell)}$ is chosen such that the eigenvalues in $\Lambda^{(\ell)}$ are descending. We decide for $k^{(\ell)} \in \{1, \dots, n-1\}$ as the dimension of the active subspace based on the eigenvalues in $\Lambda^{(\ell)}$. Recall that $E^{(\ell)} = P_1^{(\ell)\top} \Lambda P_1^{(\ell)}$ which modifies (D.22) to

$$\dot{\mathbf{Y}}^{(\ell+1)} \sim \mathcal{N}\left(0, \left(I + P_1^{(\ell)\top} \Lambda^{1/2} P_1^{(\ell)}\right)^{-1}\right) \quad (\text{D.34})$$

$$= \mathcal{N} \left(0, P_1^{(\ell)\top} \left(I + \Lambda^{1/2} \right)^{-1} P_1^{(\ell)} \right). \quad (\text{D.35})$$

To get the distribution of $\dot{\mathbf{Z}}^{(\ell+1)} = (\mathbf{Z}^{(\ell)} | \mathbf{Y}^{(\ell)} = \mathbf{y})$, we need to look at the distribution of

$$W^{(\ell)\top} \mathbf{X}^{(\ell)} \sim \mathcal{N} \left(0, W^{(\ell)\top} \Sigma^{(\ell)} W^{(\ell)} \right). \quad (\text{D.36})$$

The covariance matrix of $W^{(\ell)\top} \mathbf{X}^{(\ell)}$ computes to

$$W^{(\ell)\top} \Sigma^{(\ell)} W^{(\ell)} = P^{(\ell)\top} W^\top \Sigma^{(\ell)} W P^{(\ell)} \quad (\text{D.37})$$

$$= P^{(\ell)\top} \begin{pmatrix} \left(I + \Lambda_{1:K^{(\ell)}}^{1/2} \right)^{-1} & 0 \\ 0 & I \end{pmatrix} P^{(\ell)} \quad (\text{D.38})$$

Note that $\mathbf{Y}^{(\ell)}$ and $\mathbf{Z}^{(\ell)}$ are again independent. Hence, it holds that

$$\begin{aligned} \dot{\mathbf{Z}}^{(\ell+1)} &= (\mathbf{Z}^{(\ell)} | \mathbf{Y}^{(\ell)} = \mathbf{y}) = \mathbf{Z}^{(\ell)} \\ &= W_2^{(\ell)\top} \mathbf{X}^{(\ell)} = \begin{pmatrix} 0 & I \end{pmatrix} W^{(\ell)\top} \mathbf{X}^{(\ell)} \\ &\sim \mathcal{N} \left(0, P_2^{(\ell)\top} \begin{pmatrix} \left(I + \Lambda_{1:K^{(\ell)}}^{1/2} \right)^{-1} & 0 \\ 0 & I \end{pmatrix} P_2^{(\ell)} \right). \end{aligned} \quad (\text{D.39})$$

Finally, we get that

$$\mathbf{X}^{(\ell+1)} = W^{(\ell)} \begin{pmatrix} \dot{\mathbf{Y}}^{(\ell+1)} \\ \dot{\mathbf{Z}}^{(\ell+1)} \end{pmatrix} \sim \mathcal{N} \left(0, \Sigma^{(\ell+1)} \right), \quad (\text{D.40})$$

where

$$\Sigma^{(\ell+1)} = W^{(\ell)} \begin{pmatrix} P_1^{(\ell)\top} \left(I + \Lambda^{1/2} \right)^{-1} P_1^{(\ell)} & 0 \\ 0 & P_2^{(\ell)\top} \begin{pmatrix} \left(I + \Lambda_{1:K^{(\ell)}}^{1/2} \right)^{-1} & 0 \\ 0 & I \end{pmatrix} P_2^{(\ell)} \end{pmatrix} W^{(\ell)\top} \quad (\text{D.41})$$

$$= W^{(\ell)} P^{(\ell)\top} \begin{pmatrix} \left(I + \Lambda_{1:K^{(\ell+1)}}^{1/2} \right)^{-1} & 0 \\ 0 & I \end{pmatrix} P^{(\ell)} W^{(\ell)\top} \quad (\text{D.42})$$

$$= W \begin{pmatrix} \left(I + \Lambda_{1:K^{(\ell+1)}}^{1/2} \right)^{-1} & 0 \\ 0 & I \end{pmatrix} W^\top \quad (\text{D.43})$$

for a natural number $K^{(\ell+1)}$ with $K^{(\ell)} \leq K^{(\ell+1)} \leq n$. ■

Remark. The step in (D.42) is crucial and important to understand. There are two different scenarios that can occur:

1. **All eigenvalues selected by $P_1^{(\ell)}$ were already selected before.**

In this case, $K^{(\ell+1)} = K^{(\ell)}$ and thus $\Lambda_{1:K^{(\ell+1)}} = \Lambda_{1:K^{(\ell)}}$.

2. **Some eigenvalues selected by $P_1^{(\ell)}$ were *not* selected before.**

Then, $\Delta K^{(\ell,\ell+1)} := K^{(\ell+1)} - K^{(\ell)} > 0$ is the number of newly selected eigenvalues and $\Lambda_{1:K^{(\ell)}}$ is accordingly expanded with them to $\Lambda_{1:K^{(\ell+1)}}$. Note that $\Delta K^{(\ell,\ell+1)}$ depends on both, the permutation matrix $P^{(\ell)}$ and the dimension of the active subspace $k^{(\ell)}$.

Notation

n	Dimension of a problem
\mathbf{N}	Set of natural numbers $\{1, 2, 3, \dots\}$
\mathbf{N}_0	$\{0\} \cup \mathbf{N}$
\mathbf{R}	Set of real numbers
$\mathbf{R}_{\geq 0}$	Set of non-negative real numbers
I/I_n	Identity matrix (of dimension $n \times n$)
$\mathcal{U}(\mathcal{X})$	Uniform distribution on a Borel set \mathcal{X}
$\mathcal{N}(\mathbf{m}, C)$	Normal distribution with mean \mathbf{m} and covariance matrix C
\mathcal{G}	Parameter-to-QoI map
x, \mathbf{x}	Scalar-/vector-valued input parameter/variable
X, \mathbf{X}	Scalar-/vector-valued random variable
x_i	i -th component of vector \mathbf{x}
\mathbf{x}_i	i -th vector in a sequence of vectors
\mathbf{e}_i	i -th unit vector
$(x_j)_{j \in J}$	Sequence with index set J
$x_{k:l}$	Set of elements x_k, x_{k+1}, \dots, x_l , $k \leq l$, from a sequence $(x_j)_{j \in \mathbf{N}}$
$D_{1:k}$	Upper left submatrix of dimension $k \times k$ of a diagonal matrix D
$C(\mathcal{X}, \mathbf{R})$	Space of continuous functions mapping from \mathcal{X} to \mathbf{R}
$C^k(\mathcal{X}, \mathbf{R})$	Space of k times continuously differentiable functions mapping from \mathcal{X} to \mathbf{R}
$L^p(S, \mathcal{F}, \mu)$	Space of L^p functions defined on a measure space (S, \mathcal{F}, μ)
$H^k(\mathcal{X}, \mathbf{R})$	Space of L^2 functions from an open subset $\mathcal{X} \subseteq \mathbf{R}^n$ to \mathbf{R} with L^2 integrable weakly derivatives up to order k
$\mathcal{B}(\mathcal{X})$	Borel σ -algebra on a topological space \mathcal{X}
\mathbf{P}	Probability measure
\mathbf{E}	Expectation value (w.r.t. \mathbf{P})
$\mathbf{P}_{\mathbf{X}}$	Probability distribution of a random variable \mathbf{X}
$\rho_{\mathbf{X}}$	Probability density function of a random variable \mathbf{X}
μ_0	Prior distribution
$\mu^{\mathbf{d}}$	Posterior distribution given data \mathbf{d}
$\mathbf{X} \sim \mu$	Random variable \mathbf{X} is distributed according to μ
$\mathbf{X} \stackrel{\text{approx.}}{\sim} \mu$	Random variable \mathbf{X} is approximately distributed according to μ
$\text{ran}(A)$	Range of an operator A
$\ker(A)$	Kernel/null space of an operator A
$\lambda_\ell(A)$	ℓ -th eigenvalue of a matrix A (for ordered eigenvalues)
$\mathbf{1}_\Omega$	Indicator function of a set Ω

