



UNIL | Université de Lausanne

Unicentre

CH-1015 Lausanne

<http://serval.unil.ch>

Year : 2023

Likelihood estimation and risk assessment for inverse problems in the geosciences

Friedli Lea

Friedli Lea, 2023, Likelihood estimation and risk assessment for inverse problems in the geosciences

Originally published at : Thesis, University of Lausanne

Posted at the University of Lausanne Open Archive <http://serval.unil.ch>

Document URN : urn:nbn:ch:serval-BIB_61BCFA95009A1

Droits d'auteur

L'Université de Lausanne attire expressément l'attention des utilisateurs sur le fait que tous les documents publiés dans l'Archive SERVAL sont protégés par le droit d'auteur, conformément à la loi fédérale sur le droit d'auteur et les droits voisins (LDA). A ce titre, il est indispensable d'obtenir le consentement préalable de l'auteur et/ou de l'éditeur avant toute utilisation d'une oeuvre ou d'une partie d'une oeuvre ne relevant pas d'une utilisation à des fins personnelles au sens de la LDA (art. 19, al. 1 lettre a). A défaut, tout contrevenant s'expose aux sanctions prévues par cette loi. Nous déclinons toute responsabilité en la matière.

Copyright

The University of Lausanne expressly draws the attention of users to the fact that all documents published in the SERVAL Archive are protected by copyright in accordance with federal law on copyright and similar rights (LDA). Accordingly it is indispensable to obtain prior consent from the author and/or publisher before any use of a work or part of a work for purposes other than personal use within the meaning of LDA (art. 19, para. 1 letter a). Failure to do so will expose offenders to the sanctions laid down by this law. We accept no liability in this respect.

Faculté des géosciences et de l'environnement
Institut des sciences de la Terre

Likelihood estimation and risk assessment for inverse problems in the geosciences

Thèse de doctorat

Titre de docteur en sciences de l'environnement

Présentée à la
Faculté des géosciences et de l'environnement
Institut des sciences de la Terre
de l'Université de Lausanne
par

Lea Friedli

Diplôme (M.Sc.) en Statistique
Institute of Mathematical Statistics and Actuarial Science
Université de Berne, Suisse

Jury

Prof. Dr. Marie-Elodie Perga, présidente du jury
Prof. Dr. Niklas Linde, directeur de thèse
Prof. Dr. David Ginsbourger, codirecteur de thèse
Prof. Dr. Klaus Holliger, expert interne
Prof. Dr. Klaus Mosegaard, expert externe

Lausanne, 2023



UNIL | Université de Lausanne
Faculté des géosciences et de l'environnement
bâtiment Géopolis bureau 4631

IMPRIMATUR

Vu le rapport présenté par le jury d'examen, composé de

Présidente de la séance publique :	Mme la Professeure Marie-Elodie Perga
Présidente du colloque :	Mme la Professeure Marie-Elodie Perga
Directeur de thèse :	M. le Professeur Niklas Linde
Co-directeur de thèse :	M. le Professeur David Ginsbourger
Expert interne :	M. le Professeur Klaus Holliger
Expert externe :	M. le Professeur Klaus Mosegaard

Le Doyen de la Faculté des géosciences et de l'environnement autorise l'impression de la thèse de

Madame Lea FRIEDLI

*Titulaire d'un
Master of Science in Statistics
de l'Université de Bern*

intitulée

**LIKELIHOOD ESTIMATION AND RISK ASSESSMENT FOR INVERSE
PROBLEMS IN THE GEOSCIENCES**

Lausanne, le 25 octobre 2023

Pour le Doyen de la Faculté des géosciences et de
l'environnement

Professeure Marie-Elodie Perga

Contents

List of Figures	v
List of Tables	vii
List of Symbols	ix
List of Abbreviations	xi
Summary	xiii
Résumé	xv
Résumé grand public	xvii
1 Introduction	1
1.1 Preface	1
1.2 Inversion problems in the geosciences	2
1.2.1 Bayesian inference	5
1.2.2 Petrophysical relationships	5
1.2.3 Hyperparameters	8
1.2.4 Risk assessment	9
1.3 Posterior sampling approaches	11
1.3.1 Markov chain Monte Carlo	11
1.3.2 Sequential Monte Carlo	13
1.4 Likelihood estimation	14
1.4.1 Latent variable model	14
1.4.2 Correlated pseudo-marginal method	15
1.4.3 Linearized Gaussian approximation	17
1.5 Rare event estimation	17
1.5.1 Sequential Monte Carlo	18
1.5.2 Energy-based models	19
1.6 Objectives	20
1.7 Outline	21
2 Lithological tomography with the correlated pseudo-marginal method	25
2.1 Introduction	26

2.2	Methodology	30
2.2.1	Latent variable model	30
2.2.2	Bayesian Inference with Markov Chain Monte Carlo	31
2.2.3	(Correlated) pseudo-marginal method	33
2.2.4	Baseline inversion methods	36
2.2.5	Performance assessment	38
2.3	Results	39
2.3.1	Data and inversion setting	39
2.3.2	Linear physics	41
2.3.3	Non-linear physics	45
2.4	Discussion	50
2.5	Conclusions	52
2.6	Appendix	53
2.6.1	DREAM algorithms and prior-sampling proposals	53
2.6.2	Analytical posterior PDF and importance density for linear physics	54
3	Solving geophysical inversion problems with intractable likelihoods: Linearized Gaussian approximations versus the correlated pseudo-marginal method	57
3.1	Introduction	58
3.2	Methodology	60
3.2.1	Latent variable model	60
3.2.2	Bayesian inference and intractable likelihoods	61
3.2.3	Gaussian approximation of the intractable likelihood	62
3.2.4	Correlated pseudo-marginal method	63
3.2.5	Performance assessment	64
3.3	Case study	65
3.3.1	Synthetic data generation	65
3.3.2	Inversion setting and prior assumptions	66
3.3.3	Results	67
3.4	Sensitivity analysis	68
3.4.1	Likelihood estimation	69
3.4.2	Inversion	70
3.5	Discussion	73
3.6	Conclusions	77
4	Inference of geostatistical hyperparameters with the correlated pseudo-marginal method	79
4.1	Introduction	80
4.2	Methodology	84
4.2.1	Random effects model	84
4.2.2	Bayesian inference with Markov chain Monte Carlo	85

4.2.3	Pseudo-marginal and correlated pseudo-marginal method	87
4.2.4	Rejection sampling	90
4.2.5	Performance assessment	90
4.3	Test case 1: Hydraulic conductivity field	91
4.3.1	Data and inversion setting	92
4.3.2	Results	95
4.4	Test case 2: Fracture aperture fields	97
4.4.1	Data and inversion setting	98
4.4.2	Results	101
4.5	Discussion	104
4.6	Conclusions	107
4.7	Appendix	109
4.7.1	Importance sampling for CPM	109
4.7.2	Complementary figures concerning test case 1	111
5	Rare event probability estimation for groundwater inverse problems with Sequential Monte Carlo methods	115
5.1	Introduction	116
5.2	Methodology	119
5.2.1	Notation	119
5.2.2	Bayesian inversion	119
5.2.3	From Sequential Monte Carlo to PostRisk-SMC	120
5.3	1D flow example	128
5.3.1	Synthetic setting	128
5.3.2	Results	129
5.4	2D transport example	134
5.4.1	Problem setting	134
5.4.2	Results	135
5.5	Discussion	139
5.6	Conclusions	142
6	An energy-based model approach to rare event probability estimation	145
6.1	Introduction	146
6.2	Methodology	149
6.2.1	Problem setting	149
6.2.2	Energy-based model approach	150
6.2.3	Alternative rare event probability estimation methods	155
6.3	Illustrative test examples	156
6.3.1	Analytical contamination example	157
6.3.2	Four-branch function	160
6.3.3	Load capacity example	163

6.4	Discussion	165
6.5	Conclusion	167
6.6	Appendix	168
6.6.1	Relation to maximum likelihood estimation	168
7	Conclusions and outlook	171
7.1	Conclusions	171
7.2	Outlook	173
7.2.1	Bayesian inference	173
7.2.2	Likelihood estimation	174
7.2.3	Rare event estimation	175
	Bibliography	179
	Acknowledgements	197
	Curriculum Vitae	199

List of Figures

1.1	Illustrations of exemplary geophysical data	4
1.2	Illustration of (hydro)geophysical inversion	7
1.3	Illustration of hyperparameter estimation	9
1.4	Illustration of risk assessment	11
1.5	Illustrations of the Markov chain Monte Carlo method and the Sequential Monte Carlo method for posterior sampling	12
1.6	Illustration of the pseudo-marginal method for inversion problems	16
1.7	Illustration of the Sequential Monte Carlo method for rare events	18
1.8	Illustration of the energy-based model approach	19
2.1	Synthetic data	41
2.2	Analytical posterior linear physics	42
2.3	Results linear physics	43
2.4	Histograms comparing samples from the analytical posterior and samples from the different inversion methods (linear physics)	44
2.5	KL-divergences with respect to the analytical posterior for the different inversion methods (linear physics)	45
2.6	Variance of the log-likelihood ratio estimator with and without importance sampling (non-linear physics)	47
2.7	Results non-linear physics	48
2.8	Logarithmic scores for the different inversion methods (non-linear physics)	50
3.1	Flow chart illustrating the LinGau method at iteration j	62
3.2	Flow chart illustrating the CPM method with importance sampling at iteration j	64
3.3	The synthetic “true” model adapted from Friedli et al. 2022	66
3.4	Estimated posterior means obtained with the different inversion methods and corresponding pixel-wise logarithmic scores	68
3.5	Exemplary GPR ray paths with and without the petrophysical prediction error	69
3.6	Synthetic “true” models for the layered test cases	71
3.7	Estimators of log-likelihood of the true porosity field in the layered setting as a function of petrophysical prediction uncertainty	72
3.8	Estimates of the marginal posterior PDFs for the layered test cases	74
4.1	Flow chart illustrating the CPM method with importance sampling at iteration j	89
4.2	Synthetic data (first test case)	92

4.3	Variance of the log-likelihood ratio estimator (first test case)	94
4.4	Flow chart illustrating the simplified MH (assuming ergodicity) procedure for the first test case at iteration j	95
4.5	Posterior samples obtained with the different inversion methods (first test case)	97
4.6	Synthetic data (second test case)	99
4.7	Importance sampling (second test case)	100
4.8	Flow chart illustrating the homogeneous inversion procedure for the second test case at iteration j	101
4.9	Posterior samples obtained with the different inversion methods (second test case)	102
4.10	Posterior of the horizontal equivalent log-hydraulic transmissivity (second test case)	104
4.11	Log-likelihood values obtained with brute-force sampling of the latent aperture field (second test case)	108
4.12	Posterior samples for the remaining hyperparameters (first test case, ergodic) .	112
4.13	Posterior samples for the remaining hyperparameters (first test case, non-ergodic)	113
5.1	Illustration of the SMC method for posterior inference	122
5.2	Flow chart illustrating the SMC method for posterior inference	124
5.3	Illustration of the SMC method for rare events	125
5.4	Flow chart illustrating the SMC method for rare event estimation	126
5.5	Work flow of the PostRisk-SMC method	127
5.6	Synthetic data for the 1D flow example	129
5.7	Results for the 1D flow example with the PostRisk-SMC method	130
5.8	Illustration of the bias resulting from the adaptively determined threshold sequence within the PostRisk-SMC method for the 1D flow example	131
5.9	Impact of the configuration choices within the PostRisk-SMC method for the 1D flow example	132
5.10	Results for the 1D flow example with the MH method	132
5.11	Synthetic data for the 2D transport example	136
5.12	Rare event estimation for the 2D transport example with the PostRisk-SMC method (without inversion)	137
5.13	Particle mean fields for the 2D transport example with the PostRisk-SMC method	138
5.14	Rare event estimation for the 2D transport example with the PostRisk-SMC method (with inversion)	139
6.1	Illustration of the energy-based model approach	151
6.2	PDFs of the quantity of interest in the three illustrative examples	156
6.3	Analytical contamination example: Choice of $p_{\text{ref}}(r)$	158
6.4	Analytical contamination example: Optimization schedules	159
6.5	Four-branch example: Impact of $p_{\text{ref}}(r)$	162
6.6	Load capacity example: Form of the bias potential	164

List of Tables

2.1	Overview of the inversion methods applied on the latent variable model	36
2.2	Overview of the results obtained for the linear test example with the different inversion approaches and proposal mechanisms	46
2.3	Summary of the results obtained for the non-linear test example with the various inversion approaches and the two proposal mechanisms	49
3.1	Summary of the results for the Gaussian porosity field example	67
3.2	Summary of the results obtained for the study targeting a layered porosity field	75
4.1	Table summarizing the results for the first test case obtained with the different inversion methods for the ergodic and the non-ergodic data settings	96
4.2	Table summarizing the results for the second test case obtained with the CPM method and the inversion assuming homogeneity	103
5.1	Table summarizing the different trials of the PostRisk-SMC and MH method applied to the 1D flow test case	133
5.2	Table summarizing the different trials of the PostRisk-SMC and MH method applied to the 2D transport test case	138
6.1	Comparison of the EBM method and subset sampling for the analytical contamination example	160
6.2	Comparison of the EBM method and subset sampling for the four-branch example	162
6.3	Table summarizing the comparison of EBM and BUS for the load capacity example	165

List of Symbols

Terms separated by forward slashes are used as synonyms.

Bayesian inversion / Bayesian inverse problem

$\boldsymbol{\theta}$	Target model parameters / properties / variables
$\mathcal{G}(\cdot)$	Forward operator
\mathbf{y}	Noise-contaminated measurements / observations / data
$\boldsymbol{\varepsilon}_{\mathcal{O}}$	Measurement / observation / data error
$p(\boldsymbol{\theta} \mathbf{y})$	Posterior probability density function
$p(\mathbf{y} \boldsymbol{\theta})$	Likelihood function
$p(\boldsymbol{\theta})$	Prior probability density function

Latent variable model / random effects model

X	Latent variable(s)
$\mathcal{F}(\cdot)$	Petrophysical relationship
$\boldsymbol{\varepsilon}_{\mathcal{P}}$	Petrophysical prediction error

Rare event estimation

$R = \mathcal{R}(\boldsymbol{\theta})$	Quantity of interest
$\boldsymbol{\theta} \mapsto \mathcal{R}(\boldsymbol{\theta})$	Relationship linking the model parameters to the quantity of interest
$p_{R Y}(r \mathbf{y})$	Posterior probability density function of the quantity of interest
$\mathcal{R}(\boldsymbol{\theta}) \geq T$	Rare event / hazard
$\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq T)$	Rare event probability
$\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq T \mathbf{y})$	Rare event probability in inversion setting

General

Random variables are represented by uppercase letters, and their corresponding realizations are denoted by lowercase letters. When working with functions, we typically use lowercase letters.

List of Abbreviations

CDF	Cumulative distribution function
CPM	Correlated pseudo-marginal (method)
EBM	Energy-based model
i.i.d.	Independent and identically distributed
IS	Importance sampling
KL	Kullback Leibler (divergence)
KSD	Kernelized Stein discrepancy
LH	Likelihood (function)
LinGau	Linearized Gaussian approximation
logS	Logarithmic Score
MCMC	Markov chain Monte Carlo
MH	Metropolis–Hastings
pCN	Preconditioned Crank–Nicolson (proposals)
PDF	Probability density function
PM	Pseudo-marginal (method)
PostRisk	Posterior Risk (Sequential Monte Carlo; PostRisk-SMC)
PPE	Petrophysical prediction error
RS	Rejection sampling
SMC	Sequential Monte Carlo

Summary

In the geosciences, inversion problems emerge from the need to estimate environmental variables from indirect measurements. This thesis concentrates on the inference of geophysical or hydrogeological properties using geophysical or hydrological measurements. To enable uncertainty quantification, we employ a probabilistic inversion approach within a Bayesian framework, focusing on the posterior probability density function of the model parameters given the observed data. We address the following two research questions: (1) How can we efficiently solve Bayesian inverse problems involving intractable likelihood functions? (2) How can the combination of inversion and risk assessment methods enhance the accuracy of rare event probability estimation? The likelihood function is intractable in latent variable models, where the relationship between target model parameters and measurements is obscured by an intermediate (latent) variable. We consider cases involving uncertain petrophysical relationships and hyperparameter estimation, respectively. While the former is concerned with estimating hydrogeological parameters from geophysical data by treating the intermediate geophysical properties as latent variables, the latter targets hyperparameters (such as mean, standard deviation and integral scales) by considering the local properties of the field as latent variables. Performing inversion in both situations necessitates the estimation of an intractable likelihood function. To address this challenge, we employ two methods: the correlated pseudo-marginal method, which involves Monte Carlo averaging over samples of the latent variable, and a Gaussian approximation based on local linearization of the geophysical forward operator. In the context of petrophysical uncertainty, we find that the correlated pseudo-marginal method allows for accurate estimation, even in scenarios with high petrophysical uncertainty. In contrast, the less computationally intensive linearized Gaussian approach becomes gradually less accurate as the uncertainty in the petrophysical relationship increases. In the realm of hyperparameter estimation, we concentrate on the correlated pseudo-marginal method, showcasing its capacity to enhance the accuracy of hyperparameter estimation. Our second research question focuses on rare event estimation, for which the main objective is not to determine the posterior distribution of model parameters, but to characterize the distribution of a quantity of interest depending on these parameters through a non-linear relationship. Specifically, we aim to calculate the probability of this quantity assuming critical values, representing a rare event. To handle the associated challenges of rare event estimation, we employ two approaches: Sequential Monte Carlo combined with subset sampling and an energy-based model approach utilizing a bias potential. We evaluate the performance of the proposed methods using problems related to groundwater hazards and illustrative examples from the engineering literature. Encouragingly, both approaches demonstrate the ability to accurately estimate rare event probabilities smaller than one in a million. In summary, this thesis presents methodological advances to improve the computational efficiency and realism of probabilistic Bayesian

inversion approaches for targeted estimation of properties and attributes of environmental systems.

Keywords Bayesian inversion, geophysics, hydrology, hydrogeophysics, Markov chain Monte Carlo, Sequential Monte Carlo, latent variable model, likelihood estimation, importance sampling, rare events, energy-based models

Résumé

Dans le domaine des géosciences, les problèmes inverses émergent de la nécessité d'estimer des variables environnementales à partir de mesures indirectes. Cette thèse se concentre sur l'inférence de propriétés géophysiques ou hydrogéologiques à partir de mesures géophysiques ou hydrologiques. Pour permettre une quantification des incertitudes associées nous employons une approche d'inversion probabiliste dans un cadre bayésien, en nous concentrant sur la distribution a posteriori des paramètres du modèle compte tenu des données observées. Nous abordons les deux questions de recherche suivantes: (1) Comment adapter efficacement des méthodes d'inversion bayésiennes en présence de variables latentes? (2) Comment combiner les approches d'inversion et d'évaluation des risques pour améliorer l'exactitude dans l'estimation de probabilités d'événements rares? La fonction de vraisemblance est difficile à cerner dans les modèles à variables latentes, où la relation entre les paramètres du modèle cible et les mesures est opacifiée par les incertitudes résiduelles sur lesdites variables latentes. Nous examinons des cas impliquant respectivement des relations pétrophysiques incertaines et l'estimation d'hyperparamètres. Alors que le premier cas concerne l'estimation de paramètres hydrogéologiques à partir de données géophysiques en traitant les propriétés géophysiques intermédiaires comme des variables latentes, le second cas cible des hyperparamètres (tels que la moyenne, l'écart-type et les échelles d'intégrales) en considérant cette fois les propriétés locales du champ comme des variables latentes. L'inversion nécessite dans les deux situations l'estimation d'une fonction de vraisemblance complexe. Pour relever ce défi, nous utilisons deux méthodes: la méthode pseudo-marginale corrélée, qui implique une moyenne de Monte Carlo sur des échantillons de la variable latente, et une approximation gaussienne basée sur la linéarisation locale de la réponse géophysique ciblée. Dans le contexte de l'inversions en présence de relations pétrophysiques incertaines, nous constatons que la méthode pseudo-marginale corrélée permet une estimation précise, même dans des scénarios à forte incertitude pétrophysique. En revanche, l'approche gaussienne linéarisée, moins gourmande en ressources computationnelles, perd progressivement en précision à mesure que l'incertitude de la relation pétrophysique augmente. Dans le cas de l'estimation des hyperparamètres, nous nous concentrons sur la méthode pseudo-marginale corrélée et constatons qu'elle peut permettre d'améliorer l'estimation des hyperparamètres. Notre deuxième question de recherche porte sur l'estimation d'événements rares, pour laquelle l'objectif principal n'est pas de déterminer la distribution postérieure des paramètres du modèle, mais plutôt de caractériser la distribution d'une quantité d'intérêt dépendant de manière non linéaire de ces paramètres. Plus précisément, nous cherchons à calculer la probabilité que cette quantité prenne une valeur critique, qui représente un événement rare. Pour relever les défis associés à l'estimation des événements rares, nous utilisons deux approches : d'une part, une méthode de Monte Carlo séquentielle combinée à un sous-échantillonnage et, d'autre part, une approche fondée sur des modèles génératifs reposant sur le concept d'énergie libre (Energy-based models; EBM)

et utilisant un potentiel de biais. Nous évaluons les performances des méthodes proposées en utilisant des problèmes simplifiés liés aux risques de contamination des eaux souterraines et des exemples tirés de la littérature de l'estimation de probabilité d'évènements rares en ingénierie. Il est encourageant de constater que les deux approches permettent d'estimer avec précision des probabilités d'évènements rares d'occurrence inférieure à un sur un million. En résumé, cette thèse présente des avancées méthodologiques visant à améliorer l'efficacité computationnelle et le réalisme des approches d'inversion probabiliste bayésienne pour l'estimation ciblée de propriétés et attributs des systèmes environnementaux.

Résumé grand public

Les études géoscientifiques impliquent souvent la résolution de problème dits “inverses”, dans lesquels des paramètres et propriétés environnementaux sont estimés sur la base de mesures disponibles et d’hypothèses de départ sur les phénomènes étudiés. Cette thèse se concentre sur l’utilisation de données géophysiques et hydrogéologiques pour déduire les caractéristiques géophysiques (par exemple, la conductivité électrique), géologiques (par exemple, la teneur en argile) ou hydrogéologiques (par exemple, la saturation en eau) dans la subsurface. Lors de la résolution de problèmes inverses, il est important d’avoir conscience qu’il existe toujours une incertitude inhérente aux résultats obtenus. Pour y remédier, nous utilisons une approche d’inversion probabiliste qui fournit une gamme de solutions possibles (dite distribution postérieure bayésienne). Dans ce contexte, cette thèse se concentre sur deux questions de recherche principales: (1) Comment résoudre efficacement les problèmes d’inversion si les paramètres d’intérêt ne sont pas directement liés aux mesures? (2) Comment combiner les approches d’inversion et d’évaluation des risques pour prédire la probabilité d’occurrence des événements dangereux? L’inversion probabiliste repose sur une fonction de vraisemblance, qui quantifie la probabilité qu’un ensemble spécifique de valeurs des paramètres du modèle soit à l’origine des données observées. Dans certains modèles, il est difficile d’évaluer la fonction de vraisemblance en raison de la présence d’une variable latente intermédiaire, qui brouille la relation directe entre les paramètres du modèle et les mesures. Nous considérons d’abord un scénario dans lequel les variables hydrogéologiques d’intérêt ne sont qu’indirectement liées aux mesures géophysiques par l’intermédiaire de propriétés géophysiques latentes. Dans un second scénario d’intérêt, nous nous intéressons à l’estimation d’hyperparamètres tels que la moyenne et la structure de corrélation du domaine cible lorsque les mesures ne sont qu’indirectement liées à ces hyperparamètres par le biais de propriétés locales du domaine. Le calcul direct de la fonction de vraisemblance n’est possible dans aucun de ces deux scénarios. Pour surmonter cette difficulté, nous investiguons deux méthodes : la méthode pseudo-marginale corrélée et une approximation gaussienne linéarisée. Alors que la méthode pseudo-marginale corrélée implique un échantillonnage et un calcul de moyenne sur les variables latentes, l’approximation gaussienne linéarisée repose sur une approximation locale. Dans nos cas d’essai, la méthode pseudo-marginale corrélée apparaît comme une approche généralement applicable et précise. En revanche, l’approximation gaussienne linéarisée est plus efficace sur le plan du calcul, mais au prix d’une perte de précision. Dans notre deuxième question de recherche, nous nous concentrons sur l’estimation d’événements rares, où l’objectif principal n’est pas d’estimer les paramètres du modèle eux-mêmes, mais évaluer un risque qui dépend de ces paramètres. Un exemple d’évènement dangereux étudié en hydrogéologie est la contamination d’un aquifère, un processus qui est influencé par les propriétés hydrauliques entre la source de contamination et l’aquifère. Lorsqu’il s’agit d’événements rares, il est nécessaire de recourir à des approches spécialisées pour parvenir à

une estimation précise de la probabilité d'occurrence. Tout d'abord, nous considérons une méthode de Monte Carlo séquentielle combinée à un échantillonnage par sous-ensembles. Dans cette approche, nous générons des échantillons de manière itérative et les rapprochons progressivement de la région critique associée à l'événement rare en question. Par la suite, nous introduisons une approche basée sur le concept d'énergie libre (Energy-based models; EBM) en utilisant un potentiel de biais pour forcer l'échantillonnage dans une région qui est pertinente pour la quantité d'intérêt. De manière prometteuse, les deux approches montrent de bonnes performances dans l'estimation précise des probabilités d'événements rares, même lorsqu'ils se produisent à un taux inférieur à un sur un million. En résumé, cette thèse présente des avancées méthodologiques visant à améliorer l'efficacité computationnelle et le réalisme des approches d'inversion probabiliste pour l'estimation ciblée de propriétés et attributs des systèmes environnementaux.

Chapter 1

Introduction

1.1 Preface

Geoscience is a comprehensive research field dedicated to the study of our planet. Its primary focus is to understand the processes and interactions that shape the Earth, its natural resources and its environment. The geosciences are vital in addressing important issues such as climate change, assessing and minimizing hazards, managing resources, and promoting sustainable development. Through explorations of the past, analyses of the present, and predictions about the future of our planet, geoscientific research plays an essential role in enhancing our collective comprehension of the Earth. Furthermore, it helps to ensure the long-term sustainability of our planet, benefiting future generations.

While geosciences also encompass the study of surface and atmospheric processes, an important research domain focuses on the subsurface environment. The subsurface, hidden beneath our feet, plays a very important role in different geological, hydrological and ecological processes. One important characteristic of the subsurface is its critical role as a reservoir for groundwater, which serves as a vital source of drinking water, irrigation, and ensures the proper functioning of ecosystems (Gorelick and Zheng 2015). According to a study conducted by UNICEF in 2021, more than 1.42 billion people worldwide live in areas with high or extremely high water insecurity, including 450 million children (UNICEF 2023). These concerning numbers underline the immediate need for proactive measures to address the water crisis. By exploring the subsurface, we gain important knowledge about aquifer characteristics, water availability and the sustainable handling of water resources (Ajami et al. 2008). Moreover, a thorough investigation of the subsurface provides scientists and policymakers with the necessary understanding and tools to address issues related to water pollution. Another important aspect of the subsurface is its role in energy exploration and extraction (McCartney et al. 2016). The subsurface contains reserves of fossil fuels, such as coal, oil and natural gas, as well as mineral resources. Understanding the geology and properties of the subsurface is essential for locating and extracting these valuable energy and mineral resources, which are still vital for various industries. When it comes to sustainable energy solutions, the subsurface plays a fundamental role in the field of geothermal energy (Barbier 2002). Geothermal power systems utilize the heat stored within the Earth to generate electricity or provide heating and cooling. Understanding geothermal reservoirs and their characteristics is essential for effectively utilizing this renewable energy source. Furthermore,

the subsurface also plays a crucial role in the storage of waste and hazardous materials (Krauskopf 2013). Deep geological repositories are designed to safely store nuclear waste and other hazardous substances, minimizing their potential impact on the environment and human health. Studying the subsurface helps identifying suitable locations and assessing the long-term safety and stability of such repositories.

The subsurface environment exhibits significant heterogeneity and hosts complex and interconnected processes across various spatial and temporal scales. While valuable information about subsurface structures and processes can be derived from data sources such as borehole measurements, outcrops, laboratory analysis of field samples, geophysical data and hydrogeological experiments, it is important to acknowledge that information is incomplete (Linde et al. 2017). It is critical to recognize and properly account for uncertainties in order to ensure accurate and reliable interpretations of subsurface data and enable robust decision-making processes. When studying subsurface systems, inversion is commonly used to formulate and solve problems (Tarantola 2005). Inverse theory envelops a collection of mathematical methodologies used to gain insights about system properties and states from observations (Menke 2018). The practice of inferring parameter values based on observations has a long history in quantitative science, however, the first formal approaches emerged in the period between 1760 and 1810 (Stigler 1986). During that period, significant advancements were made in addressing two fundamental problems: estimating the Earth's shape using geodetic data and determining planetary and comet orbits based on astronomical observations (Tarantola 2006).

1.2 Inversion problems in the geosciences

In the geosciences, inversion problems arise out of the need to estimate unknown properties or parameters describing environmental systems. While inversion problems occur in a variety of subfields within the geosciences, this thesis mainly focuses on applications in geophysics (e.g., Parker 1994; Tarantola 2005, Menke 2018), hydrogeology (e.g., McLaughlin and Townley 1996; Carrera et al. 2005) and hydrogeophysics (e.g., Rubin and Hubbard 2005; Linde and Doetsch 2016). We are concerned with inversion problems defined for estimating geophysical, geological or hydrogeological properties and state variables using geophysical or hydrological measurements.

Geophysical inversion benefits from geophysical data being sensitive to the physical properties (e.g., electrical resistivity) of subsurface materials and allows us to enhance our understanding of geological processes and testing scientific hypotheses (Linde et al. 2015). In the field of applied geophysics, Reynolds (2011) identifies key families of geophysical methods, including potential field methods, seismology, electrical methods, and electromagnetic methods. Figure 1.1 illustrates two examples of geophysical data: cross-borehole ground penetrating radar (GPR) and electrical resistivity tomography (ERT). Similarly, hydraulic data provides information about flow of fluids and transport of matter within the subsurface. Hydraulic testing initially focused on pumping tests, which were primarily designed for permeable formations in the context of water supply assessments. However, with in-

creasing concerns regarding waste migration and the need to understand flow dynamics in low-permeability formations, interest has expanded to include testing procedures suitable for such conditions (Domenico and Schwartz 1997). One example is slug tests, where a discrete volume of water is rapidly added or removed from a well, and the subsequent hydraulic head response is measured.

The starting point to solve an inverse problem is to parameterize the natural system of interest using parameters or state variables $\boldsymbol{\theta} = (\theta_1, \dots, \theta_L)$. One exemplary way to accomplish this is by employing a discretization approach, where the model domain is divided into a finite number of cells. Each cell represents the value of the property of interest within that specific region. Subsequently, all relevant prior information regarding the natural system is gathered and compiled. This includes for instance expert knowledge providing insights into the characteristics of the system. The next step is concerned with the relationship between the target model parameters $\boldsymbol{\theta}$ and the observations $\mathbf{y} = (y_1, \dots, y_T)$. It is commonly expressed using a forward solver $\mathcal{G} : \mathbb{R}^L \rightarrow \mathbb{R}^T$ which numerically approximates physical laws. Deriving accurate forward solvers assuming specific conditions is the concern of numerical modeling, establishing mathematical models and algorithms that accurately predict the behavior and outcomes of a system based on given inputs and assumptions (Menke 2018). Under the assumption that the measurements include observational noise $\boldsymbol{\varepsilon}_{\mathcal{O}}$ but no model errors, we write,

$$\mathbf{Y} = \mathcal{G}(\boldsymbol{\theta}) + \boldsymbol{\varepsilon}_{\mathcal{O}}. \quad (1.1)$$

In this thesis, we refer to random variables and random vectors with upper-case letters and to realizations thereof with lower-case letters. When solving the inverse problem, the goal is to identify parameter values $\boldsymbol{\theta}$ that best match the measured data \mathbf{y} and any prior or regularization constraints.

Inverse problems are widely recognized as being ill-posed, meaning that they lack unique solutions and can be highly sensitive to small changes in the input data (Backus 1970; Tikhonov and Arsenin 1977). In underdetermined problems with insufficient information in the data \mathbf{y} to determine all the unknowns of the system, there is an infinite number of solutions that fit the data even in the absence of data noise. In overdetermined data-rich settings, measurement errors, simplified numerical forward models and assumptions about conceptual models introduce inherent uncertainty into the inversion problem. Furthermore, not all geoscientific relationships are unique (Tarantola and Valette 1982a). The latter for instance is the case with gravitational data as an infinity of different density models lead to identical gravitational fields. It is common to distinguish between two sources of parameter uncertainties (Tartakovsky 2013): epistemic and aleatory. Epistemic uncertainty arises from incomplete knowledge or understanding of the underlying system or phenomenon and can be reduced through the collection of additional data or improvement of the model itself. The latter involves refining the assumptions, equations and parameters used in the model to better align with the real-world system. On the other hand, aleatory uncertainty refers to inherent randomness that arises from natural variability in space, time, or individual behavior. This distinction between epistemic and aleatory uncertainty is also reflected in other works, such as Helton (1994), which differentiates between stochastic and subjective uncertainty, and

Hoffman and Hammonds (1994), which classifies uncertainty as Type A (aleatory) and Type B (epistemic).

Inversion problems are commonly tackled using either of two fundamental approaches: deterministic and probabilistic inversion. In the deterministic approach, solving the inversion problem is approached as an optimization task, where gradients typically play a crucial role in guiding the optimization procedure. Conventional deterministic inversion methods typically enable a unique solution of the inversion problem by relying on regularization constraints (Constable et al. 1987). Although they are popular due to their simplicity and computational efficiency, they generally rely on constraints lacking geological justification (Ellis and Oldenburg 1994). Furthermore, deterministic approaches do not offer a reliable evaluation of uncertainty (Linde et al. 2017). Probabilistic inversion, in contrast, offers a range of possible solutions. For geophysical problems, it gained significant prominence with Tarantola and Valette (1982a) and Tarantola and Valette (1982b) introducing the very general concept of states of information. Within probabilistic methods, the model parameters are treated as random variables and the aim is to determine their distribution (Menke 2018). Thereby, accurate estimation of errors and uncertainty quantification are ensured (Jackson and Matsu'Ura 1985). We rely on probabilistic inversion based on a classical Bayesian framework, which has been widely applied to inversion problems in the geosciences.

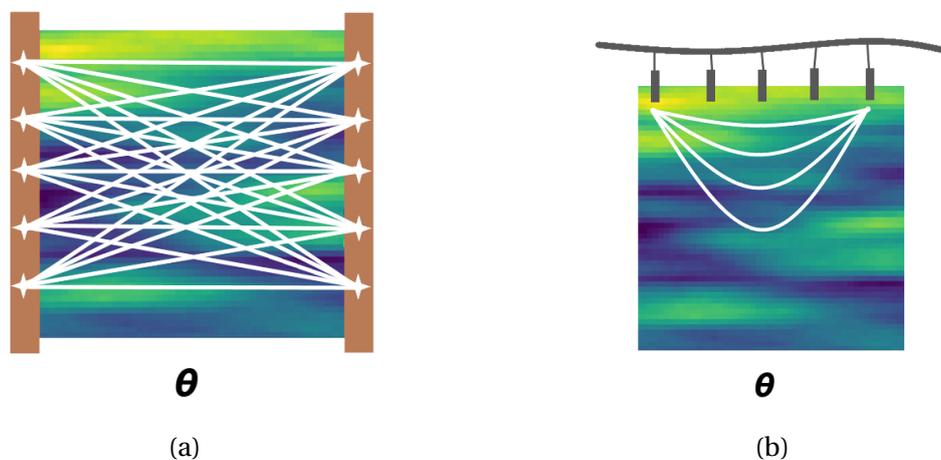


Figure 1.1: Illustrations of exemplary geophysical data: (a) Cross-borehole ground penetrating radar (GPR) employing a configuration of five transmitters and receivers, resulting in a total of 25 first-arrival travel times providing integrative measurements along the ray paths. (b) Electrical resistivity tomography (ERT) involving the utilization of two electrodes to inject electrical currents into the ground and the resulting potential differences for the electrodes located in-between being measured. While GPR data are sensitive to both electrical conductivity and permittivity, ERT data are only sensitive to electrical conductivity. The physics underlying these methods is non-linear, implying that the ray-paths and the current patterns depend on the subsurface structure.

1.2.1 Bayesian inference

In the Bayesian approach, we begin by specifying a prior probability density function (PDF) $\boldsymbol{\theta} \mapsto p(\boldsymbol{\theta})$, which captures all the prior knowledge and assumptions about the parameters $\boldsymbol{\theta}$. This prior distribution represents our initial beliefs or uncertainties about the parameter values. We consider a finite-dimensional parameter space and a prior PDF, which is absolutely continuous with respect to the Lebesgue measure. To incorporate the available data \mathbf{y} , we utilize the likelihood function $\boldsymbol{\theta} \mapsto p(\mathbf{y}|\boldsymbol{\theta})$, which quantifies the probability of obtaining the observed data given specific parameter values. To combine the prior distribution and the likelihood function, we apply Bayes' theorem to derive the posterior PDF,

$$p(\boldsymbol{\theta}|\mathbf{y}) = \frac{p(\boldsymbol{\theta})p(\mathbf{y}|\boldsymbol{\theta})}{p(\mathbf{y})}. \quad (1.2)$$

The denominator $p(\mathbf{y})$ is given by the so-called model evidence, which is assumed to be positive. The posterior PDF provides a probabilistic description of the parameter uncertainties, accounting for the compatibility between the model predictions and the observed data. Bayes' theorem is named after the Reverend Thomas Bayes, who was a statistician, a philosopher and a theologian. During the 1740s, Bayes made a significant statement related to probability theory, which laid the foundation for what is now known as Bayes' theorem. However, it was only after his death that the theorem was published in 1763, in a work titled "An Essay towards solving a Problem in the Doctrine of Chances" (Bayes 1763).

Bayesian inversion can handle complex and nonlinear forward models, it naturally accounts for parameter dependencies encompassing in particular covariance structures and ensures uncertainty quantification for the estimates. Furthermore, it enables the incorporation of complex prior knowledge. While the possibility to work with realistic priors is advantageous, the choice of the prior distribution is also challenging (Malinverno and Briggs 2004) as it has a strong influence on the estimation (Scales and Tenorio 2001). Its selection often includes uncertainty, as do the assumptions about the forward model and the data error in the likelihood function. To account for these uncertainties, "empirical Bayes" and "hierarchical Bayes" approaches are applied (Malinverno and Briggs 2004). Furthermore, Bayesian model selection utilizes the model evidence to choose between different conceptual prior models (e.g., Linde 2014; Hoegge et al. 2019). While these research topics are fascinating and have significant implications in the field, they go beyond the scope of this thesis.

1.2.2 Petrophysical relationships

In the classical Bayesian inversion problem, the data \mathbf{y} are related to the model parameters $\boldsymbol{\theta}$ through a forward model $\boldsymbol{\theta} \mapsto \mathcal{G}(\boldsymbol{\theta})$. However, in geophysical inversion problems, such relationships often only exist between geophysical properties and geophysical data. Yet, the interest of most geophysical investigations extend beyond the sole inference of geophysical property models, for instance, to provide valuable insights and constraints on (hydro)geological parameters and state variables (Fig. 1.2a). These parameters include essential characteristics such as clay fraction, mineral composition, water saturation or salinity. Even if the interpretation of such properties or state variables of interest remains the primary

objective, geophysical inverse theory has traditionally focused on evaluating the accuracy and uncertainty associated with the inferred geophysical properties (e.g., Parker 1994; Menke 2018; Tarantola 2005; Aster et al. 2018). To infer (hydro)geological properties and variables, crucial aspects involve the petrophysical (rock physics) relationships. These relationships establish the connection between geophysical properties and the desired (hydro)geological targets (e.g., Hinnell et al. 2010; Kowalsky et al. 2005).

Hydrogeophysics, drawing upon geophysical data, offers valuable insights into hydrological processes and the underlying subsurface structures that govern them (Rubin and Hubbard 2005). Defining an appropriate petrophysical relationship is a significant challenge in hydrogeophysics (Binley et al. 2015), as parameter values and analytical forms of the relationship can vary significantly among different lithologies (Hubbard and Rubin 2005). Consequently, petrophysical relationships are often inherently uncertain (e.g., Mavko et al. 2020). Nevertheless, in many hydrogeophysical inversion studies, the petrophysical relationship is assumed to be perfect provided that the right parameter values are used (e.g., Lochbühler et al. 2014; Kowalsky et al. 2005). The assumption of a perfect petrophysical relationship can introduce bias, too narrow uncertainty bounds and excessively variable parameter estimates (Brunetti and Linde 2018). Brunetti and Linde (2018) identify three sources of uncertainty in the petrophysical relationship: model uncertainty, parameter uncertainty, and prediction uncertainty (Fig. 1.2b). While model and parameter uncertainties result from the selection of the petrophysical model and its parameter values, prediction uncertainty arises from scatter and bias around the calibrated petrophysical model. Within the scope of this thesis, we are primarily addressing petrophysical prediction uncertainty.

Most studies accounting for petrophysical prediction uncertainty are relying on a two-step approach: first, geophysical properties are estimated using deterministic gradient-based inversions and second, the parameters of interest are derived using uncertain petrophysical relationships (e.g., Chen et al. 2001; Mukerji et al. 2001; González et al. 2008; Grana and Della Rossa 2010; Shahraeeni and Curtis 2011). The resulting estimates can be misleading if the difference in resolution of the inversion results and the petrophysical relationships is ignored (Day-Lewis et al. 2005). Furthermore, the two-step approach prevents the method from accounting for prior constraints on the (hydro)geological target variable (Ferré et al. 2009) and for physical constraints such as conservation of mass, continuity and momentum.

An alternative to the two-step approach is the use of coupled inversions, where hydrogeological properties are directly targeted through the inversion of geophysical data (e.g., Hinnell et al. 2010; Kowalsky et al. 2005). This approach encounters challenges, however, as the intermediate geophysical property is unobservable (latent), leading to what is known as a latent variable model. The likelihood of observing the geophysical data given the proposed hydrogeological parameters in such a latent variable model is often intractable. To overcome this difficulty, one strategy is to infer the joint posterior distribution of the hydrogeological and geophysical parameters. The concept of lithological tomography, as pioneered by Bosch (1999), introduced an approach to estimate the joint posterior by incorporating geophysical data, geological prior knowledge and uncertain petrophysical relationships. The original formulation of lithological tomography suffers, unfortunately, from inefficiency when dealing

with large geophysical datasets with high signal-to-noise ratios and significant petrophysical uncertainty. Brunetti and Linde (2018) suggested an alternative method inferring the petrophysical prediction error alongside the target parameters. This approach offers significant efficiency improvements compared to Bosch’s original formulation. Nevertheless, it faces challenges due to its high dimensional parameter space and the strong posterior correlation between the target variable and the petrophysical prediction error. To circumvent these challenges, we consider two methods to estimate the intractable likelihood function of the geophysical data given the (hydro)geological parameters: the correlated pseudo-marginal method and a linearized Gaussian approximation (Section 1.4, Chapters 2 and 3). With an accurate estimation of the likelihood, we can directly infer the (hydro)geological model parameters of interest from the geophysical data.

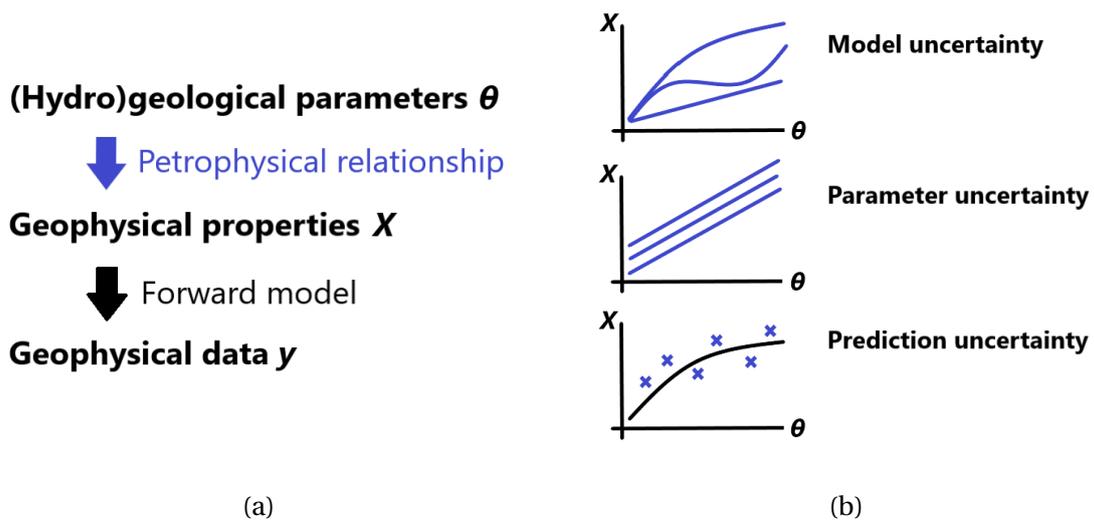


Figure 1.2: Illustration of (hydro)geophysical inversion: (a) The model parameters θ cannot directly be related to the geophysical measurements y , but only via the geophysical properties X which are related to θ via a petrophysical relationship. The latter may encompass various sources of uncertainty, as indicated in (b).

1.2.3 Hyperparameters

Geological and physical heterogeneity exist across various scales. When it comes to numerical forward solvers, a commonly employed approach is nonetheless to assume a “Representative Elementary Volume” (REV; Hill 1963). The REV represents a scale at which smaller-scale heterogeneity is averaged out and, assuming a unique upscaled value to it, offers a practical yet imperfect representation of reality, potentially resulting in misleading predictions of subsurface properties (Day-Lewis et al. 2017; Jougnot et al. 2018; Shakas and Linde 2015). The impact of a limited resolution in numerical simulations depends greatly on the specific physics involved. For certain applications, such as predicting gravimetric measurements or groundwater levels, different resolutions may not significantly alter the accuracy of the results. However, for simulations that incorporate more complex processes, such as, full-waveform modeling or tracer transport, the limitations imposed by a limited resolution can have a significant impact on the reliability and accuracy of the predictions (Linde et al. 2017).

In the context of Bayesian inversion problems, we are interested in estimating geostatistical hyperparameters for random fields representing hydrogeological or geophysical properties of the model domain. These hyperparameters include important statistical quantities as the mean, standard deviation and integral scales, that are crucial for describing the underlying random field. More specifically, we aim to estimate the posterior PDF of the hyperparameters using hydrogeological or geophysical measurements performed on the model domain. Although the geostatistical literature presents many studies concerning the estimation of hyperparameters using direct data, there has been notably less research on addressing the problem using indirect data. Inferring hyperparameters becomes particularly challenging in the so-called non-ergodic setting, where data averaging is performed over a scale that is smaller or similar to the scale of heterogeneity. In such cases, the data not only depend on the hyperparameters but also on the random field realization on which measurements are performed. This means that variations in the magnitudes and locations of high and low property values result in distinct data responses. In the context of geostatistics, a non-ergodic setting implies that there exists no analytical relationships linking the data to the upscaled hyperparameters (Fig. 1.3).

One way around this problem is to focus on the inference of both the hyperparameters and the local properties of the field (e.g. Kitanidis 1995; Hansen et al. 2012; Hansen et al. 2013a Hansen et al. 2013b; Zhao and Luo 2021; Wang et al. 2022), rather than solely focusing on inferring the hyperparameters. However, if the focus mainly lies in the estimation of the hyperparameters, a full inversion along with the local properties of the field (typically involving many thousands of unknowns) brings unnecessary computational demands. On the other hand, making simplified assumptions about the field (e.g., ergodicity or homogeneity) leads to errors in the estimates of the hyperparameters (e.g., Visentini et al. 2020; Shakas et al. 2018). To address these challenges, we treat the local property field as latent variables. This means that we solely target the hyperparameters while considering the underlying local variations as unobservable variables. Once again, this approach necessitates estimating the intractable likelihood function of the data given the hyperparameters. To accomplish this, we employ the correlated pseudo-marginal method (Section 1.4, Chapter 4).

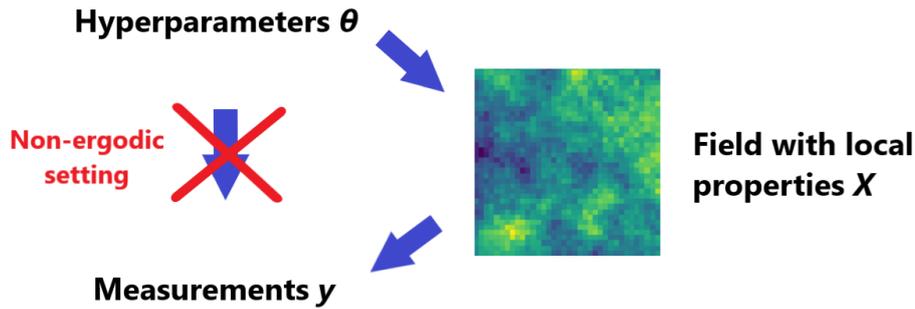


Figure 1.3: Illustration of hyperparameter estimation: In the non-ergodic setting, the hyperparameters θ cannot directly be related to the measurements y , but only via the local properties of the field X .

1.2.4 Risk assessment

When making decisions involving the geosciences, but also in other domains such as finance or engineering, it is common to identify and minimize risks by acknowledging potential hazards. Although risk assessment also considers the consequences of a hazard, the primary focus of this thesis is on estimating the probability of the hazard occurrence. Often, hazards tend to manifest as infrequent or rare events. In engineering, the evaluation of rare events is frequently associated with reliability analysis, which is concerned with the failure of systems that are designed to be highly reliable (Beck and Zuev 2015). In the domain of hydrogeology, risk assessment plays a critical role in the context of groundwater management. For instance, by determining the probability of hazardous contaminants reaching a groundwater well, we contribute in assessing the potential risk of groundwater contamination (e.g, Winter and Tartakovsky 2008, Fig. 1.4). For such real life applications, rare event estimation cannot rely on analytical formulas and specialized estimation approaches are required.

A considerable amount of research has been devoted to rare event estimation, including asymptotic approximation techniques such as first-order reliability methods (e.g., Hasofer and Lind 1974), extreme value theory (e.g., Brodin and Klüppelberg 2008), failure analysis techniques (e.g., Tartakovsky 2007) and stochastic sampling approaches. In the context of rare event estimation, conventional Monte Carlo simulation methods are often computationally inefficient (Lahkim and Garcia 1999), making them unsuitable for accurate estimation. As a result, more efficient sampling methods are required to overcome these limitations and obtain reliable estimates of rare event probabilities. These methods aim to reduce the computational burden by selectively sampling regions or scenarios where rare events are more likely to occur. By focusing computational efforts on these critical regions, the efficiency of the estimation process can be significantly enhanced. There are several reviews and books available that provide coverage of rare event simulation and associated Monte Carlo techniques (e.g., Bucklew 2004; Rubino and Tuffin 2009).

In this thesis, we are interested in rare event estimation in the context of an inversion problem. Instead of being interested in the posterior of the model parameters itself, we target the distribution of a quantity of interest depending on the field through a non-linear relationship $\boldsymbol{\theta} \mapsto \mathcal{R}(\boldsymbol{\theta})$. More specifically, we want to calculate the probability of exceeding a critical threshold for this quantity, denoted as $\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq T | \mathbf{y})$, which is potentially a rare event. In the scenario depicted in Figure 1.4, our objective could be to estimate the probability of a contaminant reaching the groundwater well on the right within a specified critical time frame with the transport time being determined by the hydrological characteristics $\boldsymbol{\theta}$ of the subsurface. However, our knowledge of these hydrological properties is limited and we rely on indirect information obtained from geophysical and hydrological measurements. In the field of structural reliability engineering, similar problems have been addressed by utilizing data to update the parameter knowledge and subsequently applying these updated distributions to predict rare events (e.g., Jensen et al. 2013; Sundar and Manohar 2013; Hadjidoukas et al. 2015).

In practical applications, it is common for the estimated posterior distribution to be available only in the form of a sample. Subset sampling is a method that allows for rare event estimation with a sample approximation as starting point (Straub et al. 2016). This method is based on the concept that the probability of a rare event can be expressed as the product of higher conditional probabilities, employing intermediate failure events referred to as subsets. Subset sampling has been employed in an inversion context for engineering problems to estimate the "updated robust failure probability" by Jensen et al. (2013) and Hadjidoukas et al. (2015). Also in the engineering literature, Straub (2011) introduced Bayesian Updating with Structural reliability methods for posterior inference (BUS; e.g. Straub and Papaioannou 2015). BUS can be viewed as an extension of the Monte Carlo simulation technique called rejection sampling, addressing its inefficiency by specifically targeting the acceptance event as a rare event through the application of structural reliability methods. Another approach that combines inversion and risk assessment is Bayesian Evidential Learning (BEL; Hermans et al. 2016). BEL is designed to establish the relationship between measurements and the quantity of interest by prior sampling of the model parameters (e.g., Thibaut et al. 2021). Within this thesis, we rely on a Sequential Monte Carlo approach based on subset sampling (Section 1.5.1 and Chapter 5) and a novel energy-based model approach (Section 1.5.2 and Chapter 6) to address rare event estimation within an inversion framework.

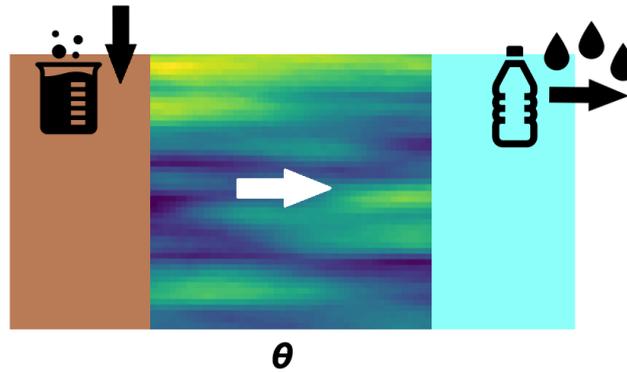


Figure 1.4: Illustration of risk assessment in a scenario concerned with potential groundwater contamination due to a hazardous substance reaching a groundwater well.

1.3 Posterior sampling approaches

In practice, the posterior PDF (as defined in Eq. 1.2) typically does not have an analytical form. Consequently, Monte Carlo sampling techniques are employed to approximate the posterior (Mosegaard and Tarantola 1995). The resulting computational cost of the Bayesian approach is widely acknowledged, particularly in scenarios involving large datasets, complex models and a high number of parameters to be estimated. Early applications of Monte Carlo sampling for geophysical inversion problems can be found in the work of Press (1968). The simplest sampling approach involves an exhaustive search of the model parameter space, evaluating the samples using a likelihood function to assess their agreement with the available measurements. A method based on this principle is rejection sampling (Ripley 2009a), applied in various instances including the work by Dorn et al. (2013). While such a brute-force approach may be feasible for a low-dimensional parameter space, it becomes computationally infeasible in a setting with thousands of model parameters (Mosegaard and Tarantola 1995). One strategy around this is to guide the sampling process using a random walk through the model space.

1.3.1 Markov chain Monte Carlo

Markov chain Monte Carlo (MCMC) is a computational method used to sample from the posterior distribution in Bayesian analysis. In MCMC, a Markov chain is constructed to iteratively generate samples from the model parameter space. As the number of samples increases, the distribution gradually converges towards the posterior distribution (e.g., Robert et al. 1999). The principle of MCMC is illustrated in Figure 1.5a. One foundational MCMC algorithm is Metropolis–Hastings (MH; Metropolis et al. 1953; Hastings 1970). This algorithm operates by iteratively proposing new sets of model parameters. The acceptance or rejection of proposed states is guided by a criterion called the acceptance ratio. This ratio is calculated by dividing the likelihood and prior values of the proposed state by the likelihood and prior values of the current state. It also incorporates a term representing the proposal distribution

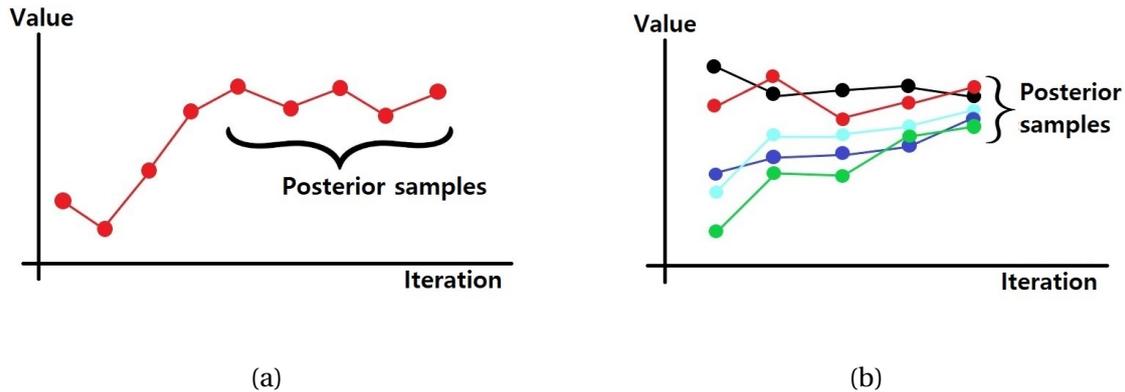


Figure 1.5: Illustrations of (a) the Markov chain Monte Carlo method and (b) the Sequential Monte Carlo method for posterior sampling.

used to generate the new state. If the acceptance ratio is high, it is likely that the proposed state is accepted and the Markov chain moves to this new state. However, if the ratio is low, the probability of rejecting the proposed parameters and remaining in the current state is high (e.g., Tarantola 2005).