$\text{dom}(f)$	Domain of a function f
$\text{supp}(f)$	Support of a function f , i.e., the closure of the set $\{\mathbf{x} \in \text{dom}(f) \mid f(\mathbf{x}) > 0\}$
\mathcal{X}°	Interior of a set \mathcal{X}
$E_1 := E_2$	Expression E_1 is defined by expression E_2
$E_1 =: E_2$	Expression E_2 is defined by expression E_1
$:\propto / \propto:$	Definition up to a normalizing constant
$A \preceq B$	Matrix $B - A$ is positive semi-definite
$\llbracket \mathbf{y}, \mathbf{z} \rrbracket_W$	$W_1 \mathbf{y} + W_2 \mathbf{z}$ for $W = (W_1 \ W_2)$

References

- [1] Anaconda. *Anaconda Software Distribution*. <https://anaconda.com/>.
- [2] Microsoft. *Python Extension for Visual Studio Code*. <https://marketplace.visualstudio.com/items?itemName=ms-python.python>.
- [3] Microsoft. *Visual Studio Code*. <https://code.visualstudio.com/>.
- [4] TeX Live. <http://tug.org/texlive/>.
- [5] Texas Advanced Computing Center. *Launcher*. <https://github.com/TACC/launcher/>.
- [6] TeXstudio. <https://www.texstudio.org/>.
- [7] WHO. *Ebola Outbreak 2014–2016*. <https://www.who.int/csr/disease/ebola/>.
- [8] WHO. *Situation Reports*. <http://apps.who.int/ebola/ebola-situation-reports/>.
- [9] WHO. *One Year into the Ebola Epidemic*. <http://www.who.int/csr/disease/ebola/one-year-report/ebola-report-1-year.pdf>, 2015.
- [10] R Core Team. *R: A Language and Environment for Statistical Computing*. <https://www.r-project.org/>, 2020.
- [11] D. Alonso-Gutiérrez and J. Bastero. *Approaching the Kannan-Lovász-Simonovits and Variance Conjectures*, volume 2131. Springer, 2015.
- [12] J. E. Andrade, Q. Chen, P. H. Le, C. F. Avila, and T. M. Evans. On the Rheology of Dilative Granular Media: Bridging Solid- and Fluid-like Behavior. *Journal of the Mechanics and Physics of Solids*, 60(6):1122–1136, 2012.
- [13] Ö. Asar, D. Bolin, P. J. Diggle, and J. Wallin. Linear Mixed-Effects Models for Non-Gaussian Repeated Measurement Data. *arXiv preprint arXiv:1804.02592*, 2018.
- [14] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and Geometry of Markov Diffusion Operators*, volume 348. Springer Science & Business Media, 2014.
- [15] G. Bal, I. Langmore, and Y. Marzouk. Bayesian Inverse Problems with Monte Carlo Forward Models. *Inverse Problems & Imaging*, 7(1):81, 2013.
- [16] M. V. Barbarossa, A. Dénes, G. Kiss, Y. Nakata, G. Röst, and Z. Vizi. Transmission Dynamics and Final Epidemic Size of Ebola Virus Disease Outbreaks with Varying Interventions. *PLoS ONE*, 10(7), 2015.
- [17] O. Barndorff-Nielsen. Exponentially Decreasing Distributions for the Logarithm of Particle Size. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 353(1674):401–419, 1977.
- [18] O. E. Barndorff-Nielsen. Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling. *Scandinavian Journal of Statistics*, 24(1):1–13, 1997.
- [19] O. E. Barndorff-Nielsen. Processes of Normal Inverse Gaussian Type. *Finance and Stochastics*, 2(1):41–68, 1997.
- [20] P. Bastian, F. Heimann, and S. Marnach. Generic Implementation of Finite Element Methods in the Distributed and Unified Numerics Environment (DUNE). *Kybernetika*, 46(2):294–315, 2010.

- [21] M. Bebendorf. A Note on the Poincaré Inequality for Convex Domains. *Zeitschrift für Analysis und ihre Anwendungen*, 22(4):751–756, 2003.
- [22] M. Bedard. Optimal Acceptance Rates for Metropolis Algorithms: Moving Beyond 0.234. *Stochastic Processes and their Applications*, 118(12):2198–2222, 2008.
- [23] J. M. Bernardo and A. F. Smith. *Bayesian Theory*, volume 405. John Wiley & Sons, 2009.
- [24] J. Besag. Comments on "Representations of Knowledge in Complex Systems" by U. Grenander and MI Miller. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 56:591–592, 1994.
- [25] A. Beskos, M. Girolami, S. Lan, P. E. Farrell, and A. M. Stuart. Geometric MCMC for Infinite-Dimensional Inverse Problems. *Journal of Computational Physics*, 335:327–351, 2017.
- [26] A. Beskos, A. Jasra, K. Law, Y. Marzouk, and Y. Zhou. Multilevel Sequential Monte Carlo with Dimension-Independent Likelihood-Informed Proposals. *SIAM/ASA Journal on Uncertainty Quantification*, 6(2):762–786, 2018.
- [27] M. Betancourt. The Convergence of Markov Chain Monte Carlo Methods: From the Metropolis Method to Hamiltonian Monte Carlo. *Annalen der Physik*, 531(3):1700214, 2019.
- [28] K. Beven. A Manifesto for the Equifinality Thesis. *Journal of Hydrology*, 320(1-2):18–36, 2006.
- [29] K. J. Beven. Comment on "Equifinality of Formal (DREAM) and Informal (GLUE) Bayesian Approaches in Hydrologic Modeling?" by Jasper A. Vrugt, Cajo J.F. ter Braak, Hoshin V. Gupta and Bruce A. Robinson. *Stochastic Environmental Research and Risk Assessment*, 23(7):1059–1060, 2009.
- [30] P. Billingsley. *Probability and Measure*. John Wiley & Sons, 1995.
- [31] D. Bittner, T. S. Narany, B. Kohl, M. Disse, and G. Chiogna. Modeling the Hydrological Impact of Land Use Change in a Dolomite-Dominated Karst System. *Journal of Hydrology*, 567:267–279, 2018.
- [32] D. Bittner, M. Teixeira Parente, S. Mattis, B. Wohlmuth, and G. Chiogna. Identifying Relevant Hydrological and Catchment Properties in Active Subspaces: An Inference Study of a Lumped Karst Aquifer Model. *Advances in Water Resources*, 135:103472, 2020.
- [33] Å. Björck and G. H. Golub. Numerical Methods for Computing Angles Between Linear Subspaces. *Mathematics of Computation*, 27(123):579–594, 1973.
- [34] S. G. Bobkov. Isoperimetric and Analytic Inequalities for Log-Concave Probability Measures. *The Annals of Probability*, 27(4):1903–1921, 1999.
- [35] D. Bolin. Spatial Matérn Fields Driven by Non-Gaussian Noise. *Scandinavian Journal of Statistics*, 41(3):557–579, 2014.
- [36] D. Bolin and J. Wallin. Multivariate Type G Matérn Stochastic Partial Differential Equation Random Fields. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2019.
- [37] G. E. Box. Robustness in the Strategy of Scientific Model Building. In *Robustness in Statistics*, pages 201–236. Elsevier, 1979.
- [38] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung. *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, 2015.
- [39] S. Brenner and R. Scott. *The Mathematical Theory of Finite Element Methods*, volume 15. Springer Science & Business Media, 2007.
- [40] S. Brooks, A. Gelman, G. Jones, and X.-L. Meng. *Handbook of Markov Chain Monte Carlo*. CRC press, 2011.
- [41] R. H. Buchholz. Perfect Pyramids. *Bulletin of the Australian Mathematical Society*, 45(3):353–368, 1992.
- [42] T. Bui-Thanh, C. Burstedde, O. Ghattas, J. Martin, G. Stadler, and L. C. Wilcox. Extreme-Scale UQ for Bayesian Inverse Problems Governed by PDEs. In *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis*, page 3. IEEE Computer Society Press, 2012.
- [43] T. Bui-Thanh, O. Ghattas, and D. Higdon. Adaptive Hessian-based Nonstationary Gaussian Process Response Surface Method for Probability Density Approximation with Application to Bayesian Solution of Large-Scale Inverse Problems. *SIAM Journal on Scientific Computing*, 34(6):A2837–A2871, 2012.
- [44] T. Bui-Thanh, O. Ghattas, J. Martin, and G. Stadler. A Computational Framework for Infinite-Dimensional Bayesian Inverse Problems Part I: The Linearized Case, with Application to Global Seismic Inversion. *SIAM Journal on Scientific Computing*, 35(6):A2494–A2523, 2013.

- [45] J. Carrera, A. Alcolea, A. Medina, J. Hidalgo, and L. J. Slooten. Inverse Problem in Hydrogeology. *Hydrogeology Journal*, 13(1):206–222, 2005.
- [46] R. J. Carroll, D. Ruppert, L. A. Stefanski, and C. M. Crainiceanu. *Measurement Error in Nonlinear Models: A Modern Perspective*. Chapman and Hall/CRC, 2006.
- [47] L. H. Chen. An Inequality for the Multivariate Normal Distribution. *Journal of Multivariate Analysis*, 12(2):306 – 315, 1982.
- [48] V. Chen, M. M. Dunlop, O. Papaspiliopoulos, and A. M. Stuart. Dimension-Robust MCMC in Bayesian Inverse Problems. *arXiv preprint arXiv:1803.03344*, 2018.
- [49] J.-H. Choi, S. Dai, J.-H. Cha, and Y. Seol. Laboratory Formation of Noncementing Hydrates in Sandy Sediments. *Geochemistry, Geophysics, Geosystems*, 15(4):1648–1656, 2014.
- [50] J. A. Christen and C. Fox. Markov Chain Monte Carlo using an Approximation. *Journal of Computational and Graphical Statistics*, 14(4):795–810, 2005.
- [51] K. A. Cliffe, M. B. Giles, R. Scheichl, and A. L. Teckentrup. Multilevel Monte Carlo Methods and Applications to Elliptic PDEs with Random Coefficients. *Computing and Visualization in Science*, 14(1):3–15, 2011.
- [52] P. G. Constantine. *Active Subspaces*, volume 2 of *SIAM Spotlights*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2015. Emerging Ideas for Dimension Reduction in Parameter Studies.
- [53] P. G. Constantine and P. Diaz. Global Sensitivity Metrics from Active Subspaces. *Reliability Engineering & System Safety*, 162:1–13, 2017.
- [54] P. G. Constantine, E. Dow, and Q. Wang. Active Subspace Methods in Theory and Practice: Applications to Kriging Surfaces. *SIAM Journal on Scientific Computing*, 36(4):A1500–A1524, 2014.
- [55] P. G. Constantine, A. Eftekhari, J. Hokanson, and R. A. Ward. A Near-Stationary Subspace for Ridge Approximation. *Computer Methods in Applied Mechanics and Engineering*, 326:402–421, 2017.
- [56] P. G. Constantine, C. Kent, and T. Bui-Thanh. Accelerating Markov Chain Monte Carlo with Active Subspaces. *SIAM Journal on Scientific Computing*, 38(5):A2779–A2805, 2016.
- [57] R. D. Cook. SAVE: a Method for Dimension Reduction and Graphics in Regression. *Communications in Statistics – Theory and Methods*, 29(9-10):2109–2121, 2000.
- [58] R. D. Cook. *Regression Graphics: Ideas for Studying Regressions through Graphics*, volume 482. John Wiley & Sons, 2009.
- [59] R. D. Cook et al. Fisher Lecture: Dimension Reduction in Regression. *Statistical Science*, 22(1):1–26, 2007.
- [60] R. D. Cook and L. Ni. Sufficient Dimension Reduction via Inverse Regression. *Journal of the American Statistical Association*, 100(470):410–428, 2005.
- [61] R. D. Cook and S. Weisberg. Sliced Inverse Regression for Dimension Reduction: Comment. *Journal of the American Statistical Association*, 86(414):328–332, 1991.
- [62] A. F. Cortesi, P. G. Constantine, T. E. Magin, and P. M. Congedo. Forward and Backward Uncertainty Quantification with Active Subspaces: Application to Hypersonic Flows Around a Cylinder. *Journal of Computational Physics*, 407:109079, 2020.
- [63] S. L. Cotter, M. Dashti, J. C. Robinson, and A. M. Stuart. Bayesian Inverse Problems for Functions and Applications to Fluid Mechanics. *Inverse Problems*, 25(11):115008, 2009.
- [64] S. L. Cotter, G. O. Roberts, A. M. Stuart, and D. White. MCMC Methods for Functions: Modifying Old Algorithms to Make Them Faster. *Statistical Science*, pages 424–446, 2013.
- [65] T. Cui, G. Detommaso, and R. Scheichl. Multilevel Dimension-Independent Likelihood-Informed MCMC for Large-Scale Inverse Problems. *arXiv preprint arXiv:1910.12431*, 2019.
- [66] T. Cui, C. Fox, and M. O’sullivan. Bayesian Calibration of a Large-Scale Geothermal Reservoir Model by a New Adaptive Delayed Acceptance Metropolis Hastings Algorithm. *Water Resources Research*, 47(10), 2011.
- [67] T. Cui, C. Fox, and M. J. O’Sullivan. A Posteriori Stochastic Correction of Reduced Models in Delayed-Acceptance MCMC, with Application to Multiphase Subsurface Inverse Problems. *International Journal for Numerical Methods in Engineering*, 118(10):578–605, 2019.

- [68] T. Cui, K. J. Law, and Y. M. Marzouk. Dimension-Independent Likelihood-Informed MCMC. *Journal of Computational Physics*, 304:109–137, 2016.
- [69] T. Cui, J. Martin, Y. M. Marzouk, A. Solonen, and A. Spantini. Likelihood-Informed Dimension Reduction for Nonlinear Inverse Problems. *Inverse Problems*, 30(11):114015, 28, 2014.
- [70] T. Cui, Y. Marzouk, and K. Willcox. Scalable Posterior Approximations for Large-Scale Bayesian Inverse Problems via Likelihood-Informed Parameter and State Reduction. *Journal of Computational Physics*, 315:363–387, 2016.
- [71] T. Cui, Y. M. Marzouk, and K. E. Willcox. Data-Driven Model Reduction for the Bayesian Solution of Inverse Problems. *International Journal for Numerical Methods in Engineering*, 102(5):966–990, 2015.
- [72] A. Damianou and N. Lawrence. Deep Gaussian Processes. In *Artificial Intelligence and Statistics*, pages 207–215, 2013.
- [73] M. Dashti and A. M. Stuart. The Bayesian Approach to Inverse Problems. *Handbook of Uncertainty Quantification*, pages 1–118, 2016.
- [74] R. Dawe and S. Thomas. A Large Potential Methane Source–Natural Gas Hydrates. *Energy Sources, Part A*, 29(3):217–229, 2007.
- [75] E. A. de Souza Neto, D. Peric, and D. R. Owen. *Computational Methods for Plasticity: Theory and Applications*. John Wiley & Sons, 2011.
- [76] A. Dedner, B. Flemisch, and R. Klöforn. *Advances in DUNE*. Springer, 2012.
- [77] C. Deusner, N. Bigalke, E. Kossel, and M. Haeckel. Methane Production from Gas Hydrate Deposits through Injection of Supercritical CO₂. *Energies*, 5(7):2112–2140, 2012.
- [78] P. Diaz, P. Constantine, K. Kalmbach, E. Jones, and S. Pankavich. A Modified SEIR Model for the Spread of Ebola in Western Africa and Metrics for Resource Allocation. *Applied Mathematics and Computation*, 324:141–155, 2018.
- [79] J. L. Doob. *Stochastic Processes*, volume 101. New York Wiley, 1953.
- [80] J. M. Drake, R. B. Kaul, L. W. Alexander, S. M. O’Regan, A. M. Kramer, J. T. Pulliam, M. J. Ferrari, and A. W. Park. Ebola Cases and Health System Demand in Liberia. *PLoS Biology*, 13(1), 2015.
- [81] N. R. Draper and H. Smith. *Applied Regression Analysis*, volume 326. John Wiley & Sons, 1998.
- [82] M. M. Dunlop, M. A. Girolami, A. M. Stuart, and A. L. Teckentrup. How Deep Are Deep Gaussian Processes? *The Journal of Machine Learning Research*, 19(1):2100–2145, 2018.
- [83] DVWK. Ermittlung Der Verdunstung Von Land- Und Wasserflächen. *DVWK-Merkblatt*, 238/1996, 1996.
- [84] D. J. Earl and M. W. Deem. Parallel Tempering: Theory, Applications, and New Perspectives. *Physical Chemistry Chemical Physics*, 7(23):3910–3916, 2005.
- [85] E. Eberlein. Application of Generalized Hyperbolic Lévy Motions to Finance. In *Lévy Processes*, pages 319–336. Springer, 2001.
- [86] B. Efron and R. J. Tibshirani. *An Introduction to the Bootstrap*. CRC press, 1994.
- [87] K. Engeland and L. Gottschalk. Bayesian Estimation of Parameters in a Regional Hydrological Model. *Hydrology and Earth System Sciences Discussions*, 6(5):883–898, 2002.
- [88] G. Evensen. Sequential Data Assimilation with a Nonlinear Quasi-Geostrophic Model using Monte Carlo Methods to Forecast Error Statistics. *Journal of Geophysical Research: Oceans*, 99(C5):10143–10162, 1994.
- [89] H. P. Flath, L. C. Wilcox, V. Akçelik, J. Hill, B. van Bloemen Waanders, and O. Ghattas. Fast Algorithms for Bayesian Uncertainty Quantification in Large-Scale Linear Inverse Problems based on Low-Rank Partial Hessian Approximations. *SIAM Journal on Scientific Computing*, 33(1):407–432, 2011.
- [90] P. Fleury, B. Ladouche, Y. Conroux, H. Jourde, and N. Dörfli. Modelling the Hydrologic Functions of a Karst Aquifer under Active Water Management—The Lez Spring. *Journal of Hydrology*, 365(3-4):235–243, 2009.
- [91] D. Ford and P. D. Williams. *Karst Hydrogeology and Geomorphology*. John Wiley & Sons, 2013.
- [92] M. Fornasier, K. Schnass, and J. Vybiral. Learning Functions of Few Arbitrary Linear Parameters in High Dimensions. *Foundations of Computational Mathematics*, 12(2):229–262, 2012.