The proposal scheme used in the MH algorithm is crucial as it must carefully balance the trade-off between exploration, which ensures a thorough search of the parameter space, and exploitation, which allows the algorithm to make use of promising regions while also avoid getting trapped in local optima (e.g., Tarantola 2005). To deal with this challenge, researchers have proposed various approaches based on adaptive proposal schemes. Generally, these approaches can be categorized into single-chain and multiple-chain methods (Vrugt 2016). One of the most commonly used single-chain methods is adaptive Metropolis operating by continuously adapting the covariance matrix of a Gaussian proposal distribution using the accepted samples of the chain (Haario et al. 2001). The use of a multivariate normal proposal distribution with an adaptive covariance matrix is effective for Gaussian-shaped target distributions. However, its exploration capabilities may be limited for multi-modal distributions (Vrugt 2016). When using a single-chain method, it can be particularly challenging to determine when convergence has been achieved. Thereby, single-chain methods face similar problems as local optimizers, as they cannot guarantee that the full parameter space has been adequately explored (Gelman and Shirley 2011). On the other hand, multiple-chain methods involve running different trajectories in parallel to explore the posterior target distribution. Using multiple chains provides protection against declaring convergence too early, and it also allows for the application of a variety of statistical tests to determine if the chains have reached convergence to a stationary distribution (Gelman and Rubin 1992). Braak (2006) introduced an adaptive Random Walk Metropolis algorithm known as Differential Evolution Markov Chain (DE-MC). It relies on multivariate proposals automatically adapting their scale and the orientation along the way to the stationary distribution. A refined version of DE-MC has been proposed with the adaptive multi-chain algorithm DREAM (DiffeREntial Evolution Adaptive Metropolis; Vrugt et al. 2008). It improves the efficiency of DE-MC by incorporating subspace sampling, where only randomly selected dimensions of the model parameter are

updated during each iteration. Additionally, outlier chain correction techniques are employed to address potential issues with chains that deviate significantly from the rest.

The curse of dimensionality is a general problem for MCMC algorithms, as the time needed for convergence increases as the number of target parameters grows (e.g., Robert et al. 2018). This can be counteracted by various approaches, the most common is to reduce the parameter space by employing dimensionality reduction techniques (Hastie et al. 2009). Generally, reducing the dimension of the model improves the computational efficiency of the inversion algorithm (Laloy et al. 2015). However, the diminished dimensionality of the problem also causes a loss of information, which can lead to an inaccurate estimation of the uncertainty (Grana et al. 2019). If for this reason dimensionality reduction techniques are omitted, it is well-advised to use model proposal schemes that preserve the prior PDF (such as preconditioned Crank–Nicolson, pCN, proposals; e.g. Cotter et al. 2013). They make the algorithm robust to the choice of discretization and enable to maintain a reasonable step size when inferring thousands of unknowns. Inversion with prior-preserving model proposals are known as extended Metropolis algorithms in geophysics (Mosegaard and Tarantola 1995).

In this thesis, we explore and evaluate several model proposal schemes. It is important to note that neither the proposal schemes we discuss, nor those presented above are exhaustive. In Chapters 2 and 3 targeting very high-dimensional parameter spaces (thousands of unknowns), we employ prior-sampling $\text{DREAM}_{(ZS)}$, a version of DREAM relying on an archive of past states (Laloy and Vrugt 2012) that is combined with a prior-preserving modification. In Chapter 4, when inferring only a limited number of hyperparameters, we employ the adaptive Metropolis algorithm introduced by Haario et al. (2001). Finally, in Chapters 5 and 6, we use standard Gaussian and pCN proposals (Cotter et al. 2013).

1.3.2 Sequential Monte Carlo

MCMC encounters challenges when employed to solve high-dimensional problems and these issues become more pronounced as the non-linearity of the forward problem increases. These difficulties can lead to insufficient exploration of the posterior distribution, manifested by Markov chains becoming trapped in local minima or being unable to transition between modes of high posterior probability (e.g., Amaya et al. 2022). Particle methods such as the Sequential Monte Carlo method offer one way around those problems.

Tempering consists in introducing a temperature variable that modifies the likelihood function (Eq. 1.2). The posterior PDFs that arise from the tempering approach are referred to as power posteriors, whereby the exponent of the likelihood corresponds to the inverse temperature. Increasing the temperature has the effect of flattening the likelihood function, which simplifies the exploration of the posterior by reducing the probability of getting trapped in local minima and facilitating the sampling of different modes. One well-known approach relying on tempering is so-called parallel tempering (Earl and Deem 2005). It involves running multiple Markov chains in parallel, each associated with a different temperature. The chains are allowed to exchange states at specific intervals, with higher temperature chains

exploring a broader space and lower temperature chains emphasizing the region of high likelihoods. While parallel tempering utilizes different temperatures for different chains, Annealed Importance Sampling (Neal 2001) gradually transforms the prior distribution to the target distribution by incrementally decreasing the temperature within one chain. In Annealed Importance Sampling, the particle approximation of the posterior is obtained through a series of importance sampling steps conducted between each pair of consecutive power posteriors. The quality of the importance sampling estimator is strongly influenced by the variance of the particle weights (Neal 2001). In Annealed Importance Sampling, this variance can grow exponentially, resulting in poor approximations of the posterior. The Sequential Monte Carlo method (SMC; Doucet et al. 2001) is a particle method also relying on sequential importance sampling. However, SMC methods go a step further by incorporating resampling steps, in which the states of the particles are resampled according to their current normalized importance weights (Del Moral et al. 2007). This resampling step helps to mitigate the growing variance in the weights and ensures a more balanced representation of the posterior among the particles.

Although the SMC method (illustrated in Figure 1.5b) is commonly employed in various fields of science and engineering, its application in geosciences has been limited (Linde et al. 2017). Recently, adaptive SMC methods have shown promising results in addressing inversion problems in geophysics (Amaya et al. 2021, Davies et al. 2023) and hydrogeology (Amaya et al. 2022). Following these studies, we apply SMC for a one-dimensional flow and a two-dimensional transport problem in Chapter 5. Thereby, we use its sequential nature not only to approximate the posterior, but also to address a rare event probability estimation problem.

1.4 Likelihood estimation

In the setups introduced in Sections 1.2.2 (petrophysical uncertainty) and 1.2.3 (hyperparameter estimation), we are concerned with a latent variable model (also referred to as random effects model) and a resulting intractable likelihood function. In this section, we introduce the corresponding notation and present the two approaches we employ to tackle the challenges associated with it.

1.4.1 Latent variable model

We consider the latent variable model (Deligiannidis et al. 2018),

$$\mathbf{X} \sim f_{\boldsymbol{\theta}}(\cdot) \quad \mathbf{Y}|\mathbf{X} \sim g_{\boldsymbol{\theta}}(\cdot|\mathbf{X}), \quad (1.3)$$

with \mathbf{X} denoting the latent variables and \mathbf{Y} the observations. This represents a setup, where the model parameters $\boldsymbol{\theta}$ are only indirectly related to the measurements \mathbf{Y} via the variables \mathbf{X} . This notation from Deligiannidis et al. (2018) with subscripts indicates that the distributions are conditioned on the random variable $\boldsymbol{\theta}$ being equal to the instance $\boldsymbol{\theta}$, whereby we do not distinguish between instance and random variable. For the sake of simplicity, we follow the

convention of Deligiannidis et al. (2018) and do not further specify the conditioning on $\boldsymbol{\theta}$ in the notation above. While we use this formulation with subscripts in Chapter 4, we denote $f_{\boldsymbol{\theta}}(\mathbf{x})$ by $p(\mathbf{x}|\boldsymbol{\theta})$ and $g_{\boldsymbol{\theta}}(\mathbf{y}|\mathbf{x})$ by $p(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta})$ in Chapters 2 and 3.

In a latent variable model, the likelihood function of the observed data \mathbf{y} given the target model parameters $\boldsymbol{\theta}$ is defined as,

$$\boldsymbol{\theta} \mapsto p(\mathbf{y}|\boldsymbol{\theta}) = \int g_{\boldsymbol{\theta}}(\mathbf{y}|\mathbf{x}) f_{\boldsymbol{\theta}}(\mathbf{x}) d\mathbf{x}, \quad (1.4)$$

and this integral is generally intractable. That is, the integral has an unknown or non-existing analytical form, which makes the direct implementation of posterior sampling algorithms impossible.

In latent variable problems concerned with petrophysical uncertainty (Section 1.2.2, Chapters 2 and 3), the target properties $\boldsymbol{\theta}$ are the (hydro)geological parameters and \mathbf{X} are the geophysical properties $\mathcal{F}(\boldsymbol{\theta}) + \boldsymbol{\varepsilon}_{\mathcal{F}}$ derived with the petrophysical relationship $\boldsymbol{\theta} \mapsto \mathcal{F}(\boldsymbol{\theta})$ including the (unobservable) petrophysical prediction error $\boldsymbol{\varepsilon}_{\mathcal{F}}$. For hyperparameter estimation (Section 1.2.3 and Chapter 4), we utilize $\boldsymbol{\theta}$ representing the target hyperparameters and \mathbf{X} the local properties of the model domain. Thereby, $\mathbf{X} \sim f_{\boldsymbol{\theta}}(\cdot)$ is a (discretized) random field with its geostatistical distribution $f_{\boldsymbol{\theta}}(\cdot)$ depending on the hyperparameters $\boldsymbol{\theta}$.

1.4.2 Correlated pseudo-marginal method

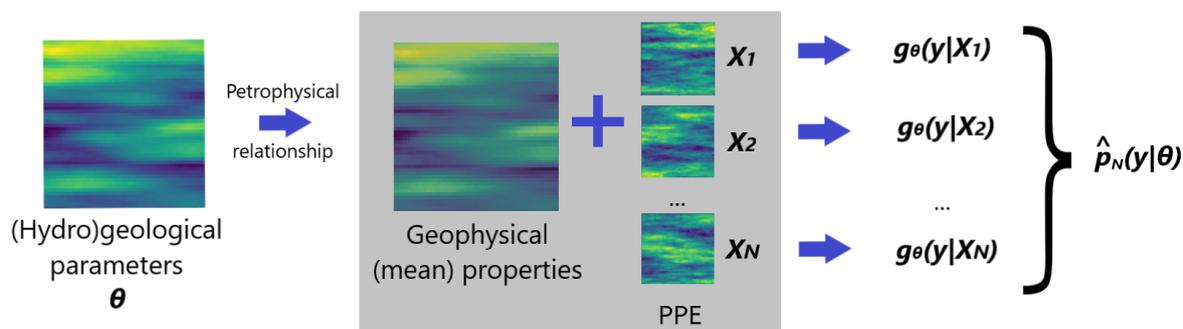
The correlated pseudo-marginal method is one approach to approximate the intractable likelihood function of latent variable models (Eq. 1.4). It is based on the pseudo-marginal approach introduced by Beaumont (2003) and studied by Andrieu and Roberts (2009). In a nutshell, it replaces the likelihood with a non-negative unbiased estimator based on Monte Carlo averaging over samples of the latent variable,

$$\hat{p}_N(\mathbf{y}|\boldsymbol{\theta}) = \frac{1}{N} \sum_{n=1}^N g_{\boldsymbol{\theta}}(\mathbf{y}|\mathbf{X}_n), \quad \mathbf{X}_n \stackrel{i.i.d.}{\sim} f_{\boldsymbol{\theta}}(\cdot). \quad (1.5)$$

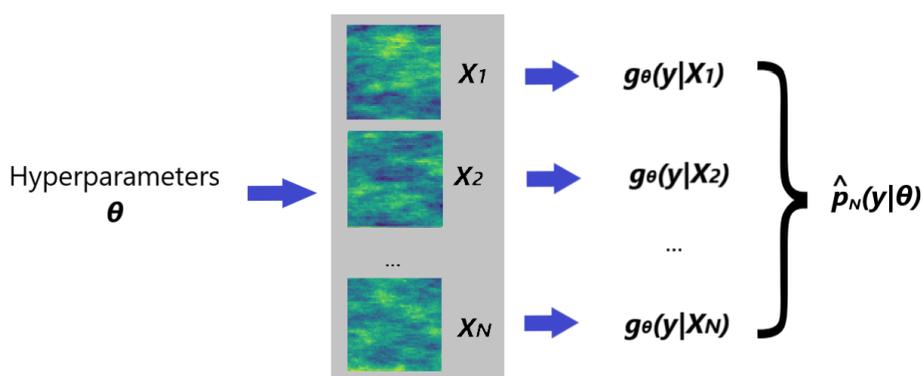
The replacement of the likelihood by such an estimator results in a MH algorithm sampling the same posterior distribution as when the true likelihood is used (Beaumont 2003). The efficiency of the pseudo-marginal method depends on the variance of the log-likelihood estimator (Doucet et al. 2015). If it is too high, the algorithm will encounter an unreasonably low acceptance rate. The variance of the estimator can be regulated by adjusting the number of latent variable samples N . However, this comes at the expense of higher computational costs as for each sample, a forward simulation is required to calculate $g_{\boldsymbol{\theta}}(\mathbf{y}|\mathbf{X}_n)$. Therefore, and especially when dealing with high-dimensional problems that involve large data sets exhibiting high signal-to-noise ratios, the utilization of importance sampling (Owen and Zhou 2000) becomes crucial when drawing samples of latent variables \mathbf{X}_n . To reduce the required number of samples while ensuring a low variance, Deligiannidis et al. (2018) introduced the correlated pseudo-marginal (CPM) correlating the draws of consecutive latent variables. The

underlying concept here is that the variance of a ratio of estimators (here the acceptance ratio of the MH) is lower when the denominator and numerator are positively correlated (Koop 1972).

In Friedli et al. (2022) (Chapter 2), we introduce and adapt the CPM method to a geophysical inversion problem involving petrophysical uncertainty. Subsequently in Friedli and Linde (2023) (Chapter 3), we compare the CPM method's performance against a linearized Gaussian approximation approach. Finally, in Friedli et al. (2023) (Chapter 4), we assess the performance of the CPM method in a setting targeting hyperparameters. In Figures 1.6a and 1.6b, illustrations of Equation (1.5) are depicted for inversion problems concerned with petrophysical uncertainty and hyperparameter estimation, respectively.



(a)



(b)

Figure 1.6: Illustration of the pseudo-marginal method for inversion problems concerned with (a) petrophysical prediction errors (PPE) and (b) hyperparameter estimation.

1.4.3 Linearized Gaussian approximation

Apart from the CPM method, we apply a novel approach to estimate the intractable likelihood function (Eq. 1.4), which is based on a Gaussian approximation and local linearization. It relies on the assumption that the latent variable follows a Gaussian distribution $\mathbf{X} \sim f_{\boldsymbol{\theta}}(\mathbf{x}) = \varphi(\mathbf{x}; \boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x)$, such that we can write $\mathbf{X} = \boldsymbol{\mu}_x + \boldsymbol{\Sigma}_x^{1/2}\mathbf{Z}$, with \mathbf{Z} denoting a random vector consisting of *i.i.d.* standard normal distributed variables (same dimension as \mathbf{X}). Then, a first-order Taylor expansion of $\mathbf{x} \mapsto \mathcal{G}(\mathbf{x})$ around $\boldsymbol{\mu}_x$ is employed,

$$\mathbf{Y} = \mathcal{G}(\boldsymbol{\mu}_x + \boldsymbol{\Sigma}_x^{1/2}\mathbf{Z}) + \boldsymbol{\varepsilon}_{\theta} \approx \mathcal{G}(\boldsymbol{\mu}_x) + \mathbf{J}_{\boldsymbol{\mu}_x} \boldsymbol{\Sigma}_x^{1/2}\mathbf{Z} + \boldsymbol{\varepsilon}_{\theta}, \quad (1.6)$$

with $\mathbf{J}_{\boldsymbol{\mu}_x}$ denoting the Jacobian (sensitivity) matrix of the forward solver corresponding to $\boldsymbol{\mu}_x$. Under Gaussian assumptions for the observational noise $p(\boldsymbol{\varepsilon}_{\theta}) = \varphi(\boldsymbol{\varepsilon}_{\theta}; \mathbf{0}, \boldsymbol{\Sigma}_Y)$, the likelihood function can subsequently be approximated by,

$$\hat{p}(\mathbf{y}|\boldsymbol{\theta}) = \varphi(\mathbf{y}; \boldsymbol{\mu}_Y, \widetilde{\boldsymbol{\Sigma}}_Y) \quad \text{with} \quad \boldsymbol{\mu}_Y = \mathcal{G}(\boldsymbol{\mu}_x) \quad \text{and} \quad \widetilde{\boldsymbol{\Sigma}}_Y = \mathbf{J}_{\boldsymbol{\mu}_x}^T \boldsymbol{\Sigma}_x \mathbf{J}_{\boldsymbol{\mu}_x} + \boldsymbol{\Sigma}_Y. \quad (1.7)$$

As the linearization is made around $\boldsymbol{\mu}_x$ and not around $\boldsymbol{\mu}_x + \boldsymbol{\Sigma}_x^{1/2}\mathbf{Z}$, errors arise when the resulting Jacobians differ. For a linear geophysical problem, there are no such approximation errors.

This novel approach initially proposed by Linde et al. (2017), which shares similarities with the delta method (Van der Vaart 2000), was first implemented in Friedli and Linde (2023) for a comparative analysis with the CPM method (Chapter 3). In this setting concerned with Gaussian petrophysical prediction errors, the approximation works well for moderate non-linearity in the physical forward solver and a low degree of petrophysical prediction uncertainty. For the latent variable model in a hyperparameter estimation setting (Chapter 4), the approach could be adapted if the local properties of \mathbf{X} are described by a Gaussian random field. However, the disparity between the two Jacobians can be substantial, particularly in non-ergodic settings. This happens as all the heterogeneity of the latent variable is included in the random part $\boldsymbol{\Sigma}_x^{1/2}\mathbf{Z}$, while the mean $\boldsymbol{\mu}_x$ is a constant field. Due to this limitation, we did not test the linearized Gaussian approach in the context of hyperparameter inference.

1.5 Rare event estimation

In the scenario considered in Section 1.2.4, we are not primarily interested in the model parameters' posterior distribution, but in a resulting quantity of interest $\boldsymbol{\theta} \mapsto \mathcal{R}(\boldsymbol{\theta})$, particularly in the rare event that this quantity exceeds a critical threshold, $\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq T | \mathbf{y})$. To tackle the inefficiency of traditional Monte Carlo sampling in estimating low probabilities, we explore two alternative approaches: the Sequential Monte Carlo method (Section 1.5.1, Chapter 5) and an energy-based model approach (Section 1.5.2, Chapter 6).

1.5.1 Sequential Monte Carlo

As other approaches targeting rare events, Sequential Monte Carlo is relying on the principle of subset sampling (Au and Beck 2001). The underlying idea of subset sampling is to express a small probability as a product of not so small conditional probabilities,

$$\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq T | \mathbf{y}) = \prod_{k=1}^K \mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq T_k | \mathcal{R}(\boldsymbol{\theta}) \geq T_{k-1} | \mathbf{y}), \quad (1.8)$$

with the increasing sequence of thresholds $\{T_0, \dots, T_K\}$ where $T_0 = -\infty$ and $T_K = T$ (C erou et al. 2012). When applying the SMC method within an inversion setting, we start the rare event probability estimation process with a particle approximation of the posterior, which is then transferred towards the critical value of $\mathcal{R}(\boldsymbol{\theta})$. In each step of the process, the particles exhibiting the lowest $\mathcal{R}(\boldsymbol{\theta})$ are filtered out, while the remaining particles are resampled and subsequently propagated within the posterior, with the additional requirement that $\mathcal{R}(\boldsymbol{\theta}) \geq T_k$. This procedure ensures that only particles meeting the threshold condition continue to contribute to the subsequent iterations. An illustration of this process is provided in Figure 1.7.

The performance of the SMC method for rare event estimation purposes is influenced by the choice of intermediate thresholds. If the thresholds increase slowly, the conditional probabilities tend to be large, requiring fewer particles for their estimation. Conversely, if the thresholds increase rapidly, more particles are needed to achieve accurate results. Consequently, a well-working trade-off between the conditional probabilities and the number of particles has to be found (Au and Beck 2001). To address this challenge, C erou et al. (2012) propose the use of an adaptive sequence of thresholds based on quantiles. However, employing adaptive thresholds comes at a slight cost of accuracy, as it introduces a small positive bias in the rare event probability estimate (C erou et al. 2012). A possible solution to address this issue is to re-run the algorithm using the optimized sequence that has been identified.

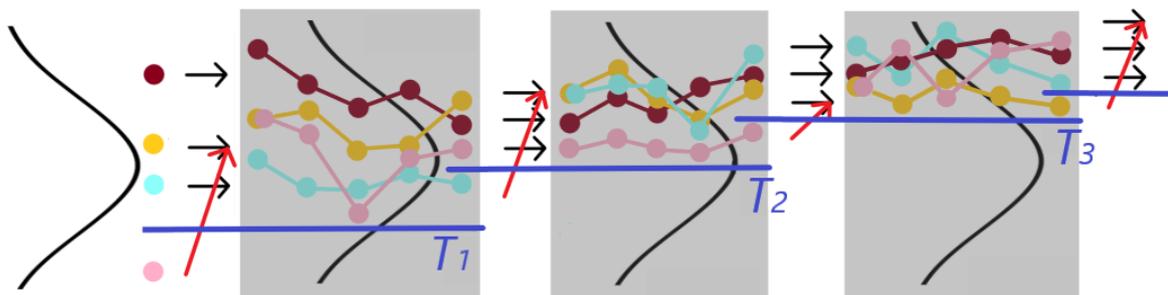


Figure 1.7: Illustration of the SMC method for rare events. We depict the first three thresholds for an example with four particles and four steps to propagate the particles within the subsets.

1.5.2 Energy-based models

As a second approach to address rare event estimation, we consider energy-based models (EBM; e.g., Lelièvre et al. 2010) and write the marginal posterior distribution of the quantity of interest $R = \mathcal{R}(\boldsymbol{\theta})$ as an energy density function with free energy $r \mapsto F(r)$,

$$p_{R|Y}(r|\mathbf{y}) \propto \exp(-F(r)). \quad (1.9)$$

To estimate $F(r)$ and $p_{R|Y}(r|\mathbf{y})$, one approach is to utilize transformed posterior samples. However, since this follows a typical Monte Carlo approach, it would require an impractically large number of samples to accurately cover the low probability regions of interest. To address this issue and enhance sampling in the region of interest $r \geq T$, we aim to sample according to a predefined distribution $p_{\text{ref}}(r)$ that assigns most of its mass to this specific region. To achieve this, we introduce the bias potential $r \mapsto V(r)$ and the corresponding PDF $r \mapsto p_V(r)$,

$$p_V(r) = \frac{\exp(-(F(r) + V(r)))}{\int \exp(-(F(s) + V(s))) ds}. \quad (1.10)$$

Our approach relies on optimizing the potential $V(r)$ such that $p_V(r)$ is equal to $p_{\text{ref}}(r)$. It is straightforward that if we find $V_{\text{opt}}(r)$, we can obtain $F(r) = -\log(p_{\text{ref}}(r)) - V_{\text{opt}}(r)$, and hence the marginal posterior distribution of the quantity of interest.

This approach illustrated in Figure 1.8 draws upon the variational method proposed by Valsson and Parrinello (2014) in metadynamics. However, although we adopt the fundamental concepts of their method, we present our approach using a distinct formulation and apply it to a novel context targeting rare events. In practice, the main challenges of the EBM approach are the parameterization of the bias potential $V(r)$, the choice of $p_{\text{ref}}(r)$ and the optimization method used to estimate $V_{\text{opt}}(r)$.

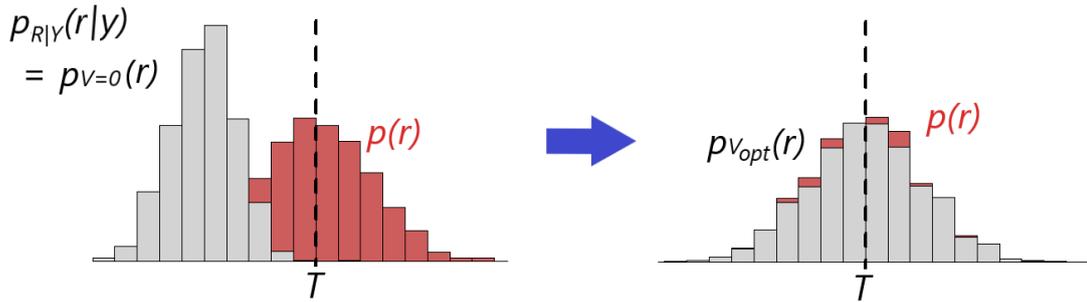


Figure 1.8: Illustration of the energy-based model approach.

1.6 Objectives

The constant progress in sensor technologies, field approaches, and numerical modeling has resulted in significant breakthroughs in addressing inversion challenges within the field of geosciences (Linde et al. 2017). However, the outcomes and predictions still contain uncertainties that need to be addressed. This thesis presents a selection of methodological advancements and their pioneering applications in the field of geoscientific inversion. Thereby, we address the following two broad research questions:

- How can statistical estimation and approximation approaches be used to obtain efficient and accurate probabilistic inversion methods in scenarios that involve intractable likelihood functions? Or more specifically, how can they handle the challenges of petrophysical uncertainty and hyperparameter estimation in non-linear inversion problems? (Chapters 2, 3 and 4)
- How can modern statistical methods be utilized to integrate inversion and risk assessment in order to address problems where the primary focus is on the probability of a rare event of interest, rather than the posterior distribution? (Chapters 5 and 6)

The implementation of the MH and most other MCMC algorithms relies on the computation of the likelihood function $p(\mathbf{y}|\boldsymbol{\theta})$. In a latent variable model, where no direct relationship between the model parameters $\boldsymbol{\theta}$ and the measurements \mathbf{y} exists, this likelihood function is typically intractable (Section 1.4.1). In this context, we present two methods addressing the estimation of the likelihood function: the correlated pseudo-marginal method (Section 1.4.2) and a linearized Gaussian approach (Section 1.4.3). We are concerned with latent variable models in two contexts: petrophysical uncertainty (Section 1.2.2) and hyperparameter estimation (Section 1.2.3).

Effective decision making in geosciences heavily relies on the assessment of potential risks. When confronted with uncertain systems, it becomes crucial to adequately propagate the quantification of uncertainty into the risk assessment (Section 1.2.4). This enables a comprehensive understanding of the potential consequences and guides decision makers towards making informed choices. Rare events pose challenges in risk assessment and traditional statistical approaches often fail to capture tails of the uncertainty distributions. To address this, specialized modeling techniques for rare event simulation are employed. Within this context, we present two approaches addressing rare event estimation under posterior uncertainty: The PostRisk-SMC method (Section 1.5.1) and a novel energy-based model approach (Section 1.5.2).

1.7 Outline

This thesis is divided into two main parts with each focusing on different aspects of inverse problem settings: likelihood estimation (Chapters 2, 3 and 4) and risk assessment (Chapters 5 and 6). Thereby, this thesis includes work that has been published in peer-reviewed journals (Chapters 2, 3 and 4) or manuscripts that will soon be submitted for publication (Chapters 5 and 6). I took main responsibility for the entire process of each paper, encompassing the development of the methodology, code implementation and authorship of the initial drafts.

Chapter 2 (Friedli et al. 2022) investigates lithological tomography, which involves inferring (hydro)geological parameters from geophysical data by treating intermediate geophysical properties as latent variables. We employ the correlated pseudo-marginal method to estimate the resulting intractable likelihood function. As test cases, we consider the inference of porosity fields using crosshole ground-penetrating radar first-arrival travel times in a high-dimensional parameter space. The correlated pseudo-marginal method, combined with importance sampling and a prior-preserving proposal scheme, is found to outperform current state-of-the-art methods by greatly enhancing posterior exploration in both linear and non-linear settings.

Chapter 3 (Friedli and Linde 2023) is a natural extension of Chapter 2 and compares the correlated pseudo-marginal method against a new approximate likelihood estimation method. This new approach utilizes a Gaussian probability density function based on local linearization of the geophysical forward operator. First, the performances of both methods are tested on the same non-linear test case as in Chapter 2. In this specific case, both methods produce nearly identical estimates as the petrophysical uncertainty is relatively low. Based on the findings of a subsequent sensitivity analysis, it is indicated that the linearized Gaussian approach, despite its computational advantages, suffers from a decrease in accuracy if the scatter in the petrophysical relationship grows. In contrast, the computationally more expensive correlated pseudo-marginal method produces accurate estimates even for settings with high petrophysical uncertainty.

Chapter 4 (Friedli et al. 2023) focuses on the application of the correlated pseudo-marginal method to another setting involving latent variables. Here, the addressed problem involves estimating the geostatistical hyperparameters of a random field of interest. This problem is particularly interesting in a non-ergodic setting, where the absence of analytical relationships between the data and the hyperparameters makes it necessary to consider the local properties of the field. In our approach, we treat the random field as latent variables and estimate the intractable likelihood of observing the data given the hyperparameters using the correlated pseudo-marginal method. We evaluate its performance in two representative inversion problems: The first data-poor problem involves inferring the hyperparameters of hydraulic conductivity fields using equivalent hydraulic conductivity data. The second data-rich problem focuses on estimating fracture aperture fields using borehole ground-penetrating radar reflection data. We find that in comparison to simplified model assumptions, the correlated pseudo-marginal method enables accurate hyperparameter estimation.

In Chapter 5, the focus shifts to risk assessment and rare events, where the primary interest lies in a critical quantity that depends on the unknown model parameters θ . First, we switch the posterior sampling method from MCMC to SMC. However, relying only on a particle approximation of the posterior PDF is unsuitable when determining rare event probabilities. To address this, we apply the SMC method a second time to perform subset sampling. The resulting PostRisk-SMC method is tested in both a one-dimensional flow and a two-dimensional transport example. Through the one-dimensional example, it becomes evident that the PostRisk-SMC method allows us to estimate rare event probabilities as low as one in a billion. The two-dimensional example showcases the method's capability for rare event probability estimation in a more realistic and complex setting. Both examples display how the PostRisk-SMC approach enables an accurate assessment and simulation of very rare events while also providing a particle-based approximation of the posterior distribution.

Also Chapter 6 is concerned with rare event estimation. In this study, we use an energy density function to represent the quantity of interest's posterior distribution, characterized by a free energy function. To enable efficient estimation of the free energy, we employ concepts from energy-based models and introduce and estimate a bias potential. Through three illustrative test examples, we demonstrate that the EBM approach, when properly configured, provides stable rare event probability estimates even in scenarios where the probability of occurrence is less than one in a million. Furthermore, we demonstrate that the EBM approach is applicable for both rare event estimation scenarios involving the prior or the posterior distribution.

Finally, Chapter 7 summarizes the conclusions drawn from this thesis and provides an overview of potential future research directions based on the discoveries made.

Chapter 2

Lithological tomography with the correlated pseudo-marginal method

Lea Friedli, Niklas Linde, David Ginsbourger and Arnaud Doucet

Published in *Geophysical Journal International*:

Friedli, L., Linde, N., Ginsbourger, D., and Doucet, A. (2022). Lithological tomography with the correlated pseudo-marginal method. *Geophysical Journal International*, 228(2):839–856.

Abstract

We consider lithological tomography in which the posterior distribution of (hydro)geological parameters of interest is inferred from geophysical data by treating the intermediate geophysical properties as latent variables. In such a latent variable model, one needs to estimate the intractable likelihood of the (hydro)geological parameters given the geophysical data. The pseudo-marginal method is an adaptation of the Metropolis–Hastings algorithm in which an unbiased approximation of this likelihood is obtained by Monte Carlo averaging over samples from, in this setting, the noisy petrophysical relationship linking (hydro)geological and geophysical properties. To make the method practical in data-rich geophysical settings with low noise levels, we demonstrate that the Monte Carlo sampling must rely on importance sampling distributions that well approximate the posterior distribution of petrophysical scatter around the sampled (hydro)geological parameter field. To achieve a suitable acceptance rate, we rely both on (1) the correlated pseudo-marginal method, which correlates the samples used in the proposed and current states of the Markov chain, and (2) a model proposal scheme that preserves the prior distribution. As a synthetic test example, we infer porosity fields using crosshole ground-penetrating radar (GPR) first-arrival travel times. We use a (50×50) -dimensional pixel-based parameterization of the multi-Gaussian porosity field with known statistical parameters, resulting in a parameter space of high dimension. We demonstrate that the correlated pseudo-marginal method with our proposed importance sampling and prior-preserving proposal scheme outperforms current state-of-the-art methods in both linear and non-linear settings by greatly enhancing the posterior exploration.

2.1 Introduction

Geophysical investigations are rarely performed with the sole aim of inferring distributed subsurface models of geophysical properties. Rather, the underlying motivation is often to gain knowledge and constraints on other properties (e.g., permeability, clay fraction or mineral composition) and state variables (e.g., water saturation, salinity, temperature) of interest. Geophysical inverse theory has traditionally focused on assessing the resolution and uncertainty of inferred geophysical properties (e.g., Parker 1994; Menke 2018; Tarantola 2005; Aster et al. 2018), while interpretation procedures in terms of properties or state variables of interest have received less attention. This is changing in hydrogeophysics (Binley et al. 2015), for instance, where it is now well-established that dedicated inversion approaches are needed when using geophysical data to gain knowledge about hydrogeological properties and state variables (e.g., Kowalsky et al. 2005). For example, when inferring hydraulic conductivity by observing geophysical observables sensitive to water content or salinity during a tracer test experiment (Linde and Doetsch 2016). However, these considerations have general validity and relevance for exploration and more fundamental geophysical studies. In a mantle context, for instance, one example concerns the inference of thermo-chemical constraints from seismological observations as reviewed by Zunino et al. (2016).

Multiple inversion frameworks have been proposed that combine hydrogeological and geophysical data in order to build predictive hydrogeological models (e.g., Ferré et al. 2009; Linde and Doetsch 2016). A critical aspect of such frameworks relates to how geophysical properties (sensed by geophysical data) are linked to hydrogeological target properties and variables of interest through petrophysical (rock physics) relationships. Brunetti and Linde (2018) distinguish between three sources of uncertainty related to petrophysical relationships: model uncertainty, parameter uncertainty and prediction uncertainty. While the first two refer to uncertainty in the choice of the appropriate petrophysical model and its parameter values, the latter is related to scatter and bias around the calibrated petrophysical model. In hydrogeophysical inversion studies targeting hydrogeological properties or state variables of interest, we note that the petrophysical relationship is often assumed to be perfect (deterministic) with known or unknown parameter values (e.g., Lochbühler et al. 2014; Kowalsky et al. 2005). However, ignoring petrophysical prediction uncertainty and its spatial correlation patterns results in bias, too narrow uncertainty bounds and overly variable hydrogeological parameter estimates (Brunetti and Linde 2018). Unfortunately, analytical solutions to such inverse problems are available only when considering linear forward models and petrophysical relationships under the assumption of Gaussian distributions (Tarantola 2005; Bosch 2004). Geophysical applications, however, often involve non-linear physics and non-linear petrophysical relationships (e.g., Mavko et al. 2020).

Inversion approaches that account for petrophysical prediction uncertainty are often based on a two-step procedure: geophysical properties are first estimated using deterministic gradient-based inversions and then converted into parameters of interest using uncertain petrophysical relationships (e.g., Chen et al. 2001; Mukerji et al. 2001; González et al. 2008; Grana and Della Rossa 2010; Shahraeeni and Curtis 2011). The results of such a two-step approach can be misleading if neglecting the spatially-varying and typically much lower resolution of smoothness-constrained geophysical inversion models compared with the scale at which petrophysical relationships are developed (core or borehole logging scale) (Day-Lewis et al. 2005). Furthermore, with such an approach it is next to impossible to ensure that the geophysical inversion accounts for the prior constraints on the (hydro)geological target variable (Ferré et al. 2009) and physical constraints such as conservation of mass, continuity and momentum. Moreover, for a deterministic inversion setting, Bosch (2004) showed that with a non-linear petrophysical relation, the two-step approach is an inherent approximation.

As an alternative to the two-step approach, coupled inversions directly target hydrogeological properties by inversion of geophysical data (e.g., Hinnell et al. 2010; Kowalsky et al. 2005). They are often formulated within a Bayesian framework whereby one seeks to characterize the posterior probability density function (PDF) of hydrogeological parameters θ given geophysical data \mathbf{y} . Since it is often impossible to sample directly from the posterior PDF $p(\theta|\mathbf{y})$ of interest, Markov chain Monte Carlo (MCMC) methods, such as the Metropolis–Hastings method (MH; Hastings 1970; Metropolis et al. 1953), are used. Since the intermediate variable, the geophysical property \mathbf{X} , connecting observations and target variables is unobservable (latent), one speaks of a latent variable model. In this study, we consider a setup where the latent geophysical property is given by $\mathbf{X} = \mathcal{F}(\theta) + \epsilon_{\mathcal{F}}$, with $\theta \mapsto \mathcal{F}(\theta)$ representing the deterministic component of a petrophysical relationship and $\epsilon_{\mathcal{F}}$ the petrophysical prediction

error. Assuming an integrable and centered petrophysical prediction error $\epsilon_{\mathcal{P}}$, $\mathcal{F}(\boldsymbol{\theta})$ stands for the expected value of the latent variable \mathbf{X} . The geophysical data is given by $\mathbf{Y} = \mathcal{G}(\mathbf{X}) + \epsilon_{\mathcal{O}}$ with $\mathbf{x} \mapsto \mathcal{G}(\mathbf{x})$ denoting the geophysical forward solver and $\epsilon_{\mathcal{O}}$ describing the observational noise.

For a latent variable model as the one described above, the likelihood of observing the geophysical data given the proposed hydrogeological parameters, $p(\mathbf{y}|\boldsymbol{\theta}) = \int p(\mathbf{y}, \mathbf{x}|\boldsymbol{\theta}) d\mathbf{x}$, is often intractable. In the present context, this implies that the integral has an unknown or non-existing analytical form, which makes the direct implementation of the MH and related algorithms impossible. One way to circumvent this difficulty is to instead infer the joint posterior PDF $(\boldsymbol{\theta}, \mathbf{x}) \mapsto p(\boldsymbol{\theta}, \mathbf{x}|\mathbf{y})$ of the hydrogeological and geophysical parameters from which $p(\boldsymbol{\theta}|\mathbf{y})$ is readily obtained by marginalization. Lithological tomography as introduced by Bosch (1999) pioneered such an approach to estimate the joint posterior by combining geophysical data, geological prior knowledge and uncertain petrophysical relationships. Within lithological tomography, pairs of the target and latent variables are proposed using marginal sampling of $\boldsymbol{\theta}$ and conditional sampling of \mathbf{X} . Then, these pairs are accepted or rejected with $p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x})$, used in the acceptance ratio of the MH algorithm (where $p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x}) = p(\mathbf{y}|\mathbf{x})$ is valid for our latent variable model). In Bosch (1999), the conditional PDF $p(\mathbf{x}|\boldsymbol{\theta})$ to sample \mathbf{X} is given by a multivariate Gaussian distribution based on a suitable petrophysical relationship. In practice, this is achieved by adding brute force Monte Carlo realizations of the petrophysical prediction error $\epsilon_{\mathcal{P}}$ to the output of $\mathcal{F}(\boldsymbol{\theta})$ at each iteration of the MCMC chain (i.e., Bosch et al. 2007). Linde et al. (2017) suggest that such an implementation is inefficient when considering large geophysical datasets with high signal-to-noise ratios and significant petrophysical uncertainty. The reason is that brute force Monte Carlo sampling of the petrophysical prediction error using $p(\mathbf{x}|\boldsymbol{\theta})$ induces high variability in the values taken by the likelihood function $p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x})$, even for the same $\boldsymbol{\theta}$, which could lead to prohibitively low acceptance rates even in the limiting case when the MCMC model proposal scale for $\boldsymbol{\theta}$ goes to zero.

Brunetti and Linde (2018) proposed an alternative approach to sample from the joint posterior PDF $p(\boldsymbol{\theta}, \mathbf{x}|\mathbf{y})$. In their method referred to herein as full inversion, the petrophysical prediction error $\epsilon_{\mathcal{P}}$ is parameterized and treated as the other unknowns within the MH algorithm. That is, the MH proposal mechanism draws new realizations of both the target variable $\boldsymbol{\theta}$ and the petrophysical prediction error $\epsilon_{\mathcal{P}}$, which combined also lead to a realization of the latent variable \mathbf{X} used to calculate the likelihood function $p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x})$. Brunetti and Linde (2018) presented a convincing performance of the full inversion approach with clear improvements in efficiency compared with the original formulation of lithological tomography by Bosch (1999). Nonetheless, the full inversion method suffers from high dimensionality, and the strong (posterior) correlation between $\epsilon_{\mathcal{P}}$ and $\boldsymbol{\theta}$ makes standard MCMC inversions inefficient (e.g., Deligiannidis et al. 2018).

In this study, we evaluate an inversion method targeting directly the marginal posterior $p(\boldsymbol{\theta}|\mathbf{y})$ by approximating the intractable likelihood $p(\mathbf{y}|\boldsymbol{\theta}) = \int p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x}) p(\mathbf{x}|\boldsymbol{\theta}) d\mathbf{x}$. In the pseudo-marginal (PM) method introduced by Beaumont (2003) and studied by Andrieu and Roberts (2009), the true likelihood is replaced with a non-negative unbiased estimator resulting in a

MH algorithm sampling the same target distribution as when using the true likelihood. In their work, Beaumont (2003) and Andrieu and Roberts (2009) use an unbiased likelihood estimator based on Monte Carlo averaging over samples of the latent variable. In our setting with the latent variable $\mathbf{X} = \mathcal{F}(\boldsymbol{\theta}) + \boldsymbol{\varepsilon}_{\mathcal{D}}$, we note that the original lithological tomography approach of Bosch (1999) is closely related to the pseudo-marginal method. In the original lithological tomography method targeting the joint posterior PDF $p(\boldsymbol{\theta}, \mathbf{x}|\mathbf{y})$, the MCMC chains store the conditional draws of the latent variables together with the target variables, and the target posterior PDF $p(\boldsymbol{\theta}|\mathbf{y})$ is obtained by marginalization. The PM method applied with one draw of the latent variable leads to equivalent results in terms of the marginal posterior PDF. In the PM method, the draws of the latent variable are not stored but only used to estimate the likelihood $p(\mathbf{y}|\boldsymbol{\theta})$. Using only one sample of the latent variable in the PM method typically leads to impractically-low acceptance rates due to the high variability of the ratio of log-likelihood estimators. To achieve an efficient algorithm, the standard deviation of the log-likelihood estimator needs to be around 1.2-1.5 (Doucet et al. 2015), which is ensured by increasing the number of samples and applying importance sampling schemes. In the context of state-space models, the number of Monte Carlo samples used in the likelihood estimator needs to increase linearly with the number of observations, which becomes impractical in data-rich applications (Deligiannidis et al. 2018). To obtain low-variance log-likelihood ratio approximations with a smaller number of Monte Carlo samples, Deligiannidis et al. (2018) introduced the correlated pseudo-marginal (CPM) method by which the draws of latent variables used in the denominator and numerator in the likelihood ratio are correlated. Both the PM and CPM methods are general in that they allow for non-linear and non-Gaussian assumptions, but their implementation and applicability in data-rich high-dimensional geophysical settings remain untested.

Inferring hundreds or thousands of parameters with a MH algorithm is challenging as the number of iterations needed for convergence grows with the number of target parameters (e.g., Robert et al. 2018). To ensure adequate performance in such settings, it is crucial to equip the algorithm with a well-working proposal scheme. In the context of Gaussian random fields with high dimension, Cotter et al. (2013) demonstrated that standard random walk MCMC algorithms leads to strong dependence on the discretization of the target field and highly inefficient algorithms. Their proposed solution lies in preserving the prior PDF within the proposal scheme such that the acceptance probability of model proposals only depends on the likelihood ratio. This type of proposal schemes was explored in geophysics by Mosegaard and Tarantola (1995), in what is often referred to as the extended Metropolis algorithm. In a high-dimensional target space, the extended Metropolis approach still needs an efficient model proposal scheme (Ruggeri et al. 2015). Following Brunetti and Linde (2018), we use the adaptive multi-chain algorithm DREAM_(ZS) (DiffeRential Evolution Adaptive Metropolis using an archive of past states) by Laloy and Vrugt (2012), which is widely used in various geophysical inversion studies (e.g., Bikowski et al. 2012; Rosas-Carbajal et al. 2014; Hunziker et al. 2017). We adapt herein the DREAM_(ZS)'s formulation in order to accommodate prior-preserving model proposals.

As an exemplary problem, we consider inference of high-dimensional multi-Gaussian porosity fields using crosshole ground-penetrating radar (GPR) first-arrival travel times. We con-

sider both a linear straight-ray solver, to enable comparisons with analytical solutions, and a more physically-based non-linear eikonal solver. We compare the results obtained by our prior-sampling-based proposal and importance-sampling-based implementation of the (correlated) pseudo-marginal method with standard model proposals and without importance sampling. Furthermore, we compare against the original lithological tomography formulation, full inversion and MCMC inversions that simply ignore the presence of petrophysical prediction uncertainty. With these examples, we will demonstrate that our implementation of the CPM method is outperforming the other inversion methods by greatly enhancing the posterior exploration.

This paper is structured as follows. Section 2.2 introduces the methodology by discussing Bayesian inference in the context of high-dimensional settings, presenting the inversion approaches considered and the tools employed for performance assessment. Section 2.3 presents the two test examples with linear and non-linear physics. The results and wider implications are discussed in Section 2.4, followed by conclusions in Section 2.5.

2.2 Methodology

The methodology section starts by introducing the considered latent variable model (Section 2.2.1), followed by general considerations concerning Bayesian inference and MCMC in high-dimensional settings (Section 2.2.2). The correlated pseudo-marginal method and our IS procedure are introduced in Section 2.2.3 and baseline methods used for comparative purposes are presented in Section 2.2.4. Finally, Section 2.2.5 presents the performance assessment metrics used to evaluate the results.

2.2.1 Latent variable model

We consider a latent variable model where the unobservable variable $\mathbf{X} = (X_1, X_2, \dots, X_L)$ is related to the d target parameters $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_d)$ and the T measurements $\mathbf{y} = (y_1, y_2, \dots, y_T)$. We write

$$\mathbf{Y} = \mathcal{G}(\mathbf{X}) + \boldsymbol{\varepsilon}_{\mathcal{O}} = \mathcal{G}(\mathcal{F}(\boldsymbol{\theta}) + \boldsymbol{\varepsilon}_{\mathcal{P}}) + \boldsymbol{\varepsilon}_{\mathcal{O}}, \quad (2.1)$$

for $\mathcal{G} : \mathbb{R}^L \rightarrow \mathbb{R}^T$ and $\mathcal{F} : \mathbb{R}^d \rightarrow \mathbb{R}^L$ with errors $\boldsymbol{\varepsilon}_{\mathcal{O}}$ and $\boldsymbol{\varepsilon}_{\mathcal{P}}$. In our setting, $\mathbf{x} \mapsto \mathcal{G}(\mathbf{x})$ describes the physical forward solver with $\boldsymbol{\varepsilon}_{\mathcal{O}}$ denoting the observational noise and $\boldsymbol{\theta} \mapsto \mathcal{F}(\boldsymbol{\theta})$ represents the petrophysical relationship with $\boldsymbol{\varepsilon}_{\mathcal{P}}$ denoting the petrophysical prediction error (PPE). We assume both errors to be Gaussian such that the distribution of $\mathbf{X}|\boldsymbol{\theta}$ can be represented with the PDF $p(\mathbf{x}|\boldsymbol{\theta}) = \varphi_L(\mathbf{x}; \mathcal{F}(\boldsymbol{\theta}), \boldsymbol{\Sigma}_P)$ and the one of $\mathbf{Y}|\boldsymbol{\theta}, \mathbf{X}$ with the PDF $p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x}) = \varphi_T(\mathbf{y}; \mathcal{G}(\mathbf{x}), \boldsymbol{\Sigma}_Y)$, with the notation $\varphi_M(\cdot; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denoting the PDF of a M -variate Normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.

2.2.2 Bayesian Inference with Markov Chain Monte Carlo

In Bayes' theorem, the posterior probability density function (PDF) $p(\boldsymbol{\theta}|\mathbf{y})$ of the model parameters $\boldsymbol{\theta}$ given the measurements \mathbf{y} is specified by

$$p(\boldsymbol{\theta}|\mathbf{y}) = \frac{p(\boldsymbol{\theta})p(\mathbf{y}|\boldsymbol{\theta})}{p(\mathbf{y})}, \quad (2.2)$$

with the prior PDF $p(\boldsymbol{\theta})$ of the model parameters, the likelihood function $p(\mathbf{y}|\boldsymbol{\theta})$ and the evidence $p(\mathbf{y})$. Generally, there is no analytical form of the posterior PDF.

If the posterior PDF can be evaluated pointwise up to a normalizing constant, MCMC methods can be used to generate posterior samples. The basic idea of MCMC algorithms is to construct a Markov chain with the posterior PDF of interest as its stationary distribution (see e.g., Robert et al. 1999). MCMC algorithms iteratively propose new values for the states of the Markov chain that are accepted or rejected with a prescribed probability. One foundational MCMC algorithm is Metropolis–Hastings (MH; Metropolis et al. 1953; Hastings 1970). It proceeds as follows at iteration j : First, using the model proposal density $q(\cdot|\boldsymbol{\theta}^{(j-1)})$, a new set of states $\boldsymbol{\theta}^{(j)}$ is proposed. Then, the acceptance probability,

$$\alpha_{MH}(\boldsymbol{\theta}^{(j-1)}, \boldsymbol{\theta}^{(j)}) = \min \left\{ 1, \frac{q(\boldsymbol{\theta}^{(j-1)}|\boldsymbol{\theta}^{(j)})p(\boldsymbol{\theta}^{(j)}|\mathbf{y})}{q(\boldsymbol{\theta}^{(j)}|\boldsymbol{\theta}^{(j-1)})p(\boldsymbol{\theta}^{(j-1)}|\mathbf{y})} \right\} \quad (2.3)$$

$$= \min \left\{ 1, \frac{q(\boldsymbol{\theta}^{(j-1)}|\boldsymbol{\theta}^{(j)})p(\boldsymbol{\theta}^{(j)})p(\mathbf{y}|\boldsymbol{\theta}^{(j)})}{q(\boldsymbol{\theta}^{(j)}|\boldsymbol{\theta}^{(j-1)})p(\boldsymbol{\theta}^{(j-1)})p(\mathbf{y}|\boldsymbol{\theta}^{(j-1)})} \right\}, \quad (2.4)$$

is calculated and the proposed $\boldsymbol{\theta}^{(j)}$ is accepted (if $\alpha_{MH}(\boldsymbol{\theta}^{(j-1)}, \boldsymbol{\theta}^{(j)}) \geq V$) or rejected (if $\alpha_{MH}(\boldsymbol{\theta}^{(j-1)}, \boldsymbol{\theta}^{(j)}) < V$) on the basis of a draw of a uniformly distributed random variable $V \sim \text{Unif}([0, 1])$. If the proposed $\boldsymbol{\theta}^{(j)}$ is rejected, the old state of the chain is kept and $\boldsymbol{\theta}^{(j)} = \boldsymbol{\theta}^{(j-1)}$.

Within the MH algorithm, we need to evaluate the likelihood function $\boldsymbol{\theta} \mapsto p(\mathbf{y}|\boldsymbol{\theta})$ in order to compute the acceptance probability. In our latent variable model (see Section 2.2.1), the likelihood is given by,

$$p(\mathbf{y}|\boldsymbol{\theta}) = \int p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x})p(\mathbf{x}|\boldsymbol{\theta})d\mathbf{x}, \quad (2.5)$$

and the integral has generally no analytical form. In Sections 2.2.3 and 2.2.4, we present three methods to circumvent the difficulties of an intractable likelihood function.

Model parameterization and proposal scheme

We consider test examples targeting a Gaussian random field $GRF(\mu_\theta(\cdot), C_\theta(\cdot, \cdot))$ with known mean $\mu_\theta(\cdot)$ and covariance function $C_\theta(\cdot, \cdot)$. We parameterize the target field θ using a regular 2D grid of size $D \times D$ (such that $d = D^2$ for the notation introduced in Section 2.2.1) with positions $\mathcal{B} = \{b_1, b_2, \dots, b_{D^2}\}$:

$$\theta \sim \mathcal{N}_{D^2}(\boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta), \text{ with } \boldsymbol{\mu}_\theta = (\mu_\theta(g_i))_{1 \leq i \leq D^2} \text{ and } \boldsymbol{\Sigma}_\theta = (C_\theta(g_i, g_j))_{1 \leq i, j \leq D^2}, \quad (2.6)$$

with $\mathcal{N}_{D^2}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denoting the D^2 -variate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. We use a high-dimensional pixel-based parameterization of the target field, $\theta = \boldsymbol{\mu}_\theta + \boldsymbol{\Sigma}_\theta^{1/2} \mathbf{Z}$, where \mathbf{Z} is a D^2 -dimensional random vector consisting of *i.i.d.* standard-normal distributed variables. To infer the target field, we need to estimate the \mathbf{Z} -variables. Similar to Ruggeri et al. (2015), we do not apply any further dimensionality reduction of the parameter space beyond the discretization (in contrast with, for instance, Brunetti and Linde (2018) who used the dimensionality reduction approach of Laloy et al. (2015)). This is done to avoid distorted posterior PDF estimates that may arise in response to a reduction of the parameter space. Furthermore, we seek to evaluate performance in a challenging high-dimensional setting with thousands of unknowns.

When inferring model parameters with the MH algorithm, it is crucial to choose the model proposal scale well. If the model proposal steps are too large, the acceptance rate is low and the Markov chain needs many iterations until convergence. If the step-width is too small, the exploration of the parameter space is very slow and the Markov chain will similarly need many iterations until convergence (see Section 2.2.5 for the assessment of convergence). To deal with this challenge of tuning the proposal scale of each model parameter, we use the adaptive multi-chain algorithm DREAM_(ZS) (DiffeREntial Evolution Adaptive Metropolis using an archive of past states) by Laloy and Vrugt (2012) for which details can be found in Appendix 2.6.1.

MCMC algorithms generally suffer from the curse of dimensionality as the number of iterations needed for convergence increases with the number of target parameters (e.g., Robert et al. 2018). In the context of Gaussian random fields, Cotter et al. (2013) show that MCMC methods based on standard random walk proposals lead to strong dependencies on the discretization of the target field and to inefficient algorithms when employed in high dimensions. For a given proposal scale, refining the grid representing the random field leads to a decreasing acceptance rate with zero as the limiting value for an infinite number of unknowns. To make MCMC algorithms robust to discretization and maintain a reasonable stepsize when inferring thousands of unknowns, they propose model proposal schemes such as the pCN (preconditioned Crank-Nicholson) that preserve the prior PDF. For a target variable \mathbf{Z} with a Standard-Normal prior, the proposal of a standard random walk method is given by $\mathbf{Z}^{(j)} = \mathbf{Z}^{(j-1)} + \gamma\zeta$, with γ being the step size and $\zeta \sim \mathcal{N}(0, 1)$, respectively. Instead, the pCN proposal scheme uses $\mathbf{Z}^{(j)} = \sqrt{1 - \gamma^2} \mathbf{Z}^{(j-1)} + \gamma\zeta$, ensuring that $\mathbf{Z}^{(j)}$ remains standard-normally distributed. Cotter et al. (2013) show that proposal schemes preserving the prior PDF lead to (1) algorithms that mix more rapidly and (2) the convergence being insensitive to the discretization of the target field.

We note that the idea of defining a model proposal scheme preserving the prior distribution was proposed more than 25 years ago in geophysics by Mosegaard and Tarantola (1995). This approach is often referred to as the extended Metropolis algorithm and has mainly been explored in the context of inversion with complex geostatistical prior models (a detailed description of the method can be found in Hansen et al. (2012)). Defining a proposal density $q(\cdot|\boldsymbol{\theta}^{(j-1)})$ such that the MCMC algorithm samples the prior PDF in the absence of data implies that $\frac{q(\boldsymbol{\theta}^{(j-1)}|\boldsymbol{\theta}^{(j)})}{q(\boldsymbol{\theta}^{(j)}|\boldsymbol{\theta}^{(j-1)})} = \frac{p(\boldsymbol{\theta}^{(j-1)})}{p(\boldsymbol{\theta}^{(j)})}$ holds true, with the implication that the MH acceptance-ratio of Equation (2.3) is reduced to the likelihood ratio,

$$\alpha_{MH}(\boldsymbol{\theta}^{(j-1)}, \boldsymbol{\theta}^{(j)}) = \min\left\{1, \frac{p(\mathbf{y}|\boldsymbol{\theta}^{(j)})}{p(\mathbf{y}|\boldsymbol{\theta}^{(j-1)})}\right\}. \quad (2.7)$$

The extended Metropolis approach still needs an efficient model proposal scheme (Ruggeri et al. 2015), which is why we use $\text{DREAM}_{(ZS)}$ in this work. In the case of a Gaussian-distributed prior, the standard $\text{DREAM}_{(ZS)}$ proposal scheme does not generate samples that preserve the prior distribution. In order to adapt extended Metropolis to $\text{DREAM}_{(ZS)}$, we rely on a transformation of the variables to the Uniform space (details in Appendix 2.6.1). This transformation makes it possible to create a proposal mechanism which unites (1) the efficiency of the $\text{DREAM}_{(ZS)}$ proposals with (2) the robustness of the prior-preserving proposals. In what follows, our proposal scheme using the uniform transform will be referred to as prior-sampling $\text{DREAM}_{(ZS)}$ proposals, while the standard proposal scheme of $\text{DREAM}_{(ZS)}$ will be referred to as standard $\text{DREAM}_{(ZS)}$ proposals. We stress that both prior-sampling $\text{DREAM}_{(ZS)}$ and standard $\text{DREAM}_{(ZS)}$ target the same posterior PDF, but the former is expected to be more efficient.