- [93] C. R. Fox and G. Ülkümen. Distinguishing Two Dimensions of Uncertainty. *Perspectives on Thinking, Judging, and Decision Making*, pages 21–35, 2011.
- [94] X. Gai and M. Sánchez. A Geomechanical Model for Gas Hydrate-Bearing Sediments. *Environmental Geotechnics*, 4(2):143–156, 2017.
- [95] GBA. Geologische Bundesländerkarten, Geologische Bundesanstalt Österreich. https://gisgba.geologie.ac.at/ArcGIS/rest/services/image/AT_GBA_GK100_200/ImageServer/. April 18th, 2018.
- [96] A. Gelman, G. O. Roberts, and W. R. Gilks. Efficient Metropolis Jumping Rules. *Bayesian Statistics*, 5(599-608):42, 1996.
- [97] M. Giese, T. Reimann, V. Bailly-Comte, J.-C. Maréchal, M. Sauter, and T. Geyer. Turbulent and Laminar Flow in Karst Conduits under Unsteady Flow Conditions: Interpretation of Pumping Tests by Discrete Conduit-Continuum Modeling. *Water Resources Research*, 54(3):1918–1933, 2018.
- [98] M. B. Giles. Multilevel Monte Carlo Methods. *Acta Numerica*, 24:259–328, 2015.
- [99] W. R. Gilks, S. Richardson, and D. Spiegelhalter. *Markov Chain Monte Carlo in Practice*. Chapman and Hall/CRC, 1995.
- [100] M. Girolami and B. Calderhead. Riemann Manifold Langevin and Hamiltonian Monte Carlo Methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(2):123–214, 2011.
- [101] A. Gittens and J. A. Tropp. Tail Bounds for All Eigenvalues of a Sum of Random Matrices. *arXiv preprint arXiv:1104.4513*, 2011.
- [102] S. A. Glantz, B. K. Slinker, and T. B. Neillands. *Primer of Applied Regression and Analysis of Variance*, volume 309. McGraw-Hill New York, 1990.
- [103] A. Glaws, P. G. Constantine, and R. D. Cook. Inverse Regression for Ridge Recovery: a Data-Driven Approach for Parameter Reduction in Computer Experiments. *Statistics and Computing*, 2019.
- [104] J. Glimm, S. Hou, Y. Lee, D. Sharp, and K. Ye. Solution Error Models for Uncertainty Quantification. *Contemporary Mathematics*, 327:115–140, 2003.
- [105] A. Golightly, D. A. Henderson, and C. Sherlock. Delayed Acceptance Particle MCMC for Exact Inference in Stochastic Kinetic Models. *Statistics and Computing*, 25(5):1039–1055, 2015.
- [106] G. H. Golub and C. F. Van Loan. *Matrix Computations (3rd Ed.)*. Johns Hopkins University Press, Baltimore, MD, USA, 1996.
- [107] P. J. Green and A. Mira. Delayed Rejection in Reversible Jump Metropolis–Hastings. *Biometrika*, 88(4):1035–1053, 2001.
- [108] L. Gross. Logarithmic Sobolev Inequalities. *American Journal of Mathematics*, 97(4):1061–1083, 1975.
- [109] S. Gupta, C. Deusner, M. Haeckel, R. Helmig, and B. Wohlmuth. Testing a Thermo-Chemo-Hydro-Geomechanical Model for Gas Hydrate-Bearing Sediments using Triaxial Compression Laboratory Experiments. *Geochemistry, Geophysics, Geosystems*, 18(9):3419–3437, 2017.
- [110] S. Gupta, R. Helmig, and B. Wohlmuth. Non-Isothermal, Multi-Phase, Multi-Component Flows through Deformable Methane Hydrate Reservoirs. *Computational Geosciences*, 19(5):1063–1088, 2015.
- [111] H. Haario, M. Laine, A. Mira, and E. Saksman. DRAM: Efficient Adaptive MCMC. *Statistics and Computing*, 16(4):339–354, 2006.
- [112] H. Haario, E. Saksman, and J. Tamminen. Adaptive Proposal Distribution for Random Walk Metropolis Algorithm. *Computational Statistics*, 14(3):375–395, 1999.
- [113] H. Haario, E. Saksman, J. Tamminen, et al. An Adaptive Metropolis Algorithm. *Bernoulli*, 7(2):223–242, 2001.
- [114] P. Hacker. Hydrologisch-hydrogeologische Untersuchungen im Bereich des Glashüttenberges zur Frage des engeren Schutzgebietes für die Kerschbaumer-Quelle. *ARC Seibersdorf Research GmbH*, 2003.
- [115] J. Hadamard. Sur Les Problèmes Aux Dérivées Partielles Et Leur Signification Physique. *Princeton University Bulletin*, pages 49–52, 1902.
- [116] C. Hager and B. Wohlmuth. Nonlinear Complementarity Functions for Plasticity Problems with Frictional Contact. *Computer Methods in Applied Mechanics and Engineering*, 198(41-44):3411–3427, 2009.
- [117] C. Hager and B. I. Wohlmuth. Semismooth Newton Methods for Variational Problems with Inequality Constraints. *GAMM-Mitteilungen*, 33(1):8–24, 2010.

- [118] M. Hairer, A. M. Stuart, S. J. Vollmer, et al. Spectral Gaps for a Metropolis–Hastings Algorithm in Infinite Dimensions. *The Annals of Applied Probability*, 24(6):2455–2490, 2014.
- [119] A. Hartmann, J. A. Barberá, and B. Andreo. On the Value of Water Quality Observations for Karst Model Parameterization. *Hydrology and Earth System Sciences*, 21:5971–5985, 2017.
- [120] A. Hartmann, N. Goldscheider, T. Wagener, J. Lange, and M. Weiler. Karst Water Resources in a Changing World: Review of Hydrological Modeling Approaches. *Reviews of Geophysics*, 52(3):218–242, 2014.
- [121] A. Hartmann, T. Wagener, A. Rimmer, J. Lange, H. Brielmann, and M. Weiler. Testing the Realism of Model Structures to Identify Karst System Processes using Water Quality and Quantity Signatures. *Water Resources Research*, 49(6):3345–3358, 2013.
- [122] A. Hartmann, M. Weiler, T. Wagener, J. Lange, M. Kralik, F. Humer, N. Mizyed, A. Rimmer, J. Barberá, B. Andreo, et al. Process-based Karst Modelling to Relate Hydrodynamic and Hydrochemical Characteristics to System Properties. *Hydrology and Earth System Sciences*, 17(8):3305–3321, 2013.
- [123] W. K. Hastings. Monte Carlo Sampling Methods using Markov Chains and Their Applications. *Biometrika*, 57(1):97–109, 1970.
- [124] P. Hennig, M. A. Osborne, and M. Girolami. Probabilistic Numerics and Uncertainty in Computations. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 471(2179):20150142, 2015.
- [125] W. R. Henson, R. de Rooij, and W. Graham. What Makes a First-Magnitude Spring?: Global Sensitivity Analysis of a Speleogenesis Model to Gain Insight into Karst Network and Spring Genesis. *Water Resources Research*, 54(10):7417–7434, 2018.
- [126] T. S. Hogue, L. A. Bastidas, H. V. Gupta, and S. Sorooshian. Evaluating Model Performance and Parameter Behavior for Varying Levels of Land Surface Model Complexity. *Water Resources Research*, 42(8), 2006.
- [127] J. T. Holodnak, I. C. Ipsen, and R. C. Smith. A Probabilistic Subspace Bound with Application to Active Subspaces. *SIAM Journal on Matrix Analysis and Applications*, 39(3):1208–1220, 2018.
- [128] T. Homma and A. Saltelli. Importance Measures in Global Sensitivity Analysis of Nonlinear Models. *Reliability Engineering & System Safety*, 52(1):1–17, 1996.
- [129] P. L. Houtekamer and H. L. Mitchell. Data Assimilation using an Ensemble Kalman Filter Technique. *Monthly Weather Review*, 126(3):796–811, 1998.
- [130] J. Huang and D. Griffiths. Return Mapping Algorithms and Stress Predictors for Failure Analysis in Geomechanics. *Journal of Engineering Mechanics*, 135(4):276–284, 2009.
- [131] J. D. Hunter. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, 9(3):90–95, 2007.
- [132] J. M. Huttunen and J. P. Kaipio. Approximation Errors in Nonstationary Inverse Problems. *Inverse Problems & Imaging*, 1(1):77–93, 2007.
- [133] M. Hyodo, Y. Li, J. Yoneda, Y. Nakata, N. Yoshimoto, and A. Nishimura. Effects of Dissociation on the Shear Strength and Deformation Behavior of Methane Hydrate-Bearing Sediments. *Marine and Petroleum Geology*, 51:52–62, 2014.
- [134] M. Hyodo, Y. Nakata, N. Yoshimoto, and T. Ebinuma. Basic Research on the Mechanical Behavior of Methane Hydrate-Sediments Mixture. *Soils and Foundations*, 45(1):75–85, 2005.
- [135] A. Jakeman and G. Hornberger. How much Complexity is Warranted in a Rainfall-Runoff Model? *Water Resources Research*, 29(8):2637–2649, 1993.
- [136] A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Courier Corporation, 2007.
- [137] J. L. Jefferson, J. M. Gilbert, P. G. Constantine, and R. M. Maxwell. Active Subspaces for Sensitivity Analysis and Dimension Reduction of an Integrated Hydrologic Model. *Computers & Geosciences*, 83:127–138, 2015.
- [138] M. Jirásek and Z. P. Bazant. *Inelastic Analysis of Structures*. John Wiley & Sons, 2001.
- [139] I. T. Jolliffe. *Principal Component Analysis*, volume 2. Springer New York, NY, 2002.

- [140] B. Jorgensen. *Statistical Properties of the Generalized Inverse Gaussian Distribution*, volume 9. Springer Science & Business Media, 2012.
- [141] H. Jourde, N. Mazzilli, N. Lecoq, B. Arfib, and D. Bertin. KARSTMOD: A Generic Modular Reservoir Model dedicated to Spring Discharge Modeling and Hydrodynamic Analysis in Karst. In *Hydrogeological and Environmental Investigations in Karst Systems*, pages 339–344. Springer, 2015.
- [142] D. Jukić and V. Denić-Jukić. Estimating Parameters of Groundwater Recharge Model in Frequency Domain: Karst Springs Jadro and Žrnovnica. *Hydrological Processes: An International Journal*, 22(23):4532–4542, 2008.
- [143] D. Jukić and V. Denić-Jukić. Groundwater Balance Estimation in Karst by using a Conceptual Rainfall–Runoff Model. *Journal of Hydrology*, 373(3–4):302–315, 2009.
- [144] S. J. Julier and J. K. Uhlmann. Unscented Filtering and Nonlinear Estimation. *Proceedings of the IEEE*, 92(3):401–422, 2004.
- [145] H. Kahn and A. W. Marshall. Methods of Reducing Sample Size in Monte Carlo Computations. *Journal of the Operations Research Society of America*, 1(5):263–278, 1953.
- [146] J. Kaipio and E. Somersalo. *Statistical and Computational Inverse Problems*, volume 160. Springer Science & Business Media, 2006.
- [147] O. Kallenberg. *Random Measures, Theory and Applications*. Springer, 2017.
- [148] R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 82(1):35–45, 1960.
- [149] D. Kalpić and N. Hlupić. *Multivariate Normal Distributions*, pages 907–910. International Encyclopedia of Statistical Science, Springer Berlin Heidelberg, 2011.
- [150] D. Kavetski, G. Kuczera, and S. W. Franks. Bayesian Analysis of Input Uncertainty in Hydrological Modeling: 1. Theory. *Water Resources Research*, 42(3), 2006.
- [151] M. C. Kennedy and A. O’Hagan. Bayesian Calibration of Computer Models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464, 2001.
- [152] S. Kimoto, F. Oka, and T. Fushita. A Chemo-Thermo-Mechanically Coupled Analysis of Ground Deformation Induced by Gas Hydrate Dissociation. *International Journal of Mechanical Sciences*, 52(2):365–376, 2010.
- [153] L. C. Kinsey and T. E. Moore. *Symmetry, Shape and Space: An Introduction to Mathematics through Geometry*. Springer Science & Business Media, 2006.
- [154] A. Klar, K. Soga, and M. Ng. Coupled Deformation–Flow Analysis for Methane Hydrate Extraction. *Geotechnique*, 60(10):765–776, 2010.
- [155] A. Klar, S. Uchida, K. Soga, K. Yamamoto, et al. Explicitly Coupled Thermal Flow Mechanical Formulation for Gas-Hydrate Sediments. *SPE Journal*, 18(02):196–206, 2013.
- [156] A. Klenke. *Probability Theory: A Comprehensive Course*. Springer Science & Business Media, 2013.
- [157] S. Kotz, T. Kozubowski, and K. Podgórski. *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*. Springer Science & Business Media, 2012.
- [158] H. W. Kuhn and A. W. Tucker. Nonlinear programming. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, Calif., 1951. University of California Press.
- [159] S. Kullback. *Information Theory and Statistics*. Courier Corporation, 1997.
- [160] S. Kullback and R. A. Leibler. On Information and Sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951.
- [161] D. Labat, R. Ababou, and A. Mangin. Linear and Nonlinear Input/Output Models for Karstic Springflow and Flood Prediction at Different Time Scales. *Stochastic Environmental Research and Risk Assessment*, 13(5):337–364, 1999.
- [162] D. Labat, R. Ababou, and A. Mangin. Rainfall–Runoff Relations for Karstic Springs. Part i: Convolution and Spectral Analyses. *Journal of Hydrology*, 238(3–4):123–148, 2000.
- [163] E. Laloy, B. Rogiers, J. A. Vrugt, D. Mallants, and D. Jacques. Efficient Posterior Exploration of a High-Dimensional Groundwater Model from Two-Stage Markov Chain Monte Carlo Simulation and Polynomial Chaos Expansion. *Water Resources Research*, 49(5):2664–2682, 2013.