2.2.3 (Correlated) pseudo-marginal method

Pseudo-marginal method

Beaumont (2003) shows that a MH algorithm using a non-negative unbiased estimator of the likelihood samples the same target distribution as when using the true likelihood. He exploits this property by estimating the likelihood in Equation (2.5) on the basis of Monte Carlo averaging over samples of the latent variable \mathbf{X} . Andrieu and Roberts (2009) adopt this approach in their pseudo-marginal (PM) method and provide a theoretical analysis of the scheme. When one brute force Monte Carlo sample of the latent variable is drawn in each MCMC iteration without importance sampling (c.f., the original lithological tomography by Bosch (1999); see Section 2.2.4), the algorithm is likely to suffer from a low acceptance rate due to the high variability of the log-likelihood estimator. This is due to the fact that a likelihood estimator given by $p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{X})$ takes very different values depending on the draw of the latent variable \mathbf{X} , even for the same $\boldsymbol{\theta}$. This occurs as the scatter ($\boldsymbol{\varepsilon}_{\mathcal{P}}$) has a strong effect on the data response, and hence, the likelihood. To improve the efficiency, Beaumont (2003)

and Andrieu and Roberts (2009) use many samples drawn by importance sampling (IS; e.g. Owen and Zhou 2000). Consequently, they propose the following unbiased estimator of the likelihood $p(\mathbf{y}|\boldsymbol{\theta})$,

$$\hat{p}_N(\mathbf{y}|\boldsymbol{\theta}) = \frac{1}{N} \sum_{n=1}^N w(\mathbf{y}|\boldsymbol{\theta}, \mathbf{X}_n), \quad \text{with} \quad w(\mathbf{y}|\boldsymbol{\theta}, \mathbf{X}_n) = \frac{p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{X}_n)p(\mathbf{X}_n|\boldsymbol{\theta})}{m(\mathbf{X}_n|\boldsymbol{\theta})}, \quad (2.8)$$

where $\mathbf{X}_n \stackrel{i.i.d.}{\sim} m(\cdot|\boldsymbol{\theta})$ for $n = 1, 2, \dots, N$ with $m(\cdot|\boldsymbol{\theta})$ being the importance density function. More details about the importance sampling procedure will follow below.

Correlated pseudo-marginal method

For the PM method to be efficient, the number of samples N used in the likelihood estimator (Eq. 2.8) should be selected such that the variance of the log-likelihood ratio estimator is low enough (Doucet et al. 2015). If it is too high, the algorithm will suffer from an impractically low acceptance rate. In the state-space model context, this implies that N needs to scale linearly with T leading to a computational cost of order T^2 at every MCMC iteration, which can be prohibitively expensive for large T (Deligiannidis et al. 2018). To reduce the computational cost, Deligiannidis et al. (2018) introduced the correlated pseudo-marginal (CPM) method by which the draws of latent variables used in the denominator and numerator of the likelihood ratio estimators are correlated. The underlying idea is that the variance of a ratio of estimators is lower if they are positively correlated (Koop 1972). Assuming that the latent variable \mathbf{X} is standard-normal distributed, the CPM method proposes (in iteration j) a realization of the n -th latent variable draw by means of pre-conditioned Crank-Nicholson proposals,

$$\mathbf{X}_n^{(j)} = \rho \mathbf{X}_n^{(j-1)} + \sqrt{1 - \rho^2} \boldsymbol{\epsilon}, \quad \text{with } \rho \in (0, 1) \text{ and } \boldsymbol{\epsilon} = (\epsilon_1, \epsilon_2, \dots, \epsilon_L), \epsilon_i \stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1). \quad (2.9)$$

The assumption that the latent variable has a standard-normal distribution hardly limits the general applicability of the CPM method, since there exist transformations from numerous distributions that will allow proposals to act on Gaussian distributions (e.g. Chen et al. 2018). We stress that if the proposed $\boldsymbol{\theta}^{(j)}$ with $\mathbf{X}_n^{(j)}$ is rejected by the CPM algorithm, we keep $\mathbf{X}_n^{(j)} = \mathbf{X}_n^{(j-1)}$ as for $\boldsymbol{\theta}^{(j)} = \boldsymbol{\theta}^{(j-1)}$.

Compared to standard MCMC algorithms, the CPM method requires two additional parameters: the latent variable sample size N and the correlation parameter ρ . To achieve optimal performance, the parameters should be chosen such that the variance of the log-likelihood ratio estimator for a fixed target variable $\boldsymbol{\theta}$,

$$R = \log\left(\hat{p}_N^{(j)}(\mathbf{y}|\boldsymbol{\theta})\right) - \log\left(\hat{p}_N^{(j-1)}(\mathbf{y}|\boldsymbol{\theta})\right), \quad (2.10)$$

takes values between 1.0 and 2.0 in regions with high probability mass (Deligiannidis et al. 2018). Here, $\hat{p}_N^{(j)}(\mathbf{y}|\boldsymbol{\theta})$ and $\hat{p}_N^{(j-1)}(\mathbf{y}|\boldsymbol{\theta})$ refer to the likelihood estimators (Eq. 2.8) obtained with the accepted latent variable of iteration $j - 1$ and the proposed (and not necessarily accepted) latent variable of iteration j , that is, the likelihood estimators used in the acceptance ratio of the MH algorithm. In order to choose the parameter values, we first fix the number of

samples N at a value that is smaller than the number of available parallel processors. Then, we evaluate different ρ and estimate corresponding values of $\text{Var}(R)$ for a fixed $\boldsymbol{\theta}$ in a region with high posterior probability mass (e.g., chosen based on initial MCMC runs).

Importance sampling procedure

For high-dimensional problems with large data sets exhibiting high signal-to-noise ratios, it is necessary to use importance sampling when drawing samples of latent variables to be used within the likelihood-estimator (Eq. 2.8). This is a consequence of the integrand $p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x})$ in Equation (2.5) having a peak in a region of \mathbf{X} having small probability under $p(\mathbf{x}|\boldsymbol{\theta})$. Importance sampling proceeds by sampling from a so-called importance distribution given by the PDF $\mathbf{x} \mapsto m(\mathbf{x}|\boldsymbol{\theta})$ that preferentially generates samples with high $p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x})p(\mathbf{x}|\boldsymbol{\theta})$. Furthermore, the support of the importance distribution must include all values \mathbf{x} , for which $p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x})p(\mathbf{x}|\boldsymbol{\theta}) > 0$ (Owen and Zhou 2000). It holds,

$$\int p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x})p(\mathbf{x}|\boldsymbol{\theta})d\mathbf{x} = \int \frac{p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x})p(\mathbf{x}|\boldsymbol{\theta})}{m(\mathbf{x}|\boldsymbol{\theta})}m(\mathbf{x}|\boldsymbol{\theta})d\mathbf{x}, \quad (2.11)$$

leading to the unbiased importance sampling estimate of the likelihood given in Equation (2.8). To ensure minimal variance of the estimator, we seek $\mathbf{x} \mapsto m(\mathbf{x}|\boldsymbol{\theta})$ to be nearly proportional to $\mathbf{x} \mapsto p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x})p(\mathbf{x}|\boldsymbol{\theta})$ as recalled in Owen and Zhou (2000) referring to the results of Kahn and Marshall (1953). Since $p(\mathbf{x}|\boldsymbol{\theta}, \mathbf{y}) \propto p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x})p(\mathbf{x}|\boldsymbol{\theta})$, it is sensible to base the importance density on $\mathbf{x} \mapsto p(\mathbf{x}|\boldsymbol{\theta}, \mathbf{y})$.

Within a latent variable model with a non-linear physical forward solver (Section 2.2.1), we can not derive the exact expression for $p(\mathbf{x}|\boldsymbol{\theta}, \mathbf{y})$. Here, we derive local approximations of this posterior by relying on linearization. To do so, we use a linearization of the map $\mathbf{x} \mapsto \mathcal{G}(\mathbf{x})$ around $\mathbf{x}_{\text{lin}} = \mathcal{F}(\boldsymbol{\theta}_{\text{lin}}) + \boldsymbol{\varepsilon}_{\mathcal{F}_{\text{lin}}}$ based on a first-order expansion,

$$\mathcal{G}(\mathbf{x}) = \mathcal{G}(\mathbf{x}_{\text{lin}} + \mathbf{x} - \mathbf{x}_{\text{lin}}) \approx \mathcal{G}(\mathbf{x}_{\text{lin}}) + \mathbf{J}_{\mathbf{x}_{\text{lin}}}(\mathbf{x} - \mathbf{x}_{\text{lin}}), \quad (2.12)$$

with $\mathbf{J}_{\mathbf{x}_{\text{lin}}}$ being the Jacobian matrix of the forward solver corresponding to \mathbf{x}_{lin} . Ideally, \mathbf{x}_{lin} should be given by a realization of the latent variable similar to the one the algorithm is currently exploring. By approximating $p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x})$ with $\tilde{p}(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x}) = \varphi_T(\mathbf{y}; \mathcal{G}(\mathbf{x}_{\text{lin}}) + \mathbf{J}_{\mathbf{x}_{\text{lin}}}(\mathbf{x} - \mathbf{x}_{\text{lin}}), \boldsymbol{\Sigma}_Y)$ and, applying $p(\mathbf{x}|\boldsymbol{\theta}) = \varphi_L(\mathbf{x}; \mathcal{F}(\boldsymbol{\theta}), \boldsymbol{\Sigma}_P)$ and the relationships between marginal and conditional Gaussians out of Bishop and Nasrabadi (2006) given in Appendix 2.6.2, we get,

$$\begin{aligned} \tilde{p}(\mathbf{x}|\boldsymbol{\theta}, \mathbf{y}) &= \varphi_L(\mathbf{x}; \boldsymbol{\mu}_{IS}, \boldsymbol{\Sigma}_{IS}), \text{ with} & (2.13) \\ \boldsymbol{\mu}_{IS} &= \boldsymbol{\Sigma}_{IS} \left(\mathbf{J}_{\mathbf{x}_{\text{lin}}}^T \boldsymbol{\Sigma}_Y^{-1} (\mathbf{y} - (\mathcal{G}(\mathbf{x}_{\text{lin}}) - \mathbf{J}_{\mathbf{x}_{\text{lin}}} \mathbf{x}_{\text{lin}})) + \boldsymbol{\Sigma}_P^{-1} \mathcal{F}(\boldsymbol{\theta}) \right), \\ \boldsymbol{\Sigma}_{IS} &= (\boldsymbol{\Sigma}_P^{-1} + \mathbf{J}_{\mathbf{x}_{\text{lin}}}^T \boldsymbol{\Sigma}_Y^{-1} \mathbf{J}_{\mathbf{x}_{\text{lin}}})^{-1}, \end{aligned}$$

for an approximation of $p(\mathbf{x}|\boldsymbol{\theta}, \mathbf{y})$. To incorporate importance sampling within the CPM method, we need to correlate the draws of latent variables. To achieve this, we rely on the fact that a realization of the latent variable \mathbf{X} can be generated with $\boldsymbol{\mu}_{IS} + \boldsymbol{\Sigma}_{IS}^{1/2} \mathbf{Z}_P$, where \mathbf{Z}_P is standard Gaussian distributed in \mathbb{R}^L . Using this representation, we can correlate the (standard-normal distributed) \mathbf{Z}_P -variables using Equation (2.9).

2.2.4 Baseline inversion methods

We present now the inversion approaches used for comparison with the CPM method. These include a method ignoring the petrophysical prediction errors and two approaches (original formulation of lithological tomography without importance sampling and full inversion) accounting for the PPEs by inferring the joint posterior PDF $p(\boldsymbol{\theta}, \mathbf{x}|\mathbf{y})$ of the target and latent variables. An overview of all inversion methods (including CPM) is given in Table 2.1.

Table 2.1: Overview of the inversion methods applied on the latent variable model introduced in Section 2.2.1; a box around a letter indicates that this parameter is saved as a target variable of the MH algorithm. For the proposal scheme we use both standard and prior-sampling DREAM_(ZS) proposals for all methods.

Method	Proposal scheme	Latent variable(s)	Likelihood $\hat{p}(\mathbf{y} \boldsymbol{\theta})$
No PPE: Ignore PPE	$\boldsymbol{\theta}^{(j)}$	$\mathbf{X}^{(j)} = \mathcal{F}(\boldsymbol{\theta}^{(j)})$	$\varphi_T(\mathbf{y}; \mathcal{G}(\mathbf{X}^{(j)}), \Sigma_Y)$
Full inversion: Infer PPE	$\boldsymbol{\theta}^{(j)}, \boldsymbol{\varepsilon}_{\mathcal{F}}^{(j)}$	$\mathbf{X}^{(j)} = \mathcal{F}(\boldsymbol{\theta}^{(j)}) + \boldsymbol{\varepsilon}_{\mathcal{F}}^{(j)}$	$\varphi_T(\mathbf{y}; \mathcal{G}(\mathbf{X}^{(j)}), \Sigma_Y)$
LithTom: Infer PPE	$\boldsymbol{\theta}^{(j)}$	$\mathbf{X}^{(j)} \sim \varphi_L(\cdot; \mathcal{F}(\boldsymbol{\theta}^{(j)}), \Sigma_P)$	$\varphi_T(\mathbf{y}; \mathcal{G}(\mathbf{X}^{(j)}), \Sigma_Y)$
LithTom IS: Infer PPE	$\boldsymbol{\theta}^{(j)}$	$\mathbf{X}^{(j)} \sim \varphi_L(\cdot; \boldsymbol{\mu}_{IS}, \boldsymbol{\Sigma}_{IS})$	$\frac{\varphi_T(\mathbf{y}; \mathcal{G}(\mathbf{X}^{(j)}), \Sigma_Y) \varphi_L(\mathbf{X}^{(j)}; \mathcal{F}(\boldsymbol{\theta}^{(j)}), \Sigma_P)}{\varphi_L(\mathbf{X}^{(j)}; \boldsymbol{\mu}_{IS}, \boldsymbol{\Sigma}_{IS})}$
(C)PM no IS: Sample out PPE	$\boldsymbol{\theta}^{(j)}$	$\mathbf{X}^{(j)} = (\mathbf{X}_1^{(j)}, \dots, \mathbf{X}_N^{(j)})$ $\mathbf{X}_n^{(j)} \overset{i.i.d.}{\sim} \varphi_L(\cdot; \mathcal{F}(\boldsymbol{\theta}^{(j)}), \Sigma_P)$ CPM: Correlation $\mathbf{X}_n^{(j-1)}$	$\frac{1}{N} \sum_{n=1}^N \varphi_T(\mathbf{y}; \mathcal{G}(\mathbf{X}_n^{(j)}), \Sigma_Y)$
(C)PM IS: Sample out PPE	$\boldsymbol{\theta}^{(j)}$	$\mathbf{X}^{(j)} = (\mathbf{X}_1^{(j)}, \dots, \mathbf{X}_N^{(j)})$ $\mathbf{X}_n^{(j)} \overset{i.i.d.}{\sim} \varphi_L(\cdot; \boldsymbol{\mu}_{IS}, \boldsymbol{\Sigma}_{IS})$ CPM: Correlation $\mathbf{X}_n^{(j-1)}$	$\frac{1}{N} \sum_{n=1}^N \frac{\varphi_T(\mathbf{y}; \mathcal{G}(\mathbf{X}_n^{(j)}), \Sigma_Y) \varphi_L(\mathbf{X}_n^{(j)}; \mathcal{F}(\boldsymbol{\theta}^{(j)}), \Sigma_P)}{\varphi_L(\mathbf{X}_n^{(j)}; \boldsymbol{\mu}_{IS}, \boldsymbol{\Sigma}_{IS})}$

Ignore petrophysical prediction errors

This inversion method (no PPE) ignores the presence of petrophysical prediction errors in the MH algorithm. For the latent variable model introduced in Section 2.2.1, this results in an approximation of the likelihood function with the Gaussian PDF $\hat{p}(\mathbf{y}|\boldsymbol{\theta}) = \varphi_T(\mathbf{y}; \mathcal{G}(\mathcal{F}(\boldsymbol{\theta})), \Sigma_Y)$, where the forward response $\mathcal{G}(\mathcal{F}(\boldsymbol{\theta}))$ is simulated without accounting for PPEs. The method is included in the comparison as it is commonly used in practice as discussed by Brunetti and Linde (2018).

Lithological Tomography

One way to consider PPEs while circumventing the difficulty of an intractable likelihood function is to infer the joint posterior PDF $(\boldsymbol{\theta}, \mathbf{x}) \mapsto p(\boldsymbol{\theta}, \mathbf{x}|\mathbf{y})$ of the hydrogeological and geophysical parameters. Lithological tomography (Bosch 1999) pursues this strategy and uses a factorization of the joint posterior PDF as $p(\boldsymbol{\theta}, \mathbf{x}|\mathbf{y}) \propto p(\boldsymbol{\theta})p(\mathbf{x}|\boldsymbol{\theta})p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x})$, where $p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x}) = p(\mathbf{y}|\mathbf{x})$ is valid for our setting. To sample from this posterior PDF, Bosch (1999) proceeds as follows: First, realizations from the joint prior of $\boldsymbol{\theta}$ and \mathbf{X} are created by marginal sampling of $\boldsymbol{\theta}$ and conditional sampling of \mathbf{X} . Then, the pairs of model proposals are accepted or rejected with $p(\mathbf{y}|\mathbf{x})$, used in the acceptance ratio of the MH algorithm. In practice, this means that brute force Monte Carlo realizations (no importance sampling) of the petrophysical prediction error $\boldsymbol{\varepsilon}_{\mathcal{F}}$ are added to the output of the petrophysical relationship $\mathcal{F}(\boldsymbol{\theta})$. For our latent variable model, this results in an approximation of the likelihood function with $\hat{p}(\mathbf{y}|\boldsymbol{\theta}) = \varphi_T(\mathbf{y}; \mathcal{G}(\mathbf{x}), \Sigma_Y)$, where the latent variable $\mathbf{X} = \mathcal{F}(\boldsymbol{\theta}) + \boldsymbol{\varepsilon}_{\mathcal{F}}$ is obtained with a draw of $\boldsymbol{\varepsilon}_{\mathcal{F}}$ from the multivariate Gaussian with PDF $\varphi_L(\cdot; 0, \Sigma_P)$.

Full Inversion

The full inversion approach infers the joint posterior PDF by treating the latent variables analogously to the other unknowns. In the context of our latent variable model (Section 2.2.1), this means that in iteration j of the MH, not only a new $\boldsymbol{\theta}^{(j)}$ but also a new $\boldsymbol{\varepsilon}_{\mathcal{F}}^{(j)}$ is proposed by the algorithm's proposal scheme. Then the likelihood function $p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x}) = \varphi_T(\mathbf{y}; \mathcal{G}(\mathbf{x}), \Sigma_Y)$ is calculated using $\mathbf{X}^{(j)} = \mathcal{F}(\boldsymbol{\theta}^{(j)}) + \boldsymbol{\varepsilon}_{\mathcal{F}}^{(j)}$. Brunetti and Linde (2018) applied full inversion to infer porosity fields by inversion of crosshole GPR first-arrival travel times, that is, to a setting similar to ours. For the parametrization of the porosity field of interest, they used a spectral representation combined with the dimensionality reduction approach of Laloy et al. (2015). Brunetti and Linde (2018) achieved convincing results and improvements compared to standard lithological tomography without importance sampling. Nevertheless, full inversion is expected to suffer from high dimensionality and strong correlation among the latent and target variables as the two sets of variables are treated as being independent within the proposal scheme (e.g., Deligiannidis et al. 2018).

2.2.5 Performance assessment

To assess the performance of the different inversion approaches, we primarily focus on the exploration of the posterior PDF. The reason for this will become clear in the results section (Section 2.3). To declare convergence, we use the \hat{R} -statistic of Gelman and Rubin (1992) that compares the within-chain variance with the between-chain variance for the second half of the MCMC chains. The general convention is that convergence is declared once this statistic is smaller or equal to 1.2 for all model parameters. Since we deal with a high-dimensional parameter space with thousands of unknowns, we relax this condition slightly and declare convergence if 99 % of the parameters satisfy this criterion. When an algorithm is considered convergent, we compare the resulting posterior samples with those of the other approaches.

For the test case with linear physics in Section 2.3.2, we compare the results with the analytical solution of the posterior PDF $p(\boldsymbol{\theta}|\mathbf{y})$. For these comparisons, we use histograms and the Kullback–Leibler divergence (KL - divergence; Kullback and Leibler 1951). The KL - divergence between two PDFs $p_1(\cdot)$ and $p_2(\cdot)$ is defined as,

$$\text{KL}(p_1||p_2) = \int p_1(x) \log\left(\frac{p_1(x)}{p_2(x)}\right) dx. \quad (2.14)$$

To obtain the PDF of the estimated posterior, we can use the MCMC samples to either (1) make a kernel density estimate or to (2) estimate the mean and variance for a Gaussian approximation (Krüger et al. 2021). Here we use the second option since the posterior is Gaussian. If the PDFs $p_1(\cdot)$ and $p_2(\cdot)$ are Gaussians with $p_1 = \mathcal{N}(\mu_1, \sigma_1^2)$ and $p_2 = \mathcal{N}(\mu_2, \sigma_2^2)$, the expression of the KL-divergence reduces to,

$$\text{KL}(p_1||p_2) = \log\left(\frac{\sigma_2}{\sigma_1}\right) + \frac{\sigma_1^2 + (\mu_1 - \mu_2)^2}{2\sigma_2^2} - \frac{1}{2}. \quad (2.15)$$

A KL-divergence of zero indicates that the two PDFs are equal and it increases as the distributions diverge from each other.

For the test example with non-linear physics in Section 2.3.3, there is no analytical solution to compare with. Hence, we compare the estimated posterior distribution with a single value (the known true porosity at each pixel). We achieve this by applying so-called scoring rules (Gneiting and Raftery 2007) assessing the accuracy of a predictive PDF $\boldsymbol{\theta} \mapsto \hat{p}(\boldsymbol{\theta})$ with respect to a true value $\boldsymbol{\theta}$. Scoring rules are functions that assign a numerical score for each prediction-observation pair $(\hat{p}, \boldsymbol{\theta})$, with a smaller score indicating a better prediction. They assess both the statistical consistency between predictions and observations (calibration) and the sharpness of the prediction. We use the logarithmic score (logS; Good 1952) defined by $\text{logS}(\hat{p}, \boldsymbol{\theta}) = -\log \hat{p}(\boldsymbol{\theta})$ that is related to the Kullback–Leibler divergence (Gneiting and Raftery 2007). As for the linear case, we use the MCMC samples to obtain a Gaussian approximation of the estimated posterior PDF. The logarithmic score favours predictive PDFs under which the true value has high probability. We supplement this measure with two simpler ones: the number of pixels in which the true porosity value was in the range of the posterior samples and the standard deviation of the estimated posterior PDF.

We also consider the acceptance rates (AR) and the integrated autocorrelation time (IACT). We aim for an acceptance rate of 15% - 30% as proposed by Vrugt (2016). The IACT of the chain $\{\boldsymbol{\theta}^{(j)}; j = 1, 2, \dots\}$ is defined as $1 + 2 \sum_{l=0}^{\infty} \text{Corr}(\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(1+l)})$. In practice, the estimated autocorrelation for large values of l is noisy such that we need to truncate the sum. Following Gelman et al. (2013), we truncate the sum when two successive autocorrelation estimates are negative. We renounce from discussing the CPU time as it depends strongly on the chosen forward model and discretization as well as on other parameters pertaining to the computing equipment.

2.3 Results

We consider the problem of inferring the porosity distribution using crosshole GPR first-arrival travel times. We first address a test case with linear physics (straight-rays) to allow for comparison with analytical solutions and then one with non-linear physics (eikonal solver) to address a more challenging and physically-based setup. Our examples are synthetic and the water-saturated porosity field is described by a multi-Gaussian random field.

2.3.1 Data and inversion setting

Synthetic data generation

Our considered subsurface domain is $7.2 \text{ m} \times 7.2 \text{ m}$ and we use 25 equidistant GPR transmitters located on the left side and 25 receivers on the right side of the model domain, resulting in 625 first-arrival travel times. The transmitter-receiver layout is depicted in Figure 2.1c. As introduced in Section 2.2.2, we assume the porosity field to be a Gaussian random field $GRF(\mu_{\theta}(\cdot), C_{\theta}(\cdot, \cdot))$. We use $\mu_{\theta}(\cdot) = 0.39$ and an exponential covariance function $C_{\theta}(\cdot, \cdot)$. For the latter, we use a sill of $2e^{-4}$ and geometric anisotropy where the main, horizontal direction has an integral scale of 4.5 m and the integral scale ratio between the horizontal and vertical direction is 0.13. We use a (50×50) -dimensional pixel-based parameterization of the porosity field; the true synthetically generated field is shown in Figure 2.1a. Note that porosity is a positive quantity bounded between zero and one while a Gaussian prior distribution has a full support. The Gaussian prior is used here to ensure an analytical solution in the linear physics case. Given the presented mean and the sill, it is extremely unlikely that a porosity value outside the physical boundaries is generated. In other settings, one could use a transform of the porosity (e.g., as in Bosch 2004) or choose a bounded distribution.

To predict the dielectric constant $\boldsymbol{\kappa}$, we use the complex refractive index model (CRIM; Roth et al. 1990),

$$\sqrt{\boldsymbol{\kappa}} = \sqrt{\kappa_s} + (\sqrt{\kappa_w} - \sqrt{\kappa_s})\boldsymbol{\theta}, \quad (2.16)$$

where κ_w and κ_s are the dielectric constants of water [81] and mineral grains [5], respectively. The resulting slowness field (which in our case is the latent variable \mathbf{X}) depicted in Figure 2.1c is given by,

$$\mathbf{x} = \sqrt{c^{-2}\boldsymbol{\kappa}} + \boldsymbol{\varepsilon}_{\mathcal{P}} = \frac{1}{c}(\sqrt{\kappa_s} + (\sqrt{\kappa_w} - \sqrt{\kappa_s})\boldsymbol{\theta}) + \boldsymbol{\varepsilon}_{\mathcal{P}}, \quad (2.17)$$

where c is the speed of light in vacuum [0.3 m/ns]. This specifies the petrophysical relationship to be linear with $\boldsymbol{\theta} \mapsto \mathcal{F}(\boldsymbol{\theta}) = \frac{1}{c}(\sqrt{\kappa_s} + (\sqrt{\kappa_w} - \sqrt{\kappa_s})\boldsymbol{\theta})$. We add a petrophysical prediction error (PPE) $\boldsymbol{\varepsilon}_{\mathcal{P}}$ that is a realization of a centred GRF over a regular 2D grid of size 50×50 . We are assuming that the PPE field (depicted in Figure 2.1b) has an exponential covariance function $C_P(\cdot, \cdot)$ with a sill of $2.1e^{-2}$ and the same correlation structure as the porosity field. The dependency of the slowness on the value of the porosity and the PPE is indicated in Figure 2.1d. Finally, the resulting 625 GPR first - arrival travel times are calculated with (i) a linear (straight-ray) forward solver referred to as \mathcal{G}_s and (ii) a non-linear (eikonal) forward solver referred to as \mathcal{G}_e (the *time2D* solver of Podvin and Lecomte (1991)), such that,

$$\mathbf{y} = \mathcal{G}(\mathbf{x}) + \boldsymbol{\varepsilon}_{\mathcal{O}}, \quad (2.18)$$

with *i.i.d.* centered normal observational noise $\boldsymbol{\varepsilon}_{\mathcal{O}}$ with standard deviation of 1 ns. The two sets of travel times are depicted in Figure 2.1e.

Inversion settings and prior assumptions

All considered inversion methods (Sections 2.2.3 and 2.2.4) are implemented with prior-sampling and standard DREAM_(ZS) proposals using the same parameter settings of the DREAM_(ZS) algorithm with four MCMC chains running in parallel. For the prior on porosity, we use the Gaussian PDF $p(\boldsymbol{\theta}) = \varphi_{2500}(\boldsymbol{\theta}; \boldsymbol{\mu}_{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})$ assuming the mean $\boldsymbol{\mu}_{\boldsymbol{\theta}}$ and covariance structure $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ to be known (the same values as for the data generation). Using a pixel-based parameterization of the field, we infer the 2500-dimensional vector \mathbf{Z} defining the porosity by $\boldsymbol{\theta} = \boldsymbol{\mu}_{\boldsymbol{\theta}} + \boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{1/2}\mathbf{Z}$, with \mathbf{Z} having a multivariate standard-normal prior PDF. The full inversion has to estimate another 2500 \mathbf{Z}_P -variables for the PPE field leading to a total of 5000 inferred parameters. For the PPE $\boldsymbol{\varepsilon}_{\mathcal{P}}$ we also use a Gaussian prior PDF $p(\boldsymbol{\varepsilon}_{\mathcal{P}}) = \varphi_{2500}(\boldsymbol{\varepsilon}_{\mathcal{P}}; \mathbf{0}, \boldsymbol{\Sigma}_P)$ with known covariance structure $\boldsymbol{\Sigma}_P$, leading to a Gaussian prior PDF for the slowness field (for fixed porosity) given by $p(\mathbf{x}|\boldsymbol{\theta}) = \varphi_{2500}(\mathbf{x}; \mathcal{F}(\boldsymbol{\theta}), \boldsymbol{\Sigma}_P)$. For the likelihood function, we assume that the 625-dimensional vector describing the observational noise $\boldsymbol{\varepsilon}_{\mathcal{O}}$ has a Gaussian distribution with zero mean and diagonal covariance matrix $\boldsymbol{\Sigma}_Y$; the standard deviation is assumed to be 1 ns as in the data generation process.

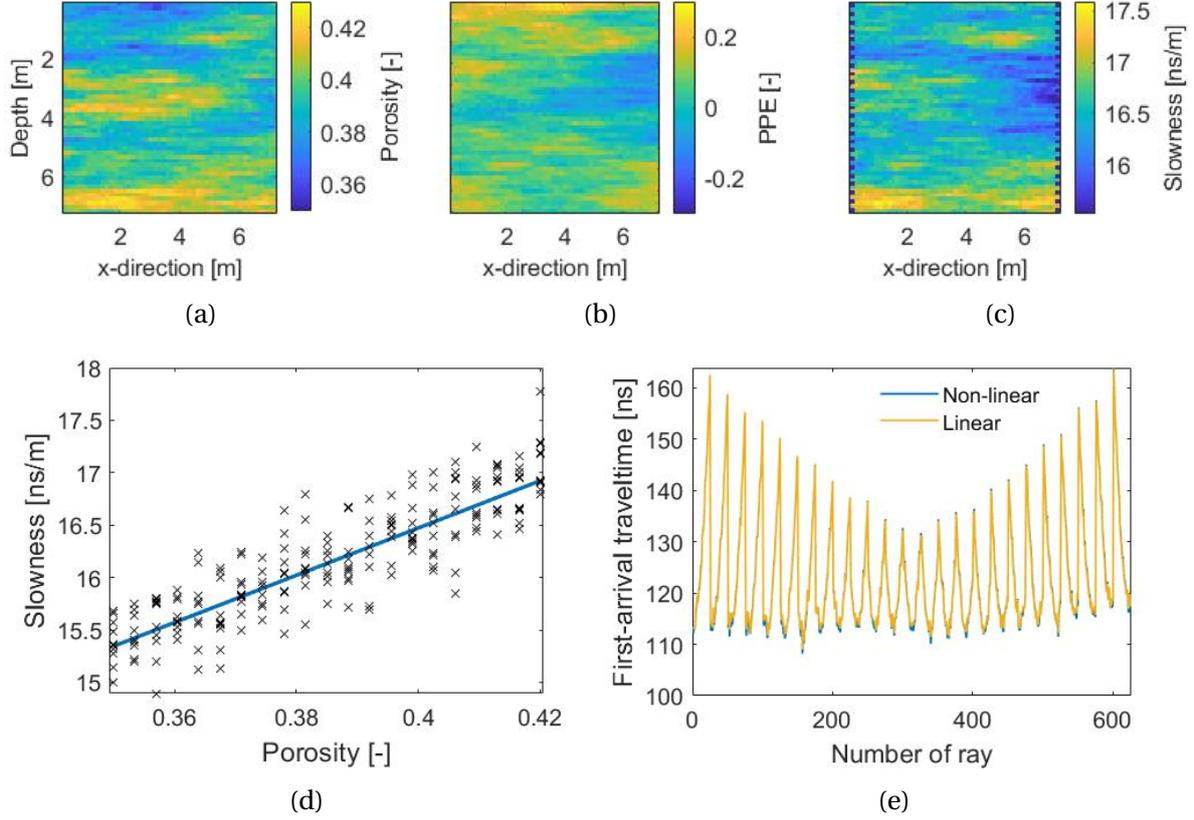


Figure 2.1: (a) Porosity field θ , (b) PPE field ϵ_{θ} , (c) slowness field \mathbf{x} with transmitter-receiver layout, (d) dependency of slowness on porosity obtained without (line) and with (scatter) PPE and (e) noise-contaminated first-arrival travel times \mathbf{y} for the linear and the non-linear forward solver corresponding to the true synthetic model.

2.3.2 Linear physics

To enable comparisons of the inferred posterior PDFs with the analytical solution for $p(\theta|\mathbf{y})$, we first consider the case of linear physics. Then,

$$\mathbf{y} = \mathcal{G}_s(\mathbf{x}) + \epsilon_{\theta} = \mathbf{J}_s \mathbf{x} + \epsilon_{\theta}, \quad (2.19)$$

with \mathbf{J}_s being the Jacobian (i.e., forward operator) of the linear forward solver. The analytical posterior PDF can be derived as detailed in Appendix 2.6.2. Figure 2.2a shows the posterior mean and Figures 2.2b - 2.2d depict three draws from the posterior distribution.

When employing the PM and CPM method in this setting of large datasets with low noise, it is crucial to use a well-chosen importance sampling for the latent variable. As introduced in Section 2.2.3, it is sensible to use $\mathbf{x} \mapsto p(\mathbf{x}|\theta, \mathbf{y})$ as a basis for the importance density. As long as we are in the linear Gaussian case, we can derive the analytical expression for this posterior (Appendix 2.6.2), resulting in a zero-variance importance sampling density (Owen and Zhou 2000). Since it then does not make sense to use multiple importance density samples (the

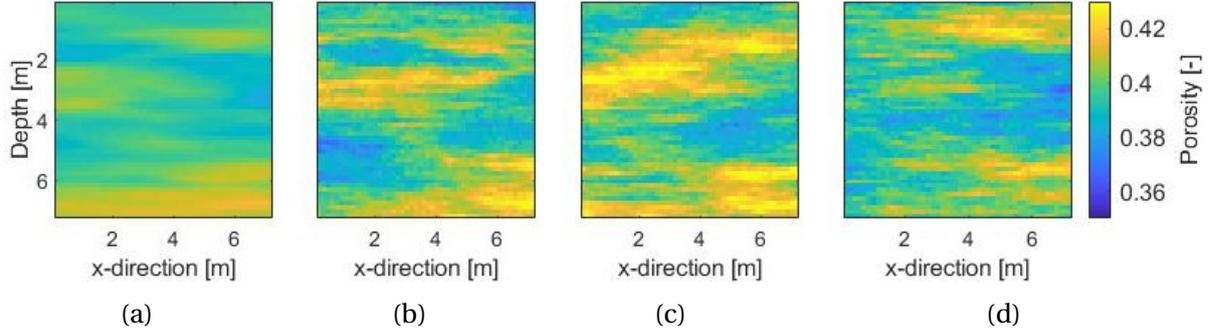


Figure 2.2: (a) Analytical posterior mean of $p(\boldsymbol{\theta}|\mathbf{y})$ for the linear test example and (b) - (d) three realizations of the analytical posterior distribution.

importance weights are constant), we combine in this linear case importance sampling with PM using $N=1$ (original lithological tomography algorithm enhanced with importance sampling that we will hereafter refer to as LithTom IS). We note that using the exact formula for the importance sampling corresponds to having access to the exact likelihood $p(\boldsymbol{\theta}|\mathbf{y})$. The use of larger N is considered in Section 2.3.3 for the case of non-linear physics. This linear setting for which analytical solutions are available serves mainly (1) to demonstrate the necessity of a well-working importance sampling distribution, (2) to investigate the exploration capabilities of MCMC-based inversion approaches that estimate the intractable likelihood using Monte Carlo samples (lithological tomography, PM and CPM methods) and (3) to compare the performances of the prior-sampling and standard $\text{DREAM}_{(ZS)}$ proposal mechanisms.

Figure 2.3 presents the estimated posterior means of the porosity field obtained when applying the no PPE (Fig. 2.3a), the full inversion (Fig. 2.3b) and the LithTom IS (Fig. 2.3c) with standard $\text{DREAM}_{(ZS)}$ proposals, as well as for LithTom IS with prior-sampling $\text{DREAM}_{(ZS)}$ proposals (Fig. 2.3d). These are the cases for which we reached convergence of the chains. The porosity field obtained with the inversion ignoring PPEs has, as expected (Brunetti and Linde 2018), a higher variance. Visually, all other estimates are very similar in terms of structure and magnitude with respect to the analytical posterior mean in Figure 2.2a. The estimated posterior mean of LithTom IS with the prior-sampling $\text{DREAM}_{(ZS)}$ proposals has a slightly lower variance than for standard $\text{DREAM}_{(ZS)}$ proposals. The ARs (Table 2.2) for standard $\text{DREAM}_{(ZS)}$ proposals are the highest for LithTom IS, while the method ignoring PPEs and full inversion have lower ARs. Classical lithological tomography without importance sampling leads to an AR of less than 0.1 % such that, in practice, it is unfeasible to reach convergence. Applying the CPM method without IS for $N=50$ and $\rho=0.95$ also results in an only slightly larger AR (roughly 0.2 %), thereby, highlighting the need for importance sampling for the considered problem. Since less than 5 % of the parameters converged after 200'000 iterations, we renounce from showing further results for the CPM and PM method without IS. The method ignoring PPEs and the full inversion using prior-sampling $\text{DREAM}_{(ZS)}$ proposals suffer from very low ARs and did not reach convergence after 200'000 iterations. Table 2.2 shows the number of iterations needed for the 99th percentile of the parameters' \hat{R} -statics to be below 1.2. It also shows the IACTs of the cell in the very middle of the porosity field for all inversion approaches reaching convergence within 200'000 MCMC iterations. We

observe that the iterations needed for convergence and the IACT of the LithTom IS method with prior-sampling $\text{DREAM}_{(ZS)}$ proposals are the lowest.

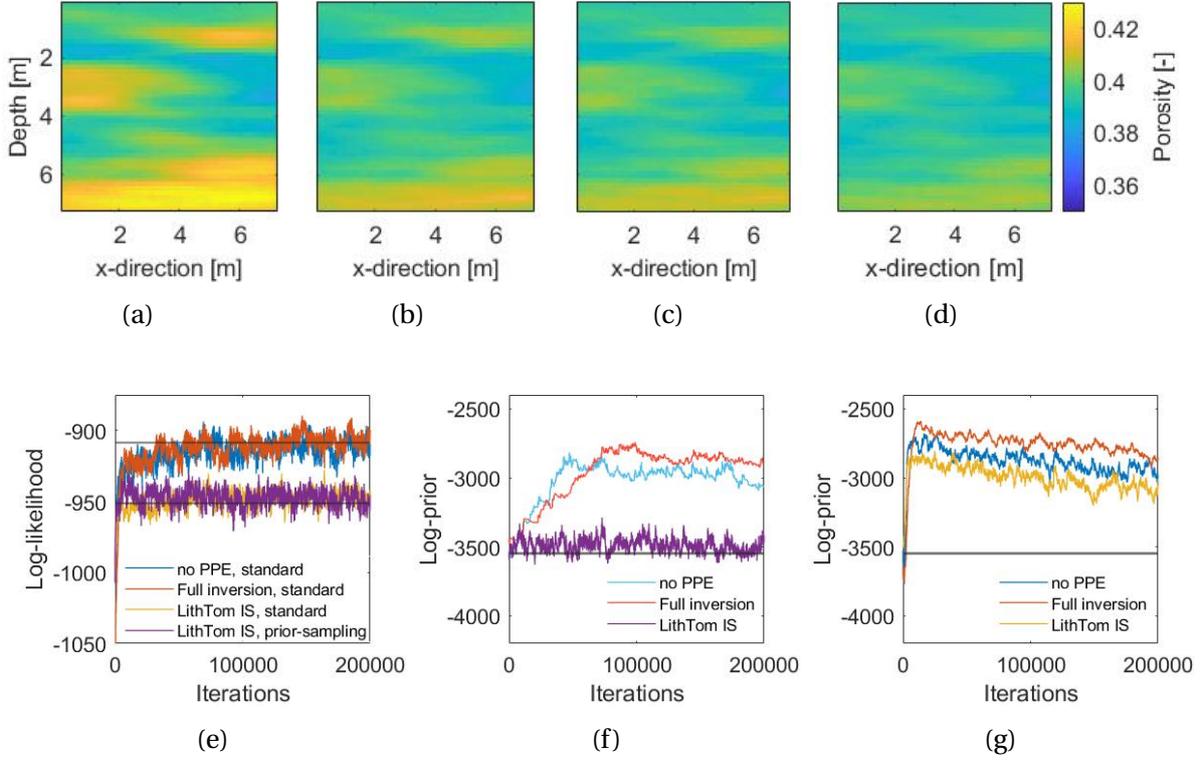


Figure 2.3: Estimated posterior means of the porosity field θ obtained for the linear test example with standard $\text{DREAM}_{(ZS)}$ proposals and (a) the algorithm ignoring PPEs, (b) the full inversion, (c) the LithTom IS method and with prior-sampling $\text{DREAM}_{(ZS)}$ proposals and (d) the LithTom IS method. (e) Corresponding log-likelihood values, black lines represent the values of $p(\mathbf{y}|\theta, \mathbf{x})$ and $p(\mathbf{y}|\theta)$ for the true porosity field θ (and the true \mathbf{X} in the former). (f) Logarithmically transformed prior probabilities for the posterior samples obtained with prior-sampling $\text{DREAM}_{(ZS)}$ proposals and (g) standard $\text{DREAM}_{(ZS)}$ proposals; the black lines depict the prior probability of the true porosity field.

Figure 2.3e shows the evolving log-likelihood values. When ignoring PPEs or performing the full inversion, the chains converge to much higher log-likelihoods than for the LithTom IS method. This is expected as they rely on the likelihood $p(\mathbf{y}|\theta, \mathbf{x})$ (where $\mathbf{X} = \mathcal{F}(\theta) + \epsilon_{\mathcal{D}}$, with $\epsilon_{\mathcal{D}} = 0$ for the algorithm ignoring PPEs), while LithTom IS estimates $p(\mathbf{y}|\theta) = \int p(\mathbf{y}|\theta, \mathbf{x})p(\mathbf{x}|\theta)d\mathbf{x}$. This example highlights that LithTom IS broadens the likelihood function. Figures 2.3f and 2.3g show the prior probabilities (logarithmically transformed) for the posterior samples obtained with the three different inversion approaches using the two alternative proposal schemes. We observe that the LithTom IS method using prior-sampling $\text{DREAM}_{(ZS)}$ proposals (Fig. 2.3f) is the only approach for which the prior probability of the true porosity field is sampled. All other methods and proposal scheme combinations sample

porosity fields with higher prior probabilities than the true field (black solid line). Practically speaking, this implies for these cases that none of the posterior samples are close to the true model. Furthermore, the corresponding prior probabilities show a trend of slowly decreasing values raising doubts about the ergodicity of the MCMC chains.

To compare the posterior PDFs with the analytical solution, we consider first histograms for an exemplary position in the porosity field and the KL-divergences of the whole field. We only show the results of the method and proposal-scheme combinations that converged within the considered 200'000 iterations. The histograms are depicted in Figure 2.4 with samples from the analytical posterior PDF (light grey) and samples from the respective inversion method (blue) for the pixel in the very middle of the model domain. The corresponding KL-divergences for all pixels are shown in Figure 2.5. The histogram and the KL-divergences of the method ignoring PPEs (with standard $\text{DREAM}_{(ZS)}$; Figures 2.4a and 2.5a) indicate that the approach suffers from biased estimates and an underestimation of the posterior variance. The posterior samples obtained with the full inversion method (with standard $\text{DREAM}_{(ZS)}$ proposals; Figures 2.4b and 2.5b) better represent the analytical posterior PDF, but there is still a significant underestimation of the posterior variance. The histogram obtained with the LithTom IS approach using standard $\text{DREAM}_{(ZS)}$ proposals (Figure 2.4c) is very similar to the one of the analytical posterior. The corresponding six-fold decreases of the KL-divergence (Figure 2.5c) compared with full inversion confirm the significant improvements of the exploration capabilities of this approach. An even better representation of the analytical posterior was obtained with the LithTom IS approach when using prior-sampling $\text{DREAM}_{(ZS)}$ proposals. This is indicated by the histogram in Figure 2.4d and by a further two-fold decrease of the KL-divergence in Figure 2.5d. An overview of the mean KL-divergences is given in Table 2.2.

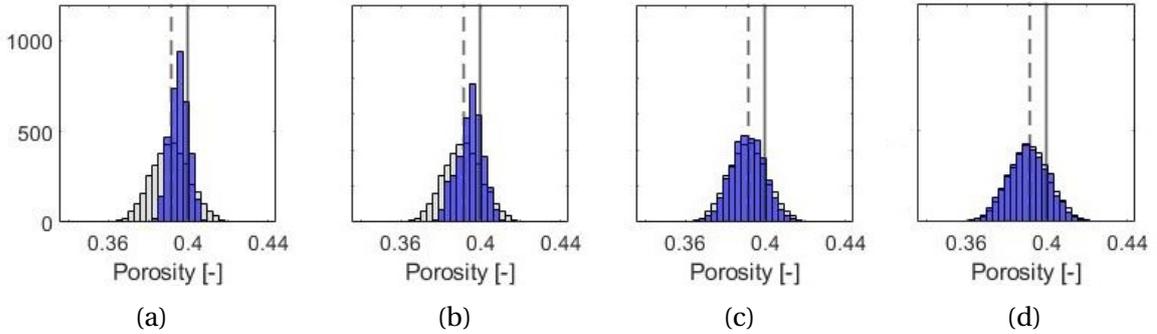


Figure 2.4: Histograms comparing samples from the analytical posterior PDF $p(\boldsymbol{\theta}|\mathbf{y})$ (light grey) for the linear test example and samples from the respective inversion method (blue), the solid line depicts the true value of the porosity in the very middle of the model domain and the dashed line indicates the analytical posterior mean (a) no PPE and standard $\text{DREAM}_{(ZS)}$ proposals, (b) full inversion and standard $\text{DREAM}_{(ZS)}$ proposals, (c) LithTom IS and standard $\text{DREAM}_{(ZS)}$ proposals and (d) LithTom IS and prior-sampling $\text{DREAM}_{(ZS)}$ proposals.

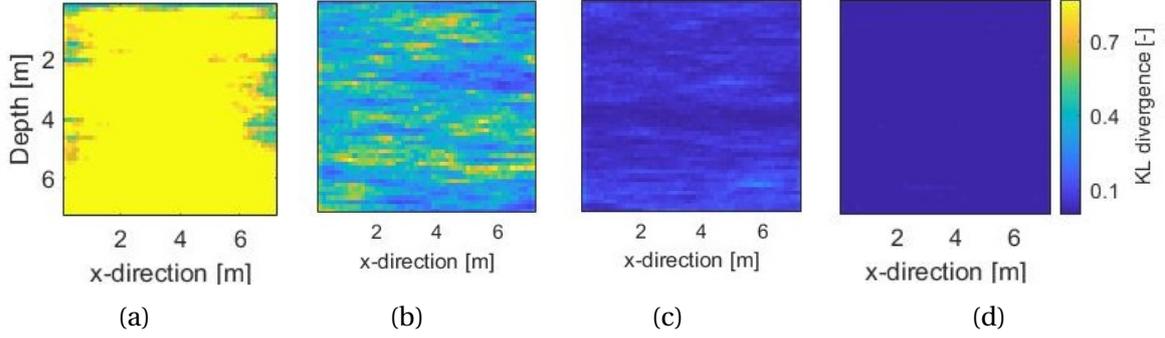


Figure 2.5: KL-divergences with respect to the analytical posterior PDF $p(\boldsymbol{\theta}|\mathbf{y})$ for the linear test example (a) no PPE and standard DREAM_(ZS) proposals, (b) full inversion and standard DREAM_(ZS) proposals, (c) LithTom IS and standard DREAM_(ZS) proposals and (d) LithTom IS and prior-sampling DREAM_(ZS) proposals.

This linear example has been used to show that importance sampling and prior-preserving proposal schemes are essential to obtain meaningful results in our considered high-dimensional setting. For this example, one can get accurate results using LithTom IS alone. The next section dealing with the non-linear case will serve to demonstrate the benefits of the CPM method in non-linear settings.

2.3.3 Non-linear physics

We now consider a non-linear test case in which the 625 arrival times are generated with the eikonal 2D traveltime solver *time2D* of Podvin and Lecomte (1991) such that,

$$\mathbf{y} = \mathcal{G}_\varepsilon(\mathbf{x}) + \boldsymbol{\varepsilon}_\theta. \quad (2.20)$$

Given the non-linear physics, the likelihood function $p(\mathbf{y}|\boldsymbol{\theta})$ is intractable and there is no analytical expression for the posterior PDF $p(\boldsymbol{\theta}|\mathbf{y})$ to compare with. The same applies for the PDF $p(\mathbf{x}|\mathbf{y}, \boldsymbol{\theta})$ that we previously used for the importance sampling of the latent variable \mathbf{X} . Hence, as importance sampling distribution we rely on the approximation of the PDF $p(\mathbf{x}|\mathbf{y}, \boldsymbol{\theta})$ introduced in Section 2.2.3. For $\mathbf{x}_{\text{lin}} = \mathcal{F}(\boldsymbol{\theta}_{\text{lin}}) + \boldsymbol{\varepsilon}_{\mathcal{D}_{\text{lin}}} = \frac{1}{c}(\sqrt{\kappa_s} + (\sqrt{\kappa_w} - \sqrt{\kappa_s})\boldsymbol{\theta}_{\text{lin}}) + \boldsymbol{\varepsilon}_{\mathcal{D}_{\text{lin}}}$, we use the last state of the porosity field for $\boldsymbol{\theta}_{\text{lin}}$ and the previous importance sampling mean $\boldsymbol{\mu}_{IS}$ for $\boldsymbol{\varepsilon}_{\mathcal{D}_{\text{lin}}}$. To decrease computational resources, we only update the linearization every 100 MCMC iterations. Since the expression is approximate, we further inflate the importance sampling covariance matrix $\boldsymbol{\Sigma}_{IS}$ by multiplying $\boldsymbol{\Sigma}_Y$ with a factor. After initial testing, we found that 1.2 yielded the best performance.

Table 2.2: Overview of the results obtained for the linear test example with the different inversion approaches and proposal mechanisms: The acceptance rates (**AR**), convergence (**Conv**) showing the number of iterations needed for the 99th percentile of the parameters' \hat{R} -statistics to be below 1.2 (or the percentage of parameters with a \hat{R} -statistics below 1.2 if the the inversion did not converge), the mean KL-divergence (**KL-div**) and the integrated autocorrelation time (**IACT**) for the cell in the very middle of the porosity field θ .

Method	Proposal	Parameter	AR	Conv	KL-div	IACT
No PPE	Standard	-	10 \nearrow 20 %	104'000	1.957	3'850
LithTom	Standard	$N = 1, \rho = 0$	< 0.1 %	-, 0 %	-	-
CPM no IS	Standard	$N = 10, \rho = 0.95$	0.1 %	-, 3 %	-	-
	Standard	$N = 50, \rho = 0.95$	0.2 %	-, 4 %	-	-
Full inversion	Standard	-	10 \nearrow 20 %	150'000	0.354	6'900
LithTom IS	Standard	$N = 1, \rho = 0$	20 \nearrow 30 %	78'000	0.063	2'750
no PPE	Prior-sampling	-	1 - 2 %	-, 35 %	-	-
Full inversion	Prior-sampling	-	1 - 2 %	-, 14 %	-	-
LithTom IS	Prior-sampling	$N = 1, \rho = 0$	13 %	76'000	0.003	1'700

Figure 2.6 depicts the dependence of the variance of the log-likelihood ratio estimator R (Eq. 2.10) on the correlation parameter ρ for $N = 1$, $N = 10$ and $N = 50$ samples of the latent variable \mathbf{X} (with θ being fixed at a region with high posterior probability mass). Figure 2.6a depicts estimates when drawing the realizations of the latent variable proportionally to its prior distribution $p(\mathbf{x}|\theta)$ and Figure 2.6b for the case where the latent variable is sampled with importance sampling. The two plots highlight three fundamental aspects of the CPM method in our geophysical setting. First, it is crucial to use a well chosen importance sampling for the latent variable draws, since for a correlation of, say, $\rho = 0$, the variance of the log likelihood ratio estimator can be reduced from values between 10'000 and 1'000'000 (using sampling from prior) to values between 3 and 31 (using importance sampling). Second, increasing the number of draws of latent variables (N) decreases the variance of the log-likelihood ratio estimator further and, third, this is also achieved by increasing the amount of correlation (ρ) used for two subsequent draws of latent variables. The variance for $\rho = 1$ is equal to zero for all parameter settings (as we use the same values for $\mathbf{X}^{(j-1)}$ and $\mathbf{X}^{(j)}$). Without importance sampling, we could still obtain a variance of the log-likelihood ratio estimator between 1 and 2 as recommended by Deligiannidis et al. (2018), but with the need of a very high N or a ρ very close to 1. In practice, this would either result in excessively high computational costs or slow mixing in the draws of the latent variables.

Due to the high variances displayed in Figure 2.6a and since the pseudo-marginal approaches without importance sampling have already proven to be highly inefficient in the linear case (Table 2.2), we now restrict ourselves only to CPM implementations involving IS. In stark contrast to the linear case, the LithTom IS approach ($N = 1, \rho = 0$) leads to a highly inefficient algorithm, as the variance of R around 30 is much higher than the upper recommended threshold of 2.0. For the CPM method, we set the number of samples to 10 and the correlation to $\rho = 0.95$ as this values leads to a variance of the log likelihood ratio estimator in-between 1.0 and 2.0. The autocorrelation of one cell of the latent variable field is given by $\text{Corr}(X_1, X_{1+l}) = \rho^l$ for lag l with the correlation mechanism of Equation (2.9), such that for $\rho = 0.95$ roughly 100 (accepted) iterations are needed to draw an independent realization of the latent variable. In practice, the decorrelation will be slower as we only move on with accepted proposals (Section 2.2.3).

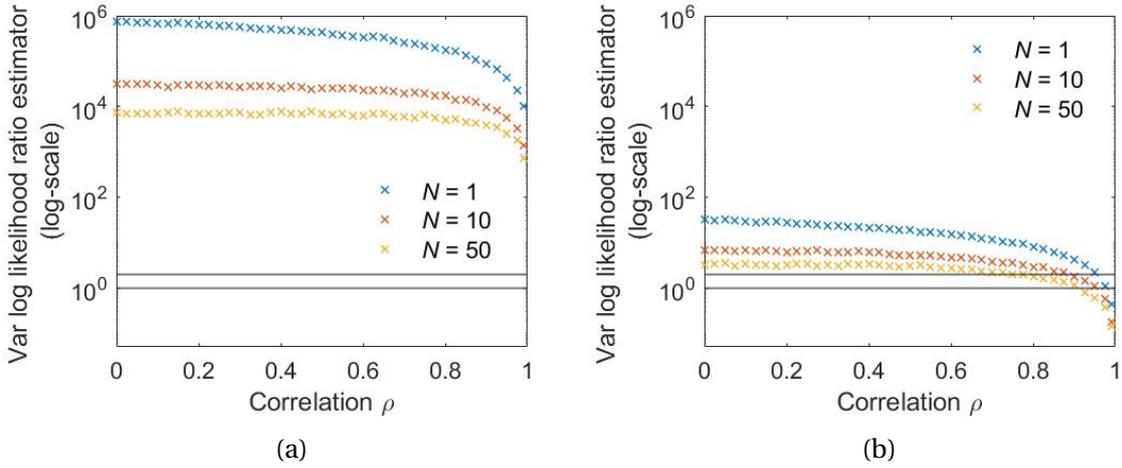


Figure 2.6: Variance of the log-likelihood ratio estimator $R = \log\left(\hat{p}_N^{(j)}(\mathbf{y}|\boldsymbol{\theta})\right) - \log\left(\hat{p}_N^{(j-1)}(\mathbf{y}|\boldsymbol{\theta})\right)$ for the non-linear test example and $\boldsymbol{\theta}$ fixed at a region with high posterior probability mass as a function of ρ (used to correlate the latent variables $\mathbf{X}^{(j)}$ and $\mathbf{X}^{(j-1)}$ as in Equation (2.9)) for $N = 1$, $N = 10$ and $N = 50$ samples of the latent variable \mathbf{X} ; the realizations of the latent variable are drawn (a) from the prior $p(\mathbf{x}|\boldsymbol{\theta})$ and (b) with importance sampling. The black lines delimit the range between 1.0 and 2.0 recommended by Deligiannidis et al. (2018).

The results for both $\text{DREAM}_{(ZS)}$ proposal schemes are shown in Figure 2.7 and Table 2.3. For the estimates of the posterior mean of the porosity field (Fig. 2.7a-2.7d), we observe similar results as in the linear case: Using prior-sampling $\text{DREAM}_{(ZS)}$ proposals results in a porosity field estimate with lower variance and using the method ignoring PPEs (Fig. 2.7a for standard $\text{DREAM}_{(ZS)}$ proposals) leads to higher variance. The highest acceptance rate is obtained with applying the CPM IS method using standard $\text{DREAM}_{(ZS)}$ proposals (Table 2.3) and the acceptance rates for prior-sampling $\text{DREAM}_{(ZS)}$ proposals are lower. The LithTom approach with IS has an AR of less than 1 % and would, therefore, require far more than 200'000 iterations to converge. Trace plots of the evolving log-likelihood values are shown in Figure 2.7e. As expected and in agreement with the linear test case (Fig. 2.3e), the methods converge to different values. As in the linear case, we find that CPM IS with prior-sampling $\text{DREAM}_{(ZS)}$ proposals is the only case providing posterior samples that match the prior probability of the true porosity field (Fig. 2.7f and 2.7g).

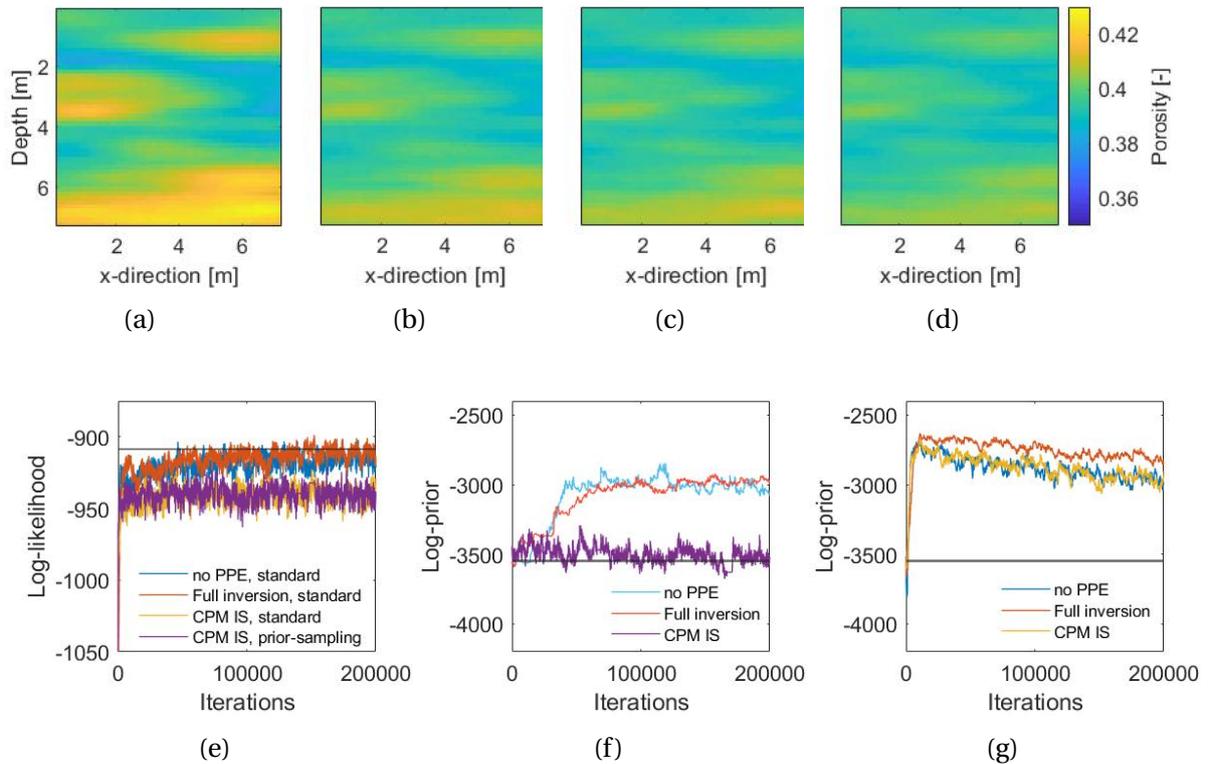


Figure 2.7: Estimates of the posterior means of the porosity field θ for the non-linear test example resulting with standard $\text{DREAM}_{(ZS)}$ proposals and (a) the algorithm ignoring PPEs, (b) the full inversion, (c) the CPM IS ($N = 10, \rho = 0.95$) method. Results for prior-sampling $\text{DREAM}_{(ZS)}$ proposals and (d) the CPM IS ($N = 10, \rho = 0.95$) method. (e) Log-likelihood functions, black line represents the value of $p(\mathbf{y}|\theta, \mathbf{x})$ for the true porosity and latent variable field. (f) Prior probabilities (logarithmically transformed) of the posterior samples obtained with prior-sampling $\text{DREAM}_{(ZS)}$ proposals and (g) standard $\text{DREAM}_{(ZS)}$ proposals; the black lines depict the corresponding value for the true porosity field.

Table 2.3: Summary of the results obtained for the non-linear test example with the various inversion approaches and the two proposal mechanisms: The acceptance rates (**AR**), the convergence (**Conv**) showing the number of iterations needed for the 99th percentile of the parameter’s \hat{R} -statics to be below 1.2 (or the percentage of parameters with a \hat{R} -statics below 1.2 if the the algorithm did not converge), the percentage of pixels in which the true porosity value lies within the range of posterior samples (θ_{true}), the mean logarithmic score (**logS**), posterior standard deviation (**Post SD**) and the integrated autocorrelation time (**IACT**) for the cell in the very middle of the porosity field θ . The CPM IS method was evaluated with the parameter choice of $N = 10$ and $\rho = 0.95$.

Method	Proposal	AR	Conv	θ_{true}	logS	Post SD	IACT
No PPE	Standard	11 \nearrow 24 %	92’000	87.2 %	3.36	5.4×10^{-3}	3’800
Full inversion	Standard	10 \nearrow 23 %	144’000	97.1 %	1.99	6.7×10^{-3}	5’150
LithTom IS	Standard	< 1 %	- , 43 %	-	-	-	-
CPM IS	Standard	12 \nearrow 24 %	90’000	99.6 %	1.56	8.3×10^{-3}	3’250
No PPE	Prior-samp	1 - 2 %	- , 29 %	-	-	-	-
Full inversion	Prior-samp	1 - 2 %	- , 13 %	-	-	-	-
CPM IS	Prior-samp	11 %	96’000	100.00 %	1.34	10.4×10^{-3}	3’300

Figure 2.8 depicts the logarithmic scores (see Section 2.2.5) comparing the true porosity values with the inferred posterior PDFs for all 2500 grid cells. We observe that the method ignoring PPEs (with standard DREAM_(ZS) proposals, Fig. 2.8a) has the highest scores (indicating the lowest accuracy). The values of the full inversion (with standard DREAM_(ZS) proposals, Fig. 2.8b) are lower, but still high. The CPM IS method with standard DREAM_(ZS) proposals (Figs. 2.8c) leads to reduced logarithmic scores that are further improved when this method is combined with prior-sampling DREAM_(ZS) proposals (Figs. 2.8d). The mean values of the logarithmic scores and other performance metrics are shown in Table 2.3. We find that the method that ignores PPEs fails to sample a range of values including the true porosity value in more than 10% of the pixels and has a mean estimated posterior standard deviation that is up to 50 % smaller than the other methods. The CPM IS method generates posterior samples with ranges that include, in more than 99 % of the pixels, the true porosity value with the percentages obtained using prior-sampling DREAM_(ZS) proposals being even higher. Finally, the full inversion does not sample the true porosity value in almost 3% of the pixels and has a reduced mean estimated posterior standard deviation by up to 40 % compared to the CPM IS method. We also note that the IACT of the CPM methods are the lowest (Table 2.3).

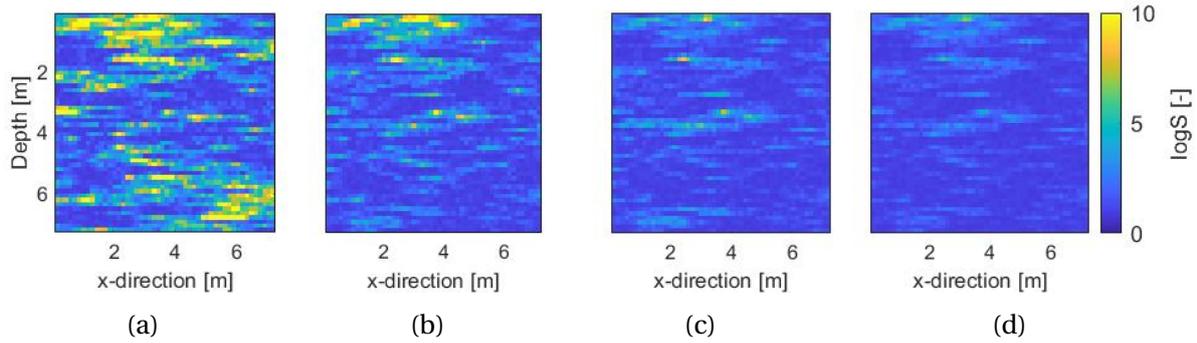


Figure 2.8: The logarithmic scores for the non-linear test case with (a) no PPE and standard $\text{DREAM}_{(ZS)}$ proposals, (b) full inversion and standard $\text{DREAM}_{(ZS)}$ proposals, (c) CPM IS and standard $\text{DREAM}_{(ZS)}$ proposals and (d) CPM IS and prior-sampling $\text{DREAM}_{(ZS)}$ proposals.

2.4 Discussion

This study showed clearly that the correlated pseudo-marginal (CPM) method, which accounts for petrophysical prediction uncertainty within the estimate of the likelihood function $p(\mathbf{y}|\boldsymbol{\theta})$, combined with importance sampling (IS) and prior-sampling MCMC proposals leads to a broader exploration of the target posterior $p(\boldsymbol{\theta}|\mathbf{y})$ than the other presented combinations of inversion methods and proposal schemes. The CPM method is an exact and general method, but it needs in the considered high-dimensional setting an efficient importance sampling and prior-sampling proposals to work well even for the case of linear physics.

In the linear setting (with available analytical solutions for the PDFs), the CPM method using importance sampling performs well using only one uncorrelated sample of the PPE (LithTom IS). In absence of importance sampling, even a high number of samples N and correlation ρ could not prevent the algorithm from being highly inefficient (Table 2.2). We find that the exploration of the posterior PDF is much improved when using the LithTom IS approach compared with full inversion (Fig. 2.4 and Fig. 2.5). Although the \hat{R} -statistic of Gelman and Rubin (1992) suggests that the full inversion algorithm (using standard $\text{DREAM}_{(ZS)}$ proposals) has converged, we demonstrate a significant underestimation of the posterior standard deviation and posterior samples with far too high prior probabilities compared with the true model (Fig. 2.3f and 2.3g). Indeed, the full inversion's high acceptance rate (for standard $\text{DREAM}_{(ZS)}$ proposals) may be mainly a consequence of local exploration combined with an adaptive MCMC expanding its archive. This (1) points out that Gelman-Rubin's \hat{R} -statistics and the acceptance rate are insufficient metrics to assess the performance of an adaptive MCMC algorithm such as $\text{DREAM}_{(ZS)}$ and (2) highlights issues with over-fitting when using adaptive MCMC. Indeed, Robert et al. (2018) warn against using adaptive MCMC methods without due caution as adaptations to the proposal scheme can lead to algorithms relying too much on previous iterations, thereby, excluding parts of the parameter space that have not yet been explored.

The need for a well-chosen importance sampling distribution is also demonstrated for the non-linear setting by analysing the variances of the log-likelihood ratio estimator (Fig. 2.6). This analysis also confirmed the strong influence of N and ρ . Since the importance sampling distribution is no longer exact in the non-linear test case, the number of samples N and the correlation ρ need to be increased. Consequently, the CPM IS method performs better (in terms of computational cost) than the PM IS method as fewer samples have to be used. For the non-linear test case, we conclude that the exploration of the posterior with the CPM IS method (especially when combined with prior-sampling DREAM_(ZS) proposals) is better than the full inversion by observing that (1) the range of the posterior samples includes more often the true porosity value while (2) the logarithmic score is lower and (3) the mean estimated posterior standard deviation is higher (Table 2.3).

We recommend to work in the full parameter space whenever possible such that any distortions in the posterior estimations due to model reductions can be avoided. The presented adaptive prior-preserving proposal scheme (prior-sampling DREAM_(ZS) proposal) is developed in the spirit of the extended Metropolis algorithm of Mosegaard and Tarantola (1995) and the pCN proposal of Cotter et al. (2013). It is a simple correction of the standard DREAM_(ZS) proposal that (1) makes the algorithm robust to the choice of the discretization of the target field and (2) maintains its capabilities to sample efficiently in complex high-dimensional parameter spaces. We find that the prior-sampling DREAM_(ZS) proposals lead to an enhanced exploration of the posterior PDF and a stable AR (Tables 2.2 and 2.3). Indeed, the CPM IS approach with prior-sampling proposals is the only one generating samples with a prior probability comparable to the one of the true porosity field (Figs. 2.3 and 2.7). Due to dependencies between latent and target variables, the full inversion with prior-sampling DREAM_(ZS) proposals suffers from a very low acceptance rate as the method does not allow for large proposal steps. This dependency is bypassed by the CPM IS, allowing larger steps for a given AR. In general, combinations of adaptive Metropolis and pCN-proposals are referred to as DIAM (dimension independent adaptive Metropolis) proposals and were introduced by Chen et al. (2016). Another way to increase the efficiency of the pCN proposal was proposed by Rudolf and Sprungk (2018) with the so-called generalized pCN-proposal (gpCN), in which the proposal scheme is tuned to have the same covariance as the target posterior distribution.

We emphasize that this study only considers synthetic data. We demonstrate that all but our method of choice (CPM IS with prior-sampling DREAM_(ZS) proposals) have severe problems in exploring the full posterior distribution even in this well-specified setting. A field demonstration of CPM IS with prior-sampling DREAM_(ZS) proposals is a natural next step. Furthermore, our entire study remains within Gaussian assumptions for the target field, petrophysical prediction uncertainty and observational noise. In the presented results, we deal only with weak non-linearity in our forward operator and assume the petrophysical relationship to be linear. In the future, it would be useful to consider test cases involving stronger non-linearity, be it through a higher variance of the slowness field or a non-linear petrophysical relationship. Stronger non-linearity would affect the accuracy of the first-order expansion used to derive the importance sampling distribution for the CPM method, implying that the approximations would become less accurate. This could lead to a decrease of efficiency that could be counter-acted by using larger N or ρ . An important topic for future

research would be to develop and assess importance sampling schemes that do not rely on Gaussian assumptions. Potential starting points could be efficient importance sampling by Richard and Zhang (2007) or multiple importance sampling introduced by Veach and Guibas (1995) and popularised by Owen and Zhou (2000).

In agreement with Brunetti and Linde (2018), we find that ignoring petrophysical prediction uncertainty leads to biased estimates and too tight uncertainty bounds. While the need for a method accounting for PPEs grows with increasing integral scale of the target field (Brunetti and Linde 2018), the ratio of the variances of the PPE, the target variable and the observational noise also influences the results. The need for a well-working importance sampling for CPM grows with increasing petrophysical prediction uncertainty and decreasing observational noise. At the same time, large petrophysical prediction uncertainty leads to a flattened likelihood function $p(\mathbf{y}|\boldsymbol{\theta})$, thereby, decreasing the variance of the likelihood estimators (assuming a well-working importance sampling) and, therefore, enhancing the efficiency of the algorithm. Our present work focuses on petrophysical prediction uncertainty for a known covariance model, but it would be possible to expand this to an unknown covariance model, an uncertain petrophysical model or uncertain model parameters.

2.5 Conclusions

We consider lithological tomography in which geophysical data are used to infer the posterior PDF of target (hydro)geological parameters. In such a latent variable model, the geophysical properties play the role of latent variables that are linked to the properties of interest through petrophysical relationships exhibiting significant scatter. Compared with the original formulation of lithological tomography that does not consider importance sampling, we make the approach more applicable to high dimensions (thousands of unknowns) and large data sets with high signal-to-noise ratios. To account for the intractable likelihood appearing in the Metropolis–Hastings algorithm in this setting, we explore the correlated pseudo-marginal (CPM) method using an importance sampling distribution and prior-sampling proposals. For the latter, we adapt the standard (adaptive) proposal scheme of DREAM_(ZS) with a prior-sampling approach, leading to a further improvement in exploration compared with standard model proposals when dealing with high-dimensional problems. We find that our implementation of the CPM method outperforms standard lithological tomography and the full inversion approach, which parameterizes and infers the posterior petrophysical prediction uncertainty. For a linear test example, the mean KL-divergence with respect to the analytical posterior can be reduced by 99 % by our implementation of the CPM method (even without using correlations) compared with full inversion. In the case of non-linear physics, we reduce the mean logarithmic score with respect to the true porosity field by up to 33 % compared with the full inversion method. The CPM method is generally applicable and accurate, but it requires a well-working importance sampling distribution (presently based on Gaussian random field theory) to be efficient. Future work with the CPM method could consider field data applications, more non-linear physics and non-linear petrophysical relationships as well as relaxing the assumptions of Gaussian random fields. Furthermore, the method’s use in coupled hydrogeophysical inversions involving hydrogeological flow and transport models would be of interest.

2.6 Appendix

2.6.1 DREAM algorithms and prior-sampling proposals

To perform a high-dimensional inversion with the MH algorithm, one needs a well-working proposal scheme. To deal with this challenge, Braak (2006) introduced an adaptive random walk MH algorithm named Differential Evolution Markov chain (DE-MC). This method runs C Markov chains in parallel, where at each iteration j , the C different realizations of the model parameters define a population $\{\mathbf{Z}_c^{(j)}; c = 1, 2, \dots, C\}$, which is used to guide new model proposals. For chain c , two chains (denoted as a and b) are drawn without replacement from the remaining set of chains. Then, the algorithm proposes a new state for the c -th chain with,

$$\mathbf{Z}_c^{(j)} = \mathbf{Z}_c^{(j-1)} + \gamma(\mathbf{Z}_a^{(j-1)} - \mathbf{Z}_b^{(j-1)}) + \zeta, \quad c \neq a \neq b \quad (2.21)$$

where γ denotes the jumping rate and ζ is a draw from $\mathcal{N}(0, s^2)$ with a small standard deviation s used to ensure that the resulting Markov chain is irreducible. By accepting or rejecting the resulting proposals with the MH-ratio of Equation (2.3), a Markov chain with the posterior PDF as its stationary distribution is obtained (Proof in Vrugt et al. 2009). This leads to an algorithm which is automatically adapting the scale and the orientation of the proposal density along the way to the stationary distribution, allowing it to provide efficient sampling on complex, high-dimensional, and multi-modal target distributions. Based on the DE-MC, Vrugt et al. (2008) introduced the adaptive multi-chain MCMC algorithm called DREAM (DiffeREntial Evolution Adaptive Metropolis). It enhances the efficiency of DE-MC by applying subspace sampling (only randomly selected dimensions of the model parameter are updated) and outlier chain correction. An excellent overview of the theory and application of the DREAM algorithm is given by Vrugt (2016). For our case study, we use the extended version DREAM_(ZS) introduced by Laloy and Vrugt (2012), as its proposal scheme using an archive of past states leads to further improved convergence and posterior exploration.

To adapt extended Metropolis to DREAM_(ZS), we rely on a transformation of the variables to the Uniform space. In our case study with Gaussian target variable $\mathbf{Z}_c^{(j)} = (Z_{c;1}^{(j)}, Z_{c;2}^{(j)}, \dots, Z_{c;D^2}^{(j)})$ sampled in chain c and iteration j , we define $U_{c;i}^{(j)} = \Phi(Z_{c;i}^{(j)})$, with $\Phi(\cdot)$ being the standard-normal cumulative distribution function (CDF), and apply the proposal mechanism of DREAM_(ZS) on this transform. Assuming that $Z_{c;i}^{(j)}$ has a standard-normal distribution, $U_{c;i}^{(j)}$ will be distributed uniformly on $[0, 1]$. The proposal scheme of DREAM_(ZS) with so-called fold boundary handling (i.e., periodic boundary conditions) ensures that the new state $U_{c;i}^{(j+1)}$ is a sample from the Uniform distribution as well. With the subsequent transformation back to the standard normal, $Z_{c;i}^{(j+1)} = \Phi^{-1}(U_{c;i}^{(j+1)})$, we hence force the algorithm to use a proposal scheme that samples from the prior PDF.

2.6.2 Analytical posterior PDF and importance density for linear physics

Assuming linear physics and petrophysics, it is possible to derive an analytical expression for the posterior PDF $p(\boldsymbol{\theta}|\mathbf{y})$ of the porosity (or other variable of interest). We consider here both relationships being linear without intercept ($\mathcal{G}(\mathbf{X}) = \mathbf{J}_s \mathbf{X}$ and $\mathcal{F}(\boldsymbol{\theta}) = \mathbf{J}_p \boldsymbol{\theta}$), however, an intercept (as the one used for $\mathcal{F}(\boldsymbol{\theta})$ in our test case; Section 2.3.1) is easily included. For the 2D grid of the porosity $\boldsymbol{\theta}$ and the latent variable \mathbf{X} , we use the following prior PDFs:

$$p(\boldsymbol{\theta}) = \varphi_{D^2}(\boldsymbol{\theta}; \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta), \quad p(\mathbf{x}|\boldsymbol{\theta}) = \varphi_L(\mathbf{x}; \mathbf{J}_p \boldsymbol{\theta}, \boldsymbol{\Sigma}_p). \quad (2.22)$$

To derive the (in this case) tractable likelihood $p(\mathbf{y}|\boldsymbol{\theta})$, we use a standard result about marginal and conditional Gaussians (Bishop and Nasrabadi 2006):

Lemma 1 Marginal and Conditional Gaussians

Assume a marginal Gaussian distribution for $\mathbf{X} \in \mathbb{R}^L$ and a conditional Gaussian distribution for $\mathbf{Y} \in \mathbb{R}^T$ given \mathbf{X} in the form

$$p(\mathbf{x}) = \varphi_T(\mathbf{x}; \boldsymbol{\mu}, \Lambda^{-1}), \quad (2.23)$$

$$p(\mathbf{y}|\mathbf{x}) = \varphi_T(\mathbf{y}; \mathbf{A}\mathbf{x} + \mathbf{b}, \mathbf{L}^{-1}), \quad (2.24)$$

with $\varphi_T(\cdot; \boldsymbol{\mu}, \mathbf{K})$ denoting the PDF of the T -variate Normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix \mathbf{K} . Then, the marginal distribution of \mathbf{Y} and the conditional distribution of \mathbf{X} given \mathbf{Y} are given by

$$p(\mathbf{y}) = \varphi_T(\mathbf{y}; \mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{L}^{-1} + \mathbf{A}\Lambda^{-1}\mathbf{A}^T) \quad (2.25)$$

$$p(\mathbf{x}|\mathbf{y}) = \varphi_L(\mathbf{x}; \boldsymbol{\Sigma}(\mathbf{A}^T\mathbf{L}(\mathbf{y} - \mathbf{b}) + \Lambda\boldsymbol{\mu}), \boldsymbol{\Sigma}) \quad (2.26)$$

where $\boldsymbol{\Sigma} = (\Lambda + \mathbf{A}^T\mathbf{L}\mathbf{A})^{-1}$.

Using the prior on the latent variable \mathbf{X} and the Gaussian likelihood $p(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta}) = \varphi_{625}(\mathbf{y}; \mathbf{J}_s \mathbf{x}, \boldsymbol{\Sigma}_Y)$, we get with Equation (2.25),

$$p(\mathbf{y}|\boldsymbol{\theta}) = \varphi_T(\mathbf{y}; \mathbf{J}_s \mathbf{J}_p \boldsymbol{\theta}, \boldsymbol{\Sigma}_Y + \mathbf{J}_s \boldsymbol{\Sigma}_p \mathbf{J}_s^T). \quad (2.27)$$

Subsequently, the analytical form of the posterior $p(\boldsymbol{\theta}|\mathbf{y})$ is derived with Equation (2.26), the prior on porosity and the expression of the likelihood $p(\mathbf{y}|\boldsymbol{\theta})$ from the last equation:

$$p(\boldsymbol{\theta}|\mathbf{y}) = \varphi_{D^2}(\boldsymbol{\theta}; \boldsymbol{\mu}_{\boldsymbol{\theta}|\mathbf{Y}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}|\mathbf{Y}}), \quad (2.28)$$

$$\boldsymbol{\mu}_{\boldsymbol{\theta}|\mathbf{Y}} = \boldsymbol{\Sigma}_{\boldsymbol{\theta}|\mathbf{Y}} \left((\mathbf{J}_s \mathbf{J}_p)^T (\boldsymbol{\Sigma}_Y + \mathbf{J}_s \boldsymbol{\Sigma}_p \mathbf{J}_s^T)^{-1} \mathbf{y} + \boldsymbol{\Sigma}_\theta^{-1} \boldsymbol{\mu}_\theta \right), \quad (2.29)$$

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}|\mathbf{Y}} = \left(\boldsymbol{\Sigma}_\theta^{-1} + (\mathbf{J}_s \mathbf{J}_p)^T (\boldsymbol{\Sigma}_Y + \mathbf{J}_s \boldsymbol{\Sigma}_p \mathbf{J}_s^T)^{-1} (\mathbf{J}_s \mathbf{J}_p) \right)^{-1} \quad (2.30)$$

For the case with linear physics, the importance density $\tilde{p}(\mathbf{x}|\boldsymbol{\theta}, \mathbf{y}) = \varphi_L(\mathbf{x}; \boldsymbol{\mu}_{IS}, \boldsymbol{\Sigma}_{IS})$ introduced in Section 2.2.3 is an exact expression for $p(\mathbf{X}|\boldsymbol{\theta}, \mathbf{y})$ and the IS mean and covariance matrix reduce to:

$$\boldsymbol{\mu}_{IS} = \boldsymbol{\Sigma}_{IS} \left(\mathbf{J}_s^T \boldsymbol{\Sigma}_Y^{-1} \mathbf{y} + \boldsymbol{\Sigma}_p^{-1} \mathcal{F}(\boldsymbol{\theta}) \right), \quad \boldsymbol{\Sigma}_{IS} = (\boldsymbol{\Sigma}_p^{-1} + \mathbf{J}_s^T \boldsymbol{\Sigma}_Y^{-1} \mathbf{J}_s)^{-1}. \quad (2.31)$$

Chapter 3

Solving geophysical inversion problems with intractable likelihoods: Linearized Gaussian approximations versus the correlated pseudo-marginal method

Lea Friedli and Niklas Linde

Published in *Mathematical Geosciences*:

Friedli, L. and Linde, N. (2023). Solving geophysical inversion problems with intractable likelihoods: Linearized Gaussian approximations versus the correlated pseudo-marginal method. *Mathematical Geosciences*, <https://doi.org/10.1007/s11004-023-10064-y>.

Abstract

A geophysical Bayesian inversion problem may target the posterior distribution of geological or hydrogeological parameters given geophysical data. To account for the scatter in the petrophysical relationship linking the target parameters to the geophysical properties, this study treats the intermediate geophysical properties as latent (unobservable) variables. To perform inversion in such a latent variable model, the intractable likelihood function of the (hydro)geological parameters given the geophysical data needs to be estimated. This can be achieved by approximation with a Gaussian probability density function based on local linearization of the geophysical forward operator, thereby, accounting for the noise in the petrophysical relationship by a corresponding addition to the data covariance matrix. The new approximate method is compared against the general correlated pseudo-marginal method, which estimates the likelihood by Monte Carlo averaging over samples of the latent variable. First, the performances of the two methods are tested on a synthetic test example, in which a multivariate Gaussian porosity field is inferred using crosshole ground-penetrating radar first-arrival travel times. For this example with rather small petrophysical uncertainty, the two methods provide near-identical estimates, while an inversion that ignores petrophysical uncertainty leads to biased estimates. The results of a sensitivity analysis are then used to suggest that the linearized Gaussian approach, while attractive due to its relative computational speed, suffers from a decreasing accuracy with increasing scatter in the petrophysical relationship. The computationally more expensive correlated pseudo-marginal method performs very well even for settings with high petrophysical uncertainty.

3.1 Introduction

This work targets a Bayesian inverse problem in which the posterior distribution of target geological or hydrogeological parameters $\boldsymbol{\theta}$ are inferred from geophysical data \mathbf{y} . Petrophysical relationships linking (hydro)geological variables (e.g., permeability, clay fraction, salinity) to geophysical properties (e.g., dielectric permittivity, electrical conductivity, magnetic susceptibility) must then be introduced (e.g., Hinnell et al. 2010; Kowalsky et al. 2005). Such relationships are often inherently uncertain (e.g., Mavko et al. 2020), however, in most hydrogeophysical inversion studies targeting hydrogeological properties, the predictive power of the petrophysical relationship is assumed to be perfect provided that the right parameter values are used (e.g., Lochbühler et al. 2014; Kowalsky et al. 2005). Brunetti and Linde (2018) show that this assumption may lead to bias, too narrow uncertainty bounds and overly variable parameter estimates.

Brunetti and Linde (2018) distinguish three sources of uncertainty in the petrophysical relationship: model, parameter and prediction uncertainty. While the first two result from uncertainty related to the choice of the petrophysical model and its parameter values, the latter arises from scatter and bias around the calibrated model. As in Brunetti and Linde (2018), only petrophysical prediction uncertainty is considered here, using a latent variable model formulation which expresses the geophysical properties as $\mathbf{X} = \mathcal{F}(\boldsymbol{\theta}) + \boldsymbol{\varepsilon}_{\mathcal{F}}$, with $\mathcal{F}(\cdot)$ being the petrophysical relationship and $\boldsymbol{\varepsilon}_{\mathcal{F}}$ the petrophysical prediction error (PPE). The inclusion

of the random effect of the PPE in the latent variable \mathbf{X} makes the likelihood function $p(\mathbf{y}|\boldsymbol{\theta})$ intractable. In this study, two alternative methods are investigated to approximate this likelihood function in a Metropolis–Hastings algorithm (MH algorithm; Metropolis et al. 1953, Hastings 1970). The first approach is a Gaussian approximation based on local linearization of the geophysical forward operator. Thereby, the effect of the noise in the petrophysical relationship is included by a corresponding addition to the data covariance matrix. This new approach which is similar to the so-called delta method (Van der Vaart 2000) was suggested by Linde et al. (2017), but it has remained untested to date. This approximate method is compared against the correlated pseudo-marginal (CPM) method of Deligiannidis et al. (2018), which is based on the pseudo-marginal (PM; Beaumont 2003; Andrieu and Roberts 2009) method using Monte Carlo sampling of the latent variable to estimate the likelihood. Friedli et al. (2022) introduced and adapted the CPM method to a geophysical setting and demonstrated that in data-rich geophysical settings with low noise levels, it is essential to both use a well-working importance sampling strategy for the draws of latent variables and to correlate the latent samples used in the proposed and current states of the Markov chain.

In Friedli et al. (2022), the CPM method is compared to the original formulation of lithological tomography (Bosch 1999) and the so-called full inversion approach of Brunetti and Linde (2018). This latter method avoids intractable likelihood functions by targeting the joint posterior PDF $(\boldsymbol{\theta}, \mathbf{x}) \mapsto p(\boldsymbol{\theta}, \mathbf{x}|\mathbf{y})$ of the hydrogeological and geophysical parameters. Within the original lithological tomography method, first the target variable is sampled using the proposal scheme of the MH and second, one realization of the latent variable is drawn with conditional sampling. This actually represents a simplified form of the PM method with only one latent variable sample and without importance sampling. The original form of lithological tomography leads to high variability in the estimate of the likelihood function, with the consequence of the algorithm often being highly inefficient (Brunetti and Linde 2018). Within the full inversion of Brunetti and Linde (2018), the latent variables are treated as additional target variables and an MH proposal scheme is used to draw new realizations of both. Friedli et al. (2022) show that this approach becomes inefficient with increasing dimensionality of the target and latent space and suffers from strong (posterior) correlations between the target and latent variables. Friedli et al. (2022) present a comparison of CPM with the original lithological tomography and full inversion approaches in a weakly non-linear setting showing that the CPM method outperforms the others by greatly enhancing the posterior exploration. While the CPM method already has been tested for geophysical inversion problems, the linearized Gaussian approach has not been applied to far and in this present study, the focus is on comparing this approximate approach against the general CPM method.

As a synthetic test case, a similar setting as in Friedli et al. (2022) is considered and multi-Gaussian porosity fields are inferred using crosshole ground-penetrating radar (GPR) first-arrival travel times. As in Friedli et al. (2022), a high-dimensional parameterization of the target porosity field is used. Subsequently, a sensitivity analysis is made to explore the performances of the linearized Gaussian approach and the CPM method as a function of increasing petrophysical prediction uncertainty. To avoid the challenges of a very-high dimensional target space in this sensitivity analysis, the complexity of the porosity field is reduced and it is assumed to be layered.

This contribution is a natural extension of the study by Friedli et al. (2022), which only considered the CPM method. While the fundamental concepts of the considered problem and the CPM method are repeated, the introduction and assessment of the linearized Gaussian approach is completely new. The manuscript is structured as follows: Section 3.2 gives a methodological overview of the considered latent variable model, Bayesian inference with intractable likelihoods, the linearized Gaussian approximation approach, the CPM method and the performance assessment metrics. Section 3.3 presents the results of our synthetic case study inferring multi-Gaussian porosity fields. In Section 3.4, the sensitivity analysis is presented. Finally, the study finishes with a discussion and conclusions in Sects. 3.5 and 3.6.

3.2 Methodology

3.2.1 Latent variable model

In the considered setting, the data vector $\mathbf{y} = (y_1, y_2, \dots, y_T)$ (geophysical data) is given by,

$$\mathbf{Y} = \mathcal{G}(\mathbf{X}) + \boldsymbol{\varepsilon}_\theta = \mathcal{G}(\mathcal{F}(\boldsymbol{\theta}) + \boldsymbol{\varepsilon}_\mathcal{F}) + \boldsymbol{\varepsilon}_\theta, \quad (3.1)$$

with $\mathbf{X} = (X_1, X_2, \dots, X_L)$ denoting the latent variable (geophysical property) and $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_d)$ the target parameters (hydrogeological parameters). The variable \mathbf{X} is referred to as a latent because it includes the PPE $\boldsymbol{\varepsilon}_\mathcal{F}$, which is unobservable (latent) but affects the observations. $\mathcal{G} : \mathbb{R}^L \rightarrow \mathbb{R}^T$ with $\boldsymbol{\varepsilon}_\theta$ and $\mathcal{F} : \mathbb{R}^d \rightarrow \mathbb{R}^L$ with $\boldsymbol{\varepsilon}_\mathcal{F}$ refer to the physical forward solver with the observational noise and the petrophysical relationship with the PPE, respectively. In what follows, random variables and random vectors are referred to with upper-case letters and realizations thereof with lower-case letters. Assuming Gaussian errors, it holds,

$$p(\mathbf{x}|\boldsymbol{\theta}) = \varphi_L(\mathbf{x}; \mathcal{F}(\boldsymbol{\theta}), \boldsymbol{\Sigma}_\mathcal{P}), \quad p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x}) = \varphi_T(\mathbf{y}; \mathcal{G}(\mathbf{x}), \boldsymbol{\Sigma}_\mathcal{Y}), \quad (3.2)$$

with $\varphi_M(\cdot; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denoting the PDF of a M -variate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. In the test example, the target parameters $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_d)$ describe a Gaussian random field parameterized on a grid of size $D \times D$ ($d = D^2$). It holds,

$$p(\boldsymbol{\theta}) = \varphi_{D^2}(\boldsymbol{\theta}; \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta), \quad (3.3)$$

and it is assumed that the mean $\boldsymbol{\mu}_\theta$ and the covariance matrix $\boldsymbol{\Sigma}_\theta$ of the target field are known.

3.2.2 Bayesian inference and intractable likelihoods

Bayesian inversion problems target the posterior probability density function (PDF) $p(\boldsymbol{\theta}|\mathbf{y})$ of the model parameters $\boldsymbol{\theta}$ given the measurements \mathbf{y} . In Bayes' theorem, this posterior PDF is given by,

$$p(\boldsymbol{\theta}|\mathbf{y}) = \frac{p(\boldsymbol{\theta})p(\mathbf{y}|\boldsymbol{\theta})}{p(\mathbf{y})}, \quad (3.4)$$

with the prior PDF $p(\boldsymbol{\theta})$ of the model parameters, the likelihood function $p(\mathbf{y}|\boldsymbol{\theta})$ and the evidence $p(\mathbf{y})$. As it is not possible to sample directly from the posterior, we use the Metropolis–Hastings algorithm (MH algorithm; Metropolis et al. 1953; Hastings 1970). At iteration j , the MH algorithm proposes a new model realization using the model proposal density $q(\cdot|\boldsymbol{\theta}^{(j-1)})$, which is then accepted or rejected based on the acceptance probability,

$$\alpha_{MH}(\boldsymbol{\theta}^{(j-1)}, \boldsymbol{\theta}^{(j)}) = \min\left\{1, \frac{q(\boldsymbol{\theta}^{(j-1)}|\boldsymbol{\theta}^{(j)})p(\boldsymbol{\theta}^{(j)})p(\mathbf{y}|\boldsymbol{\theta}^{(j)})}{q(\boldsymbol{\theta}^{(j)}|\boldsymbol{\theta}^{(j-1)})p(\boldsymbol{\theta}^{(j-1)})p(\mathbf{y}|\boldsymbol{\theta}^{(j-1)})}\right\}. \quad (3.5)$$

To implement the MH algorithm, the likelihood function $\boldsymbol{\theta} \mapsto p(\mathbf{y}|\boldsymbol{\theta})$ has to be evaluated,

$$p(\mathbf{y}|\boldsymbol{\theta}) = \int p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x})p(\mathbf{x}|\boldsymbol{\theta})d\mathbf{x}. \quad (3.6)$$

In a latent variable model, this integral has generally no analytical form, leading to an intractable likelihood function.

Proposal Scheme

When applying the MH algorithm to generate posterior samples, it is essential to choose a well-working proposal density $q(\cdot|\boldsymbol{\theta}^{(j-1)})$. Cotter et al. (2013) showed that standard random walk MCMC algorithms entail highly inefficient performance and strong dependence on the discretization when targeting high-dimensional Gaussian random fields. As a solution, they suggest proposal schemes that preserve the prior PDF, resulting in an MH algorithm for which the acceptance ratio only depends on the likelihoods. In geophysics, this proposal scheme is known as the extended Metropolis algorithm (Mosegaard and Tarantola 1995). If the target space is high-dimensional, the prior-preserving proposal scheme still needs to be chosen carefully (Ruggeri et al. 2015). Therefore, Friedli et al. (2022) introduce a prior-preserving version of the adaptive multi-chain algorithm DREAM_(ZS) (DiffeRential Evolution Adaptive Metropolis using an archive of past states; Laloy and Vrugt 2012).

3.2.3 Gaussian approximation of the intractable likelihood

The new approach to approximate the intractable likelihood (Eq. 3.6) relies on the linearized Gaussian approximation proposed by Linde et al. (2017). Here, the data covariance matrix given by the observational noise Σ_Y (Eq. 3.2) is adjusted by adding an additional contribution accounting for petrophysical prediction uncertainty. For our latent variable model, a first-order Taylor expansion of $\mathbf{x} \mapsto \mathcal{G}(\mathbf{x})$ around $\mathcal{F}(\boldsymbol{\theta})$ is used,

$$\mathbf{Y} = \mathcal{G}(\mathcal{F}(\boldsymbol{\theta}) + \boldsymbol{\varepsilon}_{\mathcal{P}}) + \boldsymbol{\varepsilon}_{\mathcal{O}} \approx \mathcal{G}(\mathcal{F}(\boldsymbol{\theta})) + \mathbf{J}_{\mathcal{F}(\boldsymbol{\theta})} \boldsymbol{\varepsilon}_{\mathcal{P}} + \boldsymbol{\varepsilon}_{\mathcal{O}}, \quad (3.7)$$

with $\mathbf{J}_{\mathcal{F}(\boldsymbol{\theta})}$ denoting the Jacobian (sensitivity) matrix of the forward solver corresponding to $\mathcal{F}(\boldsymbol{\theta})$. Due to its dependence on $\mathcal{F}(\boldsymbol{\theta})$, the sensitivity matrix is evolving between iterations (with changing $\boldsymbol{\theta}$). Using Gaussian assumptions for $p(\boldsymbol{\varepsilon}_{\mathcal{O}}) = \varphi_T(\boldsymbol{\varepsilon}_{\mathcal{O}}; \mathbf{0}, \Sigma_Y)$ and $p(\boldsymbol{\varepsilon}_{\mathcal{P}}) = \varphi_L(\boldsymbol{\varepsilon}_{\mathcal{P}}; \mathbf{0}, \Sigma_P)$ (Eq. 3.2), the likelihood function (Eq. 3.6) is approximated by,

$$\hat{p}(\mathbf{y}|\boldsymbol{\theta}) = \varphi_T(\mathbf{y}; \mu_Y, \tilde{\Sigma}_Y) \quad \text{with} \quad \mu_Y = \mathcal{G}(\mathcal{F}(\boldsymbol{\theta})) \quad \text{and} \quad \tilde{\Sigma}_Y = \mathbf{J}_{\mathcal{F}(\boldsymbol{\theta})}^T \Sigma_P \mathbf{J}_{\mathcal{F}(\boldsymbol{\theta})} + \Sigma_Y. \quad (3.8)$$

As the linearization is made around $\mathcal{G}(\mathcal{F}(\boldsymbol{\theta}))$ and not around $\mathcal{G}(\mathcal{F}(\boldsymbol{\theta}) + \boldsymbol{\varepsilon}_{\mathcal{P}})$, errors arise when the resulting Jacobians differ. For a linear geophysical problem, there are no approximation errors. Figure 3.1 shows a flow chart describing this approach at iteration j ; in what follows, this method will be referred to as LinGau.

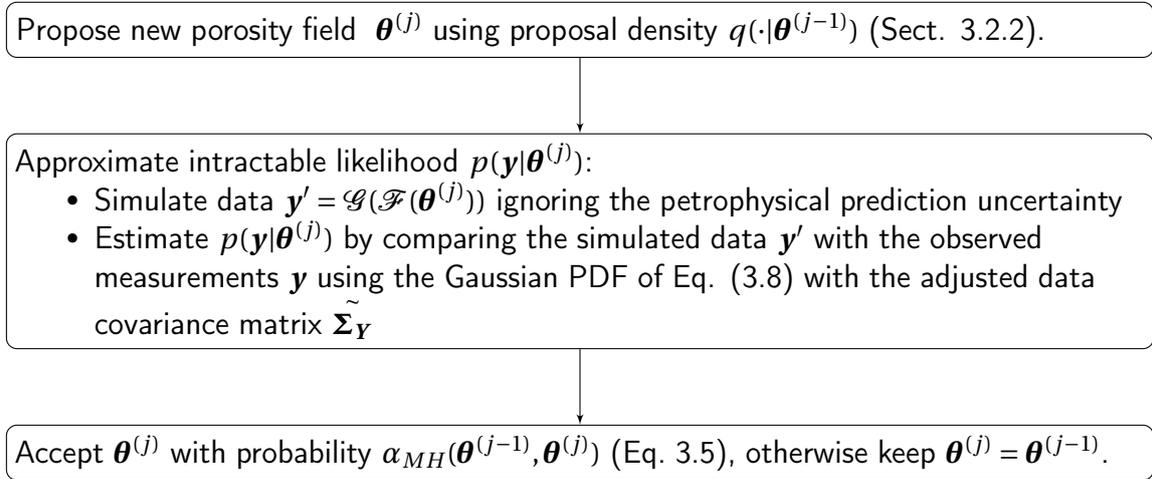


Figure 3.1: Flow chart illustrating the LinGau method at iteration j

3.2.4 Correlated pseudo-marginal method

The pseudo-marginal method (Beaumont 2003; Andrieu and Roberts 2009) estimates the intractable likelihood (Eq. 3.6) by Monte Carlo averaging over samples of the latent variable,

$$\hat{p}_N(\mathbf{y}|\boldsymbol{\theta}) = \frac{1}{N} \sum_{n=1}^N w(\mathbf{y}|\mathbf{X}_n, \boldsymbol{\theta}), \quad \text{with} \quad w(\mathbf{y}|\mathbf{X}_n, \boldsymbol{\theta}) = \frac{p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{X}_n)p(\mathbf{X}_n|\boldsymbol{\theta})}{m(\mathbf{X}_n|\boldsymbol{\theta})}, \quad (3.9)$$

where $\mathbf{X}_n \stackrel{i.i.d.}{\sim} m(\cdot|\boldsymbol{\theta})$ for $n = 1, 2, \dots, N$ with $m(\cdot|\boldsymbol{\theta})$ denoting an importance density function. This implies that after proposing a new target parameter $\boldsymbol{\theta}$, different latent variable realizations \mathbf{X}_n with the same $\boldsymbol{\theta}$ and different PPE $\boldsymbol{\epsilon}_{\mathcal{P}}$ are sampled. Then, the likelihood of each realization can be calculated and the intractable likelihood function is estimated by averaging over the obtained values. To account for the influence of importance sampling on the draws of the latent variable, weighted averaging has to be applied. Using this non-negative unbiased estimator of the likelihood leads to a MH algorithm sampling the same posterior distribution as one using the true likelihood (Beaumont 2003).

To obtain an efficient algorithm, it is crucial that the variance of the log-likelihood ratio estimator used in each MH step is low enough (Doucet et al. 2015). This can be ensured by choosing a well-working importance sampling density and by selecting a number of latent variable samples N which is high enough. Following Friedli et al. (2022), a Gaussian approximation $\mathbf{x} \mapsto \varphi_L(\mathbf{x}; \boldsymbol{\mu}_{IS}, \boldsymbol{\Sigma}_{IS})$ of $\mathbf{x} \mapsto p(\mathbf{x}|\boldsymbol{\theta}, \mathbf{y})$ is used as importance density, which implies the same linearization of the forward operator as in the LinGau approach (Eq. 3.7). An inappropriate linearization will lead to errors in the LinGau estimates, while it will only affect the efficiency of the pseudo-marginal method.

To reduce the magnitude of N , Deligiannidis et al. (2018) introduced the correlated pseudo-marginal (CPM) method by which the samples of latent variables used in the likelihood ratio estimator are correlated. Assuming a standard-normal distributed latent variable \mathbf{X} , the CPM method correlates one draw of iteration j with one of the former by,

$$\mathbf{X}^{(j)} = \rho \mathbf{X}^{(j-1)} + \sqrt{1 - \rho^2} \boldsymbol{\epsilon}, \quad \text{with} \quad \rho \in (-1, 1) \quad \text{and} \quad \boldsymbol{\epsilon} = (\epsilon_1, \epsilon_2, \dots, \epsilon_L), \epsilon_i \stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1). \quad (3.10)$$

The general applicability of the CPM method is not limited by the assumption that the latent variable has a standard-normal distribution since numerous distributions can be obtained by transformations from standard normal variates (e.g. Chen et al. 2018). The procedure of the CPM method in iteration j is illustrated in Fig. 3.2. Further details about the method can be found in Friedli et al. (2022).

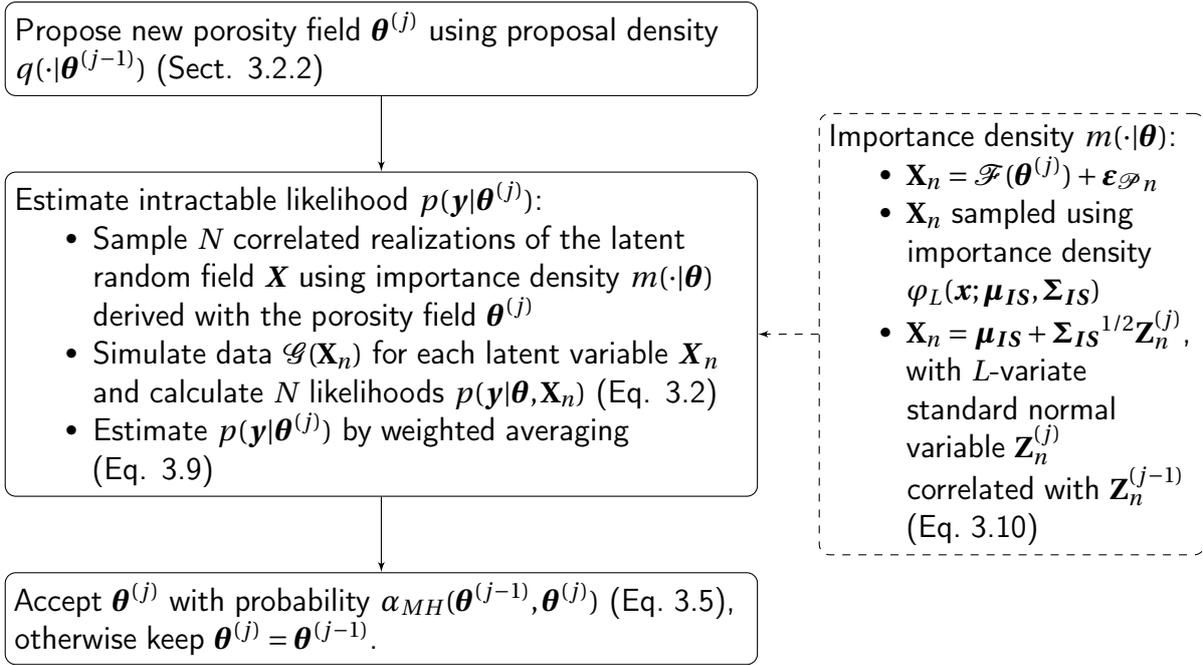


Figure 3.2: Flow chart illustrating the CPM method with importance sampling at iteration j

3.2.5 Performance assessment

The primary focus of the assessment of the inversion results is on the exploration of the posterior PDF. To declare convergence of the MCMC chains, the \hat{R} -statistic of Gelman and Rubin (1992) is used. Subsequently, the posterior samples obtained with the considered methods are compared. For a numerical assessment of the posterior estimates, the logarithmic score (logS; Good 1952) is employed. This is a so-called scoring rule (Gneiting and Raftery 2007) evaluating the accuracy of a predictive PDF $\theta \mapsto \hat{p}(\theta)$ with respect to a true value θ_{true} . The logarithmic score is defined as $\log S(\hat{p}, \theta_{\text{true}}) = -\log \hat{p}(\theta_{\text{true}})$ and kernel density estimates with manually-selected bandwidths are applied to transform the posterior samples into a PDF. Furthermore, the number of target parameters in which the true porosity value θ_{true} is in the range of the posterior samples is considered, as well as the spread of the posterior samples as quantified by their standard deviations.

3.3 Case study

The new linearized Gaussian approach is compared against the CPM method using the first case study considered by Friedli et al. (2022).

3.3.1 Synthetic data generation

The considered model domain is a water-saturated subsurface area of $7.2 \text{ m} \times 7.2 \text{ m}$. As in Friedli et al. (2022), the target porosity field is assumed to be a Gaussian random field with known mean ($\boldsymbol{\mu}_\theta = 0.39$) and exponential covariance function. For the latter, a sill of 2×10^{-4} is assumed such as geometric anisotropy with the main, horizontal direction having an integral scale of 5.4 m and the integral scale ratio between the horizontal and vertical direction being 0.13 . For the parameterization of the porosity field, a regular (50×50)-dimensional grid ($D^2 = 2,500$) is used. The "true" porosity field θ_{true} for this case study is depicted in Fig. 3.3(a).

Given the porosity $\boldsymbol{\theta}$, the dielectric constant $\boldsymbol{\kappa}$ is predicted using the complex refractive index model (CRIM; Roth et al. 1990), from which the slowness field (our latent variable \boldsymbol{X}) can be derived,

$$\boldsymbol{x} = \sqrt{c^{-2}\boldsymbol{\kappa}} + \boldsymbol{\varepsilon}_\varphi = \frac{1}{c}(\sqrt{\kappa_s} + (\sqrt{\kappa_w} - \sqrt{\kappa_s})\boldsymbol{\theta}) + \boldsymbol{\varepsilon}_\varphi, \quad (3.11)$$

with κ_w and κ_s denoting the dielectric constants of water [81] and mineral grains [5], respectively, and c referring to the speed of light in vacuum [0.3 m/ns]. A PPE $\boldsymbol{\varepsilon}_\varphi$ is added (Fig. 3.3b), which is a realization of a centred Gaussian random field over a regular two dimensional grid of size 50×50 . Thereby, the same correlation structure as for the porosity field is used with a sill of $7.11 \times 10^{-2} \text{ ns}^2/\text{m}^2$. The "true" slowness field is depicted in Fig. 3.3(c) while Fig. 3.3(d) shows a scatter plot depicting the influence of the PPE on the slowness values.

The 625 first-arrival travel times are generated using 25 equidistant GPR transmitters located on the left side and 25 receivers on the right side of the model domain (the transmitter-receiver layout is shown in Fig. 3.3c). As forward solver $\boldsymbol{y} \mapsto \mathcal{G}(\boldsymbol{y})$, the non-linear (eikonal) solver *time2D* of Podvin and Lecomte (1991) is used. The observational noise $\boldsymbol{\varepsilon}_\theta$ is assumed to be *i.i.d.* centered normal with a standard deviation of 1 ns . The noise-affected synthetic first-arrival travel times are depicted in Fig. 3.3(e).

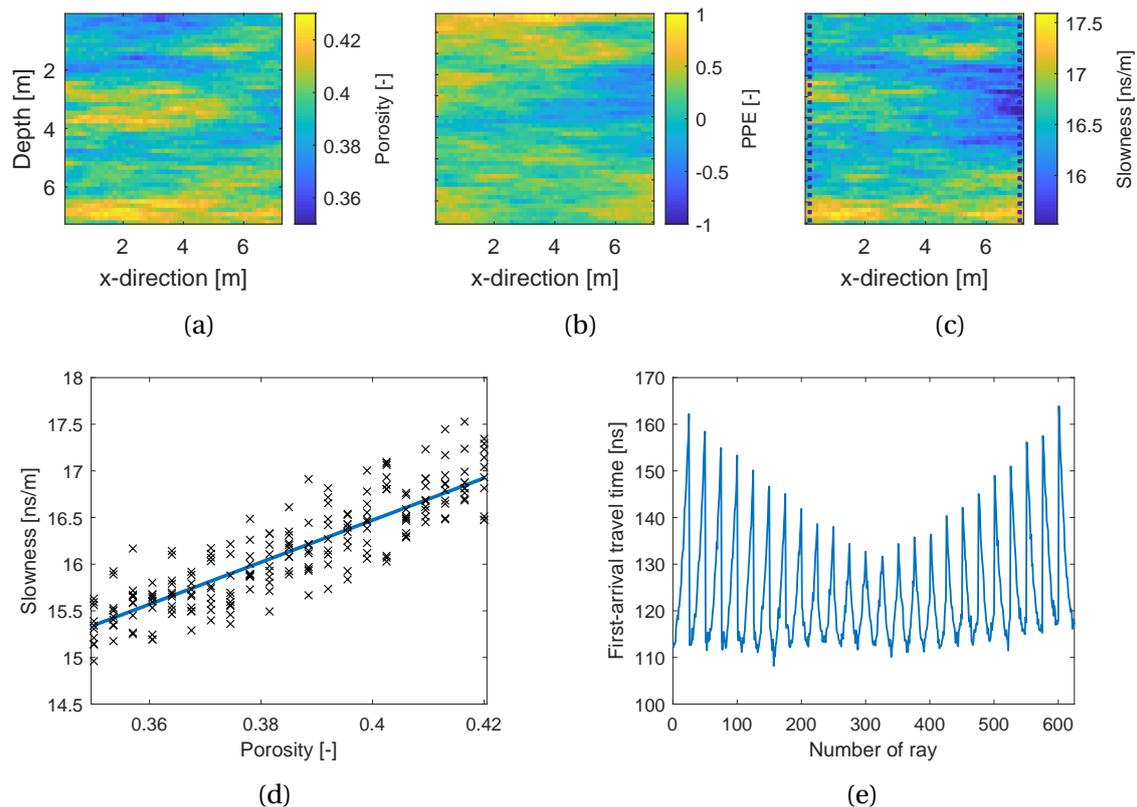


Figure 3.3: The synthetic “true” model adapted from Friedli et al. 2022: (a) porosity field θ_{true} , (b) PPE field $\epsilon_{\varphi_{\text{true}}}$, (c) slowness field x_{true} , (d) scatter plot of porosity and slowness values per grid cell, the line depicts the petrophysical relationship (Eq. 3.11) without considering PPE and (e) set of noise-affected first-arrival travel times y

3.3.2 Inversion setting and prior assumptions

As the considered target space is high-dimensional with 2,500 unknown parameters, it is crucial to choose a well-working proposal scheme for the MH algorithm. Due to its convincing performance in Friedli et al. (2022), this study relies on prior-preserving DREAM_(ZS) proposals (Sect. 3.2.2) and four MH chains are run in parallel. For both the LinGau and the CPM method, the linearization used to adjust the data covariance matrix and the importance sampling, respectively, is updated every tenth MCMC iteration. For the CPM method, a configuration of $N = 10$ and $\rho = 0.95$ is used as this choice guarantees an appropriate variance of the log-likelihood ratio estimator (see Friedli et al. 2022).