- [164] J. Latz. On the Well-Posedness of Bayesian Inverse Problems. *SIAM/ASA Journal on Uncertainty Quantification*, 8(1):451–482, 2020.
- [165] K. Law, A. Stuart, and K. Zygalakis. *Data Assimilation*. Springer, 2015.
- [166] K. J. Law. Proposals which Speed Up Function-Space MCMC. *Journal of Computational and Applied Mathematics*, 262:127–138, 2014.
- [167] O. Le Maître and O. M. Knio. *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics*. Springer Science & Business Media, 2010.
- [168] D. Leao Jr, M. Fragoso, and P. Ruffino. Regular Conditional Probability, Disintegration of Probability and Radon Spaces. *Proyecciones (Antofagasta)*, 23(1):15–29, 2004.
- [169] J. Lee, F. M. Francisca, J. C. Santamarina, and C. Ruppel. Parametric Study of the Physical Properties of Hydrate-Bearing Sand, Silt, and Clay Sediments: 2. Small-Strain Mechanical Properties. *Journal of Geophysical Research: Solid Earth*, 115(B11), 2010.
- [170] J. Lee, T. S. Yun, J. Santamarina, and C. Ruppel. Observations related to Tetrahydrofuran and Methane Hydrates for Laboratory Studies of Hydrate-Bearing Sediments. *Geochemistry, Geophysics, Geosystems*, 8(6), 2007.
- [171] J. A. Lee and M. Verleysen. *Nonlinear Dimensionality Reduction*. Springer Science & Business Media, 2007.
- [172] S. H. Lee and W. Chen. A Comparative Study of Uncertainty Propagation Methods for Black-Box-Type Problems. *Structural and Multidisciplinary Optimization*, 37(3):239, 2009.
- [173] Y. T. Lee and S. S. Vempala. The Kannan-Lovász-Simonovits Conjecture. *Current Developments in Mathematics*, 2017(1):1–36, 2017.
- [174] J. Legrand, R. F. Grais, P.-Y. Boelle, A.-J. Valleron, and A. Flahault. Understanding the Dynamics of Ebola Epidemics. *Epidemiology & Infection*, 135(4):610–621, 2007.
- [175] R. LeVeque. *Finite Difference Methods for Ordinary and Partial Differential Equations*. Society for Industrial and Applied Mathematics, 2007.
- [176] B. Li, T. Bengtsson, and P. Bickel. Curse-of-Dimensionality Revisited: Collapse of Importance Sampling in Very Large Scale Systems. Technical report, University of California, Berkeley, 2005.
- [177] B. Li and S. Wang. On Directional Regression for Dimension Reduction. *Journal of the American Statistical Association*, 102(479):997–1008, 2007.
- [178] B. Li, H. Zha, F. Chiaromonte, et al. Contour Regression: A General Approach to Dimension Reduction. *The Annals of Statistics*, 33(4):1580–1616, 2005.
- [179] K.-C. Li. Sliced Inverse Regression for Dimension Reduction. *Journal of the American Statistical Association*, 86(414):316–327, 1991.
- [180] K.-C. Li. On Principal Hessian Directions for Data Visualization and Dimension Reduction: Another Application of Stein’s Lemma. *Journal of the American Statistical Association*, 87(420):1025–1039, 1992.
- [181] C. Lieberman, K. Willcox, and O. Ghattas. Parameter and State Model Reduction for Large-Scale Statistical Inverse Problems. *SIAM Journal on Scientific Computing*, 32(5):2523–2542, 2010.
- [182] T. I. Lin and J. C. Lee. Bayesian Analysis of Hierarchical Linear Mixed Modeling using the Multivariate t Distribution. *Journal of Statistical Planning and Inference*, 137(2):484–495, 2007.
- [183] W. A. Link and M. J. Eaton. On Thinning of Chains in MCMC. *Methods in Ecology and Evolution*, 3(1):112–115, 2012.
- [184] J. S. Liu. *Monte Carlo Strategies in Scientific Computing*. Springer Science & Business Media, 2008.
- [185] Y. Ma and L. Zhu. A Review on Dimension Reduction. *International Statistical Review*, 81(1):134–150, 2013.
- [186] J. Martin, L. C. Wilcox, C. Burstedde, and O. Ghattas. A Stochastic Newton MCMC Method for Large-Scale Statistical Inverse Problems with Application to Seismic Inversion. *SIAM Journal on Scientific Computing*, 34(3):A1460–A1487, 2012.
- [187] J. Martinec. The Degree-Day Factor for Snowmelt Runoff Forecasting. *IUGG General Assembly of Helsinki, IAHS Commission of Surface Waters*, 51:468–477, 1960.

- [188] Y. Marzouk, T. Moselhy, M. Parno, and A. Spantini. *Sampling via Measure Transport: An Introduction*. Handbook of Uncertainty Quantification, Springer International Publishing, 2016.
- [189] A. Masui, H. Haneda, Y. Ogata, and K. Aoki. Effects of Methane Hydrate Formation on Shear Strength of Synthetic Methane Hydrate Sediments. *The Fifteenth International Offshore and Polar Engineering Conference*, 8:364–369, 2005.
- [190] N. Mazzilli, V. Guinot, H. Jourde, N. Lecoq, D. Labat, B. Arfib, C. Baudement, C. Danquigny, L. Dal Soglio, and D. Bertin. KarstMod: A Modelling Platform for Rainfall-Discharge Analysis and Modelling dedicated to Karst Systems. *Environmental Modelling & Software*, 2017.
- [191] W. McKinney. Data Structures for Statistical Computing in Python. In *Proceedings of the 9th Python in Science Conference*, pages 56–61, 2010.
- [192] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.
- [193] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Springer Science & Business Media, 1993.
- [194] A. Mira. Ordering and Improving the Performance of Monte Carlo Markov Chains. *Statistical Science*, pages 340–350, 2001.
- [195] K. Miyazaki, A. Masui, Y. Sakamoto, K. Aoki, N. Tenma, and T. Yamaguchi. Triaxial Compressive Properties of Artificial Methane-Hydrate-Bearing Sediment. *Journal of Geophysical Research: Solid Earth*, 116(B6), 2011.
- [196] K. Miyazaki, A. Masui, N. Tenma, Y. Ogata, K. Aoki, T. Yamaguchi, Y. Sakamoto, et al. Study on Mechanical Behavior for Methane Hydrate Sediment based on Constant Strain-Rate Test and Unloading-Reloading Test under Triaxial Compression. *International Journal of Offshore and Polar Engineering*, 20(01), 2010.
- [197] D. C. Montgomery, E. A. Peck, and G. G. Vining. *Introduction to Linear Regression Analysis*, volume 821. John Wiley & Sons, 2012.
- [198] A. Morawiec. *Orientations and Rotations*. Springer, 2003.
- [199] G. Moridis, T. Collett, M. Pooladi-Darvish, S. Hancock, C. Santamarina, R. Boswel, T. Kneafsey, J. Rutqvist, M. Kowalsky, M. Reagan, et al. Challenges, Uncertainties, and Issues Facing Gas Production from Gas-Hydrate Deposits. *SPE Reservoir Evaluation and Engineering*, 14(1):76–112, 2011.
- [200] G. J. Moridis, T. S. Collett, R. Boswell, M. Kurihara, M. T. Reagan, C. Koh, E. D. Sloan, et al. Toward Production from Gas Hydrates: Current Status, Assessment of Resources, and Simulation-based Evaluation of Technology and Potential. *SPE Reservoir Evaluation & Engineering*, 12(05), 2009.
- [201] M. Morzfeld, X. T. Tong, and Y. M. Marzouk. Localization for MCMC: Sampling High-Dimensional Posterior Distributions with Local Structure. *Journal of Computational Physics*, 380:1–28, 2019.
- [202] I. J. Myung. Tutorial on Maximum Likelihood Estimation. *Journal of Mathematical Psychology*, 47(1):90–100, 2003.
- [203] C. J. Oates and T. J. Sullivan. A Modern Retrospective on Probabilistic Numerics. *Statistics and Computing*, 29(6):1335–1351, 2019.
- [204] J. T. Oden. Adaptive Multiscale Predictive Modelling. *Acta Numerica*, 27:353–450, 2018.
- [205] T. E. Oliphant. *A Guide to NumPy*, volume 1. Trelgol Publishing USA, 2006.
- [206] D. Orrell, L. Smith, J. Barkmeijer, and T. Palmer. Model Error in Weather Forecasting. *Nonlinear Processes in Geophysics*, 8:357–371, 2001.
- [207] A. B. Owen. *Monte Carlo Theory, Methods and Examples*. 2013.
- [208] A. O’Sullivan and M. Christie. Simulation Error Models for Improved Reservoir Prediction. *Reliability Engineering & System Safety*, 91(10-11):1382–1389, 2006.
- [209] M. D. Parno and Y. M. Marzouk. Transport Map Accelerated Markov Chain Monte Carlo. *SIAM/ASA Journal on Uncertainty Quantification*, 6(2):645–682, 2018.
- [210] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

- [211] B. Peherstorfer and Y. Marzouk. A Transport-based Multifidelity Preconditioner for Markov Chain Monte Carlo. *Advances in Computational Mathematics*, 45(5-6):2321–2348, 2019.
- [212] E. Pinero, M. Marquardt, C. Hensen, M. Haeckel, and K. Wallmann. Estimation of the Global Inventory of Methane Hydrates in Marine Sediments using transfer Functions. *Biogeosciences (BG)*, 10(2):959–975, 2013.
- [213] S. Pinkert. The Lack of True Cohesion in Hydrate-Bearing Sands. *Granular Matter*, 19(3):57, 2017.
- [214] S. Pinkert and J. Grozic. Prediction of the Mechanical Response of Hydrate-Bearing Sands. *Journal of Geophysical Research: Solid Earth*, 119(6):4695–4707, 2014.
- [215] S. Pinkert, J. Grozic, and J. Priest. Strain-Softening Model for Hydrate-Bearing Sands. *International Journal of Geomechanics*, 15(6):04015007, 2015.
- [216] A. Pinkus. *Ridge Functions*, volume 205. Cambridge University Press, 2015.
- [217] R.-E. Plessix. A Review of the Adjoint-State Method for Computing the Gradient of a Functional with Geophysical Applications. *Geophysical Journal International*, 167(2):495–503, 2006.
- [218] J. A. Priest, E. V. Rees, and C. R. Clayton. Influence of Gas Hydrate Morphology on the Seismic Velocities of Sands. *Journal of Geophysical Research: Solid Earth*, 114(B11), 2009.
- [219] M. Quiroz, M.-N. Tran, M. Villani, and R. Kohn. Speeding up MCMC by Aelayed Acceptance and Data Subsampling. *Journal of Computational and Graphical Statistics*, 27(1):12–22, 2018.
- [220] T. Reimann and M. E. Hill. MODFLOW-CFP: A New Conduit Flow Process for MODFLOW–2005. *Groundwater*, 47(3):321–325, 2009.
- [221] D. Revuz. *Markov Chains*, volume 11. Elsevier, 2008.
- [222] C. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer Science & Business Media, 2013.
- [223] G. O. Roberts, A. Gelman, W. R. Gilks, et al. Weak Convergence and Optimal Scaling of Random Walk Metropolis Algorithms. *The Annals of Applied Probability*, 7(1):110–120, 1997.
- [224] G. O. Roberts and J. S. Rosenthal. Optimal Scaling of Discrete Approximations to Langevin Diffusions. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 60(1):255–268, 1998.
- [225] G. O. Roberts and J. S. Rosenthal. Examples of Adaptive MCMC. *Journal of Computational and Graphical Statistics*, 18(2):349–367, 2009.
- [226] G. O. Roberts, R. L. Tweedie, et al. Exponential Convergence of Langevin Distributions and Their Discrete Approximations. *Bernoulli*, 2(4):341–363, 1996.
- [227] E. Rosero, Z.-L. Yang, T. Wagener, L. E. Gulden, S. Yatheendradas, and G.-Y. Niu. Quantifying Parameter Sensitivity, Interaction, and Transferability in Hydrologically Enhanced Versions of the Noah Land Surface Model over Transition Zones during the Warm Season. *Journal of Geophysical Research: Atmospheres*, 115(D3), 2010.
- [228] R. J. Rossi. *Mathematical Statistics: An Introduction to Likelihood based Inference*. John Wiley & Sons, 2018.
- [229] G. Rozza. Fundamentals of Reduced Basis Method for Problems Governed by Parametrized PDEs and Applications. In *Separated Representations and PGD-based Model Reduction*, pages 153–227. Springer, 2014.
- [230] T. M. Russi. *Uncertainty Quantification with Experimental Data and Complex System Models*. PhD thesis, University of California, Berkeley, 2010.
- [231] J. Rutqvist. Status of the TOUGH-FLAC Simulator and Recent Applications Related to Coupled Fluid Flow and Crustal Deformations. *Computers & Geosciences*, 37(6):739–750, 2011.
- [232] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. *Global Sensitivity Analysis: The Primer*. John Wiley & Sons, 2008.
- [233] J. C. Santamarina and C. Ruppel. The Impact of Hydrate Saturation on the Mechanical, Electrical, and Thermal Properties of Hydrate-Bearing Sand, Silts, and Clay. *Geophysical Characterization of Gas Hydrates*, 14:373–384, 2010.
- [234] S. Särkkä. *Bayesian Filtering and Smoothing*, volume 3. Cambridge University Press, 2013.
- [235] M. Sauter, T. Geyer, A. Kovács, and G. Teutsch. Modellierung der Hydraulik von Karstgrundwasserleitern—Eine Übersicht. *Grundwasser*, 11(3):143–156, 2006.

- [236] S. M. Schennach. Recent Advances in the Measurement Error Literature. *Annual Review of Economics*, 8:341–377, 2016.
- [237] G. Schwarz. Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2):461–464, 1978.
- [238] P. Seshadri, S. Yuchi, and G. T. Parks. Dimension Reduction via Gaussian Ridge Functions. *SIAM/ASA Journal on Uncertainty Quantification*, 7(4):1301–1322, 2019.
- [239] C. Sherlock, A. Golightly, and D. A. Henderson. Adaptive, Delayed-Acceptance MCMC for Targets with Expensive Likelihoods. *Journal of Computational and Graphical Statistics*, 26(2):434–444, 2017.
- [240] J. C. Simo and T. J. Hughes. *Computational Inelasticity*, volume 7. Springer Science & Business Media, 2006.
- [241] V. Sivellev, D. Labat, N. Mazzilli, N. Massei, and H. Jourde. Dynamics of the Flow Exchanges between Matrix and Conduits in Karstified Watersheds at Multiple Temporal Scales. *Water*, 11(3):569, 2019.
- [242] E. D. Sloan. Gas Hydrates: Review of Physical/Chemical Properties. *Energy & Fuels*, 12(2):191–196, 1998.
- [243] R. C. Smith. *Uncertainty Quantification: Theory, Implementation, and Applications*, volume 12. SIAM series on Computational Science & Engineering, 2013.
- [244] R. L. Smith. Some Interlacing Properties of the Schur Complement of a Hermitian Matrix. *Linear Algebra and its Applications*, 177:137–144, 1992.
- [245] I. M. Sobol. Sensitivity Estimates for Nonlinear Mathematical Models. *Mathematical Modelling and Computational Experiments*, 1(4):407–414, 1993.
- [246] I. M. Sobol. Global Sensitivity Indices for Nonlinear Mathematical Models and Their Monte Carlo Estimates. *Mathematics and Computers in Simulation*, 55(1-3):271–280, 2001.
- [247] A. Spantini, A. Solonen, T. Cui, J. Martin, L. Tenorio, and Y. Marzouk. Optimal Low-Rank Approximations of Bayesian Linear Inverse Problems. *SIAM Journal on Scientific Computing*, 37(6):A2451–A2487, 2015.
- [248] B. Sprungk. On the Local Lipschitz Stability of Bayesian Inverse Problems. *Inverse Problems*, 2020.
- [249] M. L. Stein. *Interpolation of Spatial Data: Some Theory for Kriging*. Springer Science & Business Media, 2012.
- [250] R. Storn. On the Usage of Differential Evolution for Function Optimization. In *Proceedings of North American Fuzzy Information Processing*, pages 519–523. IEEE, 1996.
- [251] R. Storn and K. Price. Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. *Journal of Global Optimization*, 11(4):341–359, 1997.
- [252] A. M. Stuart. Inverse Problems: A Bayesian Perspective. *Acta Numerica*, 19:451–559, 2010.
- [253] T. J. Sullivan. *Introduction to Uncertainty Quantification*, volume 63 of *Texts in Applied Mathematics*. Springer, 2015.
- [254] N. Sultan, P. Cochonat, M. Canals, A. Cattaneo, B. Dennielou, H. Haffidason, J. Laberg, D. Long, J. Mienert, F. Trincardi, et al. Triggering Mechanisms of Slope Instability Processes and Sediment Failures on Continental Margins: A Geotechnical Approach. *Marine Geology*, 213(1-4):291–321, 2004.
- [255] N. Sultan, P. Cochonat, J.-P. Foucher, and J. Mienert. Effect of Gas Hydrates Melting on Seafloor Slope Instability. *Marine Geology*, 213(1-4):379–401, 2004.
- [256] M. Teixeira Parente, S. Mattis, S. Gupta, C. Deusner, and B. Wohlmuth. Efficient Parameter Estimation for a Methane Hydrate Model with Active Subspaces. *Computational Geosciences*, 23(2):355–372, 2019.
- [257] M. Teixeira Parente, J. Wallin, and B. Wohlmuth. Generalized Bounds for Active Subspaces. *Electronic Journal of Statistics*, 14(1):917–943, 2020.
- [258] M. Teixeira Parente, D. Bittner, S. A. Mattis, G. Chiogna, and B. Wohlmuth. Bayesian Calibration and Sensitivity Analysis for a Karst Aquifer Model using Active Subspaces. *Water Resources Research*, 55(8):7086–7107, 2019.
- [259] M. Thiemann, M. Trosset, H. Gupta, and S. Sorooshian. Bayesian Recursive Parameter Estimation for Hydrologic Models. *Water Resources Research*, 37(10):2521–2535, 2001.
- [260] C. W. Thornthwaite. An Approach toward a Rational Classification of Climate. *Geographical Review*, 38(1):55–94, 1948.