For the prior on porosity, a Gaussian PDF $p(\theta) = \varphi_{2500}(\theta; \mu_{\theta}, \Sigma_{\theta})$ with known mean μ_{θ} and covariance structure Σ_{θ} (the same values as for the data generation) is assumed. Then, the 2,500-dimensional vector Z defining the porosity by $\theta = \mu_{\theta} + \Sigma_{\theta}^{1/2} Z$ is inferred, with Z having a multivariate standard-normal prior PDF. For the PPE ϵ_{φ} also a Gaussian prior PDF $p(\epsilon_{\varphi}) = \varphi_{2500}(\epsilon_{\varphi}; 0, \Sigma_{\varphi})$ with known covariance structure Σ_{φ} is used. For the likelihood function, centred independent Gaussian measurement errors ϵ_{θ} with a standard deviation of 1 ns as in the data generation process are assumed.

The LinGau and CPM methods are compared with an inversion ignoring the petrophysical prediction uncertainty (No PPE). In this case, the intractable likelihood $p(\mathbf{y}|\boldsymbol{\theta})$ is estimated by $\varphi_T(\mathbf{y}; \mathcal{G}(\mathcal{F}(\boldsymbol{\theta})), \boldsymbol{\Sigma}_Y)$. As prior-preserving DREAM_(ZS) proposals lead to an unfeasible low acceptance rate in this case, standard DREAM_(ZS) proposals are employed.

3.3.3 Results

Figures 3.4(a) and 3.4(b) display the posterior mean estimates of the LinGau and the CPM method, respectively. Both images look very similar and the structural resemblance to the true porosity field in Fig. 3.3(a) is high. Both methods need about 100,000 MCMC iterations to converge and have a very similar acceptance rate (AR) of 10 %. They also lead to a similar performance in terms of posterior exploration as both methods sample the true porosity value in all of the pixels and have very similar median standard deviations and logarithmic scores (Figs. 3.4d and 3.4e, Table 3.1). By comparing the estimated posterior means of LinGau and CPM (Figs. 3.4a and 3.4b) with the one of the inversion ignoring the petrophysical prediction uncertainty (Fig. 3.4c), it is found that the mean estimate of the inversion ignoring the petrophysical prediction uncertainty has larger amplitudes even if the mean estimates are structurally similar. Its posterior exploration is less extensive, leading to a higher median logarithmic score (Fig. 3.4f), roughly half the median posterior standard deviation and many pixels that never sample the corresponding true porosity value (about one eighth of the pixels; Tab. 3.1). For this example exhibiting weak non-linearity and rather small petrophysical prediction uncertainty, it is concluded that the performance for LinGau and CPM are similar. On the other hand, ignoring petrophysical prediction uncertainty leads to biased estimates and too small uncertainty bounds.

Table 3.1: Summary of the results for the Gaussian porosity field example: The acceptance rates (**AR**), the convergence (**Conv**) showing the number of iterations needed for the 99th percentile of the parameter's \hat{R} -statics to be below 1.2, the percentage of pixels in which the true porosity value $\boldsymbol{\theta}_{\text{true}}$ lies within the range of posterior samples, the median logarithmic score (**logS**) and the median posterior standard deviation (**Post SD**).

Method	AR	Conv	$\boldsymbol{\theta}_{\text{true}}$	logS	Post SD
LinGau	10 %	108,000	100.00 %	1.10	10.7×10^{-3}
CPM	10 %	96,000	100.00 %	1.16	10.5×10^{-3}
No PPE	15 %	104,000	87.24 %	2.45	5.2×10^{-3}

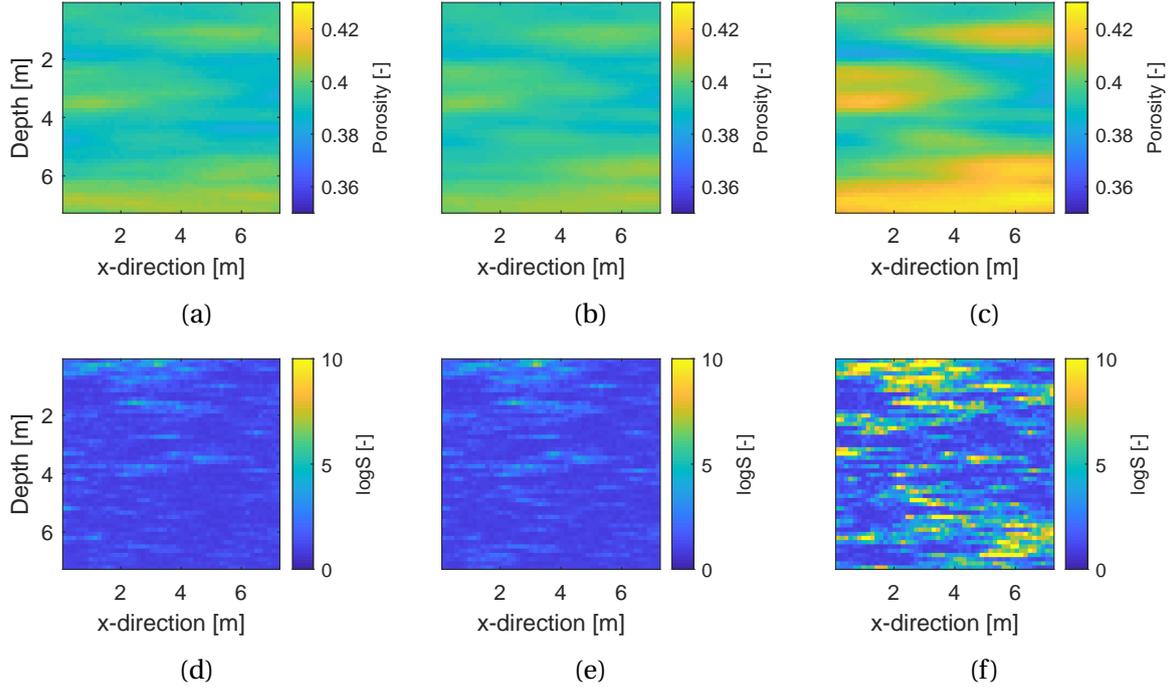


Figure 3.4: Estimated posterior means of the porosity field θ obtained with (a) the LinGau method, (b) the CPM method and (c) an inversion ignoring the petrophysical prediction uncertainty. Corresponding pixel-wise logarithmic scores assessing the estimated posterior PDFs for (d) the LinGau method, (e) the CPM method and (f) the inversion ignoring the petrophysical prediction uncertainty

3.4 Sensitivity analysis

In the previous synthetic test case, both the LinGau and the CPM method perform comparably well. As the computational cost of the LinGau method is lower (no need for N repeated sampling of the latent variable at each iteration), the use of the LinGau method would be recommended in such a setting. However, the degree of non-linearity in the geophysical forward operator is rather low for this test case. This is illustrated in Fig. 3.5 showing exemplary ray paths for the true slowness field \mathbf{x}_{true} (Fig. 3.5a) and the slowness field based on the true porosity field θ_{true} but ignoring the PPE (Fig. 3.5b): most ray paths are close to linear and they are very similar for both fields. Since the LinGau approximation of the likelihood (Eq. 3.8) relies on a first-order Taylor expansion of the physical forward solver (Eq. 3.7), it deteriorates when adding PPE ϵ_{φ} realizations lead to different ray-paths than the field around which the linearization is made. For a linear geophysical relationship, the first-order Taylor expansion is exact for any degree of Gaussian petrophysical prediction uncertainty.

In practice, the question is how to decide, for a given setting, if the LinGau approximation of the likelihood is accurate enough. To shed light on this, a sensitivity analysis exploring the performances of the LinGau and CPM methods for different levels of petrophysical prediction

uncertainty is performed. To make the comparison more didactic and to avoid unrelated challenges associated with a very-high dimensional target space, the porosity field is assumed to be layered. Generally, the same setup as in Section 3.3 is considered, but only with 13 transmitters and receivers (169 data points). The observational error is assumed to have a standard deviation of 1 ns.

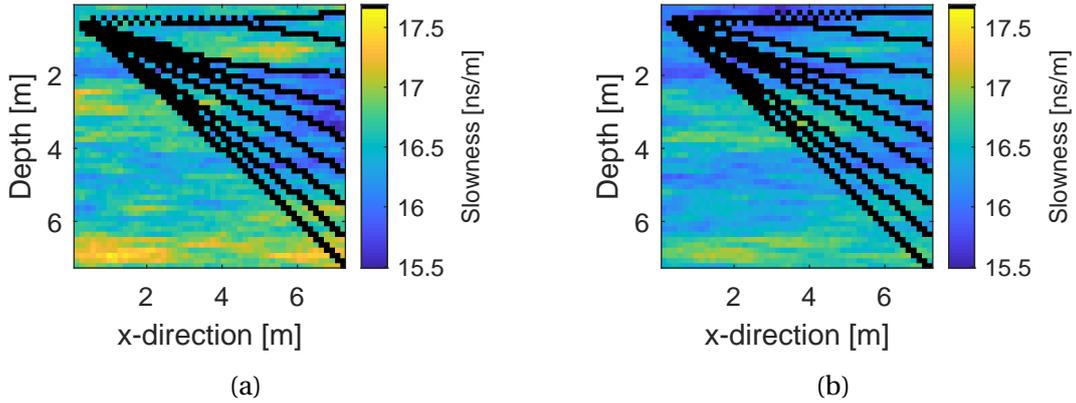


Figure 3.5: Exemplary GPR ray paths for (a) the true slowness field \mathbf{x}_{true} of the test case presented in Section 3.3 and (b) the slowness field resulting from the true porosity field $\boldsymbol{\theta}_{\text{true}}$ when ignoring the PPE $\boldsymbol{\varepsilon}_{\mathcal{P}}$

3.4.1 Likelihood estimation

The "true" porosity field $\boldsymbol{\theta}_{\text{true}} = (\theta_1, \dots, \theta_{10})$ is generated by assuming 10 horizontal layers of equal thickness and drawing independently from a Gaussian distribution with mean 0.3 and standard deviation 0.03 (Fig. 3.6a). The resulting slowness field (Eq. 3.11) is distorted with a layered PPE field having zero mean and independent layers with standard deviation σ_P . For each of sixteen different standard deviation values σ_P ranging from 0.0 to 1.5, one hundred data sets are generated using the same porosity field $\boldsymbol{\theta}_{\text{true}}$ (Fig. 3.6a) but different realizations of the PPE $\boldsymbol{\varepsilon}_{\mathcal{P}}$. Thereby, the true log-likelihood value $p(\mathbf{y}|\boldsymbol{\theta}_{\text{true}})$ of $\boldsymbol{\theta}_{\text{true}}$ has a different value depending on the realization of the PPE and the observational noise, even for the same σ_P . Two exemplary PPE fields and resulting slowness fields with $\sigma_P = 0.5$ and $\sigma_P = 1.0$ are depicted in Figs. 3.6(c)+(d) and 3.6(e)+(f), respectively.

For each of the one hundred data sets per value of σ_P , $p(\mathbf{y}|\boldsymbol{\theta}_{\text{true}})$ is approximated by using the LinGau approach. The corresponding values and their mean for the different σ_P are shown in Fig. 3.7(b). Those values are compared with the log-likelihood estimates obtained under the assumption of no PPE (Fig. 3.7c). Eventually, the aim is to assess these estimates by comparing them to the true log-likelihood values $p(\mathbf{y}|\boldsymbol{\theta}_{\text{true}})$ of the corresponding data

set. Due to the intractability inherited by the latent variable model, the analytical solution remains elusive and the unbiased importance sampling estimate of Equation (3.9) employing 1,000 realizations of the PPE (Fig. 3.7a) is applied. Due to the high number of samples, a well-specified importance density and the low dimensionality of the problem at hand, this estimator is close to the true log-likelihood value. To show that this assumption is valid, the standard deviation of the importance sampling estimator of $p(\mathbf{y}|\boldsymbol{\theta}_{\text{true}})$ for both the exemplary settings with $\sigma_P = 0.5$ and $\sigma_P = 1.0$ of Fig. 3.6 is evaluated (indicated with red symbols in Fig. 3.7). For both $\sigma_P = 0.5$ and $\sigma_P = 1.0$, the standard deviation is below one for log-likelihood values around -260.

With the underlying assumption of the estimators in Fig. 3.7(a) being close to the true value, they are compared with the ones of the LinGau method (Fig. 3.7d, darkgrey triangles). The absolute errors in the log-likelihood estimation of LinGau increase with increasing σ_P . While the mean absolute error in the log-likelihood estimator for $\sigma_P = 0.5$ is around 30, it grows to 150 for $\sigma_P = 1.0$ and to 400 for $\sigma_P = 1.5$. However, in comparison to the errors of the method ignoring the PPE (Fig. 3.7d, lightgrey crosses), the errors of the LinGau approach are comparably small. When ignoring the PPE completely, the mean absolute error for $\sigma_P = 0.5$ is about 400, for $\sigma_P = 1.0$ it is 1,700 and for $\sigma_P = 1.5$ even 5,000. Thereby, the method ignoring the PPE almost always underestimates the true log likelihood values $p(\mathbf{y}|\boldsymbol{\theta}_{\text{true}})$ as it cannot account for the true PPE and therefore gives reduced likelihood to the true porosity field $\boldsymbol{\theta}_{\text{true}}$. The same holds true for the LinGau approach with increasing σ_P as the method accounts for an increasingly wrong PPE.

The LinGau approximation relies on a first-order Taylor expansion of the physical forward solver (Eq. 3.7). This approximation deteriorates with increasing degree of PPE as can be seen in Fig. 3.7(e) depicting the root mean square errors (RMSE) of the Taylor expansions for the data sets with increasing σ_P . While the mean of the RMSEs is comparable to the observational noise for $\sigma_P = 0.5$, it is twice as large for $\sigma_P = 1.0$. To establish the influence of the discussed errors on the inversion results, the setups introduced in Fig. 3.6 that are indicated with red symbols in Fig. 3.7 are considered: The first with $\sigma_P = 0.5$ employs rather small errors, however, the standard deviation of the PPE is twice as high as in the case-study in Section 3.3. In the second setting with $\sigma_P = 1.0$ the errors are doubled and a realization is targeted where the error in the LinGau likelihood approximation is especially high (about 460 while the mean error for $\sigma_P = 1.0$ is 150).

3.4.2 Inversion

Using the two layered synthetic data sets (Fig. 3.6), the MH algorithms are run with three chains in parallel. Due to the simplicity of the problem, a basic Gaussian random walk is used as proposal scheme, within which, for comparison purposes, a step width (standard deviation) of 0.005 is applied for all methods. Furthermore, the linearizations used in the LinGau and CPM methods are updated at every MCMC iteration to prevent any errors resulting from less frequent updates. For the CPM method, importance sampling and a configuration of $N = 3, \rho = 0.9$ for the first and $N = 50, \rho = 0.975$ for the second setting is used. For the

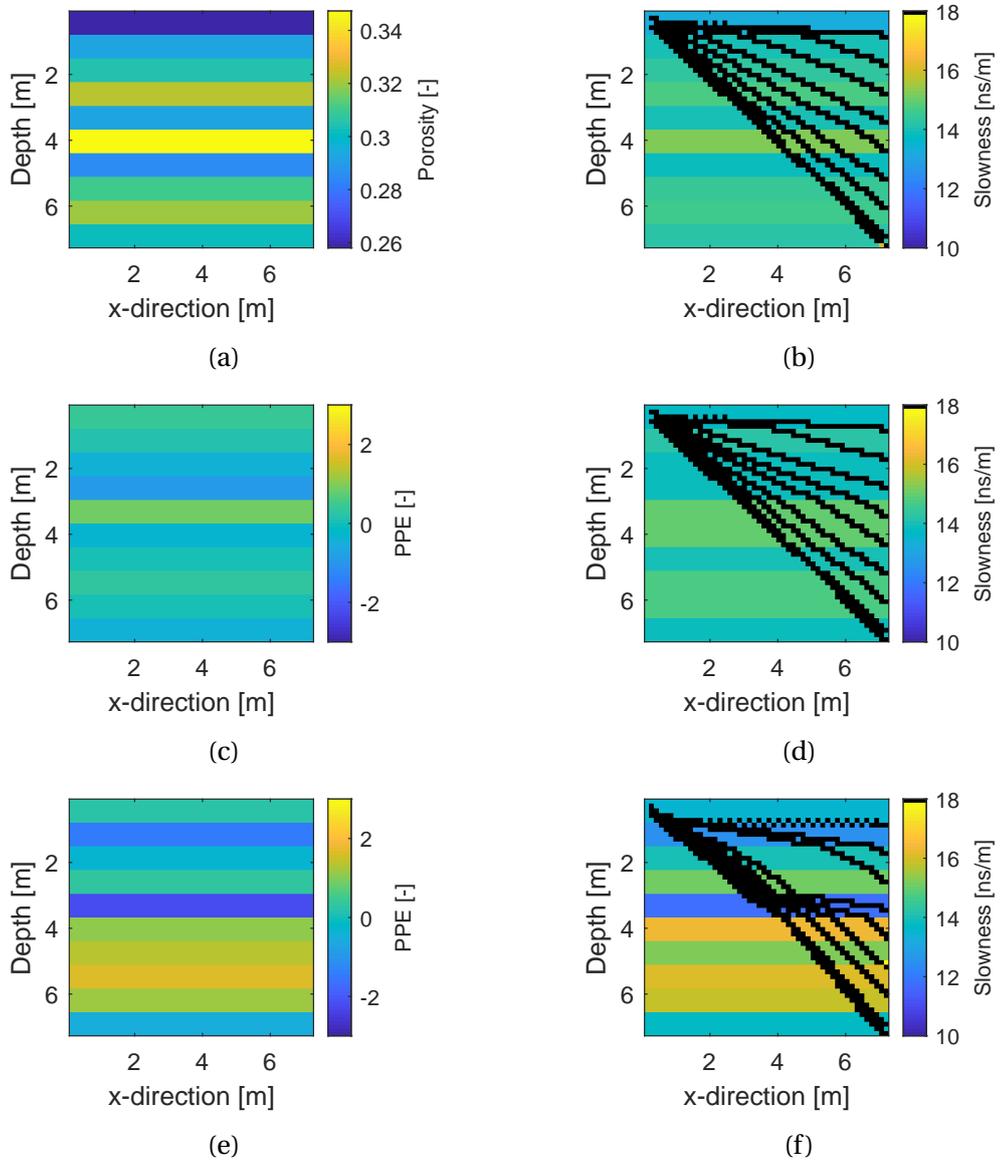


Figure 3.6: Synthetic “true” models for the layered test cases with the (a) same porosity field θ_{true} and different PPE for (c) Setting 1 and (e) Setting 2. The right column shows the slowness fields with exemplary ray paths for (b) the true porosity θ_{true} without adding PPE, (d) Setting 1 and (f) Setting 2

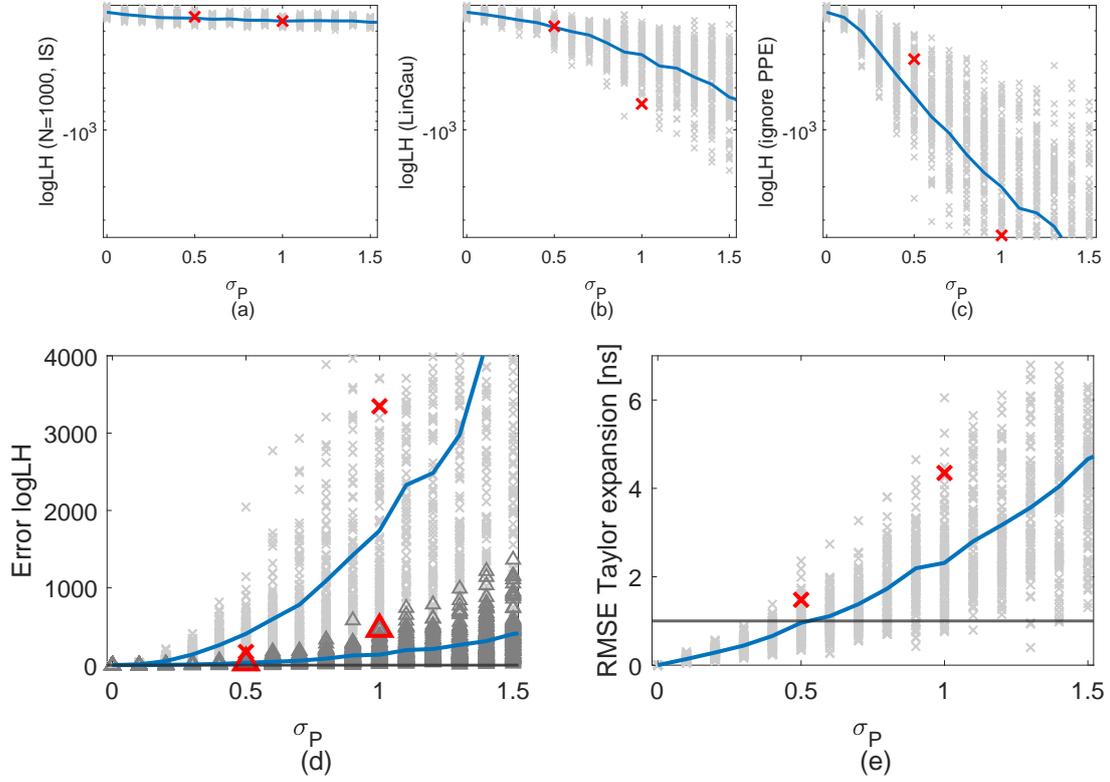


Figure 3.7: Estimators of log-likelihood $p(\mathbf{y}|\boldsymbol{\theta}_{\text{true}})$ of the true porosity field in the layered setting as a function of petrophysical prediction uncertainty (σ_P): (a) importance sampling estimate using 1,000 samples of the PPE, (b) LinGau estimate and (c) estimate ignoring the PPE. For each value of σ_P , the same underlying porosity field $\boldsymbol{\theta}_{\text{true}}$ is used with 100 different realizations of the PPE, leading to 100 data sets each. The crosses indicate the values for each data set and the solid lines their mean. (d) The mean absolute difference between the values of (a) and (b) are shown in darkgrey triangles and (a) and (c) in lightgrey crosses. (e) The RMSEs in the first-order Taylor expansion (Eq. 3.7) of the data sets; the black horizontal line shows the standard deviation of the observational noise (1 ns). The red symbols in (a) - (e) indicate the errors of the settings used in the subsequent inversion examples; thereby in (d), the red triangles refer to the errors obtained when using LinGau and the red crosses refer to the errors obtained when ignoring PPE.

layers of the porosity field, independent Gaussian prior PDFs with mean 0.3 and standard deviation 0.03 and for the PPE layers, independent centred Gaussian priors with a standard deviation of 0.5 and 1.0 are assumed, respectively. Finally, for the likelihoods uncorrelated Gaussian observational noise with a standard deviation of 1 ns is assumed.

The estimated marginal posterior PDFs for both settings are shown in Fig. 3.8 for three of the ten layers. These results are representative of the other layers, but a summary of the performance for all layers is for completeness provided in Table 3.2. For the first setting (Figs. 3.8a-c), the modes of all considered methods are located around rather similar values. While the method ignoring the PPE generates samples that do not include the true value of the porosity in half of the layers, both LinGau and CPM capture the true value in all of the ten layers (Table 3.2). However, while the LinGau method already doubles the standard deviation of the posterior samples compared to the method ignoring the PPE, the CPM method leads to a further doubling compared to LinGau (Table 3.2). Compared to the method ignoring the PPE, the LinGau method reduces the median logarithmic score from 3.53 to -2.37 and the CPM method reduces it even further to -2.71.

For the second setting (Figs. 3.8d-f), rather different posterior estimates are obtained for the three methods. The method ignoring the PPE generates posterior samples which are far from the true porosity values and with a small standard deviation (Table 3.2). The realizations obtained with the LinGau method have a twice as high standard deviation but are also centered far from the true values. The CPM method yields posteriors three times wider than those obtained with LinGau. Thereby, while the posterior samples obtained with the CPM method include the true values of the porosity in all layers, the LinGau method misses them for two layers and the method ignoring the PPE in seven layers. The median logarithmic score of the CPM method (-2.00) is distinctly lower than the one of the LinGau method (-1.35), which in turn is dramatically lower than for the method ignoring the PPE (378.77).

3.5 Discussion

In this work, non-linear geophysical inversion problems involving uncertain petrophysical relationships are targeted. Two different approaches to account for the corresponding intractable likelihood function are explored: a linearized Gaussian approximation (LinGau) and the correlated pseudo-marginal (CPM) method. The performance of these two methods for examples with increasing petrophysical prediction uncertainty is assessed. This work is a continuation of Friedli et al. (2022) as it is the first time the LinGau method is employed and compared with the CPM method. A synthetic crosshole travel-time tomography is first considered, with the aim of inferring a water-saturated multivariate Gaussian porosity field in a situation of only moderate non-linearity and petrophysical prediction uncertainty (adapted from Friedli et al. 2022). The results obtained with the LinGau and the CPM methods are very similar (Fig. 3.4 and Table 3.1), thereby, indicating that the approximate LinGau method works well in this setting. The LinGau method is then to be preferred as it only necessitates one forward simulation for each MCMC chain and iteration, while CPM with $N = 10$ (number

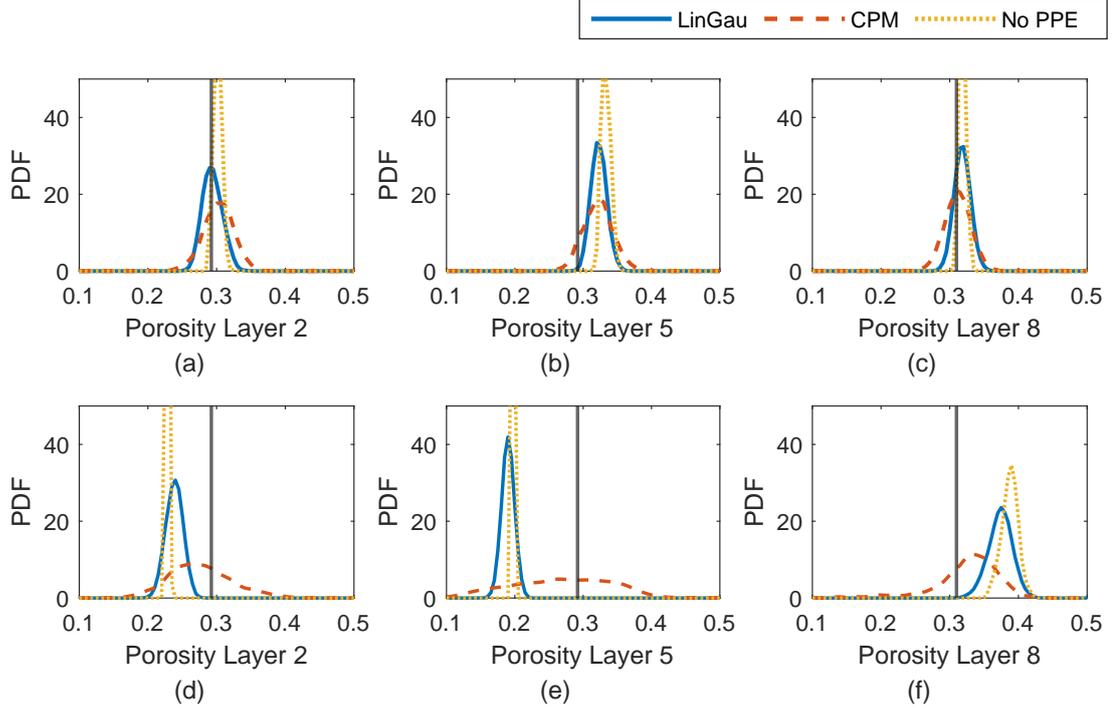


Figure 3.8: Estimates of the marginal posterior PDFs for the layered test cases: Setting 1 for (a) layer 2, (b) layer 5 and (c) layer 8. Setting 2 for (d) layer 2, (e) layer 5 and (f) layer 8. The solid vertical lines indicate the true porosity values

of latent variable samples) uses ten times as many forward simulations. If parallelization with $N = 10$ more processors are employed, then the computational time can still be made similar to the LinGau approach even if the overall computational cost is significantly higher. For this test case, one iteration for one MCMC chain on a standard laptop takes about 0.3 seconds for LinGau and 2.1 seconds for CPM ($N = 10$ and no parallelization). Importantly, this example demonstrated that an inversion that ignores petrophysical prediction uncertainty (the most common case in the literature) leads to biased estimates and underestimations of the widths of the posterior estimates by a factor of two.

In Section 3.4, the effect of the degree of petrophysical prediction uncertainty on the likelihood estimation and inversion results is studied for the different methods. To simplify the comparison, layered porosity and PPE fields are used. By comparing log-likelihood estimates, it is shown that the errors in the LinGau method increase with increasing degree of petrophysical prediction uncertainty, even if these errors are much smaller than for the method ignoring petrophysical prediction uncertainty (Fig. 3.7). By selecting one data set with twice as high petrophysical prediction uncertainty as in the previous multivariate Gaussian example ($\sigma_P = 0.5$, Fig. 3.6c) and one with four times as high uncertainty ($\sigma_P = 1.0$, Fig. 3.6e), the influence of the likelihood estimation errors on the inversion results is investigated. For the first setting (moderate degree of PPE), the LinGau method performs less well than CPM (underestimation of posterior uncertainty by a factor of two; higher logS scores). However,

Table 3.2: Summary of the results obtained for the study targeting a layered porosity field: the acceptance rates (**AR**), convergence with respect to the \hat{R} -statics (**Conv**), the number of layers in which the true porosity value lies within the range of posterior samples (θ_{true}), the median logarithmic score (**logS**) and the median posterior standard deviation (**Post SD**).

Method	Setting	AR	Conv	θ_{true}	logS	Post SD
LinGau		40 %	2,500	10/10	-2.37	11.8×10^{-3}
CPM	1 ($\sigma_P = 0.5$)	40 %	12'000	10/10	-2.71	21.9×10^{-3}
No PPE		2 %	13,000	5/10	3.53	4.2×10^{-3}
LinGau		40 %	6,000	8/10	-1.35	15.0×10^{-3}
CPM	2 ($\sigma_P = 1.0$)	30 %	51'000	10/10	-2.00	46.4×10^{-3}
No PPE		2 %	20,000	3/10	378.77	7.2×10^{-3}

employing the LinGau method still enables a reasonable approximation of the posterior modes of the layers. Although the true log-likelihood value $p(\mathbf{y}|\theta_{\text{true}})$ is underestimated by about 30 with LinGau (Fig. 3.7d), the RMSE of the Taylor approximation is roughly on the same order of magnitude as the observational error (Fig. 3.7e). That is different in the second setting with a higher degree of PPE, where the true log-likelihood value $p(\mathbf{y}|\theta_{\text{true}})$ is underestimated by about 460 and the RMSE of the Taylor approximation is four times as high as the observational noise. The resulting deterioration in performance for the LinGau method is drastic: underestimation of the posterior uncertainty by a factor of three and twice as high difference in the logS score to CPM than for the first setting. The growing error in the LinGau estimate is due to the Jacobian $\mathbf{J}_{\mathcal{F}(\theta)}$ being increasingly different when considering or not considering the PPE (Figs. 3.6b and 3.6f). For the CPM method, the growing petrophysical prediction uncertainty is accompanied by a very important increase in the posterior standard deviation of the samples (Figs. 3.8d-f), guaranteeing that the method samples the true porosity values even if the mode of the posterior may be located at the wrong place. Even if the LinGau method still provides much better results than the common approach of ignoring PPE altogether, the linearization on which this method is based (Eq. 3.7) is unable to account for non-linear effects associated with specific PPE realizations. In contrast, the CPM method only relies on the linearization to derive an importance sampling distribution and errors in this distribution will only lead to slower convergence of the CPM method while still targeting the true posterior distribution.

The approximation error of the first-order Taylor expansion used in the LinGau method (Eq. 3.7) grows with increasing non-linear effects related to the PPE. By comparing the forward operators for latent variables with $(\mathcal{F}(\boldsymbol{\theta}) + \boldsymbol{\varepsilon}_{\mathcal{P}})$ and without considering the PPE $\boldsymbol{\varepsilon}_{\mathcal{P}}$ ($\mathcal{F}(\boldsymbol{\theta})$), this relation can be established (as in Figs. 3.5 and 3.6). It is seen that high standard deviations in the $\boldsymbol{\varepsilon}_{\mathcal{P}}$ have an adverse effect on the approximation accuracy by highly influencing the forward operator. If the PPE $\boldsymbol{\varepsilon}_{\mathcal{P}}$ strongly influences the Jacobian (as in Fig. 3.6e), caution is advised when applying the LinGau method (Fig. 3.7d). To choose between the LinGau and the CPM method, this study recommends to investigate the RMSE of the first-order Taylor expansion (Fig. 3.7e) and to compare it with the observational noise. For a setting when the RMSE is significantly lower than the observational error (as for $\sigma_P = 0.25$ in Sect. 3.3), the recommendation is to use the LinGau method due to its lower computational cost (the effect of the petrophysical prediction uncertainty is incorporated into the likelihood function and no importance sampling of latent variables is needed). If the RMSE is in the same order of magnitude as the observational error (as for $\sigma_P = 0.5$ in Sect. 3.4), the LinGau method can be applied if moderate errors in the posterior estimation are acceptable. If the RMSE is clearly higher than the observational error (as for $\sigma_P = 1.5$ in Sect. 3.4), the exact CPM method should be used. If this is computationally too expensive, one could consider accounting for the approximation error in the LinGau method by treating it as a model error term (see e.g. Hansen et al. 2014).

The approximation underlying the LinGau approach could easily be incorporated in deterministic inversion methods (using gradients), while this is impossible for CPM. On the other hand, while the LinGau method requires Gaussian assumptions for the PPE, the CPM method can sample the latent variables from a variety of possible distributions. But in practice the CPM method needs a well-working importance sampling strategy to be efficient. For the presented test cases, a decreasing acceptance rate for the CPM method is observed with increasing petrophysical prediction uncertainty if the number of latent variables N and correlation ρ are fixed. This occurs as our importance sampling scheme gets more and more inaccurate, which can be compensated by using larger N or ρ . Although the efficiency is reduced, the accuracy of the posterior samples remains the same as the importance sampling is only used to decrease the variance of the likelihood estimator. In contrast, the LinGau method does not only lose efficiency with increasing non-linearity, but also leads to overconfident and biased estimates of the target parameters as the approximation of the likelihood is getting increasingly inaccurate (Fig. 3.8).

In the presented test cases, the mean and the covariance structure of the target and petrophysical prediction error fields are assumed to be known. In a setting where they are unknown, hierarchical Bayes could be employed (as e.g. in Laloy et al. 2015, Brunetti and Linde 2018). A recent study by Friedli et al. (2023) targets a setting where the problem is formulated differently: only the posterior mean and the covariance structure are derived, while the small-scale variations in the model domain are accounted for but not inferred. In this study, a crosshole ray-based setting that is only weakly non-linear is considered. More non-linear problems such as electrical resistivity tomography or surface-based seismic refraction tomography might exhibit an even stronger sensitivity to petrophysical prediction uncertainty and the applicability of the LinGau might be reduced compared with the present study. Indeed, the

errors in the LinGau approximation are not due to the petrophysical prediction uncertainty as such, but rather how individual realizations of it affects the sensitivity patterns compared to the Taylor expansion that is performed in absence of petrophysical prediction uncertainty. Even if the focus herein is on how to account for petrophysical prediction uncertainty, both the LinGau and the CPM methods could be adapted to, for instance, account for 3-D effects in 2-D inversions, variations in porosity in tracer-test tomography targeting the permeability field, or to account for hydraulic storativity fluctuations in hydraulic tomography studies. Indeed, this type of latent variable problem arises as soon as the measured data do not only depend on the main parameters of interest but also on some other variables influencing the response.

3.6 Conclusions

This work focus on geophysical inversion problems targeting the posterior distribution of (hydro)geological parameters while accounting for uncertain petrophysical relationships and non-linear physics. The resulting intractable likelihood function is accounted for by either the linearized Gaussian approximation (LinGau) method or the correlated pseudo-marginal (CPM) method. The LinGau method, so far untested in geophysical inversion, is an approximate method that is computationally cheaper than CPM as it does not necessitate a Monte Carlo estimation of the likelihood at each MCMC iteration. In a first case study, a water-saturated multivariate Gaussian porosity field is considered for which the scatter in the petrophysical relationship and the non-linearity is comparatively small. In this setting involving crosshole first-arrival travel times, the LinGau method succeeds equally well as CPM in exploring the posterior distribution. For comparison, an inversion ignoring petrophysical uncertainty provides biased results and too narrow uncertainty estimates. In a subsequent sensitivity analysis considering layered fields, the degree of petrophysical prediction uncertainty is increased, thereby introducing increasing inaccuracies associated with the Taylor expansion on which the LinGau method is based. Consequently, the LinGau method produces increasingly inaccurate results as the petrophysical prediction uncertainty grows such that the true values are more and more often unsampled and the logarithmic scores are high. In contrast, the CPM method performs very well for all settings and accommodate the growing uncertainty in the petrophysical uncertainty, while this is only partially achieved by the LinGau method. The computationally less intensive LinGau method is attractive when the impact of the scatter of petrophysical prediction uncertainty is small compared to the observational noise. In comparison, the computationally more costly CPM method is an exact and much more general method that clearly outperforms the LinGau method when the petrophysical uncertainty grows in magnitude, but it needs an efficient importance sampling distribution to work well in practice. If the CPM method is computationally too expensive for a given application and if petrophysical uncertainty is significant, it is still better to use the LinGau method than inversions ignoring petrophysical prediction errors as the resulting results are less biased and the underestimation of posterior uncertainty is less pronounced.

Chapter 4

Inference of geostatistical hyperparameters with the correlated pseudo-marginal method

Lea Friedli, Niklas Linde, David Ginsbourger,
Alejandro Fernandez Visentini and Arnaud Doucet

Published in *Advances in Water Resources*:

Friedli, L., Linde, N., Ginsbourger, D., Visentini, A. F., & Doucet, A. (2023). Inference of geostatistical hyperparameters with the correlated pseudo-marginal method. *Advances in Water Resources*, 173: 104402.

Abstract

We consider non-linear Bayesian inversion problems targeting the geostatistical hyperparameters of a random field describing hydrogeological or geophysical properties given hydrogeological or geophysical data. This problem is of particular importance in the non-ergodic setting as there are no analytical upscaling relationships linking the data to the hyperparameters, such as, mean, standard deviation, and integral scales. Full inversion of the hyperparameters and the local properties of the field (typically involving many thousands of unknowns) brings substantial computational challenges, such that simplifying model assumptions (e.g., homogeneity or ergodicity) are typically made. To prevent the errors resulting from such simplified assumptions while also circumventing the burden of high-dimensional full inversions, we use a pseudo-marginal Metropolis–Hastings algorithm that treats the random field as latent variables. In this random effects model, the intractable likelihood of observing the data given the hyperparameters is estimated by Monte Carlo averaging over realizations of the random field. To increase the efficiency of the method, low-variance approximations of the likelihood ratio are obtained by using importance sampling and by correlating the samples used in the proposed and current steps of the Markov chain. We assess the performance of this correlated pseudo-marginal method by considering two representative inversion problems involving diffusion-based and wave-based physics, respectively, in which we infer the hyperparameters of (1) hydraulic conductivity fields using apparent hydraulic conductivity data in a data-poor setting and (2) fracture aperture fields using borehole ground-penetrating radar (GPR) reflection data in a more data-rich setting. For the first test case, we find that the correlated pseudo-marginal method generates similar estimates of the geostatistical mean as classical rejection sampling, while an inversion assuming ergodicity provides biased estimates. For the second test case, we find that the correlated pseudo-marginal method estimates the hyperparameters well, while rejection sampling is computationally unfeasible and a simplified model assuming homogeneity leads to biased estimates.

4.1 Introduction

The scale dependence of most environmental processes poses significant challenges for hydrogeological and geophysical modeling (e.g., Klemeš 1983; Blöschl and Sivapalan 1995). The governing partial differential equations (PDEs) traditionally employed to describe fluid flow, chemical or electrical transport (Neuman and Di Federico 2003) are solved at some support volume scale assumed to be a “Representative Elementary Volume” (REV; Hill 1963). That is, it is assumed that smaller-scale heterogeneity averages out and can be represented (with regard to the process under consideration) by averaged physical or chemical properties. In practice, the conditions necessary for the existence of a REV are often not met because geological media exhibit heterogeneity over a wide range of scales (Neuman and Di Federico 2003). Errors occurring when only partially accounting for or ignoring heterogeneity generally grow with the non-linearity of the physical or chemical process under study and can result in misleading predictions (e.g., Dentz et al. 2011, Yu and Michael 2022). For this reason, it is essential to characterize and account for the statistical properties of small-scale heterogeneity even when targeting mean properties.

We consider non-linear inversion problems targeting geostatistical hyperparameters (e.g., mean, standard deviation, integral scale and anisotropy factor) of a random field describing hydrogeological or geophysical properties given indirect data. This problem setting is applicable when the main properties of interest are the hyperparameters and not the local field properties. The geostatistical literature is full of studies (e.g., Rehfeldt et al. 1992, Hess et al. 1992, Bohling et al. 2016) focusing on hyperparameter estimation based on direct data (e.g., permeability data along boreholes), but much less work has considered indirect data (e.g., pressure data, tracer breakthrough data) as in the present study. In what follows, we only discuss this latter case.

One of the first approaches considering unknown hyperparameters in such an inversion setting was the quasi-linear geostatistical approach by Kitanidis (1995), which optimizes the hyperparameters along with the spatial field. Another approach enabling joint inference of a Gaussian random field and its variogram parameters relied on so-called sequential Gibbs sampling (Hansen et al. 2012, Hansen et al. 2013a Hansen et al. 2013b). Zhao and Luo (2021) applied an iterative approach based on principal components which is updating biased or unknown hyperparameters while solving a non-linear inversion problem. Recently, Wang et al. (2022) proposed an hierarchical Bayesian inversion targeting first global variables (such as hyperparameters but also physical variables) and later the posterior of the whole field (referred to as spatial variables). Note that none of these studies focus on inferring the hyperparameters only.

We rely on a Bayesian framework and infer the hyperparameters' posterior probability density function (PDF) given indirect hydrogeological or geophysical measurements. To sample from the posterior, we apply a Markov chain Monte Carlo (MCMC) method building on the Metropolis–Hastings algorithm (MH; Hastings 1970; Metropolis et al. 1953). The basic procedure of the MH algorithm in this setting is to propose iteratively a new set of hyperparameters, which are then accepted or rejected based on their prior probabilities and likelihoods.

We consider synthetic experimental setups in which the hydrogeological or geophysical data average over a random field realization that is either ergodic or non-ergodic. A random field must be stationary to be ergodic, but not vice versa. Stationarity implies that the distribution does not change with position. Ergodicity, on the other hand, implies that the field realization is much larger than the characteristic scale of heterogeneity. By the so-called ergodic setting, we consider data that average over a scale that is much larger than the field's scale of heterogeneity such that the effects of small-scale fluctuations average out. Consequently, the data do not depend on the local properties of a given random field realization, but on the hyperparameters only. By the non-ergodic setting, we refer to cases when the data averaging takes place over a scale that is smaller or comparable to the scale of heterogeneity. This implies that the data depend not only on the hyperparameters but also on the random field realization on which measurements are made. That is, variations between field realizations in terms of magnitudes and locations of high and low property values lead to different data responses as the fluctuations do not average out. Broadly speaking, such behavior is expected when the physical response is averaging over length scales that are less than some ten correlation lengths of the parameter field. In the non-ergodic setting, there are no analytical upscaling

relationships linking the data to the hyperparameters of interest. If relationships assuming ergodicity or assumptions of homogeneity are employed in such a case, bias is likely to occur in the inferred hyperparameters (e.g., Visentini et al. 2020; Shakas et al. 2018). We suggest that most measurements in hydrogeology and geophysics take place in such a non-ergodic setting.

Equivalent properties derived from measurements of one type of physics (e.g., the equivalent aperture describing fluid flow) generally do not represent the equivalent property for another type of physics (e.g., the equivalent aperture of thermal transport; e.g., Tsang 1992). This disparity occurs as soon as the underlying physics is non-linear, implying for instance that equivalent mean properties do not correspond to arithmetic mean properties (e.g., Jougnot et al. 2018, Shakas and Linde 2015). One solution to this problem that is pursued in the present study is to instead infer hyperparameters while accounting for small-scale heterogeneity. In this way, it is possible to use estimates derived from one type of physics to make predictions for another type of physics. In many ergodic settings, upscaling theory provides relevant relationships between hyperparameters and equivalent properties (e.g., Renard and De Marsily 1997, Torquato and Haslach Jr 2002, Sanchez-Vila et al. 2006), while no such relationships are available in the non-ergodic setting.

One way to infer hyperparameters in the non-ergodic setting by MCMC methods is to parameterize the field by hyperparameters and white noise to describe the local properties (as e.g. in Laloy et al. 2015, Hunziker et al. 2017 and Xiao et al. 2021). The corresponding full inversion problem involves typically many thousands of parameters, for which either an efficient MH proposal scheme has to be designed (e.g., Xiao et al. 2021) or dimensionality reduction arguments have to be invoked (e.g., Laloy et al. 2015, Rubin et al. 2010). While the first approach is very challenging (curse of dimensionality, e.g., Robert et al. 2018), the second approach may lead to biased estimates (Laloy et al. 2015). An example of the application of dimensionality reduction relevant to the current study is Shakas et al. (2018) who inferred fracture aperture distribution and geometry by combining GPR forward modeling with flow-and-transport simulations. Even if this study provided reasonable estimates of the statistical properties, it was plagued by a low acceptance rate, slow mixing of the chains and no formal convergence despite a large number of iterations.

Instead of a full inversion, we here target the hyperparameters of interest only. Since the local properties of the field influence the observations in the non-ergodic setting, the field is considered a latent (unobservable) variable. Due to the random effect the unobservable field has on the data, we speak of a random effects model. To implement a MH algorithm inferring the hyperparameters only, we have to evaluate the likelihood of observing the data given the currently proposed set of hyperparameters. In a random effects model, this likelihood has generally no analytical form and is, therefore, referred to as intractable. The pseudo-marginal (PM) method introduced by Beaumont (2003) and studied by Andrieu and Roberts (2009) relies on an unbiased estimator of this intractable likelihood function that is based on averaging over Monte Carlo samples of the latent variables. This implies that after proposing a new set of hyperparameters, different field realizations with the same hyperparameters are sampled. Then, the likelihood of each field realization can be calculated and the intractable likelihood function is estimated by averaging over the obtained values.

Beaumont (2003) demonstrates that using such a non-negative and unbiased estimator of the likelihood within the MH algorithm results in an algorithm that draws samples from the same target distribution as when using the true likelihood. In the PM method, a high variance of the log-likelihood ratio estimator has a very strong adverse impact on performance, but achieving a low variance often comes at the price of using an excessive number of samples in the Monte Carlo averaging. To obtain an efficient algorithm balancing these two aspects, it has been shown that the standard deviation of the log-likelihood estimator should be around 1.2-1.5 (Doucet et al. 2015). This can be ensured by (1) properly choosing the number of samples used in the Monte Carlo averaging and by (2) applying importance sampling to draw the realizations of the latent variables. In the context of state-space models, the number of samples has to increase linearly with the number of observations, which is computationally impractical in data-rich settings (Deligiannidis et al. 2018). To address this problem, the correlated pseudo-marginal (CPM) by Deligiannidis et al. (2018) correlates the draws of latent variables between two subsequent iterations, thereby, reducing the number of Monte Carlo draws needed to ensure low-variance log-likelihood ratio approximations.

The pseudo-marginal and correlated pseudo-marginal methods have hardly been studied in hydrogeological and geophysical settings. In Friedli et al. (2022), the CPM method was shown to outperform other competing approaches to lithological tomography (Bosch 1999), in which geophysical data are used to directly infer (hydro)geological properties of interest. Friedli et al. (2022) considered a very high dimensionality of the target and latent variables under the assumption of known hyperparameters. Here, the interest is instead placed on inferring few hyperparameters while accounting for the effects of thousands of latent variables. This leads to a very different model setting and study objectives than Friedli et al. (2022).

We assess the performance of the CPM method with two synthetic test cases in which we infer the hyperparameters describing (1) hydraulic property fields using equivalent (apparent) hydraulic conductivity data and (2) fracture aperture fields using borehole ground-penetrating radar (GPR) reflection data. The two test cases are chosen to be representative for transmission problems governed by diffusion (e.g., groundwater flow, heat transport, electrical conduction) and reflection problems governed by wave-based physics (e.g., GPR, seismics and acoustics). In the first test case, we consider a very data-poor setting and are mainly interested in the geostatistical mean of the field. By comparing the CPM results with those of an MH algorithm that replaces the forward solver with an analytical upscaling relationship that assumes ergodicity, we show that assuming a simplified model can lead to strongly biased estimates of the hyperparameters in the non-ergodic setting. We also demonstrate that the CPM results are in agreement with those obtained by rejection sampling, which is computationally feasible for this very data-poor example. In the second test case, we consider much more data and show that the CPM method provides accurate estimates of the geostatistical mean and other hyperparameters. Additionally, we show how these hyperparameters describing aperture properties inferred from GPR data allow us to predict fracture transmissivity.

This paper is structured as follows. Section 4.2 introduces the CPM methodology in the considered context. Section 4.3 presents the first test case based on measurements across a hydraulic conductivity field and Section 4.4 presents the second test case in which borehole

GPR data are used to infer the hyperparameters of fracture aperture fields. This is followed by a discussion in Section 4.5 and conclusions in Section 4.6.

4.2 Methodology

The methodology section starts by presenting the considered random effects model and the chosen notation of Gaussian random fields (Section 4.2.1). Bayesian inference and MCMC algorithms are then described (Section 4.2.2) before introducing the correlated pseudo-marginal method (Section 4.2.3) and giving a brief introduction into rejection sampling (Section 4.2.4). It ends with a description of the performance assessment metrics used to evaluate the results (Section 4.2.5).

4.2.1 Random effects model

We are interested in a random field describing hydrogeological or geophysical property distributions. A random field (spatial stochastic process) $\mathbf{X}(d, \omega)$ with $\omega \in \Omega$ is a family of random variables indexed by the spatial location $d \in \mathcal{D} \subset \mathbb{R}^2$ (Chiles and Delfiner 2012). For fixed $\omega = \omega_0$, $\mathbf{X}(\cdot, \omega_0)$ is a realization of the random field with ω referring to the “randomness” of the field. For a fixed location $d = d_0$, $\mathbf{X}(d_0, \cdot)$ is a real-valued random variable. For simplicity, in the following we write $\mathbf{X}(\cdot)$ to indicate $\mathbf{X}(\cdot, \omega)$. The “true” hydrogeological or geophysical property field is considered a realization of the underlying random field. We are interested in inferring the hyperparameters $\boldsymbol{\theta}$ parameterizing the geostatistical distribution of the random field $\mathbf{X}(\cdot)$.

We consider a Gaussian random field (GRF) $\mathbf{X}(\cdot)$ for which all finite-dimensional distributions are multivariate Gaussians (Chiles and Delfiner 2012). Its distribution is determined by the mean and the covariance function. We assume the mean function $\mu_{\boldsymbol{\theta}}(\cdot)$ to be constant even if it would be straightforward to employ a non-stationary function. For the covariance function $C_{\boldsymbol{\theta}}(\cdot, \cdot)$, we apply the powered exponential expressed here in isotropic form:

$$C_{\boldsymbol{\theta}}(b, b') = \sigma^2 \exp\left(-\left(\frac{\|b - b'\|}{I}\right)^{2H}\right), \quad (4.1)$$

whereby $\|b\| = \sqrt{b^T b}$ denotes the Euclidean norm, σ the standard deviation, I the integral scale and H the Hurst exponent (with $0 < H \leq 1$). For $H = 0.5$, the powered exponential covariance function reduces to the classical exponential covariance function and for $H = 1$ to the Gaussian (squared exponential) covariance function. We also consider geometric anisotropy (e.g., Chiles and Delfiner 2012), for which the covariance depends not only on the Euclidean distance but also on the direction between the considered positions. We assume a known anisotropy angle of 90 degrees and refer to the integral scale in the vertical direction as I^y , which, multiplied by the anisotropy factor λ , gives the integral scale in the horizontal direction I^x .

To infer the P hyperparameters $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_P)$, we have access to T measurements $\mathbf{y} = (y_1, y_2, \dots, y_T)$. As generally there exists no upscaling relationship linking the hyperparameters to the measurements, we formulate the problem with a random effects model using the latent random field $\mathbf{X}(\cdot)$:

$$\mathbf{X} \sim f_{\boldsymbol{\theta}}(\cdot) \quad \mathbf{Y}|\mathbf{X} \sim g_{\boldsymbol{\theta}}(\cdot|\mathbf{X}). \quad (4.2)$$

For the latent random field $\mathbf{X}(\cdot)$ we use a discretized representation on a $(D \times D)$ -grid, whereby we assume the grid cells to be representative elementary volumes (REV) for the governing physical process. We consider a setting in which the number of target hyperparameters P is much smaller than the number of latent variables (grid cells) D^2 . The measurements are described by the random variable $\mathbf{Y} = \mathcal{G}(\mathbf{X}) + \boldsymbol{\varepsilon}_{\mathcal{O}}$ with $\mathcal{G} : \mathbb{R}^{D^2} \rightarrow \mathbb{R}^T$ denoting the physical forward solver and $\boldsymbol{\varepsilon}_{\mathcal{O}}$ the observational noise. While \mathbf{Y} refers to the random variable, \mathbf{y} denotes the ‘‘true’’ measurements considered to be a realization of \mathbf{Y} .

Assuming the latent random field to be Gaussian, we write $f_{\boldsymbol{\theta}}(\mathbf{x}) = \varphi_{D^2}(\mathbf{x}; \boldsymbol{\mu}_{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})$ with $\varphi_{D^2}(\cdot; \boldsymbol{\mu}_{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})$ denoting the PDF of a D^2 -variate normal distribution with mean vector $\boldsymbol{\mu}_{\boldsymbol{\theta}} = (\mu_{\boldsymbol{\theta}}(b_i))_{1 \leq i \leq D^2}$ and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\theta}} = (C_{\boldsymbol{\theta}}(b_i, b_j))_{1 \leq i, j \leq D^2}$ specified by the hyperparameters $\boldsymbol{\theta}$. Furthermore, we assume the observational noise $\boldsymbol{\varepsilon}_{\mathcal{O}}$ to be Gaussian, such that $\mathbf{Y}|\mathbf{X} \sim g_{\boldsymbol{\theta}}(\cdot|\mathbf{X})$ is distributed with the PDF $g_{\boldsymbol{\theta}}(\mathbf{y}|\mathbf{x}) = \varphi_T(\mathbf{y}; \mathcal{G}(\mathbf{x}), \boldsymbol{\Sigma}_{\mathbf{Y}})$, with $\boldsymbol{\Sigma}_{\mathbf{Y}}$ being a diagonal matrix with the variance of the observational noise on its diagonal. To generate a realization of the D^2 -dimensional GRF $\mathbf{X}(\cdot)$ with mean vector $\boldsymbol{\mu}_{\boldsymbol{\theta}}$ and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$, we rely on a pixel-based parameterization,

$$\mathbf{X} = \boldsymbol{\mu}_{\boldsymbol{\theta}} + \boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{1/2} \mathbf{Z}, \quad (4.3)$$

with \mathbf{Z} denoting a D^2 -dimensional random vector consisting of *i.i.d.* standard normal distributed variables.

4.2.2 Bayesian inference with Markov chain Monte Carlo

Bayes’ theorem specifies the posterior PDF $p(\boldsymbol{\theta}|\mathbf{y})$ of the model parameters $\boldsymbol{\theta}$ conditioned on the measurements \mathbf{y} as,

$$p(\boldsymbol{\theta}|\mathbf{y}) = \frac{p(\boldsymbol{\theta})p(\mathbf{y}|\boldsymbol{\theta})}{p(\mathbf{y})} \propto p(\boldsymbol{\theta})p(\mathbf{y}|\boldsymbol{\theta}), \quad (4.4)$$

where $p(\boldsymbol{\theta})$ denotes the prior PDF of the model parameters, $p(\mathbf{y}|\boldsymbol{\theta})$ the likelihood function and $p(\mathbf{y})$ the evidence (assumed positive). If there is no analytical form of the posterior PDF but it is possible to evaluate the unnormalized entity for some value of $\boldsymbol{\theta}$, MCMC methods (see, e.g., Robert et al. 1999) can be applied to generate realizations drawn proportionally from the posterior PDF. The basic procedure behind MCMC algorithms is to propose new values for the target parameters, which are then accepted or rejected with a given probability.

The Metropolis–Hastings (MH; Metropolis et al. 1953; Hastings 1970) method is a well-known example. At iteration j , it proceeds as follows: First, new values for the target parameters $\boldsymbol{\theta}^{(j)}$ are proposed using the model proposal density $q(\cdot|\boldsymbol{\theta}^{(j-1)})$. Then, the acceptance probability,

$$\begin{aligned}\alpha_{MH}(\boldsymbol{\theta}^{(j-1)}, \boldsymbol{\theta}^{(j)}) &= \min\left\{1, \frac{q(\boldsymbol{\theta}^{(j-1)}|\boldsymbol{\theta}^{(j)})p(\boldsymbol{\theta}^{(j)}|\mathbf{y})}{q(\boldsymbol{\theta}^{(j)}|\boldsymbol{\theta}^{(j-1)})p(\boldsymbol{\theta}^{(j-1)}|\mathbf{y})}\right\} \\ &= \min\left\{1, \frac{q(\boldsymbol{\theta}^{(j-1)}|\boldsymbol{\theta}^{(j)})p(\boldsymbol{\theta}^{(j)})p(\mathbf{y}|\boldsymbol{\theta}^{(j)})}{q(\boldsymbol{\theta}^{(j)}|\boldsymbol{\theta}^{(j-1)})p(\boldsymbol{\theta}^{(j-1)})p(\mathbf{y}|\boldsymbol{\theta}^{(j-1)})}\right\},\end{aligned}\tag{4.5}$$

is calculated and the proposed $\boldsymbol{\theta}^{(j)}$ is accepted (if $\alpha_{MH}(\boldsymbol{\theta}^{(j-1)}, \boldsymbol{\theta}^{(j)}) \geq V$) or rejected (if $\alpha_{MH}(\boldsymbol{\theta}^{(j-1)}, \boldsymbol{\theta}^{(j)}) < V$) on the basis of a draw of a uniformly distributed random variable $V \sim \text{Unif}([0, 1])$. If the proposed $\boldsymbol{\theta}^{(j)}$ is rejected, the MCMC chain remains at the old position ($\boldsymbol{\theta}^{(j)} = \boldsymbol{\theta}^{(j-1)}$).