- [261] L. Tierney and A. Mira. Some Adaptive Monte Carlo Methods for Bayesian Inference. *Statistics in Medicine*, 18(17-18):2507–2515, 1999.
- [262] A. N. Tikhonov. Solution of Incorrectly Formulated Problems and the Regularization Method. *Soviet Math.*, 4:1035–1038, 1963.
- [263] A. N. Tikhonov and V. I. Arsenin. *Solutions of Ill-Posed Problems*, volume 14. Winston, Washington, DC, 1977.
- [264] X. Tong, M. Morzfeld, and Y. Marzouk. MALA-within-Gibbs Samplers for High-Dimensional Distributions with Sparse Conditional Structure. *arXiv preprint arXiv:1908.09429*, 2019.
- [265] J. F. Traub. *Information-based Complexity*. John Wiley and Sons Ltd., 2003.
- [266] J. F. Traub and A. G. Werschulz. *Complexity and Information*. Cambridge University Press, 1998.
- [267] R. Tripathy, I. Bilonis, and M. Gonzalez. Gaussian Processes with Built-in Dimensionality Reduction: Applications to High-Dimensional Uncertainty Propagation. *Journal of Computational Physics*, 321:191–223, 2016.
- [268] S. Uchida, K. Soga, and K. Yamamoto. Critical State Soil Constitutive Model for Methane Hydrate Soil. *Journal of Geophysical Research: Solid Earth*, 117(B3), 2012.
- [269] R. van Handel. Probability in High Dimension. Technical report, Princeton University, 2016.
- [270] G. van Rossum and F. L. Drake. *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA, 2009.
- [271] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. Jarrod Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. Carey, Í. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020.
- [272] J. A. Vrugt. Markov Chain Monte Carlo Simulation using the DREAM Software Package: Theory, Concepts, and MATLAB Implementation. *Environmental Modelling & Software*, 75:273–316, 2016.
- [273] J. A. Vrugt, C. Ter Braak, C. Diks, B. A. Robinson, J. M. Hyman, and D. Higdon. Accelerating Markov Chain Monte Carlo Simulation by Differential Evolution with Self-Adaptive Randomized Subspace Sampling. *International Journal of Nonlinear Sciences and Numerical Simulation*, 10(3):273–290, 2009.
- [274] J. A. Vrugt, C. J. Ter Braak, M. P. Clark, J. M. Hyman, and B. A. Robinson. Treatment of Input Uncertainty in Hydrologic Modeling: Doing Hydrology Backward with Markov Chain Monte Carlo Simulation. *Water Resources Research*, 44(12), 2008.
- [275] J. A. Vrugt, C. J. ter Braak, C. G. Diks, and G. Schoups. Hydrologic Data Assimilation using Particle Markov Chain Monte Carlo Simulation: Theory, Concepts and Applications. *Advances in Water Resources*, 51:457–478, 2013.
- [276] J. A. Vrugt, C. J. Ter Braak, H. V. Gupta, and B. A. Robinson. Equifinality of Formal (DREAM) and Informal (GLUE) Bayesian Approaches in Hydrologic Modeling? *Stochastic Environmental Research and Risk Assessment*, 23(7):1011–1026, 2009.
- [277] W. F. Waite, J. C. Santamarina, D. D. Cortes, B. Dugan, D. N. Espinoza, J. Germaine, J. Jang, J. Jung, T. J. Kneafsey, H. Shin, et al. Physical Properties of Hydrate-Bearing Sediments. *Reviews of Geophysics*, 47(4), 2009.
- [278] J. Wallin and D. Bolin. Geostatistical Modelling using Non-Gaussian Matérn Fields. *Scandinavian Journal of Statistics*, 42(3):872–890, 2015.
- [279] H. Wang and Y. Xia. Sliced Regression for Dimension Reduction. *Journal of the American Statistical Association*, 103(482):811–821, 2008.
- [280] Z. Wang, J. M. Bardsley, A. Solonen, T. Cui, and Y. M. Marzouk. Bayesian Inverse Problems with L_1 Priors: A Randomize-then-Optimize Approach. *SIAM Journal on Scientific Computing*, 39(5):S140–S166, 2017.
- [281] M. Waskom, O. Botvinnik, D. O’Kane, P. Hobson, J. Ostblom, S. Lukauskas, et al. seaborn: Statistical Data Visualization (v0.9.0), 2018.
- [282] P. Wei, Z. Lu, and J. Song. Variable Importance Analysis: A Comprehensive Review. *Reliability Engineering & System Safety*, 142:399–432, 2015.

-
- [283] C. K. Williams and C. E. Rasmussen. *Gaussian Processes for Machine Learning*, volume 2. MIT Press Cambridge, MA, 2006.
- [284] E. Wit, E. v. d. Heuvel, and J.-W. Romeijn. ‘All models are wrong...’: An Introduction to Model Uncertainty. *Statistica Neerlandica*, 66(3):217–236, 2012.
- [285] Wolfram Research, Inc. Mathematica, Version 11.3. Champaign, IL, 2018.
- [286] D. M. Wood. *Soil Behaviour and Critical State Soil Mechanics*. Cambridge University Press, 1990.
- [287] Y. Xia, H. Tong, W. K. Li, and L.-X. Zhu. An Adaptive Estimation of Dimension Reduction Space. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(3):363–410, 2002.
- [288] Z. Xu, N. Massei, I. Padilla, A. Hartmann, and B. Hu. Characterization, Modeling, and Remediation of Karst in a Changing Environment. *Environmental Earth Sciences*, 77(12):476, 2018.
- [289] J. Yang, G. O. Roberts, and J. S. Rosenthal. Optimal Scaling of Random-Walk Metropolis Algorithms on General Target Distributions. *arXiv preprint arXiv:1904.12157v3*, 2020.
- [290] Y. Yu. On Normal Variance–Mean Mixtures. *Statistics & Probability Letters*, 121:45–50, 2017.
- [291] T. S. Yun, J. C. Santamarina, and C. Ruppel. Mechanical Properties of Sand, Silt, and Clay Containing Tetrahydrofuran Hydrate. *Journal of Geophysical Research: Solid Earth*, 112(B4), 2007.
- [292] O. Zahm, P. G. Constantine, C. Prieur, and Y. M. Marzouk. Gradient-based Dimension Reduction of Multivariate Vector-Valued Functions. *SIAM Journal on Scientific Computing*, 42(1):A534–A558, 2020.
- [293] O. Zahm, T. Cui, K. Law, A. Spantini, and Y. Marzouk. Certified Dimension Reduction in Nonlinear Bayesian Inverse Problems. *arXiv preprint arXiv:1807.03712v2*, 2018.
- [294] P. Zhang, Z. Qiu, Y. Fu, and P. X.-K. Song. Robust Transformation Mixed-Effects Models for Longitudinal Continuous Proportional Data. *Canadian Journal of Statistics*, 37(2):266–281, 2009.
- [295] H. Zhou, J. J. Gómez-Hernández, and L. Li. Inverse Methods in Hydrogeology: Evolution and Recent Trends. *Advances in Water Resources*, 63:22–37, 2014.
- [296] O. C. Zienkiewicz and R. L. Taylor. *The Finite Element Method for Solid and Structural Mechanics*. Elsevier, 2005.