In order to evaluate the acceptance probability in Equation (4.5), the value of the likelihood function $\boldsymbol{\theta} \mapsto p(\mathbf{y}|\boldsymbol{\theta}^{(j)})$ has to be calculated. In a random effects model (see Section 4.2.1), the likelihood function is given by,

$$p(\mathbf{y}|\boldsymbol{\theta}) = \int g_{\boldsymbol{\theta}}(\mathbf{y}|\mathbf{x}) f_{\boldsymbol{\theta}}(\mathbf{x}) d\mathbf{x}.\tag{4.6}$$

This integral often does not admit an analytical form making the direct implementation of the MH algorithm impossible and specific algorithms such as the correlated pseudo-marginal method are needed (outlined in Section 4.2.3 below).

MCMC proposal scheme

To achieve an efficient MCMC algorithm, one needs a suitable proposal density $q(\cdot|\boldsymbol{\theta}^{(j-1)})$. Even in an inversion targeting only few parameters, one has to choose the direction and size of the model proposal steps carefully. Too large steps lead to a low acceptance rate, while too small steps lead to very slow exploration of the target space; both of these situations lead to an algorithm needing an unnecessarily large number of iterations until convergence (see Section 4.2.5 below for the assessment of convergence).

To generate model proposals, we apply the adaptive Metropolis algorithm of Haario et al. (2001), in which the covariance matrix describing the Gaussian proposal distribution is updated during the MCMC run. Despite the adaptation, the algorithm is ensured to be ergodic, although not Markovian (Haario et al. 2001). The Gaussian proposal distribution at iteration j is expressed as $q(\boldsymbol{\theta}^{(j)} | (\boldsymbol{\theta}^{(0)}, \boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(j-1)})) = \varphi_d(\boldsymbol{\theta}^{(j-1)}, \mathbf{C}^{(j)})$, with

$$\mathbf{C}^{(j)} = \begin{cases} \mathbf{C}^{(0)} & j \leq j_0 \\ s_P(\text{COV}(\boldsymbol{\theta}^{(0)}, \boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(j-1)}) + \epsilon \mathbb{I}_P) & j > j_0, \end{cases}\tag{4.7}$$

denoting the evolving covariance matrix. During the first j_0 iterations, the method uses an initial covariance matrix $\mathbf{C}^{(0)}$ selected according to available prior knowledge. After this initial period, the covariance matrix is updated with $\mathbf{C}^{(j)} = s_P(\text{COV}(\boldsymbol{\theta}^{(0)}, \boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(j-1)}) + \epsilon \mathbb{I}_P)$,

where s_P is a parameter depending on the dimension of the target space (Haario et al. (2001) use $s_P = (2.4)^2/P$ as in Gelman et al. 2013), $\epsilon > 0$ is a small constant and \mathbb{I}_P denotes the identity matrix of dimension P . To ensure an efficient calculation, Haario et al. (2001) use the recursion formula,

$$\mathbf{C}^{(j+1)} = \frac{(j-1)}{j} \mathbf{C}^{(j)} + \frac{s_P}{j} \left(j \overline{\boldsymbol{\theta}^{(j-1)}} \overline{\boldsymbol{\theta}^{(j-1)T}} - (j+1) \overline{\boldsymbol{\theta}^{(j)}} \overline{\boldsymbol{\theta}^{(j)T}} + \boldsymbol{\theta}^{(j)} \boldsymbol{\theta}^{(j)T} + \epsilon \mathbb{I}_P \right), \quad (4.8)$$

with $\overline{\boldsymbol{\theta}^{(j)}} = 1/(j+1) \sum_{i=0}^j \boldsymbol{\theta}^{(i)}$ and $\boldsymbol{\theta}^{(i)}$ considered to be column vectors.

For a target parameter θ_i with bounded support $[a, b]$, one has to make sure that the proposed value lies within the considered interval. Therefore, we apply fold boundary handling implying that a proposal which passes one boundary of the support is re-entered through the other boundary (Vrugt 2016), that is, similar to periodic boundary conditions in numerical simulations.

4.2.3 Pseudo-marginal and correlated pseudo-marginal method

In Section 4.2.2, we explained that the considered random effects model has an intractable likelihood function. The pseudo-marginal and correlated pseudo-marginal methods presented below provide a solution to this in the form of Monte Carlo estimations of the likelihood function. To illustrate the presented concepts, a flow chart describing the basic procedure of the correlated pseudo-marginal method is depicted in Figure 4.1.

Pseudo-marginal method

A MH algorithm employing a non-negative unbiased estimator of the likelihood function samples realizations of the same target distribution as one using the true likelihood (Beaumont 2003). To exploit this remarkable property, Beaumont (2003) proposes a MH algorithm estimating, at each iteration, an intractable likelihood function using Monte Carlo averaging over samples of the latent variables. This approach was termed the pseudo-marginal (PM) method and analyzed theoretically by Andrieu and Roberts (2009).

The efficiency of the PM method depends mainly on the variability of the likelihood estimator. When only one brute force Monte Carlo sample of the latent variables is used to estimate the likelihood, the algorithm is likely to suffer from a low acceptance rate caused by the high variability of the log-likelihood estimator. This happens when the likelihood estimator can take very different values for different realizations of the latent variables. In our setting, this is the case if different local properties of the latent random field $\mathbf{X}(\cdot)$ lead to very different data responses even if the hyperparameters of the fields are the same. The variance of the log-likelihood estimator can be reduced by (1) using many samples of the latent variables and (2) selecting a well-working importance sampling (IS; e.g. Owen and Zhou 2000) scheme to draw them from.

The PM method proposes the following unbiased estimator for the likelihood $p(\mathbf{y}|\boldsymbol{\theta})$ of Equation (4.6),

$$\hat{p}_N(\mathbf{y}|\boldsymbol{\theta}) = \frac{1}{N} \sum_{n=1}^N w_{\boldsymbol{\theta}}(\mathbf{y}|\mathbf{X}_n), \quad \text{with} \quad w_{\boldsymbol{\theta}}(\mathbf{y}|\mathbf{X}_n) = \frac{g_{\boldsymbol{\theta}}(\mathbf{y}|\mathbf{X}_n)f_{\boldsymbol{\theta}}(\mathbf{X}_n)}{m_{\boldsymbol{\theta}}(\mathbf{X}_n)}, \quad (4.9)$$

where $\mathbf{X}_n \stackrel{i.i.d.}{\sim} m_{\boldsymbol{\theta}}(\cdot)$ for $n = 1, 2, \dots, N$ with $m_{\boldsymbol{\theta}}(\cdot)$ denoting the importance density function.

To derive the importance density $\mathbf{x} \mapsto m_{\boldsymbol{\theta}}(\cdot)$, we follow the approach of Friedli et al. (2022). Therefore, we choose a distribution which is nearly proportional to $\mathbf{x} \mapsto g_{\boldsymbol{\theta}}(\mathbf{y}|\mathbf{x})f_{\boldsymbol{\theta}}(\mathbf{x})$ (see e.g., Owen and Zhou 2000 referring to the results of Kahn and Marshall 1953). Since it holds that $p(\mathbf{x}|\boldsymbol{\theta}, \mathbf{y}) \propto g_{\boldsymbol{\theta}}(\mathbf{y}|\mathbf{x})f_{\boldsymbol{\theta}}(\mathbf{x})$, we approximate the importance density with a Gaussian expression of $\mathbf{x} \mapsto p(\mathbf{x}|\boldsymbol{\theta}, \mathbf{y})$. For details, see Appendix A.

Correlated pseudo-marginal method

The efficiency of the PM method depends strongly on the number of latent variable samples N used to estimate the likelihood function. If this number is too low, the variability of the log-likelihood ratio estimator is likely to be high and the MH algorithm suffers from an impractically low acceptance rate (Beaumont 2003). In the context of state-space models, Deligiannidis et al. (2018) show that N needs to increase linearly with the number of data T , thereby, often implying prohibitively high computational costs. For this reason, Deligiannidis et al. (2018) adapted the PM method by correlating the draws of latent variables used in the current and proposed step of the MH algorithm. The resulting correlated pseudo-marginal method (CPM method; illustrated in Fig. 4.1) leads to a better performance as the variance of a ratio of estimators is reduced when positively correlating the estimators of the denominator and numerator (Koop 1972). For a standard normal distributed latent variable \mathbf{Z} , the CPM method draws a correlated realization of the n -th latent variable in iteration j by,

$$\mathbf{Z}_n^{(j)} = \rho \mathbf{Z}_n^{(j-1)} + \sqrt{1 - \rho^2} \boldsymbol{\epsilon}, \quad \text{with } \rho \in (0, 1) \text{ and } \boldsymbol{\epsilon} = (\epsilon_1, \epsilon_2, \dots, \epsilon_L), \epsilon_i \stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1). \quad (4.10)$$

As numerous distributions can be obtained by transformations from standard normal variates, the general applicability of the CPM method is not limited by the uncorrelated Gaussian assumption (e.g. Chen et al. 2018). For example, in our two test cases that will follow, we generate correlated Gaussian latent variables \mathbf{X} with mean $\boldsymbol{\mu}_{\boldsymbol{\theta}}$ and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ (or $\boldsymbol{\mu}_{IS}$ and $\boldsymbol{\Sigma}_{IS}$) by transforming correlated standard-normally-distributed variables \mathbf{Z} using Equation (4.3). We stress that the proposed latent variables $\mathbf{Z}_n^{(j)}$ are only saved if $\boldsymbol{\theta}^{(j)}$ is accepted, otherwise we keep $\mathbf{Z}_n^{(j)} = \mathbf{Z}_n^{(j-1)}$ as for $\boldsymbol{\theta}^{(j)} = \boldsymbol{\theta}^{(j-1)}$ in the MH algorithm.

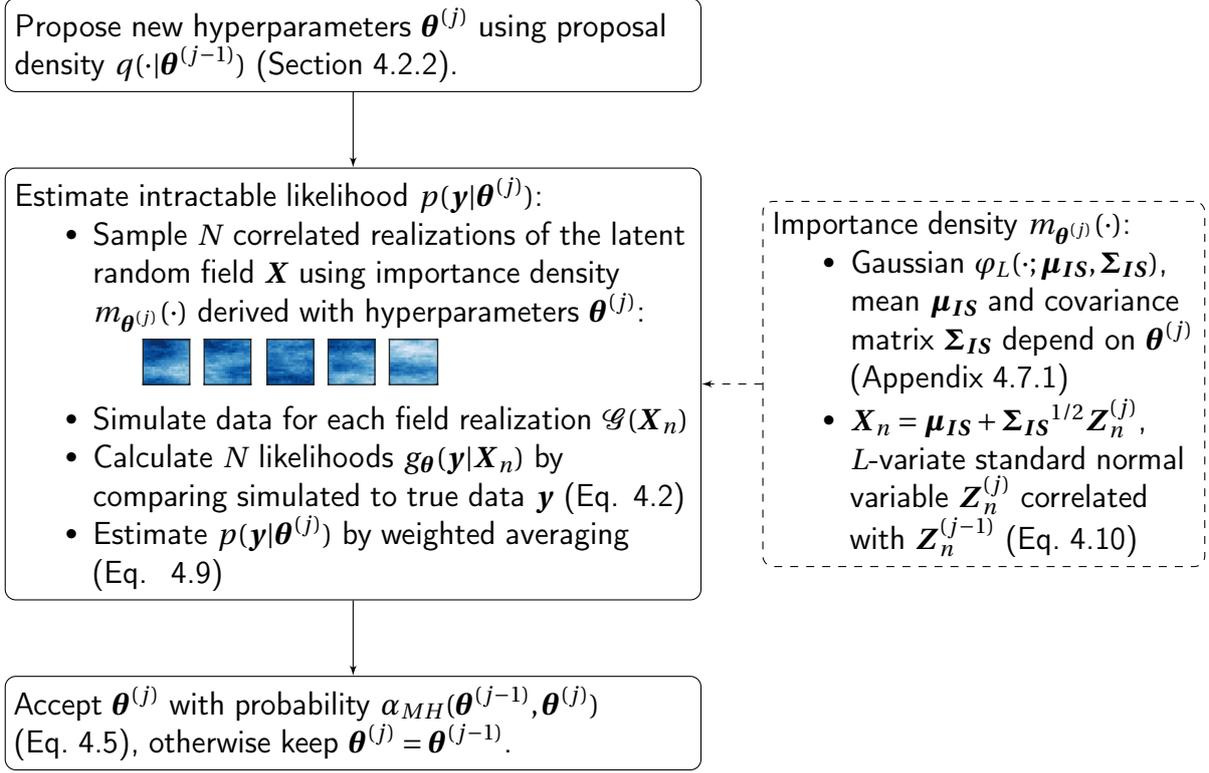


Figure 4.1: Flow chart illustrating the CPM method with importance sampling at iteration j .

The CPM method has two additional parameters compared to the standard MH algorithm: the latent variable sample size N and the correlation parameter ρ . Deligiannidis et al. (2018) propose to select N and ρ such that the variance of the log-likelihood ratio estimator,

$$W = \log\left(\hat{p}_N^{(j)}(\mathbf{y}|\boldsymbol{\theta})\right) - \log\left(\hat{p}_N^{(j-1)}(\mathbf{y}|\boldsymbol{\theta})\right), \quad (4.11)$$

takes values between 1.0 and 2.0 for $\boldsymbol{\theta}$ fixed in a region of high posterior probability mass. In practice, decreasing the variance of the estimator requires (1) more samples of the latent field or (2) a higher correlation of the samples making the exploration of the latent space slower. The range of 1.0 to 2.0 ensures a reasonable trade-off between the variance of the estimator, the exploration of the latent space (which would be slowed down by high ρ) and the computational cost (increases with increasing N). The region of $\boldsymbol{\theta}$ with high posterior mass can be chosen based on an initial MCMC run with N and ρ selected according to available prior knowledge. This choice can be inefficient, but will anyway give some first information. In practice, we first fix the number of samples N such that it is smaller than the number of available parallel processors. Then, we test a range of values for ρ and select one leading to $\text{Var}(W)$ being between 1.0 and 2.0.

4.2.4 Rejection sampling

Rejection sampling (RS; Ripley 2009a) is a basic Monte Carlo technique to generate independent samples from the posterior PDF. While it often suffers from an unfeasibly low acceptance rate, it is an exact sampling method (e.g., Robert et al. 1999) proceeding as follows:

1. Sample $\boldsymbol{\theta}$ from its prior distribution $p(\boldsymbol{\theta})$.
2. Sample u from a uniform distribution over $[0, 1]$.
3. Accept $\boldsymbol{\theta}$ if $u \leq \frac{p(\mathbf{y}|\boldsymbol{\theta})}{S_L}$, where S_L is the supremum of the likelihood function.

For our random effects model (Section 4.2.1), we estimate the intractable likelihood $p(\mathbf{y}|\boldsymbol{\theta})$ by sampling one brute-force realization \mathbf{x} of the latent variable field $\mathbf{X} \sim f_{\boldsymbol{\theta}}(\cdot)$ with hyperparameters $\boldsymbol{\theta}$,

$$\begin{aligned} \hat{p}(\mathbf{y}|\boldsymbol{\theta}) &= g_{\boldsymbol{\theta}}(\mathbf{y}|\mathbf{x}) = \varphi_T(\mathbf{y}; \mathcal{G}(\mathbf{x}), \boldsymbol{\Sigma}_Y) \\ &= \det(2\pi\boldsymbol{\Sigma}_Y)^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{y} - \mathcal{G}(\mathbf{x}))^T \boldsymbol{\Sigma}_Y^{-1}(\mathbf{y} - \mathcal{G}(\mathbf{x}))\right). \end{aligned} \quad (4.12)$$

In practice, an important challenge of RS methods is the need to estimate a tight bound S_L for the likelihood function. The most conservative choice is to assume a perfect data fit such that for our Gaussian likelihood function above we get $S_L = \det(2\pi\boldsymbol{\Sigma}_Y)^{-1/2}$, but this will typically lead to an acceptance rate being close to zero. If we assume the errors to be equal to the standard deviation of the observational noise, we get $S_L = \det(2\pi\boldsymbol{\Sigma}_Y)^{-1/2} \exp(-\frac{1}{2}T)$, which might lead to some bias as some realizations are likely to have higher likelihoods. One further possibility is to use the maximum likelihood value of the prior samples. To achieve this, RS is run by first saving all sampled prior realizations and their corresponding likelihood values. From this database, the maximum likelihood value is determined and all samples are assessed using this value. To obtain some accepted prior samples of $\boldsymbol{\theta}$ while ensuring an accurate estimate, we combine the second and the third approach and use the maximum of those two values as the supremum S_L .

4.2.5 Performance assessment

To assess if the CPM algorithm has converged, we use the \hat{R} -statistic of Gelman and Rubin (1992) comparing the within-chain variance with the between-chain variance of the second half of the MCMC chains. We follow the convention that the \hat{R} -statistic has to be smaller or equal to 1.2 for all model parameters. We also consider the acceptance rates (AR), which are aimed to be between 15% and 30% as proposed by Vrugt (2016).

We evaluate the amount of information gained by the inversion by comparing the marginal prior and posterior PDFs of the hyperparameters. This is achieved using the Kullback–Leibler divergence (KL divergence; Kullback and Leibler 1951) expressing the distance between two PDFs $z \mapsto p_1(z)$ and $z \mapsto p_2(z)$ (assumed positive) by,

$$\text{KL}(p_1||p_2) = \int p_1(z) \log\left(\frac{p_1(z)}{p_2(z)}\right) dz. \quad (4.13)$$

If $\text{KL}(p_1||p_2) = 0$, this means that the two PDFs are equal (almost everywhere), while an increasing value indicates diverging distributions. For example, for a standard normal PDF $p_2(\cdot)$, a KL divergence $\text{KL}(p_1||p_2) = 0.1$ is obtained by reducing the standard deviation within a centred standard normal $p_1(\cdot)$ to 0.7 and a KL divergence $\text{KL}(p_1||p_2) = 1$ is obtained by reducing the standard deviation to 0.23. To approximate the posterior PDFs, we apply kernel density estimation to the posterior samples (with manually adapted bandwidth).

To assess the quality of the posterior estimates, we use histograms to visually compare the marginal distributions with the true underlying values. Additionally, we evaluate the accuracy of the obtained posterior samples for each hyperparameter θ_i ($i \in \{1, 2, \dots, P\}$) numerically by applying a so-called scoring rule (Krüger et al. 2021). A scoring rule assesses the accuracy of a predictive PDF $z \mapsto p(z)$ with respect to a true value θ by accounting for both the statistical consistency between predictions and observations (calibration) and the sharpness of the prediction (Gneiting and Raftery 2007). For our test cases, we employ the logarithmic score (logS; Good 1952) defined by,

$$\text{logS}(p, \theta) = -\log p(\theta), \quad (4.14)$$

that is related to the Kullback–Leibler divergence (Gneiting and Raftery 2007). If we compare two posterior estimates, the one with the lower score is favoured. In practice (as for the KL divergence), we use a kernel density estimate of the posterior samples, which depends on the choice of the kernel and the bandwidth of the kernel smoothing window. We use a Gaussian kernel with manually adapted bandwidth. Our testings show that the choice influences the specific score values, but that the main results in terms of comparisons and conclusions are robust. If the posterior samples do not include the true value of θ , the density estimate of $p(\theta)$ can be numerically zero resulting in a logarithmic score of infinity. The logarithmic score is also available for multivariate densities, thereby, allowing evaluation of the estimated joint posterior PDFs of the hyperparameters.

4.3 Test case 1: Hydraulic conductivity field

Hydraulic conductivity is a key hydrogeological property. Particularly in contamination studies, the spatial variation of hydraulic conductivity plays an important role as it has a major influence on solute movement (Butler 2005). Visentini et al. (2020) rely on time-lapse electrical resistance data during a tracer test to demonstrate that measurements of equivalent electrical properties, at a given scale, can be used to infer hyperparameters of the hydraulic conductivity field below this scale. Here, we seek to infer the hyperparameters of a log-hydraulic conductivity field in a data-poor setting involving only horizontally- and vertically-averaged equivalent hydraulic conductivity data. With such limited data, it is tempting to ignore heterogeneity or rely on upscaling relations valid for ergodic fields, as there is little hope that the data can constrain the field or its hyperparameters well. This example is used to demonstrate that ignoring heterogeneity or assuming ergodicity leads to significant errors when estimating the geostatistical mean. Furthermore, this data-poor setting allows for comparisons with rejection sampling (Section 4.2.4), thereby, demonstrating that the CPM method targets the right posterior of the hyperparameters.

4.3.1 Data and inversion setting

We target a $1 \text{ m} \times 1 \text{ m}$ log-hydraulic conductivity field distributed according to a Gaussian random field $GRF(\mu_{\theta}(\cdot), C_{\theta}(\cdot, \cdot))$ with constant mean $\mu_{\theta}(\cdot)$ and exponential covariance function $C_{\theta}(\cdot, \cdot)$ (Eq. 4.1 with $H = 0.5$). We allow geometric anisotropy (Section 4.2.1) and denote the integral scale in the vertical direction (depth) as I^y and the anisotropy factor as λ . Together with the mean and standard deviation of the log-field, this forms the hyperparameters $\theta = (\mu, \sigma, I^y, \lambda)$. Although we are mainly interested in the mean, we infer the other hyperparameters along with it, thereby, accounting for the possible non-ergodicity of the field. The log-hydraulic conductivity field (natural logarithm) is generated on a 100×100 grid (cell size is 1 cm) using a pixel-based approach (Section 4.2.1).

Synthetic data generation

We generate noise-contaminated synthetic data in both an ergodic and a non-ergodic setting. For the ergodic setting, we create one “true” field realization from which we obtain noise-contaminated data by assuming the field to be isotropic and use $\theta = (\ln(10^{-4}), 0.5, 0.03 \text{ m}, 1)$ and for the non-ergodic, anisotropic case we choose $\theta = (\ln(10^{-4}), 1.5, 0.1 \text{ m}, 3)$. The true log-hydraulic conductivity fields are shown in Figure 4.2. Due to the discretization of the field chosen to limit the number of grid cells, the ergodic field is only nearly ergodic, implying that the generated data will vary somewhat when considering different field realizations with the true hyperparameters. In what follows, we will refer to it as ergodic except when a more specific designation is needed.

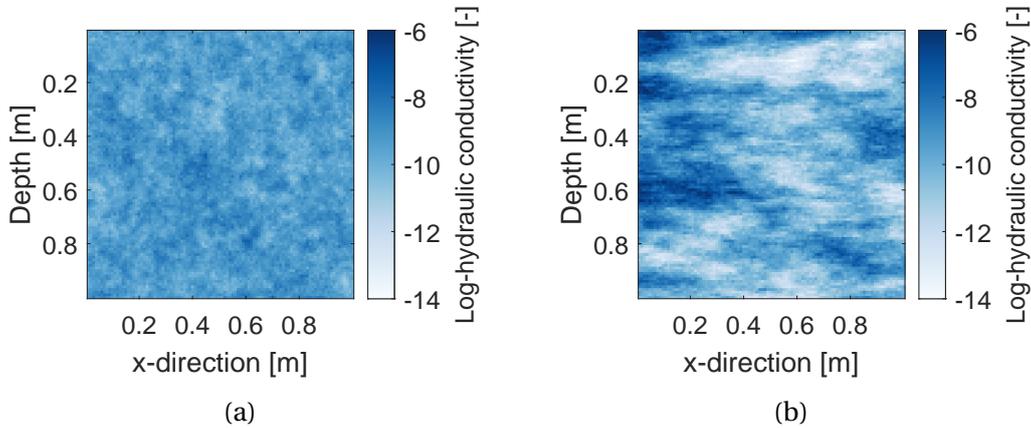


Figure 4.2: “True” log-hydraulic conductivity fields of the first test case for (a) the ergodic setting and (b) the non-ergodic setting.

For the simulated measurements, we impose a hydraulic pressure gradient along either the horizontal or the vertical direction of the target field and observe a flux across one boundary. This information can then be used to calculate the equivalent horizontal and vertical hydraulic conductivities, given by (Visentini et al. 2020),

$$K^H = \frac{1}{\Delta\Phi^H} \int_{\Gamma^H} -K(\mathbf{x}) \nabla_x \Phi^H(\mathbf{x}) d\mathbf{x}, \quad (4.15)$$

$$K^V = \frac{1}{\Delta\Phi^V} \int_{\Gamma^V} -K(\mathbf{x}) \nabla_y \Phi^V(\mathbf{x}) d\mathbf{x}, \quad (4.16)$$

where $K(\mathbf{x})$ denotes the hydraulic conductivity at position \mathbf{x} with $\mathbf{x} = (x, y)^T$ referring to the 2-D position vector. Furthermore, $\Delta\Phi^H = \Delta\Phi^V = 1$ kPa denotes the constant hydraulic pressure difference imposed along the horizontal and vertical direction, respectively and $\Phi^H(\mathbf{x})$ and $\Phi^V(\mathbf{x})$ the resulting hydraulic head. Finally, Γ^H and Γ^V refer to integration paths separating the left and right and the top and bottom boundaries, respectively.

For the ergodic and isotropic field (Fig. 4.2a), we obtain equivalent hydraulic conductivities of $K^H = K^V = 9.2 \times 10^{-5}$ m/s and for the non-ergodic anisotropic field (Fig. 4.2b) we get an equivalent horizontal hydraulic conductivity of $K^H = 6.6 \times 10^{-5}$ m/s and an equivalent vertical hydraulic conductivity of $K^V = 4.8 \times 10^{-5}$ m/s. Finally, we add *i.i.d.* relative errors $\epsilon_{\mathcal{O}}$ to the data pairs using a centred Gaussian distribution with a standard deviation given by 3% of the corresponding values.

Inversion settings and prior assumptions

The CPM method is implemented running three chains in parallel with adaptive proposals (Section 4.2.2) using an initialization period of $j_0 = 100$ where $\mathbf{C}^{(0)}$ is a diagonal matrix with (0.008, 0.008, 0.002, 0.2) along its diagonal. For the prior PDFs of the first three hyperparameters, we use Uniform distributions: for the mean μ , we use the interval $[\log(10^{-5}), \log(10^{-3})]$, a range of standard deviation σ in-between $[0, 2]$ and for the integral scale l^y we assume $[0 \text{ m}, 0.5 \text{ m}]$. To account for the anisotropy factor λ being asymmetric around one, we employ a log-Uniform distribution with boundaries $[0.1, 10]$.

To tune N and ρ in the CPM method (Section 4.2.3), we consider the variance of the log-likelihood ratio estimator W (Eq. 4.11). Figure 4.3 depicts the dependence of the variance of W on the correlation ρ for ten and fifty samples ($N = 10, 50$) of the latent variable \mathbf{X} for both the ergodic and the non-ergodic data setting. To evaluate the variances, we fix θ at values having high posterior probability and draw realizations of the field by both sampling from its prior PDF $f_{\theta}(\mathbf{x})$ (noIS) and using importance sampling (IS, Appendix 4.7.1). In the ergodic setting (Fig. 4.3a), all considered cases lead to variances of W being close to the target range between 1.0 and 2.0 recommended by Deligiannidis et al. (2018) even for $\rho = 0$. This is not surprising as in a purely ergodic setting, the realization of the random field does not influence the data. For the non-ergodic setting (Fig. 4.3b), the variance of W is up to 10^3 times higher and it is necessary to employ importance sampling. Of course, sampling from the prior could lead to variances of W being within the desired range, but the required values of N and ρ would lead to either excessively high computational costs at each iteration or

very slow mixing in the draws of the latent variables. In the limit of $\rho = 1$, the variance of W is trivially equal to zero for all settings as we use the same latent variable samples in the first and second term of W , but this would lead to biased results. Initial MCMC runs showed that very diverse values of σ , I^y and λ have high posterior probabilities in both the ergodic and non-ergodic data settings as, in both cases, non-ergodic field realizations are sampled frequently by the CPM method. To ensure a controlled variance for all values θ with high posterior probability for both data settings, we perform importance sampling and choose $N = 50$ and $\rho = 0.975$ as it is appropriate for the more challenging non-ergodic settings.

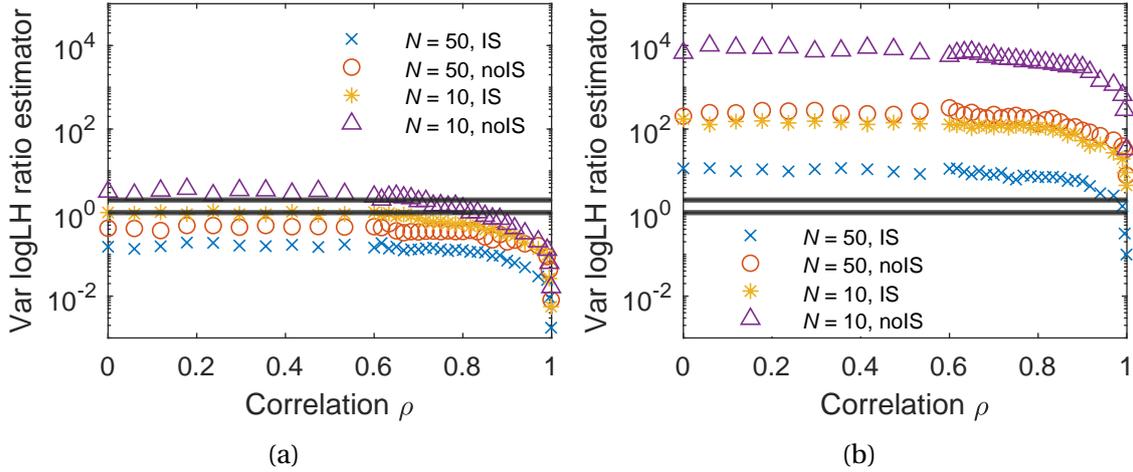


Figure 4.3: Variance of the log-likelihood ratio estimator W for the first test case (hydraulic conductivity) as a function of the correlation ρ with θ being fixed at values with high posterior probability: (a) data generated with the ergodic setting and (b) data generated with the non-ergodic setting. The different markers refer to the number of latent variable samples and if sampling is made with or without importance sampling (IS). The black horizontal lines delimit the range between 1.0 and 2.0 recommended by Deligiannidis et al. (2018).

For comparison purposes, we also run rejection sampling (RS; see Section 4.2.4) and a MH inversion assuming the parameter field to be ergodic (referred to as simplified MH; Fig. 4.4). For RS, we use the same number of field samples with corresponding forward simulations as needed by the CPM method for convergence. For the simplified MH, we rely on equations presented by Gelhar and Axness (1983) for the equivalent hydraulic conductivities in a two-dimensional anisotropic infinite domain (ensuring ergodicity) under mean uniform flow conditions:

$$K^H = K_G \left(1 + \sigma^2 \left(\frac{1}{2} - \frac{1}{1+\lambda} \right) \right), \quad K^V = K_G \left(1 + \sigma^2 \left(\frac{1}{2} - \frac{\lambda}{1+\lambda} \right) \right), \quad (4.17)$$

with K_G denoting the geometric mean of the linear hydraulic conductivity field $\mathbf{K} = \exp(\mathbf{X})$ (entry-wise exponential), which is the only parameter influencing the response for isotropic fields ($\lambda = 1$). It holds that $K_G = \exp(\bar{\mathbf{X}})$ with $\bar{\mathbf{X}}$ being the arithmetic mean of \mathbf{X} .

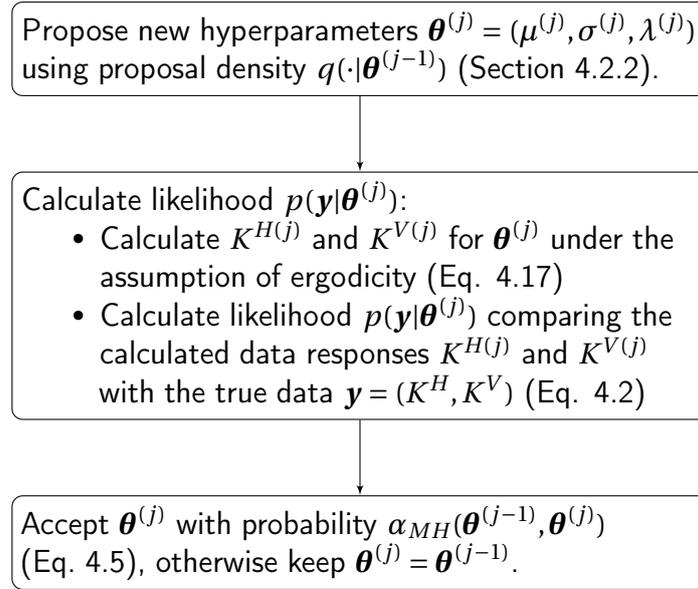


Figure 4.4: Flow chart illustrating the simplified MH (assuming ergodicity) procedure for the first test case at iteration j .

4.3.2 Results

We consider first the posterior PDF of the geostatistical mean μ obtained with the CPM method for the data generated with the ergodic setting (Fig. 4.2a). Only the samples obtained for the second half of the chains, after convergence has been declared (with respect to the \hat{R} -statistics, Table 4.1), are shown. The posterior PDF of the mean value μ in the ergodic data setting is centred around the true geostatistical mean and is clearly distinguishable from the Uniform prior PDF (Fig. 4.5a). This is confirmed by the correspondingly low logarithmic score emphasizing the accuracy of the posterior samples and the high KL divergence with respect to the prior PDF (Table 4.1). Comparison with the posterior samples obtained with RS (Fig. 4.5b) shows that both methods generate similar results with comparable KL divergences to the prior and almost equal logarithmic scores (Table 4.1). For the posterior samples obtained by assuming ergodicity (simplified MH, Fig. 4.4), we note a more compactly defined posterior than with CPM and RS with values of the mean being close to the true value (c.f., Figs. 4.5a, 4.5b and 4.5c). Still, the logarithmic score of the mean is much higher than the one obtained with CPM and RS (Table 4.1), indicating that the samples generated under ergodic assumptions are not centred around the true geostatistical mean and are overconfident. This somewhat paradoxical result is a consequence of the data setting only being nearly ergodic, demonstrating the risk of getting biased and overconfident results even when the assumption of ergodicity is nearly fulfilled. The considered measurement scale is indeed 33 times larger than the integral scale.

For the non-ergodic data setting (Fig. 4.2b), the CPM-derived posterior distribution of the mean value μ contains the true value while being shifted towards the observed equivalent properties $\log(K^H)$ and $\log(K^V)$, leading to a higher logarithmic score than in the ergodic setting (Fig. 4.5d and Table 4.1). The posterior samples obtained with RS (Fig. 4.5e) are spread

slightly wider than the ones of CPM, thereby, capturing more frequently the true value and leading to a lower KL divergence and a lower logarithmic score. We note that RS has an acceptance rate of 0.04% in this very data-poor and non-ergodic setting. In the non-ergodic setting, the simplified MH method leads to important errors in the estimated mean value (Fig. 4.5f). Indeed, the posterior samples are located around the (log-transformed) observed equivalent properties K^H and K^V and are removed from the true value of the geostatistical mean. This is reflected in a logarithmic score of infinity (see Table 4.1). Importantly, while the inversion assuming ergodicity solely samples mean values outside of the true value and has a very small posterior width, the CPM method includes the true value in the posterior samples (Figs. 4.5d and 4.5f).

For the other hyperparameters σ , I^y and λ inferred along with the mean μ , we get less well-resolved posterior estimates with both the CPM method and RS indicating that they are only weakly resolved by the available data. The corresponding plots are depicted in the Appendix 4.7.2.

Table 4.1: Table summarizing the results for the first test case (log-hydraulic conductivity) obtained with the CPM method, rejection sampling (RS) and the inversion assuming ergodicity (simplified MH; Fig. 4.4) for the ergodic (Fig. 4.2a) and the non-ergodic (Fig. 4.2b) data settings: convergence refers to the iteration in the MH methods with the \hat{R} -statistics being smaller than 1.2 for all parameters, the logarithmic score (LogS; Eq. 4.14) assesses the accuracy of the posterior samples with respect to the true value and the KL divergence (Eq. 4.13) is calculated for p_1 being the kernel density estimate gained with the (marginal) posterior samples and p_2 being the prior PDF. The bandwidth used for the marginal kernel density estimates of μ is 0.03 for CPM and RS and 0.005 for simplified MH.

Method	CPM	RS	Simp. MH	CPM	RS	Simp. MH
Dataset	Erg.	Erg.	Erg.	Non-erg.	Non-erg.	Non-erg.
Convergence	1'300	-	4'000	2'200	-	10'000
AR	15 %	0.11 %	20 %	15 %	0.04 %	15 %
LogS μ	-0.33	-0.32	17.23	1.57	1.19	Inf
KL div. μ	1.07	1.16	3.92	1.05	0.66	4.01

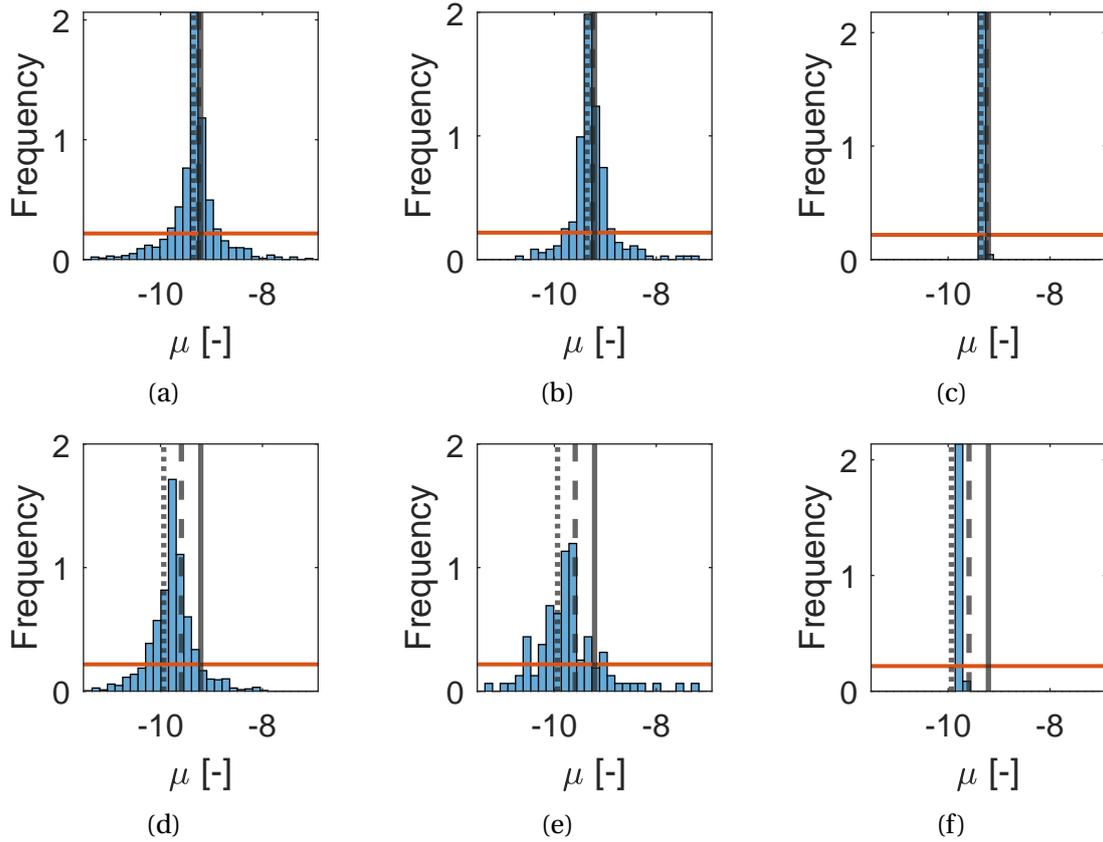


Figure 4.5: Posterior samples obtained with the CPM method, rejection sampling (RS) and an inversion assuming ergodic conditions (simplified MH, Fig. 4.4) for the mean μ in the first test case (log-hydraulic conductivity): CPM results using the data generated with the (a) ergodic setting, (d) non-ergodic setting, RS using the data generated with the (b) ergodic setting and (e) non-ergodic setting and simplified MH using the data generated with the (c) ergodic setting and (f) non-ergodic setting. The dashed line is denoting $\log(K^H)$, the dotted line $\log(K^V)$, the solid vertical line indicates the true mean value and the red horizontal line the prior PDF

4.4 Test case 2: Fracture aperture fields

Rock fractures play an important role as conduits (or barriers) for flow and solute transport. Their properties have often a major influence on hydrogeologic and geotechnical processes (National Research Council 1996), but field characterization is inherently difficult. The high contrast between the electrical properties of the filling of the fractures and the host rock (e.g., a water-filled fracture in granite host rock) leads to a very strong thin-bed response in ground penetrating radar (GPR) data. Quite remarkably, even sub-mm apertures yield measurable GPR responses even when the wavelength in the host rock may be on a metric scale. Imaging and characterization of fractures with GPR data has been studied extensively both from a theoretical perspective (e.g., Bradford and Deeds 2006; Deparis and Garambois 2009) and in controlled experiments (e.g., Grégoire and Hollender 2004; Tsoflias et al. 2015). In these studies, it is typically either assumed that the aperture and material properties do not vary

over the first Fresnel zone or that the influence of heterogeneous aperture fields average arithmetically in the acquired data. In a modeling study, Shakas et al. (2018) assess this latter simplification by exploring a deterministic inversion in which the actual aperture field is heterogeneous at small scales, while it is assumed to be homogeneous when inferred for. Despite that the data can be fitted to the noise level, they find that the estimated apertures offer only reliable approximations of the arithmetic mean of the aperture field when the correlation length of the aperture heterogeneity is larger than the first Fresnel zone. Since fractures are known to be highly heterogeneous with self-affine properties, the study by Shakas et al. (2018) suggest that many GPR-based estimations of mean apertures are biased and unreliable. They suggest that such heterogeneity needs to be explicitly accounted for, but they do not propose a solution. In this second data-rich test case, we will demonstrate how the CPM method can be used to obtain unbiased estimates of the mean aperture and statistics pertaining to the aperture field. We will then show how this information can be used to predict the equivalent hydraulic transmissivity of the fractures.

4.4.1 Data and inversion setting

We consider a $5 \text{ m} \times 5 \text{ m}$ fracture aperture field $\mathbf{X}(\cdot)$ described as an isotropic Gaussian random field $GRF(\mu_{\boldsymbol{\theta}}(\cdot), C_{\boldsymbol{\theta}}(\cdot, \cdot))$ with constant mean $\mu_{\boldsymbol{\theta}}(\cdot)$ and powered exponential covariance function $C_{\boldsymbol{\theta}}(\cdot, \cdot)$ as specified in Equation (4.1). With the CPM method, we target the mean μ , the standard deviation σ , the integral scale $I = I^x = I^y$ and the Hurst exponent H . The heterogeneous aperture field $\mathbf{X}(\cdot)$ is simulated using a pixel-based approach (Section 4.2.1) on a 50×50 -dimensional grid ($D = 50$, cells of 10 cm side-lengths).

Synthetic data generation

The fracture aperture field from which data are generated is depicted in Figure 4.6a; the true hyperparameters are $\boldsymbol{\theta} = (\mu, \sigma, I, H) = (0.5 \text{ cm}, 0.1, 0.2 \text{ m}, 0.8)$. We only consider a single fracture in a model domain of $10 \text{ m} \times 10 \text{ m} \times 10 \text{ m}$ (Fig. 4.6b). The background rock matrix is assumed to be homogeneous with a relative electrical permittivity of 9 and an electrical conductivity of 0.001 S/m. For the fracture, we assume a constant relative electrical permittivity of 81 and electrical conductivity of 0.1 S/m.

To generate the synthetic GPR reflection data, we rely on the effective-dipole method of Shakas and Linde (2015). This modeling framework combines analytical solutions for radiation in the matrix domain and dipole elements, corresponding to discretized sections of the fracture, radiating as electric dipoles modulated by the thin-bed reflection coefficients. A simple schematic of the method is represented in Figure 4.6c (adapted from Fig. 3 of Shakas et al. 2018). We use two GPR reflection traces generated with sources and receivers located 5 m away from the fracture and with offsets of 0 m and 2 m (Fig. 4.6b). The source signal is assumed to be vertically-oriented with a source spectrum consisting of a Ricker wavelet with dominant wavelength of 100 cm. With a discretization of 10 cm of the aperture field, this results in 10 elements per dominant wavelength for which highly accurate simulations are expected (Shakas et al. 2018). The responses are generated in the frequency-domain using a frequency range from zero to 300 MHz with a sampling step size of 1 MHz. As in practice,

the amplitude of the source wavelet is unknown, Shakas et al. (2018) normalize the response values in the data generation and inversion. Here, we instead introduce an unknown factor c_A by which the responses are multiplied. This factor is equal to one for the true data and it is inferred within the inversion. This extends the target variables to $\theta = (\mu, \sigma^2, l_c, H, c_A)$. Finally, for each of the 300 complex-valued numbers representing the electric field, we add independent realizations of Gaussian measurement noise ε_{θ} with a standard deviation of 3 % of the maximal value. The inversions are performed in the frequency-domain, but we present for visual purposes the two corresponding traces in the time-domain (Figure 4.6d). For completeness, we also show the smoother traces (Figure 4.6e) obtained by sampling over the same frequency range with a sampling step size of 0.1 MHz.

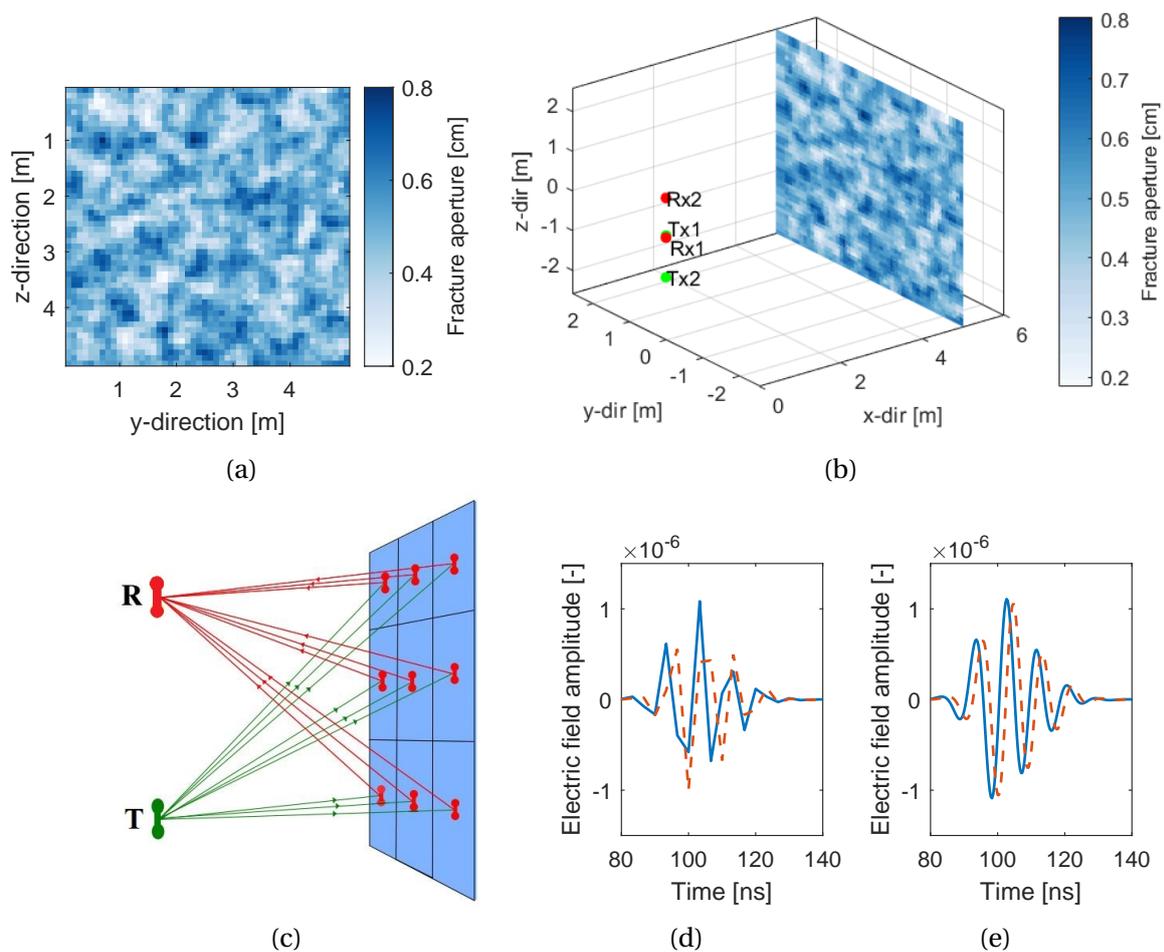


Figure 4.6: (a) “True” fracture aperture field of the second test case with $\theta = (0.5 \text{ cm}, 0.1, 0.2 \text{ m}, 0.8, 1)$. (b) Model domain with aperture field and transmitter-receiver layout. (c) Schematic of the effective-dipole forward modeling framework (adapted from Fig. 3 of Shakas et al. 2018). GPR reflection traces (time-domain) with a sampling step size of (d) 1 MHz and (e) 0.1 MHz.

Inversion settings and prior assumptions

As in the first test case, we run an adaptive Metropolis–Hastings version of the CPM method with three chains in parallel. We specify $j_0 = 500$ and $\mathbf{C}^{(0)}$ as a diagonal matrix with 0.001 on its diagonal. Furthermore, to ensure a suitable acceptance rate, we decrease the step size by 50 %. For the prior PDFs of the hyperparameters, we use Uniform distributions: for the mean μ we use Unif[0 cm, 1 cm], for the standard deviation σ we use Unif[0 cm, 0.5 cm], for the integral scale l_c we use Unif[0 m, 1 m], for the Hurst exponent H we use Unif[0.1, 1] and for the amplitude factor c_A we use Unif[0.5, 2].

The importance sampling mean $\boldsymbol{\mu}_{IS}$ for the latent aperture field \mathbf{X} when the proposed $\boldsymbol{\theta}^{(j)}$ is the true hyperparameters is depicted in Figure 4.7a (see formulas in Appendix 4.7.1). Figure 4.7b depicts the dependence of the variance of the log-likelihood ratio estimator W (Eq. 4.11) on N and ρ . The importance sampling leads to a tremendous decrease of the variance of W (e.g., for $N = 1$ and $\rho = 0$, the variance of W is reduced from 10^6 to 10^2). Furthermore, increasing the number of latent variable samples N and the correlation parameter ρ also reduces the variance of W strongly. Following Figure 4.7b, we run the CPM algorithm with $N = 5$ and $\rho = 0.975$ in combination with importance sampling.

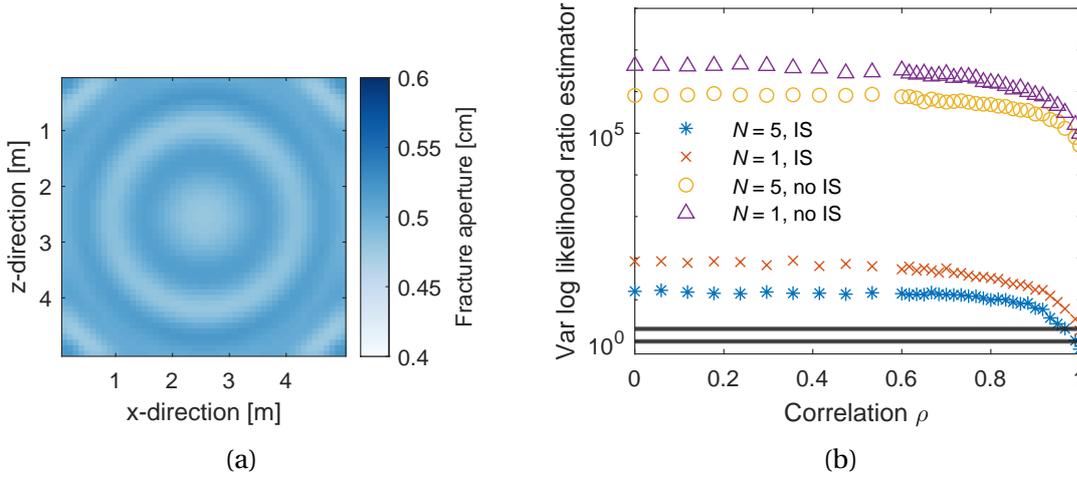


Figure 4.7: (a) Importance sampling mean $\boldsymbol{\mu}_{IS}$ of the aperture field for the true values of $\boldsymbol{\theta}$ in the second test case (Appendix 4.7.1) and (b) variance of the log-likelihood ratio estimator W as a function of the correlation ρ with $\boldsymbol{\theta}$ being fixed at values with high posterior probability. The different markers refer to different numbers N of latent variable samples drawn with or without importance sampling (IS). The black horizontal lines delimit the range between 1.0 and 2.0 recommended by Deligiannidis et al. (2018).

To place the results obtained with CPM into context, we compare them to those obtained with an inversion assuming the aperture field to be homogeneous (illustrated with a flow chart in Fig. 4.8). To achieve this, we only infer the mean aperture μ and the amplitude factor c_A , which is broadly similar to the inversion setting considered by Shakas et al. (2018).

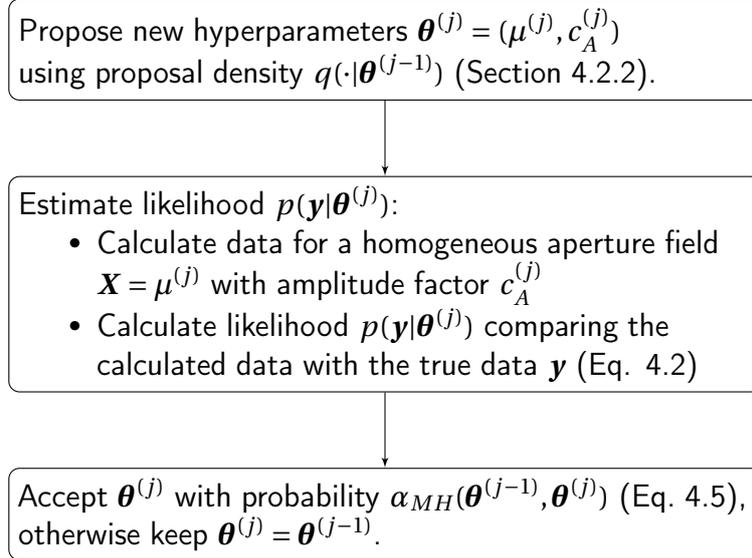


Figure 4.8: Flow chart illustrating the homogeneous inversion procedure for the second test case at iteration j .

4.4.2 Results

The estimated marginal posterior PDFs of $\theta = (\mu, \sigma^2, l_c, H, c_A)$ obtained with the CPM method are depicted in Figures 4.9a-4.9e. Convergence is reached within 10,000 iterations with respect to the \hat{R} -statistic and we display the results for the second half of the chains. The histograms depicting the posterior samples of the mean μ (Fig. 4.9a), standard deviation σ (Fig. 4.9b) and amplitude factor c_A (Fig. 4.9e) show the strongest concentration with respect to the prior and correspondingly high KL divergences (Table 4.2). The sample range for the integral scale I (Fig. 4.9c) and the Hurst exponent H (Fig. 4.9d) is equally wide as the respective prior PDFs and the corresponding KL divergences are rather small. Nonetheless, the integral scale is preferentially sampled in the region of the true value. As the values of the logarithmic score (Table 4.2) can generally not be compared between hyperparameters (different width of support), they will become of interest only in the comparison with a competing method.

Figures 4.9f and 4.9g show the histograms of the posterior samples for the mean μ and the amplitude factor c_A obtained for the inversion assuming a homogeneous aperture (Fig. 4.8). The range of the samples is very narrow with high KL divergences with respect to the prior PDF (Table 4.2) but located far from the true parameter values. This results in infinite logarithmic scores (Table 4.2). As we have seen, the CPM method accounting for heterogeneity leads

to posterior samples of the mean and amplitude factor that cover a wider range including the true values used to generate the data as reflected in lower logarithmic scores (Table 4.2). The estimates of the mean aperture μ and the amplitude factor c_A are highly correlated. Figure 4.9h shows that under the assumption of knowing $c_A = 1$, the range of the samples obtained with CPM for the mean aperture would be more narrow and shifted towards the true value of 0.5 cm. We further see that the homogeneous inversion only explores a small part (and the wrong part) of the joint posterior model space, leading to a logarithmic score of infinity for the estimated joint posterior PDF (Table 4.2). For this second data-rich test case, rejection sampling is unfeasible as the acceptance rate is below 0.001 %.

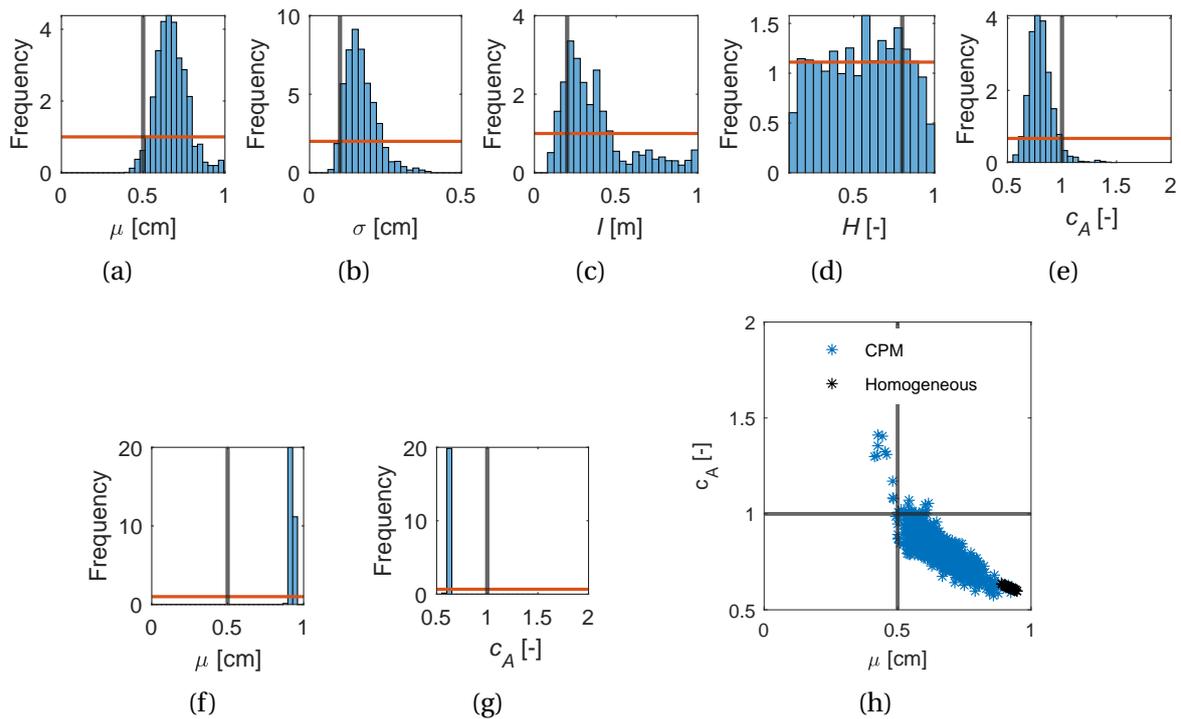


Figure 4.9: Posterior samples obtained with CPM in the second test case for the (a) mean aperture μ , (b) standard deviation σ , (c) integral scale l , (d) Hurst exponent H and (e) amplitude factor c_A . Posterior samples obtained with an inversion assuming the aperture field to be homogeneous for the (f) mean aperture μ and (g) amplitude factor c_A . Scatter plot of the sampled pairs of mean aperture μ and amplitude factor c_A for the CPM method and the homogeneous inversion. The black lines indicate the true values and the red horizontal lines the prior PDFs.

Table 4.2: Table summarizing the results for the second test case (aperture field, Fig. 4.6) obtained with the CPM method and the inversion assuming homogeneity (Fig. 4.8): convergence refers to the iteration with a \hat{R} -statistics being smaller than 1.2 for all parameters, the logarithmic score (logS; Eq. 4.14) evaluates the accuracy of the marginal and joint posterior samples with respect to the true value and the KL divergence (Eq. 4.13) is calculated for p_1 being the marginal kernel density estimate gained with the posterior samples and p_2 being the prior PDF. For the kernel density estimates of CPM, we use the following bandwidths: 0.01 (μ), 0.006 (σ), 0.02 (I), 0.04 (H), 0.01 (c_A) and 0.02 for both in the joint PDF of (μ, c_A). For the homogeneous inversion, we use 0.001 (μ), 0.001 (c_A) and 0.002 for both in the joint PDF of (μ, c_A).

Method	CPM	Homogeneous inv.
Convergence	10'000	1'000
Acceptance rate	15 %	15 %
LogS μ	0.60	Inf
LogS σ	-1.34	-
LogS I	-1.05	-
LogS H	-0.27	-
LogS c_A	0.51	Inf
LogS (μ, c_A)	-1.31	Inf
KL divergence μ	0.96	3.31
KL divergence σ	0.95	-
KL divergence I	0.35	-
KL divergence H	0.02	-
KL divergence c_A	1.29	3.90

Predictions of hydraulic transmissivity

To complement the results obtained for this GPR test case and to strengthen the link to hydrogeology, we use the aperture field estimates to derive equivalent hydraulic transmissivities. First, we use the inferred mean apertures obtained with the inversion assuming the field to be homogeneous (Fig. 4.9f). These aperture field realizations are used to derive hydraulic transmissivities at the fracture scale using the classical parallel plate model (Tsang 1992),

$$T = (1/(12\eta))\mu^3, \quad (4.18)$$

with $\eta = 8.9 \times 10^{-4}$ Pa·s denoting the dynamic viscosity (25 degree C) and μ being the inferred mean aperture values (in meters). The resulting horizontal equivalent log-hydraulic transmissivities are shown in Figure 4.10 (light gray). This result is now compared with the value obtained for the true aperture field under the assumption that the Reynolds equation is valid, implying that we can apply Equation (4.18) locally to obtain a hydraulic transmissivity field and then solve numerically for the resulting effective transmissivity at the fracture scale. The results show that the true effective hydraulic transmissivity is roughly one order of magnitude smaller and that the posterior PDF of the homogeneous inversion (light gray) is nowhere close to include this value. This is reflected in a infinite logarithmic score. We then sample

field realizations using the posterior PDFs of the hyperparameters inferred with the CPM method. The resulting equivalent log-hydraulic transmissivity values are shown in Figure 4.10 (blue). This distribution is much wider, it includes the true value, and the mean is clearly shifted towards the true value. The corresponding logarithmic score is 0.23.

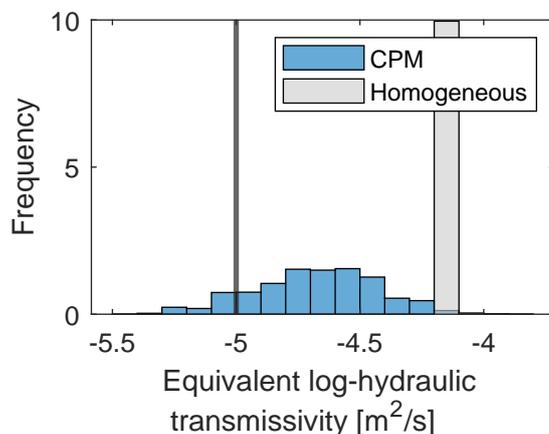


Figure 4.10: Posterior PDF of the horizontal equivalent log-hydraulic transmissivity (log10-scale) for heterogeneous aperture fields (second test case) sampled with the inferred hyperparameters of the CPM method (blue) and homogeneous aperture fields generated with the inferred mean values of the inversion assuming the field to be homogeneous (light gray). The black vertical line indicates the value corresponding to the true aperture field.

4.5 Discussion

Our two test cases presented in Sections 4.3 and 4.4 demonstrate the ability of the correlated pseudo-marginal method (CPM method; Fig. 4.1) to estimate the posterior PDF of the target field’s hyperparameters (e.g., mean, standard deviation, integral scale, Hurst exponent and anisotropy factor) while accounting for the impact of small-scale heterogeneity within the estimate of the likelihood function. We further demonstrate that inversions invoking simplified assumption such as ergodicity or homogeneity lead to biased and overconfident results such that the inferred posteriors often do not include the true values. Compared to previous inversion approaches targeting hyperparameters (e.g., Laloy et al. 2015 and Xiao et al. 2021), the CPM method infers the hyperparameters only, thereby, avoiding to infer the posterior PDF of the many thousands of latent variables. The two presented test cases cover one data-poor transmission problem governed by diffusion (e.g. electrical conduction, heat conduction, or groundwater flow) and one more data-rich reflection problem governed by wave-based physics (e.g., GPR, seismics, acoustics). The generality of these settings suggest that the CPM method has a wide applicability in hydrogeology and geophysics.

The first test case related to heterogeneous hydraulic conductivity fields concerns a very data-poor setting in which only the horizontal and vertical equivalent hydraulic conductivities are used as data points. To compare the performance of the CPM method with an inversion assuming ergodicity (referred to as simplified MH; Fig. 4.4), we consider a nearly ergodic and a non-ergodic data setting. In both settings, the geostatistical mean of the model domain can be inferred from the equivalent conductivities using the CPM method. In the ergodic setting, both the CPM method and the simplified MH lead to reasonable estimates of the geostatistical mean, with the posterior range of the CPM method being wider as its underlying assumptions are less restrictive (Figs. 4.5a and 4.5c). In the non-ergodic data setting, the simplified MH leads to important errors in the estimation of the geostatistical mean with a posterior range far from the true value (Fig. 4.5f). For the CPM method, the estimated posterior uncertainty is wider and the true value of the mean is included (Fig. 4.5d). Thereby, the logarithmic score is reduced from infinity to 1.57 when applying CPM compared with the simplified MH (Table 4.1). We conclude that even in this extremely data-poor setting, the use of simplified model assumptions leads to a substantial bias in the mean estimate and an overconfident posterior bound. For the other hyperparameters (standard deviation, integral scale and anisotropy factor), we conclude that only little information can be gained in this data-poor setting. Furthermore, we demonstrate that the CPM results are in agreement with those obtained by rejection sampling.

In the second test case concerning fracture aperture fields, we limit ourselves to a non-ergodic data setting and compare the results obtained with CPM with those of an inversion assuming the aperture field to be homogeneous (Fig. 4.8). We can consider this homogeneous inversion as either (1) an inversion inferring the geostatistical mean under simplified model assumptions or (2) an inversion targeting the equivalent GPR aperture. We show that the homogeneous assumption leads to posterior samples being located far from the true geostatistical mean value (Fig. 4.9f), demonstrating in accordance with Shakas et al. (2018) (1) that the geostatistical mean of the aperture field can be very different than the equivalent GPR aperture and (2) that inferring the geostatistical mean based on a too simple model description leads to biased estimates. Indeed, in such an inversion one appears to get increasingly certain about the wrong parameter values as more data are added or the data noise level is decreased (Brynjarsdóttir and O'Hagan 2014). In contrast, the CPM method accounting for non-ergodicity and heterogeneity by inferring additionally for the standard deviation, integral scale and Hurst exponent leads to a wider posterior including the true value of the aperture mean (Fig. 4.9a). For this second example, employing the CPM method leads to a reduction of the logarithmic score from infinity to 0.60 for the posterior estimate of the mean in comparison with the homogeneous inversion (Table 4.2). Additionally, CPM enables to infer information about other hyperparameters (standard deviation and integral scale) of the field.

Probabilistic inference of hyperparameters offers the possibility to translate from one type of equivalent property to another. We demonstrate this by predicting the equivalent log-hydraulic transmissivity at the fracture scale using the fracture aperture fields obtained in the second test case (Fig. 4.10). The predicted values for the constant aperture field inversion are obtained by applying the equivalent GPR aperture in the cubic law. When deriving hydraulic

properties from these constant fields, we assume that this equivalent GPR aperture is the same as the equivalent “cubic law aperture” (in the sense of Tsang 1992), which is the equivalent parallel plate aperture with respect to hydraulic flow properties. These predictions are very different from those obtained from the true aperture field when applying the local cubic law (Fig. 4.10). This visualises clearly that the equivalent aperture for one type of physics cannot be assumed to be the same when considering another type of physics. Actually, the equivalent aperture (in a cubic law sense) with respect to the hydraulic data of the true aperture field is 0.47 cm, a value considerably diverging from the one inferred from the GPR data when assuming homogeneity (about 0.9 cm). Using field realizations sampled with the posterior PDFs of the hyperparameters obtained by CPM lead to a wider and more accurate range of effective log-hydraulic transmissivity values (Fig. 4.10). While the logarithmic score for the transmissivity predictions obtained with the homogeneous inversion is infinity, the one obtained with CPM is 0.23. This suggests that while equivalent properties always refer to one specific kind of physics, the inference of hyperparameters enables a general description of the model domain. The CPM method is well suited to achieve this by targeting only the hyperparameters of interest, thereby, enabling probabilistic forecasts for different types of physics.

This study expands further the range of applications that the CPM method can address in geoscientific settings. While Friedli et al. (2022) used it to account for uncertainties in petrophysical relationships in the context of hydrogeophysics, we provide here a very different problem setting in which the CPM method is used to account for non-ergodicity and small-scale heterogeneities when inferring hyperparameters. In these examples, we only consider heterogeneities in two dimensions. In field applications, the data are of course affected by heterogeneities outside the 2-D plane of measurements (e.g., between boreholes) or by outer-space effects (Maurer and Friedel 2006). To further improve the estimation and uncertainty quantification in such setups, the CPM method could be employed to integrate out heterogeneities in three dimensions (in the context of the present study), or in the third out-of-plane dimension in the setting considered by Friedli et al. (2022) or in general 2-D inversions to avoid over-confident (and possibly biased) estimates. For the presented test cases we used a pixel-based representation of the Gaussian latent random field. We stress that there exist many alternative ways to represent and generate a Gaussian random field as, for example, the fast circulant embedding technique using a spectral representation by Dietrich and Newsam (1997). While such an approach offers an increased efficiency in the generation of the random field realizations, careful consideration must be given on a case-by-case basis as to whether this could be integrated into a well-working importance sampling strategy. Moreover, in settings where the correlation length is of similar size as the model domain, the embedding has to be extended and the efficiency is reduced. We assume the latent random fields to be Gaussian, simplifying the derivation of the importance sampling density. An important topic for future research would be to develop and assess suitable importance distributions in non-Gaussian settings.

The efficiency of the CPM method depends strongly on the variance of the log-likelihood ratio estimator. Especially in settings with a high number of observations with a low signal-to-noise ratio, one needs a well-working importance density when sampling the latent variables. The

relevance of a well-tuned importance sampling strategy becomes clear when comparing the number of samples needed to control the variance in the first and second test cases (Figs. 4.3 and 4.7b). For the first test case, the IS density is of only moderate quality and many samples ($N = 50$) are needed even for this data-poor setting. For the second more data-rich test case with a well-defined IS density, only a few samples ($N = 5$) are sufficient. If the determination of a well-working IS distribution is not feasible, this can be detrimental to the applicability of the CPM method. In such a scenario there is also the risk of poor exploration of the latent space, namely if the likelihood estimator depends mainly on one or two latent variable samples with a particularly beneficial small-scale structure. One solution in such a scenario is to infer some additional main features of the latent field together with the hyperparameters and then to apply the CPM method to sample out the remaining randomness of the field. This could be done using the main components of a dimensionality reduction approach and should reduce the importance of a well-tuned IS density. We leave this idea for future research.

Recently, Wang et al. (2022) proposed an hierarchical Bayesian inversion approach targeting first so-called global variables (such as hyperparameters but also physical variables) and then estimating the posterior of the whole field. For the estimation of the global variable's posterior in a non-linear setting, Wang et al. (2022) apply a machine-learning based approach and train a neural network to output the global variables given a data realization followed by kernel density estimation of the results. Such a method relies on the ability to estimate the hyperparameters by brute-force prior sampling and subsequent comparison of the resulting data with the true measurements. In strongly non-ergodic settings, this can be computationally challenging as an unrealistically high number of prior samples would be needed to obtain reasonable estimates. To illustrate this, Figure 4.11 shows the 100 highest log-likelihood values sampled from 5000 prior samples of the aperture field in the second test case (Section 4.4). We note that no sample was generated with a likelihood close to the true one (black horizontal line) implying that an unfeasible large amount of samples would be needed to guarantee accurate hyperparameter estimates. Indeed, even the highest sampled likelihood has a likelihood that is still 10^{44} times smaller than the true likelihood. In contrast, our CPM method using three chains need 10'000 iterations per chain for convergence.

4.6 Conclusions

We consider Bayesian MCMC inversions inferring hyperparameters (e.g., mean, standard deviation and integral scales) from hydrogeological or geophysical data. To achieve this is particularly challenging in the non-ergodic setting, in which the data depend on the actual geostatistical field realization under consideration and not only on the hyperparameters. To prevent errors arising when assuming homogeneity or ergodicity, we rely on the correlated pseudo-marginal method targeting the hyperparameters while integrating out the random effects of actual field realizations in the likelihood estimation. This approach has the advantage of ensuring accurate posterior estimates of hyperparameters without having to infer thousands or more parameters as needed if the whole random field would be inferred. To ensure efficiency, the correlated pseudo-marginal method employs importance sampling and correlation of the latent draws used in the proposed and current steps of the MCMC chain.

We assess the performance of this method through two synthetic test cases involving (1) diffusion-based physics in a data-poor setting targeting hydraulic properties using equivalent hydraulic conductivity data and (2) wave-based physics in a more data-rich example targeting a fracture aperture field using single-hole ground-penetrating radar (GPR) reflection data. By using these two examples that are representative of a broad range of geophysical and hydrogeological problems, we demonstrate that the correlated pseudo-marginal method provides accurate estimation of the geostatistical mean in both ergodic and non-ergodic settings. Furthermore, for all considered hyperparameters, we show that the correlated pseudo-marginal method avoids over-confident and biased posterior PDF estimates that plague inversion results obtained when assuming ergodicity or homogeneity. Estimating hyperparameters allows for a general description of property fields which is independent of the physics under consideration, thereby, allowing ultimately to use the estimated posterior PDFs to make predictions for other types of physics or experimental set-ups. This is demonstrated by transforming the fracture properties inferred by GPR data into predictions of equivalent hydraulic transmissivity at the fracture scale.

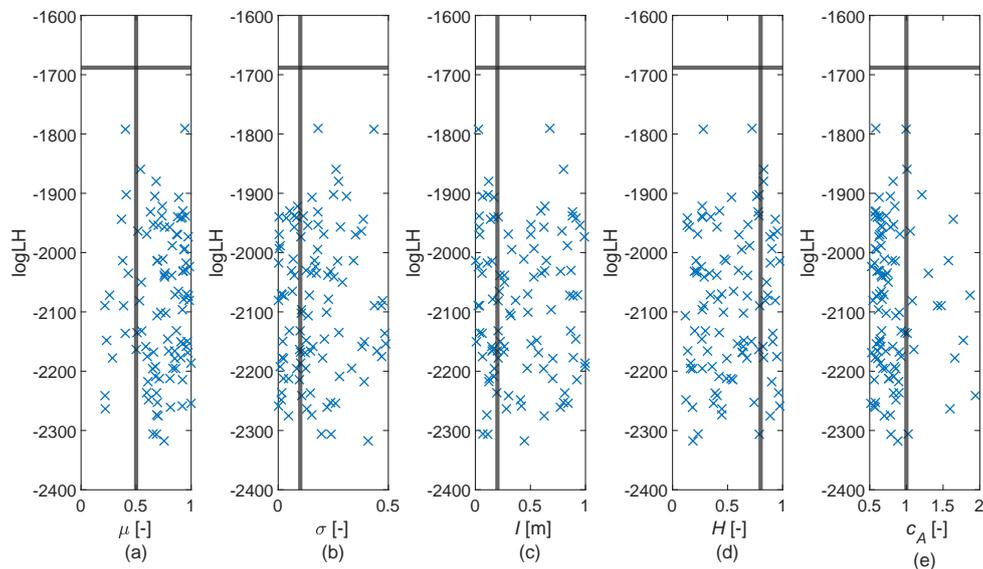


Figure 4.11: Log-likelihood values obtained with brute-force sampling of the latent aperture field \mathbf{X} in Test case 2 (Section 4.4) when drawing prior samples from the hyperparameters; for readability, we only show the samples with the 100 highest log-likelihood values (out of 5000 samples). The horizontal lines depict the true log-likelihood and the vertical lines the true values of the hyperparameters.

4.7 Appendix

4.7.1 Importance sampling for CPM

As emphasized in Section 4.2.3, it is essential that the variance of the log-likelihood ratio estimator is low to ensure efficient PM or CPM performance. Especially in settings with large data sets with high signal-to-noise ratios, the integrand $g_{\theta}(\mathbf{y}|\mathbf{x})$ in Equation (4.6) is likely to have a peak in a region of \mathbf{X} with very small probability under its prior PDF $\mathbf{x} \mapsto f_{\theta}(\mathbf{x})$. Sampling the latent variables using the prior distribution then leads to an inefficient algorithm. One remedy to this is importance sampling, where instead of the prior distribution, a so-called importance distribution given by the PDF $\mathbf{x} \mapsto m_{\theta}(\mathbf{x})$ is employed. The importance distribution is chosen such that it preferentially generates samples with high $g_{\theta}(\mathbf{y}|\mathbf{x})f_{\theta}(\mathbf{x})$ while guaranteeing that all values \mathbf{x} , for which $g_{\theta}(\mathbf{y}|\mathbf{x})f_{\theta}(\mathbf{x}) > 0$, are included in its support (Owen and Zhou 2000). It holds,

$$\int g_{\theta}(\mathbf{y}|\mathbf{x})f_{\theta}(\mathbf{x})d\mathbf{x} = \int \frac{g_{\theta}(\mathbf{y}|\mathbf{x})f_{\theta}(\mathbf{x})}{m_{\theta}(\mathbf{x})}m_{\theta}(\mathbf{x})d\mathbf{x}, \quad (4.19)$$

from which the unbiased importance sampling estimate of the likelihood function in Equation (4.9) is derived. To minimize the variance of the estimator, $\mathbf{x} \mapsto m_{\theta}(\mathbf{x})$ should be nearly proportional to $\mathbf{x} \mapsto g_{\theta}(\mathbf{y}|\mathbf{x})f_{\theta}(\mathbf{x})$ as presented in Owen and Zhou (2000) referring to the results of Kahn and Marshall (1953). Since it holds that $p(\mathbf{x}|\boldsymbol{\theta}, \mathbf{y}) \propto g_{\theta}(\mathbf{y}|\mathbf{x})f_{\theta}(\mathbf{x})$, we use for the importance density an approximation of $\mathbf{x} \mapsto p(\mathbf{x}|\boldsymbol{\theta}, \mathbf{y})$.

In Section 4.2.1, we specify $\mathbf{Y} = \mathcal{G}(\mathbf{X}) + \boldsymbol{\varepsilon}_{\theta}$ with $\mathcal{G} : \mathbb{R}^{D^2} \rightarrow \mathbb{R}^T$ being the (physical) forward solver and $\boldsymbol{\varepsilon}_{\theta}$ the observational noise. If the forward solver is non-linear, there is no exact expression for $p(\mathbf{x}|\boldsymbol{\theta}, \mathbf{y})$. For this reason, we approximate $p(\mathbf{x}|\boldsymbol{\theta}, \mathbf{y})$ by expressing the map $\mathbf{x} \mapsto \mathcal{G}(\mathbf{x})$ based on either an upscaling formula assuming an anisotropic ergodic setting (test case 1) or a linearization of the forward solver (test case 2). Following Friedli et al. (2022), we use Gaussian distributions and a lemma about marginal and conditional Gaussians; see for example Bishop and Nasrabadi (2006).

Lemma 2 *Marginal and Conditional Gaussians*

Assume a marginal Gaussian distribution for $\mathbf{X} \in \mathbb{R}^{D^2}$ and a conditional Gaussian distribution for $\mathbf{Y} \in \mathbb{R}^T$ given \mathbf{X} in the form

$$p(\mathbf{x}) = \varphi_T(\mathbf{x}; \boldsymbol{\mu}, \Lambda^{-1}), \quad (4.20)$$

$$p(\mathbf{y}|\mathbf{x}) = \varphi_T(\mathbf{y}; \mathbf{A}\mathbf{x} + \mathbf{b}, \mathbf{L}^{-1}), \quad (4.21)$$

with $\varphi_T(\cdot; \boldsymbol{\mu}, \mathbf{K})$ denoting the PDF of the T -variate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix \mathbf{K} . Then, the marginal distribution of \mathbf{Y} and the conditional distribution of \mathbf{X} given \mathbf{Y} are given by

$$p(\mathbf{y}) = \varphi_T(\mathbf{y}; \mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{L}^{-1} + \mathbf{A}\Lambda^{-1}\mathbf{A}^T), \quad (4.22)$$

$$p(\mathbf{x}|\mathbf{y}) = \varphi_L(\mathbf{x}; \Sigma(\mathbf{A}^T\mathbf{L}(\mathbf{y} - \mathbf{b}) + \Lambda\boldsymbol{\mu}), \Sigma), \quad (4.23)$$

where $\Sigma = (\Lambda + \mathbf{A}^T\mathbf{L}\mathbf{A})^{-1}$.

Test case 1: Hydraulic conductivity fields

We are concerned with the latent log-conductivity field $X \sim f_{\theta}(\mathbf{x}) = \varphi_L(\mathbf{x}; \boldsymbol{\mu}_{\theta}, \boldsymbol{\Sigma}_{\theta})$ and, to express the measurements as a linear function of the latent field, we use the log-transformed Equation (4.17) out of Sanchez-Vila et al. (2006),

$$\log(K^H) = \bar{X} + \log\left(1 + \sigma^2\left(\frac{1}{2} - \frac{1}{1+\lambda}\right)\right) \quad (4.24)$$

$$\log(K^V) = \bar{X} + \log\left(1 + \sigma^2\left(\frac{1}{2} - \frac{\lambda}{1+\lambda}\right)\right). \quad (4.25)$$

Then, we denote $\tilde{\mathbf{y}} = \log(\mathbf{y})$ and write $\tilde{\mathbf{y}} \approx \mathbf{J}\mathbf{X} + \mathbf{b}$ with \mathbf{J} being a $(2, D^2)$ -dimensional matrix with constant entries of $\frac{1}{D^2}$ and $\mathbf{b} = (b_1, b_2)$, where $b_1 = \log\left(1 + \sigma^2\left(\frac{1}{2} - \frac{1}{1+\lambda}\right)\right)$ and $b_2 = \log\left(1 + \sigma^2\left(\frac{1}{2} - \frac{\lambda}{1+\lambda}\right)\right)$. Subsequently, we approximate $g_{\theta}(\tilde{\mathbf{y}}|\mathbf{x})$ with $\tilde{g}_{\theta}(\tilde{\mathbf{y}}|\mathbf{x}) = \varphi_T(\tilde{\mathbf{y}}; \mathbf{J}\mathbf{X} + \mathbf{b}, \tilde{\boldsymbol{\Sigma}}_Y)$, where $\tilde{\boldsymbol{\Sigma}}_Y = I_2 * 0.1^2$, with I_2 denoting the two by two identity matrix. With this choice of $\tilde{\boldsymbol{\Sigma}}_Y$, we transfer the observational error to the log-space and artificially increase the uncertainty to account for the errors resulting from the ergodic assumption made to derive the IS distribution. To finally derive an approximation for $p(\mathbf{x}|\boldsymbol{\theta}, \mathbf{y})$, we use:

$$\begin{aligned} \tilde{p}(\mathbf{x}|\boldsymbol{\theta}, \tilde{\mathbf{y}}) &= \varphi_L(\mathbf{x}; \boldsymbol{\mu}_{IS}, \boldsymbol{\Sigma}_{IS}), \text{ with} & (4.26) \\ \boldsymbol{\mu}_{IS} &= \boldsymbol{\Sigma}_{IS} \left(\mathbf{J}^T \tilde{\boldsymbol{\Sigma}}_Y^{-1} (\tilde{\mathbf{y}} - \mathbf{b}) + \boldsymbol{\Sigma}_{\theta}^{-1} \boldsymbol{\mu}_{\theta} \right), \\ \boldsymbol{\Sigma}_{IS} &= (\boldsymbol{\Sigma}_{\theta}^{-1} + \mathbf{J}^T \tilde{\boldsymbol{\Sigma}}_Y^{-1} \mathbf{J})^{-1}. \end{aligned}$$

Test case 2: Fracture aperture fields

We target the fracture aperture field $X \sim f_{\theta}(\mathbf{x}) = \varphi_L(\mathbf{x}; \boldsymbol{\mu}_{\theta}, \boldsymbol{\Sigma}_{\theta})$ and locally approximate $p(\mathbf{x}|\boldsymbol{\theta}, \mathbf{y})$ by expressing the map $\mathbf{x} \mapsto \mathcal{G}(\mathbf{x})$ based on a first-order expansion around \mathbf{x}_{lin} (as Friedli et al. 2022),

$$\mathcal{G}(\mathbf{x}) = \mathcal{G}(\mathbf{x}_{\text{lin}} + \mathbf{x} - \mathbf{x}_{\text{lin}}) \approx \mathcal{G}(\mathbf{x}_{\text{lin}}) + \mathbf{J}_{\mathbf{x}_{\text{lin}}}(\mathbf{x} - \mathbf{x}_{\text{lin}}). \quad (4.27)$$

Here, $\mathbf{J}_{\mathbf{x}_{\text{lin}}}$ refers to the sensitivity (Jacobian) matrix of the forward solver corresponding to \mathbf{x}_{lin} , which is a homogeneous field with the currently proposed geostatistical mean ($\boldsymbol{\mu}_{\theta} = \theta_1$). Subsequently, we approximate $g_{\theta}(\mathbf{y}|\mathbf{x})$ with $\tilde{g}_{\theta}(\mathbf{y}|\mathbf{x}) = \varphi_T(\mathbf{y}; \mathcal{G}(\mathbf{x}_{\text{lin}}) + \mathbf{J}_{\mathbf{x}_{\text{lin}}}(\mathbf{x} - \mathbf{x}_{\text{lin}}), \boldsymbol{\Sigma}_Y)$. Again, we derive an approximation for $p(\mathbf{x}|\boldsymbol{\theta}, \mathbf{y})$ by using Equation (4.23):

$$\begin{aligned} \tilde{p}(\mathbf{x}|\boldsymbol{\theta}, \mathbf{y}) &= \varphi_L(\mathbf{x}; \boldsymbol{\mu}_{IS}, \boldsymbol{\Sigma}_{IS}), \text{ with} & (4.28) \\ \boldsymbol{\mu}_{IS} &= \boldsymbol{\Sigma}_{IS} \left(\mathbf{J}_{\mathbf{x}_{\text{lin}}}^T \boldsymbol{\Sigma}_Y^{-1} (\mathbf{y} - (\mathcal{G}(\mathbf{x}_{\text{lin}}) - \mathbf{J}_{\mathbf{x}_{\text{lin}}} \mathbf{x}_{\text{lin}})) + \boldsymbol{\Sigma}_{\theta}^{-1} \boldsymbol{\mu}_{\theta} \right), \\ \boldsymbol{\Sigma}_{IS} &= (\boldsymbol{\Sigma}_{\theta}^{-1} + \mathbf{J}_{\mathbf{x}_{\text{lin}}}^T \boldsymbol{\Sigma}_Y^{-1} \mathbf{J}_{\mathbf{x}_{\text{lin}}})^{-1}. \end{aligned}$$

Since this expression is approximate due to the linearization step, we multiply $\boldsymbol{\Sigma}_Y$ with a factor. Following Friedli et al. (2022), we use 1.2 as this choice led to a satisfactory performance.

4.7.2 Complementary figures concerning test case 1

Here we present the additional hyperparameter-plots of the posterior samples obtained for the first test case (Section 4.3). Figure 4.12 shows the results obtained for the ergodic data setting (Fig. 4.2a) and Figure 4.13 those for the non-ergodic setting (Fig. 4.2b). In the ergodic setting, the posteriors of the standard deviation σ obtained by CPM (Fig. 4.12a), RS (Fig. 4.12d) and the simplified MH (Fig. 4.12g) are as wide as the prior PDFs with the mode of the distributions being located around the right value for all three approaches. Thereby, the posterior obtained with the simplified MH is better defined than the ones of RS and CPM. The integral scale is only inferred by CPM (Fig. 4.12b) and RS (Fig. 4.12e) with both methods generating posterior samples still distributed proportionally to the Uniform prior PDF. For the anisotropy factor λ , the simplified MH clearly favours values above 1 (Fig. 4.12h) and the same holds true for RS (Fig. 4.12f). Finally, CPM (Fig. 4.12c) samples close to the prior PDF.

Employing the data generated with the non-ergodic setting, we obtain posteriors favoring correctly the horizontal layering of the field, whereby the estimates of RS (Fig. 4.13f) and the simplified MH (Fig. 4.13h) are better defined than the one of CPM (Fig. 4.13c). For I^y , we again obtain estimates close to the prior for both RS (Fig. 4.13e) and CPM (Fig. 4.13b). While CPM also samples σ close to the prior (Fig. 4.13a), the RS realizations show a tendency for higher values (Fig. 4.13d). Finally, the simplified MH samples values of σ (Fig. 4.13g) having a high concentration at incorrect values.

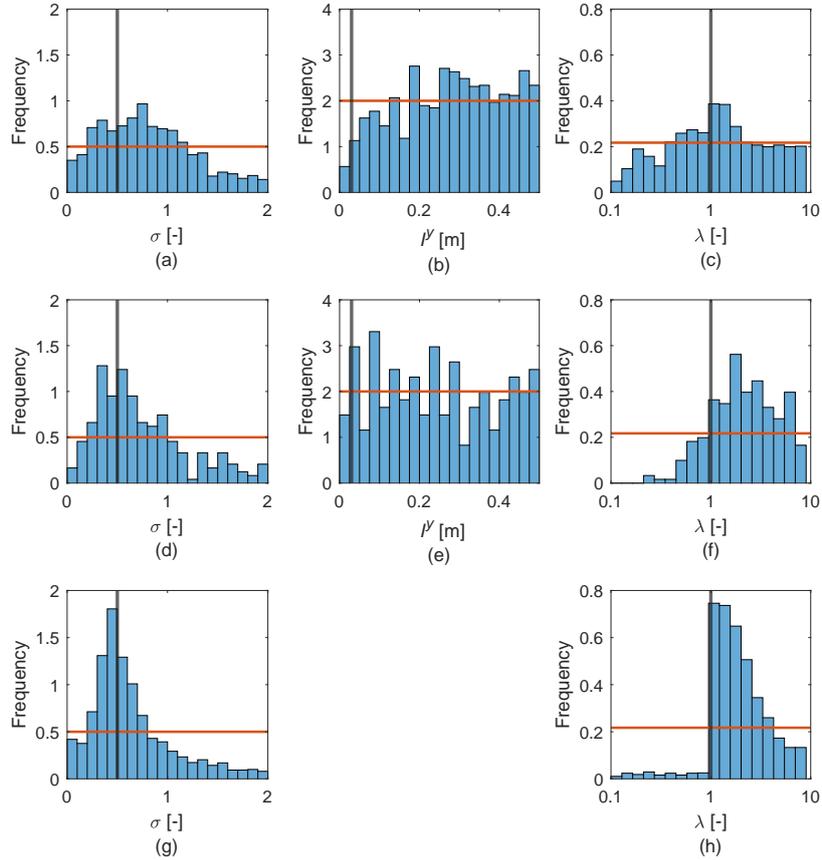


Figure 4.12: Posterior samples for the remaining hyperparameters for the first test case (log-hydraulic conductivity) using the data generated with the ergodic setting; (a) standard deviation σ , (b) integral scale I^y and (c) anisotropy factor λ with CPM. (d) Standard deviation σ , (e) integral scale I^y and (f) anisotropy factor λ with RS. (g) Standard deviation σ and (h) anisotropy factor λ with simplified MH. The solid vertical lines indicate the true hyperparameter values and the red horizontal lines the prior PDFs. Note that for the anisotropy factor λ we employ a logarithmic scale on the x-axis.

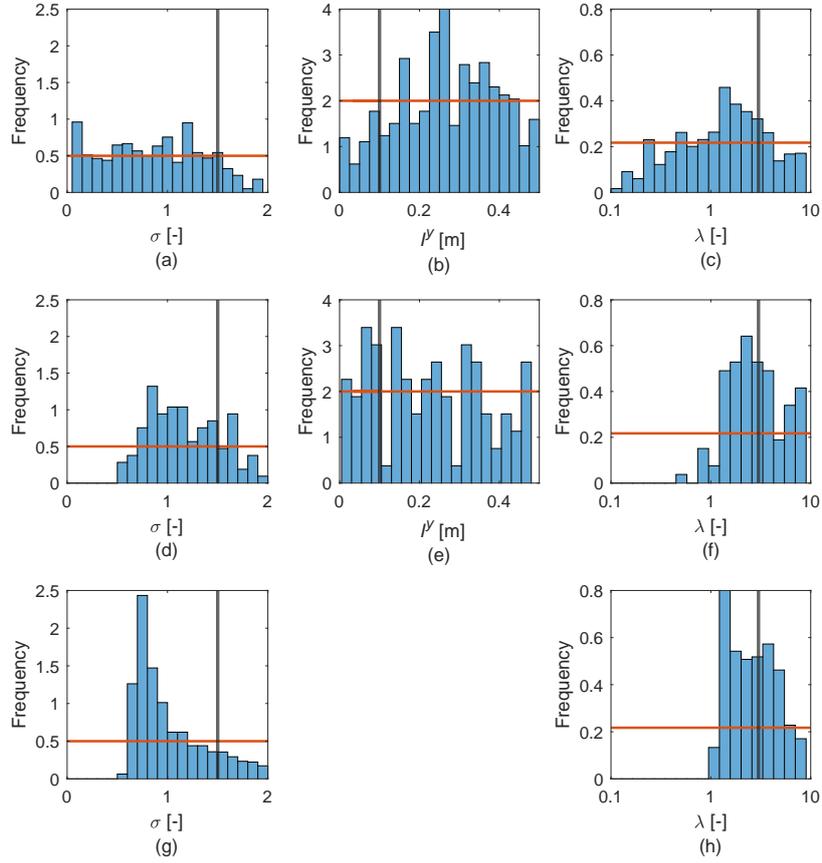


Figure 4.13: Posterior samples for the remaining hyperparameters for the first test case (log-hydraulic conductivity) using the data generated with the non-ergodic setting; (a) standard deviation σ , (b) integral scale I^y and (c) anisotropy factor λ with CPM. (d) Standard deviation σ , (e) integral scale I^y and (f) anisotropy factor λ with RS. (g) Standard deviation σ and (h) anisotropy factor λ with simplified MH. The solid vertical lines indicate the true hyperparameter values and the red horizontal lines the prior PDFs. Note that for the anisotropy factor λ we employ a logarithmic scale on the x-axis.

Chapter 5

Rare event probability estimation for groundwater inverse problems with Sequential Monte Carlo methods

Lea Friedli and Niklas Linde

Submitted to to Water Resources Research

Abstract

Bayesian inversions followed by estimations of rare event probabilities are often needed to analyse groundwater hazards. Instead of focusing on the posterior distribution of model parameters, the main interest lies then in the distribution of a specific quantity of interest contingent upon these parameters. To address the associated methodological challenges, we introduce a two-stage application of Sequential Monte Carlo methods. In the first phase, it generates particles that approximate the posterior distribution; in the second phase, it employs subset sampling techniques to assess the probability of the rare event of interest. By considering two hydrogeological problems of increasing complexity, we showcase the efficiency and accuracy of the resulting PostRisk-SMC method for rare event probability estimation related to groundwater hazards. We compare the performance of the PostRisk-SMC method with a traditional Monte Carlo approach that relies on Markov chain Monte Carlo samples. We showcase that our estimates align with those of the traditional method, provided that the rare event probability can be estimated by Monte Carlo sampling, but the coefficients of variation are notably lower for the same computational budget when targeting more rare events. Furthermore, we highlight that the PostRisk-SMC method allows estimating rare event probabilities approaching one in a billion using less than one hundred thousand forward simulations. Even if the presented examples are related to groundwater hazards, the methodology is well-suited for addressing a wide range of challenges in the geosciences and beyond.

5.1 Introduction

Decision-making processes concerning groundwater and other environmental systems are subject to uncertainty. Consequently, decision-making often involves the identification and avoidance of hazards while assessing associated risks. While a hazard represents a dangerous phenomenon itself, risk considers the resulting potential of harm for human individuals or economic assets (Ward et al. 2020). Risk assessment plays a crucial role in the context of groundwater management, as fresh and uncontaminated groundwater is a prerequisite for global water security (Famiglietti 2014) and as remediation of contaminated aquifers is extremely costly and time-consuming. Groundwater contamination and over-exploitation have not only direct adverse consequences for humans, but also for ecosystems and ecosystem services.

Our focus is a particular aspect of risk assessment, namely, the estimation of the probability of a hazard occurring. This hazard is defined by a quantity of interest that takes on critical values. For a precise analysis of hazard occurrence, it is essential to take into consideration the uncertainty associated with the parameters of a conceptual model. Hence, in the field of hydrogeology, Monte Carlo approaches for sampling uncertain hydrological model parameters have been widely employed (e.g., Lahkim and Garcia 1999, Khadam and Kaluarachchi 2003, Benekos et al. 2007, Siirila et al. 2012, Enzenhoefer et al. 2012). Such approaches can be challenging since hazards often fall under the category of rare events, requiring specialized modeling techniques to accurately represent the tail behavior of the quantity of interest. In this context, classical Monte Carlo estimation is impractical as it requires an excessively

large sample (C erou et al. 2012). One approach to mitigate the computational burden is to combine Monte Carlo methods with surrogate modeling (e.g., Li and Xiu 2010), thereby speeding up the computation time of forward evaluations. Another option is to employ importance sampling in order to focus the sampling on critical regions of the quantity of interest. However, selecting a well-working importance density for high-dimensional problems is often difficult (Au and Beck 2003a). Extreme value theory (e.g., Brodin and Kl uppelberg 2008), relying on fitting an extreme value distribution to represent the distribution of the quantity of interest, offers yet another alternative and is widely used to predict probabilities of environmental hazards such as extreme floods (Morrison and Smith 2002). Extreme value theory necessitates sizable sample sizes for distribution fitting, is contingent on the chosen distribution’s shape, and does not offer simulations of the rare event (e.g., Diebold et al. 2000). An alternative data-intensive method for estimating extremes is based on the ‘Peaks over Threshold’ technique (POT; Leadbetter 1991). In this approach, extreme events are analysed by focusing on values that exceed a certain threshold.

We perform rare event probability estimation for the case when indirect site-specific data \mathbf{y} are available (e.g., from tracer or pumping tests). We employ a Bayesian framework in which the hydrogeological parameters $\boldsymbol{\theta}$ are characterized by a posterior probability density function (PDF) $p(\boldsymbol{\theta}|\mathbf{y})$, given by the distribution of $\boldsymbol{\theta}$ conditioned on measurements \mathbf{y} . Compared to a standard Bayesian inversion problem in which the end-product is an approximation of the posterior PDF, we interrogate the distribution of a quantity of interest depending on the parameters through a non-linear relationship $\boldsymbol{\theta} \mapsto \mathcal{R}(\boldsymbol{\theta})$, for instance, in the probability of this quantity exceeding a critical threshold (considered as the hazard). In practical scenarios, the presence of non-linearity frequently precludes the availability of an analytical formula for the distribution of the quantity of interest when conditioned on the data \mathbf{y} .

In structural engineering, similar problems have been addressed by performing probabilistic updating of system parameters using dynamic data and subsequently updating the estimation of the system’s reliability (e.g., Papadimitriou et al., 2001). In this context, Straub (2011) introduced the so-called Bayesian Updating with Structural reliability methods (BUS; e.g. Straub and Papaioannou 2015). For the Bayesian analysis, BUS can be interpreted as an extension of rejection sampling (Ripley 2009b). To extend BUS for posterior rare event probability estimation, Straub et al. (2016) present an approach targeting both the posterior and the rare event by using reliability methods. A challenge of this method is the selection of the constant employed in the extended rejection sampling, as its choice can impact overall performance. In a similar approach targeting ‘updated robust reliability measures’, Jensen et al. (2013) rely on transitional MCMC (Ching and Chen 2007) to derive a set of posterior samples followed by subset sampling for the reliability analysis. A very different approach enabling the combination of inference and rare event estimation that has been explored in the geosciences is Bayesian Evidential Learning (BEL; Hermans et al. 2016), which aims to learn a direct relationship between measurements and quantity of interest by sampling from the prior distribution (e.g., Thibaut et al. 2021). For higher-dimensional parameter spaces and non-linear relationships, it can be difficult for BEL to capture the full joint distribution with a reasonable number of samples.

We propose a two-stage application of Sequential Monte Carlo (SMC; Doucet et al. 2001), which we refer to as the Posterior Risk Sequential Monte Carlo (PostRisk-SMC) method. Bayesian inversion in hydrogeology and other environmental fields is often addressed using Markov chain Monte Carlo (MCMC) methods. For high-dimensional problems with non-linear forward solvers, standard MCMC methods often have difficulties in approximating the posterior PDF within realistic computational constraints. This happens as the Markov chains may be trapped in local minima for long times or have insignificant probabilities of switching between posterior modes (e.g., Neal 2001, Amaya et al. 2022). As an alternative to MCMC, particle methods use a weighted sample of particles for posterior approximation. One such particle approach is the SMC method (Doucet et al. 2001), which simulates a sequence of power posteriors transferring the prior PDF to the posterior PDF by successively increasing the weight of the likelihood (Del Moral et al. 2007). While the SMC method is extensively used in science and engineering, it has only seen limited use in the geosciences (i.e., Vrugt et al. 2013, Linde et al. 2017). We build our PostRisk-SMC method on an adaptive version of the SMC method by Zhou et al. (2016), which automatically tunes the cooling sequence between power posteriors. Recently, adaptive SMC methods have been employed successfully for geophysical (Amaya et al. 2021 Davies et al. 2023) and hydrogeological (Amaya et al. 2022) inversion problems, demonstrating superior performance compared with state-of-the-art MCMC methods.

Relying only on a particle approximation of the posterior PDF is insufficient when estimating rare event probabilities. As a relatively small number (tens or hundreds, sometimes thousands) of particles is used in practice, this means that no particle is likely to be associated with the rare event that might, for instance, have a probability of one in a million. To address this, a new SMC formulation has emerged that specifically targets rare events by employing a sequence of nested sets pertaining to the hazard scenario. This approach relies on the fact that the small probability of the rare event can be expressed as a product of larger conditional probabilities involving the intermediate sets. Such a splitting technique was first introduced as ‘subset sampling’ by Au and Beck (2001) in the context of reliability analysis and has been applied for instance in the context of radioactive waste management (e.g., Cadini et al. 2012) and earthquake engineering (e.g., Au and Beck 2003b). In the SMC literature, subset sampling is presented by Del Moral et al. (2006) and Johansen et al. (2006). C erou et al. (2012) and Botev and Kroese (2008) extended the existing methods by using an adaptive method that optimally selects the subsets on the fly. Subset sampling has been further leveraged by employing surrogates (Bourinet et al. 2011) or by employing a multilevel approach (Ullmann and Papaioannou 2015). While all of these applications rely on uncertain parameters θ following a ‘prior’ PDF, we here adapt this approach to rare event estimation with respect to a posterior PDF that is first approximated by adaptive SMC. The resulting PostRisk-SMC method relies on the same principles as the approach of Jensen et al. (2013) but within the theoretical formulation of particle methods and SMC. While Jensen et al. (2013) consider engineering applications and dynamic data, we introduce the PostRisk-SMC in the context of hydrogeological rare event probability estimation. Furthermore, we perform resampling of the particles only occasionally (during the posterior phase), while the transitional MCMC approach applied by Jensen et al. (2013) does so in every iteration. Since resampling impacts the variance of estimates (Douc and Capp e 2005), it is usually beneficial to resample only when the variation in the particle weights becomes too high.

For comparison purposes, we consider a conventional Monte Carlo approach for the rare event probability estimation, as applied for instance by Dall’Alba et al. (2023) for risk assessment of groundwater inflow within the setting of tunnel construction. In our inversion setting, we rely on MCMC samples approximating the posterior PDF for the Monte Carlo estimation. Our first example consists of a simplified one-dimensional flow scenario where we utilize pumping tests to estimate the probability of high flow rates. Subsequently, we consider a more realistic two-dimensional flow and transport problem, focusing on assessing the probability of contamination breakthrough. The remainder of the manuscript is organized as follows: Section 5.2 gives a methodological overview of the considered setting and introduces the PostRisk-SMC method; Section 5.3 presents the one-dimensional flow example and Section 5.4 the two-dimensional transport example; finally, the study ends with a discussion and conclusions in Sections 5.5 and 5.6, respectively.

5.2 Methodology

5.2.1 Notation

We target an unknown property vector $\boldsymbol{\theta} \in \mathbb{R}^P$ representing a model domain from which we obtain measurements $\mathbf{y} \in \mathbb{R}^M$. We consider a setting where measurements are realizations of the random variable $\mathbf{Y} = \mathcal{G}(\boldsymbol{\theta}) + \boldsymbol{\varepsilon}_\theta$, with $\mathcal{G} : \mathbb{R}^P \rightarrow \mathbb{R}^M$ referring to the forward solver and $\boldsymbol{\varepsilon}_\theta$ to the observational noise. Assuming independent Gaussian observational errors, we express the likelihood as $p(\mathbf{y}|\boldsymbol{\theta}) = \varphi_M(\mathbf{y}; \mathcal{G}(\boldsymbol{\theta}), \boldsymbol{\Sigma}_Y)$, with $\varphi_M(\cdot; \mathcal{G}(\boldsymbol{\theta}), \boldsymbol{\Sigma}_Y)$ denoting the PDF of a M -variate normal distribution with the mean $\mathcal{G}(\boldsymbol{\theta})$ and the diagonal covariance matrix $\boldsymbol{\Sigma}_Y$ of the observational errors.

We consider a quantity of interest $R = \mathcal{R}(\boldsymbol{\theta})$ derived from $\boldsymbol{\theta}$ via some function $\mathcal{R} : \mathbb{R}^P \rightarrow \mathbb{R}$. More specifically, we target a rare set $A = \{\boldsymbol{\theta} \in \mathbb{R}^P : \mathcal{R}(\boldsymbol{\theta}) \in \mathcal{T}\}$ for some interval $\mathcal{T} \subseteq \mathbb{R} \cup \{\infty, -\infty\}$. If we target the exceedance probability $\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq T)$ for some real number T , we assign $\mathcal{T} = [T, \infty)$. We are interested in $\mathbb{P}(\boldsymbol{\theta} \in A|\mathbf{y})$ for $\boldsymbol{\theta}$ distributed according to the posterior PDF $p(\boldsymbol{\theta}|\mathbf{y})$ and write,

$$\mathbb{P}(\boldsymbol{\theta} \in A|\mathbf{y}) = \int_A p(\boldsymbol{\theta}|\mathbf{y}) d\boldsymbol{\theta}. \quad (5.1)$$

5.2.2 Bayesian inversion

In Bayes’ theorem, the posterior PDF is given by,

$$p(\boldsymbol{\theta}|\mathbf{y}) = \frac{p(\boldsymbol{\theta})p(\mathbf{y}|\boldsymbol{\theta})}{p(\mathbf{y})}, \quad (5.2)$$

with the prior PDF $p(\boldsymbol{\theta})$ of the model parameters, the likelihood function $p(\mathbf{y}|\boldsymbol{\theta})$ and the evidence $p(\mathbf{y})$. As in practice, it is often not possible to sample directly from the posterior when the forward solver $\boldsymbol{\theta} \mapsto \mathcal{G}(\boldsymbol{\theta})$ is non-linear, MCMC and SMC methods are applied.

Metropolis–Hastings algorithm

The most used MCMC method is the Metropolis–Hastings algorithm (MH algorithm; Metropolis et al. 1953; Hastings 1970). The MH algorithm is an iterative algorithm that, in each iteration, proposes a new set of model parameter values $\boldsymbol{\theta}_{prop}$ using the proposal density $q(\cdot|\boldsymbol{\theta}_{curr})$, which is then accepted or rejected based on the acceptance probability,

$$\alpha_{MH}(\boldsymbol{\theta}_{curr}, \boldsymbol{\theta}_{prop}) = \min \left\{ 1, \frac{q(\boldsymbol{\theta}_{curr}|\boldsymbol{\theta}_{prop}) p(\boldsymbol{\theta}_{prop}) p(\mathbf{y}|\boldsymbol{\theta}_{prop})}{q(\boldsymbol{\theta}_{prop}|\boldsymbol{\theta}_{curr}) p(\boldsymbol{\theta}_{curr}) p(\mathbf{y}|\boldsymbol{\theta}_{curr})} \right\}. \quad (5.3)$$

The choice of the proposal density $q(\cdot|\boldsymbol{\theta}_{curr})$ is crucial, as it has to balance the trade-off between exploration and exploitation. While standard Gaussian model proposals can be applied for a model space with reduced dimension, more high-dimensional parameter spaces present considerable challenges (e.g., Robert et al. 2018). To ensure robustness against different discretization choices and to maintain a reasonable step size while inferring thousands of unknowns, we rely on preconditioned Crank-Nicolson proposals that preserve the prior PDF (pCN; e.g. Cotter et al. 2013). For a target variable with a Standard-Normal prior, the pCN proposal is given by,

$$q(\cdot|\boldsymbol{\theta}_{curr}) = \sqrt{1 - \rho^2} \boldsymbol{\theta}_{curr} + \rho \boldsymbol{\zeta}, \quad (5.4)$$

with $\boldsymbol{\zeta} \sim \mathcal{N}(0, 1)$ and $\rho \in (0, 1)$, ensuring that the variable remains standard-normally distributed. The utilization of such prior-preserving proposals results in the acceptance probability being solely dependent on the likelihood values. In the field of geophysics, MCMC algorithms with model proposals that preserve the prior are known as extended Metropolis (Mosegaard and Tarantola 1995).

5.2.3 From Sequential Monte Carlo to PostRisk-SMC

In this Section, we first introduce Sequential Monte Carlo for posterior inference and Sequential Monte Carlo for rare event estimation. Subsequently, we introduce PostRisk-SMC, a novel sequential combination of both methods, designed to tackle the challenge of estimating rare event probabilities while accounting for posterior uncertainty. For the methodology of the first phase, we rely on the framework of Del Moral et al. (2007) and Zhou et al. (2016) and refer to their works for further details. Likewise, for the second phase, we follow the framework presented by C erou et al. (2012) and suggest consulting their paper for additional information.

Sequential Monte Carlo for posterior inference

Posterior estimation with the SMC method is based on a particle approximation using N particles $\{\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \dots, \boldsymbol{\theta}^{(N)}\}$ with weights $\{W^{(1)}, W^{(2)}, \dots, W^{(N)}\}$. If the particles are sampled according to the posterior, the weights are redundant and reduce to $1/N$. In practice, it is generally not possible to sample from the posterior and importance sampling using a density $\eta(\boldsymbol{\theta}|\mathbf{y})$ is applied. Importance sampling generates samples from an importance distribution that assigns higher probabilities to regions where the target distribution is expected to have most of its mass, thereby reducing the variance of estimators (e.g. Owen and Zhou 2000). To achieve a well-working importance sampling approach for the posterior PDF, one should strive for a $\eta(\boldsymbol{\theta}|\mathbf{y})$ as close as possible to $p(\boldsymbol{\theta}|\mathbf{y})$. This can be achieved by building a sequence of K PDFs $\{p_0(\boldsymbol{\theta}|\mathbf{y}), p_1(\boldsymbol{\theta}|\mathbf{y}), \dots, p_K(\boldsymbol{\theta}|\mathbf{y})\}$ with $p_0(\boldsymbol{\theta}|\mathbf{y}) = p(\boldsymbol{\theta})$ and $p_K(\boldsymbol{\theta}|\mathbf{y}) = p(\boldsymbol{\theta}|\mathbf{y})$, thus moving gradually from the prior PDF to the posterior PDF (Del Moral et al. 2007). The sequence is built on unnormalized power posteriors (Neal 2001),

$$p_k(\boldsymbol{\theta}|\mathbf{y}) = p(\mathbf{y}|\boldsymbol{\theta})^{\alpha_k} p(\boldsymbol{\theta}), \quad (5.5)$$

with $0 = \alpha_0 < \alpha_1 < \dots < \alpha_K = 1$. With increasing exponent α_k , the relative influence of the likelihood on the power posterior grows. For a smaller exponent, the exponentiated term is ‘flatter’ such that the power posterior is closer to the prior PDF. When using the importance density $\eta(\boldsymbol{\theta}|\mathbf{y})$ to sample the particles $\boldsymbol{\theta}^{(p)}$, the weights $W^{(p)}$ correspond to the normalized version of the importance weights $w^{(p)} = p(\boldsymbol{\theta}^{(p)}|\mathbf{y})/\eta(\boldsymbol{\theta}^{(p)}|\mathbf{y})$.

We start at iteration $k = 0$ with particles $\boldsymbol{\theta}_0^{(p)}$ ($p = 1, 2, \dots, N$) sampled from the prior PDF $p_0(\boldsymbol{\theta}|\mathbf{y}) = p(\boldsymbol{\theta})$ and initial weights $W_0^{(p)}$ being all equal to $1/N$. At iteration k of the SMC method, $p_k(\boldsymbol{\theta}|\mathbf{y})$ is approximated by importance sampling based on the previously estimated power posterior $p_{k-1}(\boldsymbol{\theta}|\mathbf{y})$. Therefore, the particles $\boldsymbol{\theta}_{k-1}^{(p)}$ are assigned with incremental weights,

$$w_k^{(p)} = \frac{p_k(\boldsymbol{\theta}_{k-1}^{(p)}|\mathbf{y})}{p_{k-1}(\boldsymbol{\theta}_{k-1}^{(p)}|\mathbf{y})} = \frac{p(\mathbf{y}|\boldsymbol{\theta}_{k-1}^{(p)})^{\alpha_k}}{p(\mathbf{y}|\boldsymbol{\theta}_{k-1}^{(p)})^{\alpha_{k-1}}} = p(\mathbf{y}|\boldsymbol{\theta}_{k-1}^{(p)})^{\alpha_k - \alpha_{k-1}}. \quad (5.6)$$

To account for the previous importance sampling steps, the cumulative normalized weights $W_k^{(p)}$ of the particles $\boldsymbol{\theta}_{k-1}^{(p)}$ are defined as,

$$W_k^{(p)} = \frac{W_{k-1}^{(p)} w_k^{(p)}}{\sum_{j=1}^N W_{k-1}^{(j)} w_k^{(j)}}, \quad (5.7)$$

taking into account the history of weights and normalizing them to ensure their sum equals one. The particles $\boldsymbol{\theta}_{k-1}^{(p)}$ approximating $p_{k-1}(\boldsymbol{\theta}|\mathbf{y})$ are generated by propagating each particle $\boldsymbol{\theta}_{k-2}^{(p)}$ according to a Markov kernel leaving $p_{k-1}(\boldsymbol{\theta}|\mathbf{y})$ invariant (Neal 2001). This can be achieved by employing a finite number s of MH steps (Del Moral et al. 2007). In contrast to general MCMC methods, the MH steps do not have to converge as the importance sampling weights account for any possible sampling from the wrong distribution (Del Moral et al. 2007). This SMC procedure for posterior inference is illustrated in Figure 5.1.

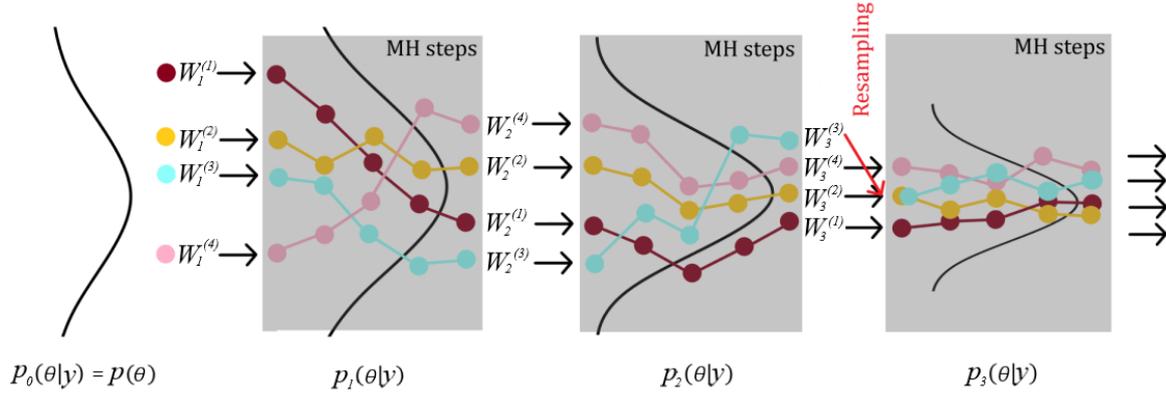


Figure 5.1: Illustration of the SMC method for posterior inference. We depict the first four power posteriors for an example with $N = 4$ particles and $s = 4$ MH steps.

When the (empirical) variance of the weights $W_k^{(p)}$ at iteration k becomes large, it is beneficial to resample the particles before propagation (Del Moral et al. 2007, Doucet and Johansen 2009). Resampling decreases the variance of the weights by discarding most particles with low weights and preferably reproducing those with high weights. Here, we use systematic resampling (Doucet and Johansen 2009). Subsequently, the weights $W_k^{(p)}$ are set to $1/N$, as the resampled particles are approximately distributed according to $p_k(\boldsymbol{\theta}|\mathbf{y})$. Resampling increases the variance of the estimator, making it wasteful if the importance weights do not exhibit significant variability (Del Moral et al. 2006). To decide when resampling is to be performed, the effective sample size (ESS; Kong et al. 1994),

$$ESS_k = \frac{\left(\sum_{p=1}^N W_{k-1}^{(p)} w_k^{(p)}\right)^2}{\sum_{p=1}^N \left(W_{k-1}^{(p)}\right)^2 \left(w_k^{(p)}\right)^2}, \quad (5.8)$$

is used. For instance, Del Moral et al. (2006) apply the decision rule of resampling if the ESS_k falls below 30 % of the number of particles N . To ensure that the final particles are a (unweighted) approximation of the posterior, we enforce a resampling step in the last iteration.

When defining the sequence of exponents α , one has to consider that too large differences between α_{k-1} and α_k lead to a large discrepancy between the power posteriors $p_{k-1}(\boldsymbol{\theta}|\mathbf{y})$ and $p_k(\boldsymbol{\theta}|\mathbf{y})$ and a subsequent poor performance of the importance sampling. However, if the difference is very small, an excessive number of steps are needed until $\alpha_k = 1$ is reached. It is natural to aim for a similar discrepancy between successive power posteriors (Zhou et al. 2016). To select the sequence of exponents α , we use the adaptive method of Zhou et al. (2016), based on the conditional effective sample size (CESS),

$$CESS_k = N \frac{\left(\sum_{p=1}^N W_{k-1}^{(p)} w_k^{(p)}\right)^2}{\sum_{p=1}^N W_{k-1}^{(p)} \left(w_k^{(p)}\right)^2}. \quad (5.9)$$

The $CESS_k$ quantifies the quality of $p_{k-1}(\boldsymbol{\theta}|\mathbf{y})$ as an importance density to estimate expectations under $p_k(\boldsymbol{\theta}|\mathbf{y})$ (Zhou et al. 2016). The $CESS$ is equal to the ESS when resampling is conducted at each iteration. Zhou et al. (2016) show that using the $CESS$ for the adaptive sequence leads to a reduction in estimator variance compared to an approach using the ESS . To define the next α_k , a binary search for the value for which the $CESS$ is the closest to a pre-defined target value $CESS^*$ is performed. The closer this target value $CESS^*$ is to N , the better the approximation, but the slower the algorithm becomes as the number of power posteriors grows. The SMC algorithm stops when α_k reaches one. Such an adaptive approach is expected to result in a more efficient algorithm compared to its non-adaptive counterpart. Importantly, it also leads to a more automated algorithm by minimizing the number of user-defined tuning parameters (Beskos et al. 2016). However, using an adaptive method for the selection of the exponents introduces a slight bias into the results. Beskos et al. (2016) explore the convergence behaviour for such adaptive approaches and establish that the output satisfies a weak law of large numbers and a central limit theorem. To indicate if we use an adaptive or fixed sequence of exponents, we specify the binary variable $ADAP$ as 1 for an adaptive and 0 for a predetermined selection. The full workflow of the SMC method for posterior inference is summarized in Figure 5.2.

Sequential Monte Carlo for rare event estimation

The SMC method can be modified to enable simulation of rare events and estimation of their probabilities by using a sequence of not-so-rare nested events (Del Moral et al. 2006 Johansen et al. 2006 Cérou et al. 2012). It is assumed that $\boldsymbol{\theta}$ is a random element on \mathbb{R}^P with probability distribution $p(\boldsymbol{\theta})$ that can be sampled from. To estimate $\mathbb{P}(\boldsymbol{\theta} \in A)$, the SMC method for rare event estimation employs a sequence of nested sets $A_k = \{\boldsymbol{\theta} \in \mathbb{R}^P : \mathcal{R}(\boldsymbol{\theta}) \in \mathcal{T}_k\}$, with $\mathbb{R}^P = A_0 \supset A_1 \supset \dots \supset A_K = A$. It holds that,

$$\mathbb{P}(\boldsymbol{\theta} \in A) = \prod_{k=1}^K \mathbb{P}(\boldsymbol{\theta} \in A_k | \boldsymbol{\theta} \in A_{k-1}). \quad (5.10)$$

If we are interested in $\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq T)$, the sequence of nested sets $A_k = \{\boldsymbol{\theta} \in \mathbb{R}^P : \mathcal{R}(\boldsymbol{\theta}) \in [T_k, \infty)\}$ corresponds to a sequence of increasing thresholds $\{T_0, \dots, T_K\}$ with $T_0 = -\infty$ and $T_K = T$. For $\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \leq T)$, we employ $A_k = \{\boldsymbol{\theta} \in \mathbb{R}^P : \mathcal{R}(\boldsymbol{\theta}) \in (-\infty, T_k]\}$ using a sequence of decreasing thresholds with $T_0 = \infty$ and $T_K = T$.

The SMC method for rare event estimation starts by initializing N particles $\boldsymbol{\theta}_0 = (\boldsymbol{\theta}_0^{(1)}, \dots, \boldsymbol{\theta}_0^{(N)})$ sampled from $p(\boldsymbol{\theta})$. The first intermediate distribution $p_{A_0}(\boldsymbol{\theta}) = p(\boldsymbol{\theta} | \boldsymbol{\theta} \in A_0)$ is equal to $p(\boldsymbol{\theta})$. To approximate the intermediate distribution $p_{A_k}(\boldsymbol{\theta}) = p(\boldsymbol{\theta} | \boldsymbol{\theta} \in A_k)$ for $k \geq 1$, each particle $\boldsymbol{\theta}_{k-1}^{(p)}$ is assigned a weight,

$$W_k^{(p)} = \begin{cases} 1/|I_k|, & \text{if } \boldsymbol{\theta}_{k-1}^{(p)} \in A_k \\ 0, & \text{otherwise,} \end{cases} \quad (5.11)$$

with $I_k = \{p : \boldsymbol{\theta}_{k-1}^{(p)} \in A_k\}$ and $|I_k|$ denoting its cardinality. Thereby, we are assuming that I_k is non-empty, otherwise the particle system dies. Subsequently, systematic resampling (Doucet

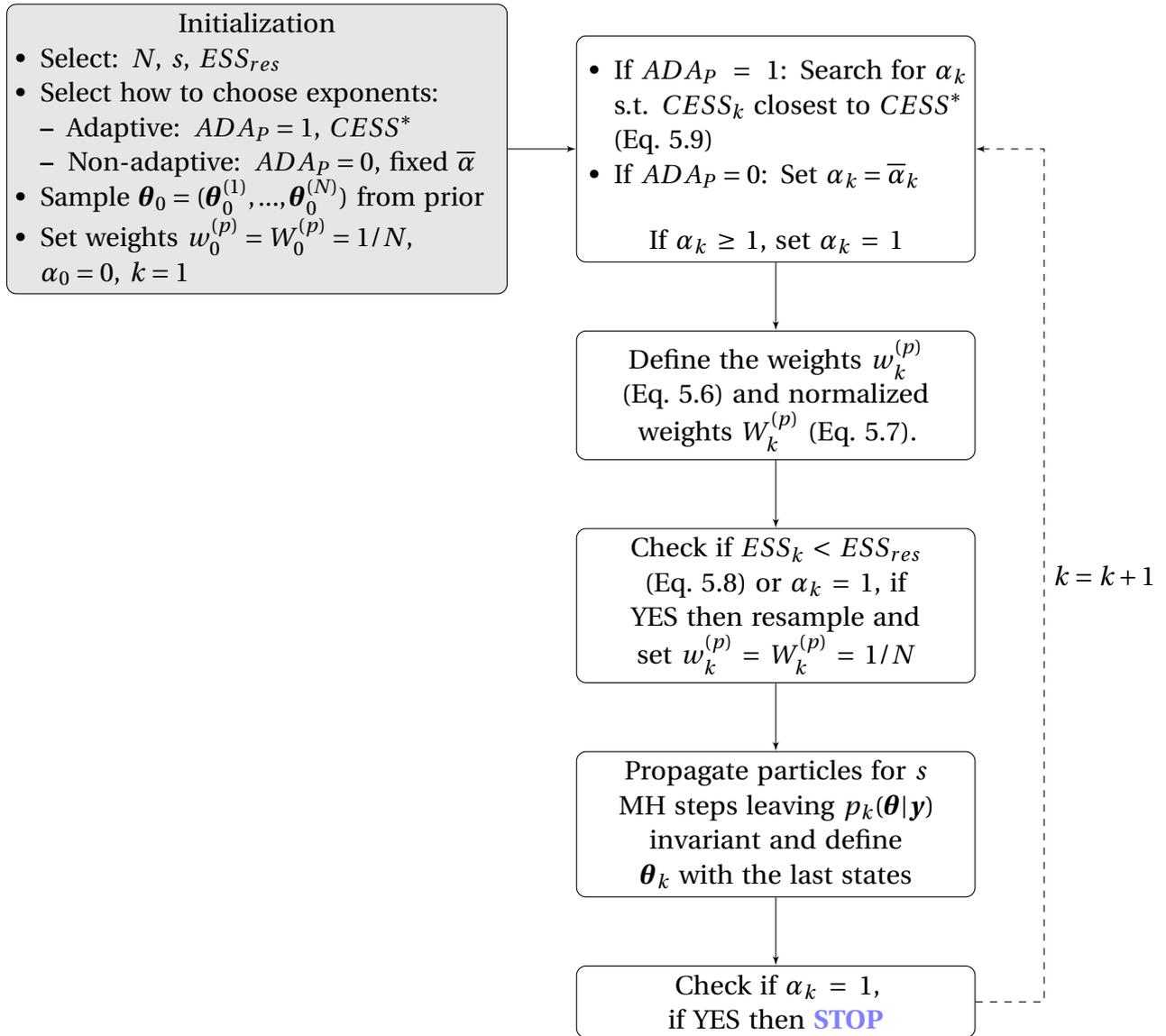


Figure 5.2: Flow chart illustrating the SMC method for posterior inference.

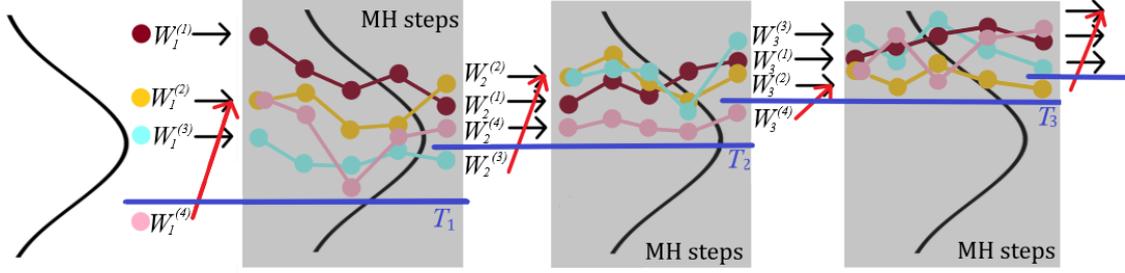


Figure 5.3: Illustration of the SMC method for rare events targeting $\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq T)$. We depict the first three thresholds for an example with $N = 4$ particles, $s = 4$ MH steps and a quantile of $\gamma = 0.25$.

and Johansen 2009) is employed such that particles which do not lie in A_k are replaced by particles that do. The resampled particles are propagated using a Markov kernel, leaving $p_{A_k}(\boldsymbol{\theta})$ invariant (C erou et al. 2012). We are considering s steps with a MH algorithm whereby a transition is only accepted if $\boldsymbol{\theta}$ stays in A_k . The procedure of SMC for rare event estimation targeting $\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq T)$ is illustrated in Figure 5.3.

We need to choose a sequence of nested sets such that $\mathbb{P}(\boldsymbol{\theta} \in A_k | \boldsymbol{\theta} \in A_{k-1})$ is reasonably high. C erou et al. (2012) detail both a fixed and an adaptive algorithm. For $A_k = \{\boldsymbol{\theta} \in \mathbb{R}^P : \mathcal{R}(\boldsymbol{\theta}) \in [T_k, \infty)\}$, an adaptive method based on quantiles of $\mathcal{R}(\cdot)$ of the particles ensures that the asymptotic variance of the estimator is minimal (see C erou et al. 2012). Utilizing the γ -quantile,

$$T_k = q_\gamma(\mathcal{R}(\boldsymbol{\theta}_{k-1})), \quad (5.12)$$

guarantees that a ratio of $(1 - \gamma)$ of the particles survive. The adaptive algorithm's stopping criterion is met when the quantile surpasses the targeted threshold, at which point the last T_K is set equal to T . Then, $\mathbb{P}(\boldsymbol{\theta} \in A)$ is estimated by multiplication of all $P_k = |I_k|/N$ for $k = 1, \dots, K$. Due to the adaptiveness of the thresholds, the resulting estimator is biased given the finite number of particles N (Au and Beck 2001). This bias is positive and becomes negligible compared to the variance of the estimator as the number of particles increases (C erou et al. 2012). To circumvent this bias, one can either re-run the algorithm with the previously optimized sequence or use a predetermined fixed sequence of thresholds. With the binary variable ADA_R we indicate if we use fixed ($ADA_R = 0$) or adaptive ($ADA_R = 1$) sequences of thresholds. The work flow of the SMC method for rare event estimation is summarized in the flow chart in Figure 5.4.

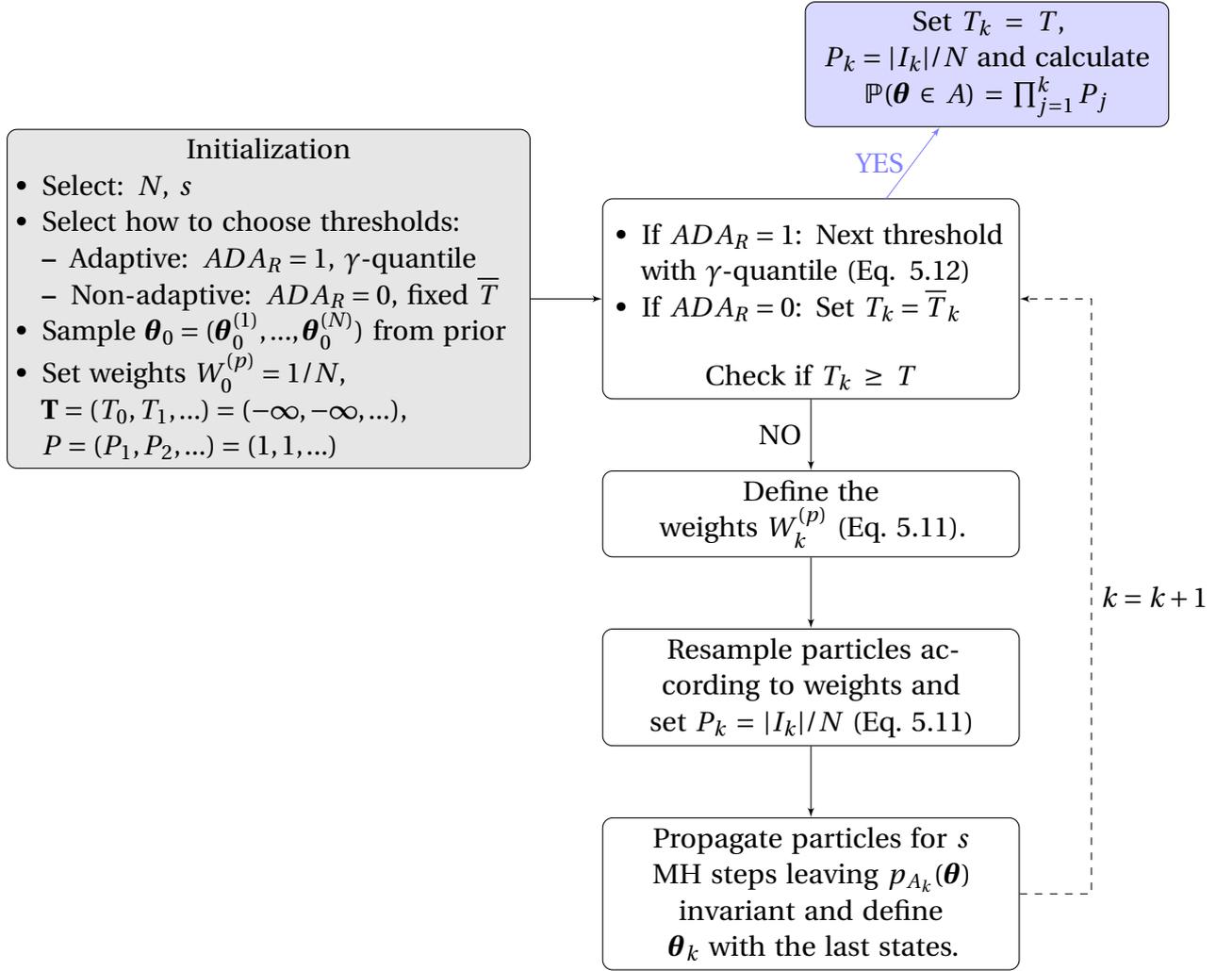


Figure 5.4: Flow chart illustrating the SMC method for rare event estimation of $\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq T)$.

Posterior Risk Sequential Monte Carlo method

To estimate $\mathbb{P}(\boldsymbol{\theta} \in A | \mathbf{y})$, we introduce a sequential combination of the two SMC methods described in the previous paragraphs (PostRisk-SMC). Let us write the k -th power posterior with respect to the subset A_k as,

$$p_k^A(\boldsymbol{\theta} | \mathbf{y}) = p(\mathbf{y} | \boldsymbol{\theta})^{\alpha_k} p(\boldsymbol{\theta}) \mathbb{1}\{\boldsymbol{\theta} \in A_k\}. \quad (5.13)$$

While the first stage of the PostRisk-SMC algorithm generates particles distributed according to the posterior by increasing the exponent of the likelihood α_k with the subset A_k being held constant as \mathbb{R}^p , the second stage shrinks the subset while leaving the exponent of the power posterior at 1. For the rare event analysis, it is crucial that we start the second phase with a unweighted particle approximation of the posterior, ensured by the resampling step in the last step of the posterior inference stage. We denote as K_P the number of intermediate power posteriors, as K_R the number of thresholds and as $K = K_P + K_R$ their sum. Additionally, we define s_P as the number of MH steps employed between each importance sampling step in

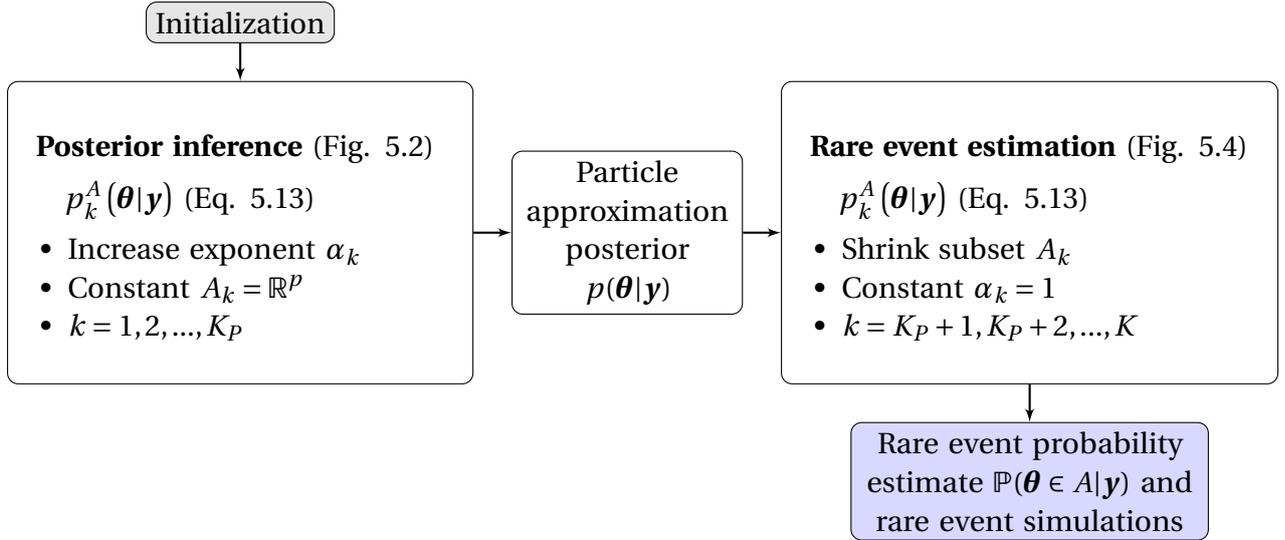


Figure 5.5: Work flow of the PostRisk-SMC method.

the posterior phase and s_R as the number between the subset sampling steps during the rare event phase. When the same number of steps is used for both, we denote it as $s = s_P = s_R$. The PostRisk-SMC method inherits the theoretical properties of the SMC methods utilized in the two stages, including any biases present in the estimators resulting from adaptive sequences of exponents and thresholds. The complete work flow of the PostRisk-SMC method is summarized in Figure 5.5.

In high-dimensional scenarios characterized by complex posterior distributions, the process of particle propagation using a limited number of MH steps can become limiting. In such contexts, the frequency of particle resampling becomes important to monitor. In the rare event probability estimation phase, this aspect becomes even more critical as frequent resampling is unavoidable. This implies the need to ensure that a sufficient number of MH steps are used to prevent particle collapse following the resampling steps. In groundwater settings where the rare event revolves around contamination hazards, the simulation of the quantity of interest often demands more computational resources than the forward model used to estimate the posterior PDF. To achieve computational speed-up under such situations (as exemplified in Section 5.4), we introduce a minor modification to the propagation step during the rare event phase of PostRisk-SMC. Instead of simulating both the forward response and quantity of interest in every step, we conduct first a series of $s s_R$ posterior steps within each of the s_R steps. Subsequently, the last state is treated as a proposal from the posterior which is accepted or rejected based on whether it falls within the current subset.

5.3 1D flow example

As a first example, we study a steady-state 1-D groundwater flow problem (diffusion equation). The chosen problem setting is influenced by a test case by Straub et al. (2016), which corresponds to the steady-state version of a test case introduced by Marzouk and Najm (2009). The fast run time of this simple toy example allows for a sensitivity analysis of the algorithmic parameters of our (PostRisk-SMC) method.

5.3.1 Synthetic setting

The model domain is the unit interval $D = [0, 1]$ m and we consider the following steady-state equation,

$$\frac{d}{dx} \left(\boldsymbol{\theta}(x) \frac{dh}{dx} \right) + b(x) = 0, \quad (5.14)$$

with hydraulic conductivity $\boldsymbol{\theta}(x)$ [m/s], source $b(x)$ [1/s] and hydraulic head $h(x)$ [m].

The log-conductivity $\log \boldsymbol{\theta}(x)$ is parameterized as a finite rank Gaussian random process expressed by,

$$\log \boldsymbol{\theta}(x) = \mu_{\log \boldsymbol{\theta}} + \sum_{i=1}^n \sqrt{w_i} v_i(x) Z_i, \quad (5.15)$$

with $\{w_i, v_i\}$ representing the first n eigenvalues and eigenfunctions from the Karhunen-Loève expansion of a Gaussian process with mean $\mu_{\log \boldsymbol{\theta}} = \log(10^{-5})$ and exponential covariance function $\kappa_{\log \boldsymbol{\theta}}(\Delta x) = \sigma^2 \exp(-\Delta x/l)$ with standard deviation $\sigma = 3$ and integral scale $l = 0.3$ m. Z_i denote independent standard normally-distributed variables. Following Straub et al. (2016), we employ a truncation after $n = 10$ terms. For the representation, we use a uniform grid with 40 intervals and under the assumption of the mean and covariance structure being known, we infer the ten first Z_i . The 'true' log-hydraulic conductivity values $\log \boldsymbol{\theta}(x)$ are depicted in Figure 5.6a.

For the measurements \mathbf{y} , the source term $b(x)$ in Equation (5.14) is modelled using sources in the cells at 0.26, 0.51 and 0.76 m with identical strengths of 0.001 1/s. The measurements are performed on the steady-state solution of $h(x)$ employing 7 sensors spaced uniformly on D excluding the endpoints. To achieve this, Equation (5.14) is solved with linear finite differences on a uniform grid employing 40 cells and boundary conditions $h(0) = h(1) = 0$ m (Langtangen and Linge 2017). Finally, the synthetically-generated measurement values are contaminated with independent Gaussian errors having a standard deviation of 0.01 m (Fig. 5.6b).

For the rare event, we consider flow from the left to the right of the model domain and define the 'hazard' as the flow rate on the right boundary exceeding a critical value of T . To calculate the flow rate, we assume a hydraulic head difference of 1 m and take the harmonic mean

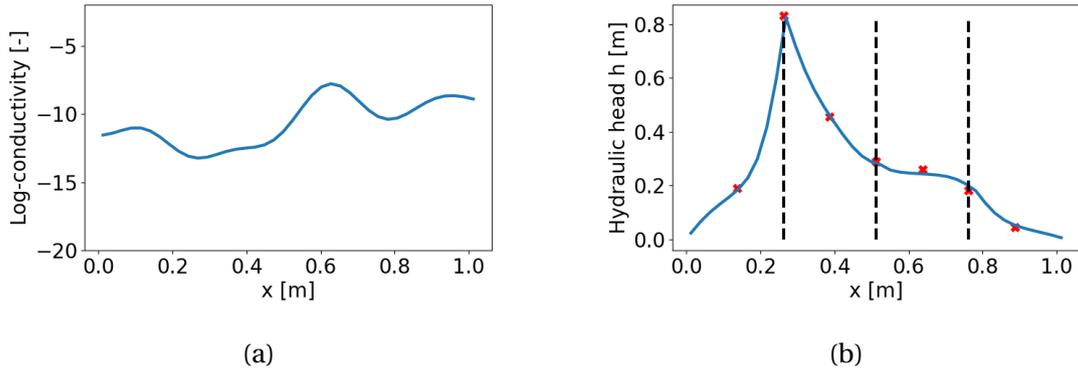


Figure 5.6: (a) 'True' log-hydraulic conductivity $\log\theta(x)$ on $D = [0, 1]$ m and corresponding (b) steady-state solution $h(x)$ (solid line) for the diffusion equation including the pumping sources (source locations dashed) and the resulting noisy measurements \mathbf{y} (crosses).

of the conductivity values. To enable a comparison with MC estimation, we consider a first value of $T^* = 9 \times 10^{-6}$ m/s; the second value of $T^{**} = 9.5 \times 10^{-6}$ m/s is selected such that it targets a rare event with probability of one in a billion.

5.3.2 Results

We employ independent normal prior PDFs for the unknown Z_i of the KL-expansion representing the log-conductivity (Eq. 5.15). For the likelihood, we assume independent Gaussian measurement errors with the same standard deviation as used in the data generation process. We compare the results of the PostRisk-SMC method with those of a standard MH algorithm employing Gaussian proposals. To ensure an acceptance rate of approximately 30 %, the step width of the proposals is adjusted accordingly, taking into account the different scales of variation in the KL components (based on initial MH runs). The same configuration of the MH algorithm is used in the MH steps employed in each iteration of the PostRisk-SMC method.

For the PostRisk-SMC method, the following parameter choices have to be made: the number of particles N , the number of MH steps s in each iteration (here $s = s_p = s_R$), the selection of the exponents α_k (Eq. 5.9), the threshold ESS_{res} below which resampling is employed (Eq. 5.8) and the selection of the thresholds T_k (Eq. 5.12). Following Del Moral et al. (2006), we fix $ESS_{res} = 0.3 \times N$ for the resampling in the initial stage of posterior inference. We start by testing a configuration of PostRisk-SMC with $N = 40$, $CESS^* = 0.99 \times N$, $\gamma = 0.05$ and $s = 40$, employing adaptive schedules for the likelihood's exponents and the thresholds. Figure 5.7 depicts resulting particle approximations of the following distributions of the log-diffusivity profile: (a) prior $p_0^A(\boldsymbol{\theta}|\mathbf{y}) = p(\boldsymbol{\theta})$, (b) posterior $p_{Kp}^A(\boldsymbol{\theta}|\mathbf{y}) = p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})$ and (c) posterior rare event $p_K^A(\boldsymbol{\theta}|\mathbf{y}) = p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})\mathbb{1}\{\mathcal{R}(\boldsymbol{\theta}) \geq T^*\}$.

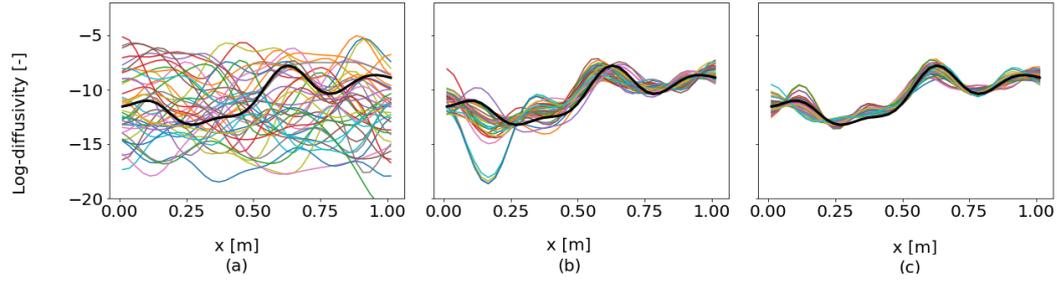


Figure 5.7: Results for the 1D flow example with the PostRisk-SMC method: Particle representation ($N = 40$) of the log-conductivity's (a) prior, (b) posterior and (c) posterior rare event (for T^*) distribution; the black lines depict the true profile and the coloured lines the particles.

To explore the level of bias introduced by the adaptive schemes for our choice of $N = 40$ particles, we re-run the algorithm using the previously defined sequences as fixed values. The range of ten resulting estimates for T^{**} are depicted in Figure 5.8a (adaptive and re-run). The adaptive runs yield a mean estimate that is approximately 200 times greater than that of the re-runs. To circumvent bias while avoiding the computational burdens associated with increasing the number of particles or performing re-runs, we adopt in what follows a fixed sequence of thresholds for the rare event estimation part ($ADA_R = 0$ in Fig. 5.4). With K_P denoting the number of intermediate power posteriors and following the flow chart in Figure 5.5, the first threshold different from minus infinity is T_{K_P+1} . For the shape of the sequence, a suitable form can be determined, for example, by conducting an initial adaptive run (Fig. 5.8b). We use a logarithmic function,

$$f_T(k) = a \log(k) + T_{K_P+1}, \quad (5.16)$$

increasing from T_{K_P+1} to T^{**} . Therefore, we set the thresholds to $T_k = f_T(k - K_P)$ for $k = K_P + 1, \dots, K$ and ensure that $T_K = f_T(K - K_P) = T^{**}$ by expressing $a = (T^{**} - T_{K_P+1}) / \log(K - K_P)$. Finally, we change the closest value of T^* to this very value. For the first threshold, we test the choices of $T_{K_P+1} = 3, 5, 7 \times 10^{-6}$. The resulting threshold sequences are depicted in Figure 5.8b, together with the adaptive sequence utilizing $\gamma = 0.05$. The range and mean of ten estimates for T^{**} obtained with the different sequences are depicted in Figure 5.8a. We note that while the adaptive sequence leads to much higher values, the ones of the re-runs and the fixed sequences with the different T_{K_P+1} are comparable.

In our specific context, where the focus is on estimating the probability of rare events and the posterior of θ is rather smooth, the bias caused by the adaptive schedule in the first stage of posterior estimation is minimal. Tests (not shown) demonstrated that even when considering T^{**} and $N = 40$, the adaptive sequence for the posterior estimation resulted in an almost identical mean estimate compared to the re-runs (less than 0.02 % difference). As a result, we continue to use an adaptive sequence of exponents for the first stage of the algorithm ($ADA_P = 1$ in Fig. 5.2).

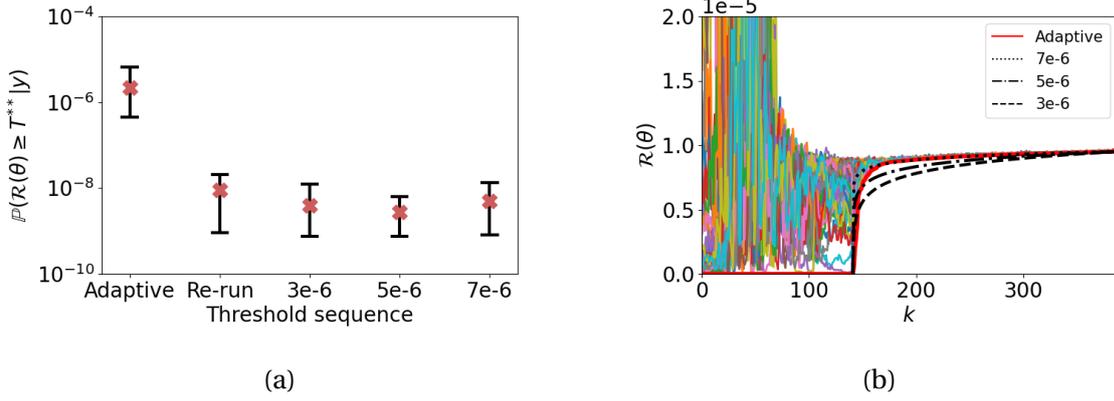


Figure 5.8: Illustration of the bias resulting from the adaptively determined threshold sequence within the PostRisk-SMC method for the 1D flow example: (a) Range of estimates $\mathbb{P}(\mathcal{R}(\theta) \geq T^{**} | \mathbf{y})$ using the different threshold sequences (ten runs each); the red crosses indicate the mean of the values and (b) evolving particle estimation of $\mathcal{R}(\theta)$ with the adaptive T_k -sequence (red) and the different fixed logarithmic sequences (black, Eq. 5.16 with different $T_{K_{P+1}}$).

We now keep $T_{K_{P+1}} = 5 \times 10^{-6}$, and explore the influence of the remaining parameters on the rare event estimation. As a baseline configuration, we use $N = 20$, $CESS^* = 0.9 \times N$ (resulting in $K_P = 40$), $K_R = 100$ and $s = 20$, requiring 55,000 forward simulations for T^{**} . Next, we multiply the computational budget by a factor of ten, allocating these extra computational resources successively to each of the parameters. This results in $N = 200$, $CESS^* = 0.9999 \times N$ (such that $K_P = 1250$), $K_R = 1330$ and $s = 200$. The resulting ranges of the rare event probability estimates for T^{**} using ten runs are depicted in Figure 5.9a and the means and coefficients of variation (COV; ratio of standard deviation to the mean) for both thresholds are summarized in Table 5.1. While the means are comparable for all configurations, it is seen that the parameter with the most impact in reducing the COV for both thresholds is the number of particles N . In this test example, the optimal $CESS^*$ only has limited influence on the variance of the rare event estimate. Still, a high-quality representation of the posterior from the first stage leads to a smaller variance of the rare event estimate. Concerning the number of MH steps, we perform additional tests with values $s = 5, 10, 20, 200, 500$ (Fig. 5.9b for T^{**}). While there is high variance in the estimates for $s = 5$, the variance seems to stabilize from a value of $s = 20$ steps. Further increasing s to 200 or 500 necessitates a considerable number of additional forward operations, but leads to a much smaller improvement in the accuracy of the rare event estimate compared to increasing the number of particles. Furthermore, in the context of parallel computation, increasing the number of particles is more efficient compared to increasing s . Finally, when testing a value of K_R smaller than 100, we observed frequent failures due to the particle system dying. On the other hand, increasing the value to $K_R = 1330$ resulted in a decrease in the COV for both thresholds. Although this decrease was more significant than the effect of increasing the number of MH steps s , it still did not match the substantial improvement achieved by increasing the number of particles.

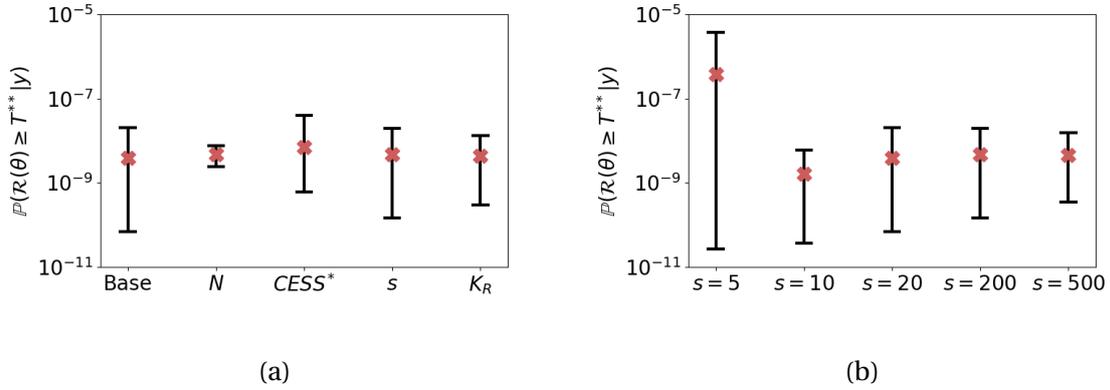


Figure 5.9: Impact of the configuration choices within the PostRisk-SMC method for the 1D flow example. (a) Range of the rare event probability estimates for T^{**} with the first bar corresponding to the base configuration and the following ones referring to the successive allocation of ten times more computational resources for either of the parameters with $N = 200$, $CESS^* = 0.9999 \times N$, $K_R = 1330$ and $s = 200$. (b) Range of the rare event probability estimates for T^{**} using different numbers of MH steps s . The red crosses in both plots indicate the mean values of the ten runs.

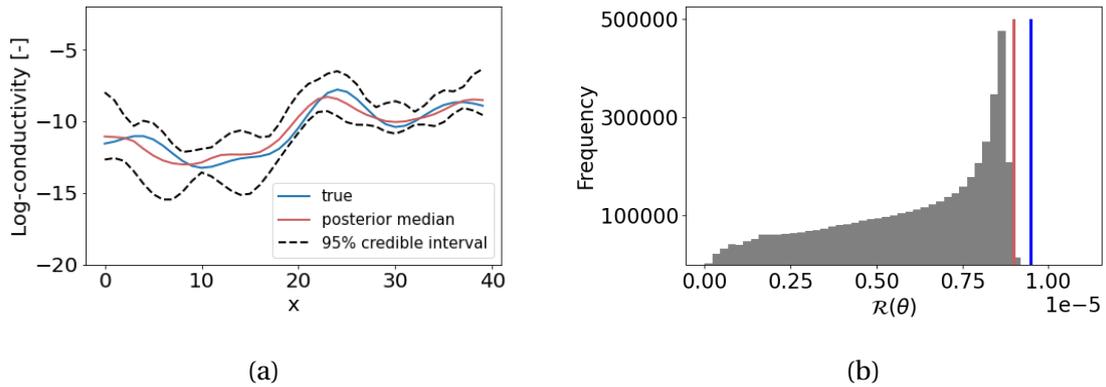


Figure 5.10: Results for the 1D flow example with the MH method: (a) Estimated posterior median (red) and credible interval (dashed) of the log-conductivity profile, together with the true profile (blue) and (b) transformed MH samples using $\theta \mapsto \mathcal{R}(\theta)$ with the thresholds of interest indicated (T^* in red and T^{**} in blue).

Table 5.1: Table summarizing the different trials of the PostRisk-SMC and MH method applied to the 1D flow test case. The second column indicates the computational budgets used for the thresholds (in terms of the total number of forward and quantity of interest simulations); the mean and COV (coefficient of variation) are calculated based on 10 estimates of $\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq T | \mathbf{y})$ for T^* and T^{**} .

	T^*/T^{**} [$\times 10^3$]	N	$\frac{CESS^*}{N}$	s	K_R	Mean T^* [$\times 10^{-3}$]	COV T^*	Mean T^{**} [$\times 10^{-9}$]	COV T^{**}
PostRisk-SMC	40/55	20	0.9	20	100	2.28	0.71	3.88	1.72
PostRisk-SMC	400/550	200	0.9	20	100	2.45	0.27	4.81	0.35
PostRisk-SMC	510/550	20	0.9999	20	100	2.60	0.49	6.91	1.66
PostRisk-SMC	400/550	20	0.9	200	100	2.91	0.65	4.77	1.08
PostRisk-SMC	255/550	20	0.9	20	1330	2.72	0.44	4.31	0.79
MH	400/550	-	-	-	-	2.45	0.25	0	-

To enable comparison with the standard MH algorithm, we run 10 chains in parallel with one million iterations each. The resulting posterior median and 95% credible interval of the estimated log-diffusivity are shown in Figure 5.10a and the resulting samples of $\mathcal{R}(\boldsymbol{\theta}) | \mathbf{y}$ in Figure 5.10b. If we would perform MH running three chains in parallel, convergence according to the potential-scale reduction factor (\hat{R} -statistics using a target value of 1.2 for all parameters and the second half of the chains; Gelman and Rubin 1992) would be declared after 140'000 iterations and the resulting estimate would be 6.44×10^{-3} for T^* and zero for T^{**} . This indicates that with the computational budget of the basic version of PostRisk-SMC as shown in Table 5.1, we are unable to obtain any reliable estimates with MH. With a higher budget of 400,000 for T^* , the mean of the ten estimates is 2.45×10^{-3} , and the COV is 0.25. The mean value matches the ones obtained with the PostRisk-SMC method. The comparable COV for the same computational budget of PostRisk-SMC ($N = 200$) is not surprising since the target probability enables enough samples in the MH chains. However, for T^{**} , all estimates obtained with MH are zero, even when using the full one million samples per chain.

Finally, we would like to highlight the power of including measurement data into this rare event estimation problem. As indicated in Figure 5.8b, for the prior distribution of the log-conductivity field ($k = 0$), $\mathcal{R}(\boldsymbol{\theta}) \geq T$ is not a rare event for the considered thresholds. Therefore, we can easily estimate $\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq T)$ under the prior using a limited number of Monte Carlo samples, which gives us 0.23 for T^* and 0.22 for T^{**} (here employing 10,000 samples). We conclude that, compared to this previous prior probability of about one quarter, the pumping test measurements lead us to the assessment that the hazard occurrence can be specified as highly unlikely, especially for T^{**} .

5.4 2D transport example

In the second test case, we infer a hydraulic transmissivity field θ using steady-state pressure data \mathbf{y} from pumping tests. For the quantity of interest $\mathcal{R}(\theta)$, we consider the release of a contaminant on the left side of the model domain and observe the breakthrough of the concentration at a location on the right side of the domain. We are examining a hypothetical scenario where the contamination is expected to no longer pose a risk beyond a pre-defined time frame. That is, the hazard materializes if we observe a breakthrough at the considered location before this time has elapsed.

5.4.1 Problem setting

The aquifer under consideration has a size of $250 \times 250 \times 5$ m and we use a discretization on a grid with $51 \times 51 \times 1$ cells. We assume the properties to be uniform in the vertical direction, thereby simplifying the problem to two spatial dimensions. For the purpose of simulating both the data and the quantity of interest, we utilize the MODFLOW package implemented in Python, specifically the FloPy library (Bakker et al. 2016).

We make the assumption that the system under investigation is confined. The unknown log-transmissivity field θ is assumed to be a Gaussian Random field (Chiles and Delfiner 2012). We assume a constant mean $\mu_{\log\theta} = \log(5 \times 10^{-5})$ with the transmissivity having units of m^2/s . For the isotropic covariance function, we employ an isotropic exponential covariance function in \mathbb{R}^2 with standard deviation $\sigma = 3$ and integral scale $l = 25$ m. In order to generate a realization of the (51×51) -dimensional Gaussian random field, we utilize a pixel-based parameterization,

$$\mathbf{X} = \boldsymbol{\mu}_\theta + \boldsymbol{\Sigma}_\theta^{1/2} \mathbf{Z}, \quad (5.17)$$

where $\boldsymbol{\Sigma}_\theta$ denotes the exponential covariance matrix and \mathbf{Z} represents a (51×51) -dimensional random vector composed of independent and identically distributed (i.i.d.) standard normal variables. The ‘true’ log-transmissivity field is depicted in Figure 5.11a.

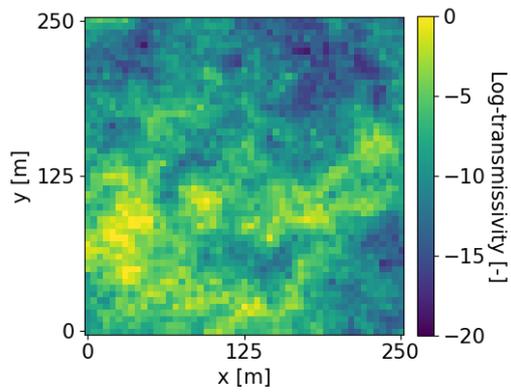
For the data \mathbf{y} , we are considering a five-spot pumping test using a pumping well located in the middle of the model domain and local measurements of the log-transmissivity field at the well locations (Fig. 5.11b). For the pumping test, we consider a fixed hydraulic head at the left (2.5 m) and right (0 m) sides of the domain, no-flow boundaries on the other boundaries and pump with a rate of $5 \times 10^{-4} \text{ m}^3/\text{s}$. For the data collection, we consider the steady-state of the system and measure the hydraulic head in four wells centered in the middle of the four quadrants of the domain. For the generation of the synthetic data, we add independent Gaussian observational errors with a standard deviation of 0.02 m. For the local measurements in the five wells, we assume a Gaussian measurement error with a standard deviation of 0.1 (log-scale). Then, we employ standard results for conditional Gaussian random fields, resulting in a mean and covariance matrix in Equation (5.17), which are conditioned on the local measurements and include their error.

For the rare event, we examine a scenario where a contaminant is released on the left side of the model domain, while monitoring the concentration of the contaminant on the right side. Our primary focus lies in determining the time of breakthrough $\mathcal{R}(\boldsymbol{\theta})$ in a critical area in the middle of the right side of the model domain. The hazard is specified as a breakthrough before 60 days ($T = 60$ days), with the breakthrough being specified as the concentration being higher or equal than 1 mg/l. To simulate this, we assume a constant concentration of 1 g/l on the left side, along with a fixed hydraulic head difference of 2.5 m between the left and right sides (as for the data collection). Additionally, we maintain a constant porosity of 0.3, an effective molecular diffusion coefficient of 10^{-9} m²/s, a longitudinal dispersivity of 1 m, and a ratio of the transverse to the longitudinal dispersivity of 0.1. Figure 5.11c illustrates the concentration distribution after 60 days from the start of the injection for the true field, and Figure 5.11d visualizes the corresponding breakthrough front.

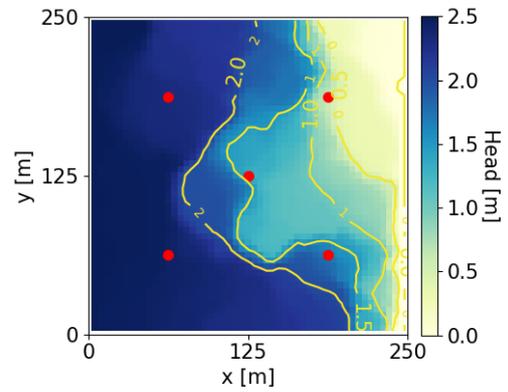
5.4.2 Results

We first investigate the occurrence of a contamination breakthrough without incorporating the data. Given the resource-intensive nature of the transport simulations, we adhere to a computational limit of approximately 15,000 evaluations of $\mathcal{R}(\cdot)$. When using the PostRisk-SMC method for this setting, we only employ the second phase and use $N = 40$ particles and $s_R = 10$ MH steps per subset (Fig. 5.4). For the sequence of thresholds, we employ a decreasing logarithmic sequence ranging from $T_1 = 3500$ days down to 100 days, utilizing 30 steps (according to Equation 5.16 with $K_P = 0$). As the conditional probability during the last steps becomes lower and the risk of the particle system dying is particularly high, we adapt the sequence to steps of five days from 100 days down to the 60 days of interest, leading to $K_R = 38$. For the propagation of the particles with MH, we use pCN proposals (Eq. 5.4) initialized with a $\rho = 1$ (independent proposals), which is then geometrically decreased by a factor of 0.9 in each subset. In Figure 5.12, we provide visual representations of three illustrative log-hydraulic transmissivity field realizations extracted from the final subset where $\mathcal{R}(\boldsymbol{\theta}) \leq 60$ days. These examples are accompanied by their respective contamination fields. Figure 5.13a displays the mean transmissivity field of the particles. Running ten repetitions of the PostRisk-SMC method, we obtain a mean of 0.71×10^{-4} and a COV of 0.37 for $\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \leq 60 \text{ d})$ (Table 5.2). With prior sampling and Monte Carlo estimation for the same computational budget, we obtain a mean of 0.87×10^{-4} and a COV of 0.60. While the Monte Carlo approach includes zero in the range of the ten probability estimates, the PostRisk-SMC method specifies the probability as being at least 0.24×10^{-4} .

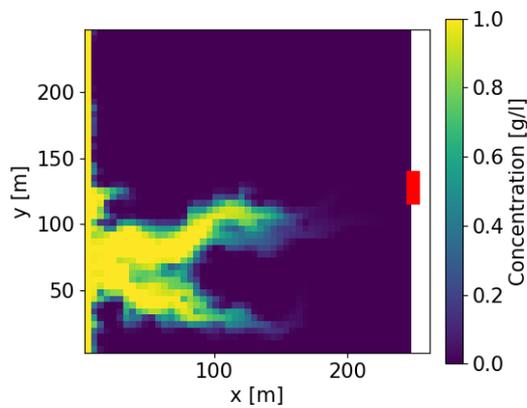
We now consider the data. Figure 5.11d demonstrates that the hazard is occurring for the true log-hydraulic transmissivity field and we are interested to see if the integration of the local and pumping measurements helps to reflect this by increasing the rare event probability estimate. For the posterior inference part of PostRisk-SMC, we use a configuration with $N = 40$, $CESS^*/N = 0.99$ (leading to $K_P = 100$) and $s_P = 100$ MH steps per iteration (Fig. 5.2). A particle estimate of the posterior mean is depicted in Figure 5.13b. For the rare event phase of PostRisk-SMC, we implement the adaptation outlined in Section 5.2.3, wherein we conduct $ss_R = 100$ posterior steps within each of the $s_R = 10$ MH steps during the rare event phase of the algorithm. This implies that for every subset, we need to assess $\mathcal{R}(\cdot)$ ten



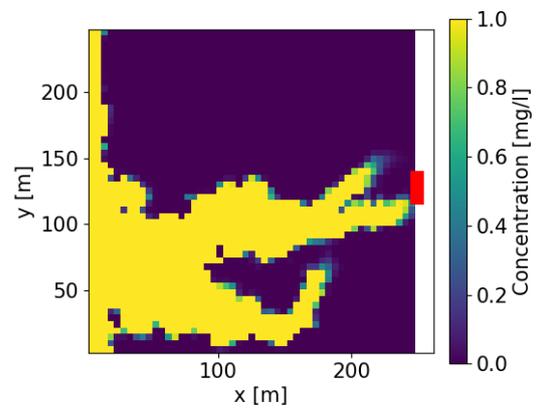
(a)



(b)



(c)



(d)

Figure 5.11: (a) 'True' log-hydraulic transmissivity field and corresponding (b) hydraulic heads resulting from the steady-state pumping test, the red dots indicate the well locations, (c) contamination field and (d) breakthrough front after 60 days.

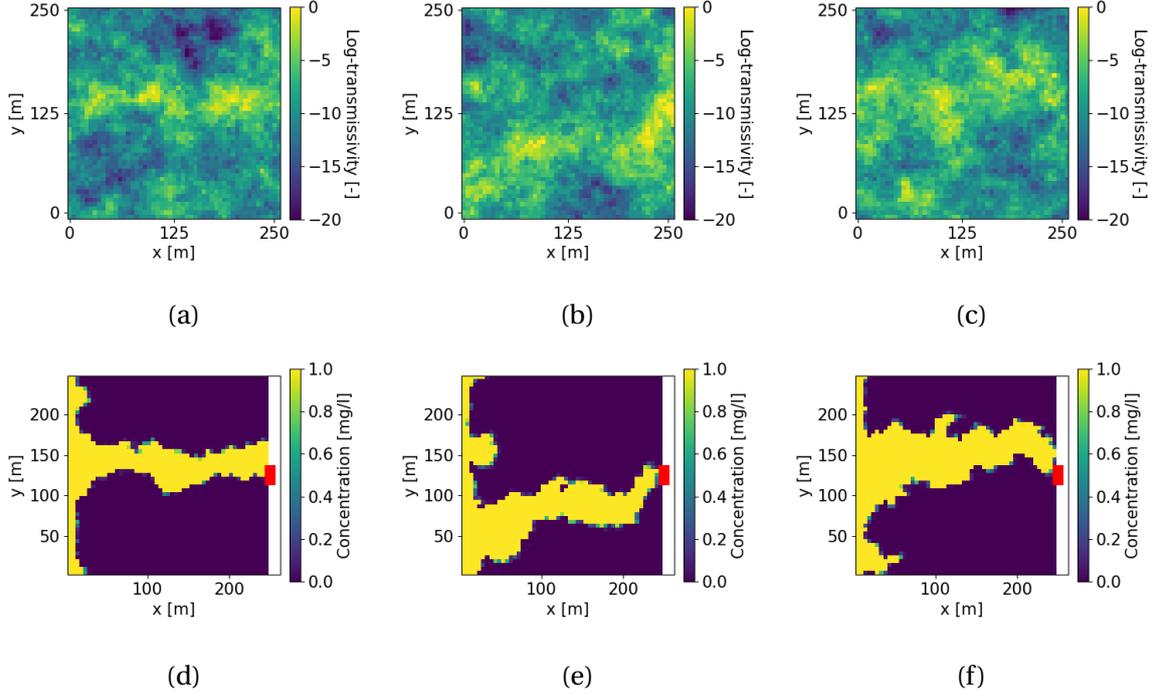


Figure 5.12: Rare event estimation for the 2D transport example with the PostRisk-SMC method (without inversion): (a-c) log-hydraulic transmissivity field examples from the final subset with $\mathcal{R}(\boldsymbol{\theta}) \leq 60$ days and (d-f) their corresponding contamination breakthrough fronts.

times and $\mathcal{G}(\cdot)$ one thousand times. We use the same sequence of thresholds with $K_R = 38$ as described above. In total, this results in $N \times (K_P \times s_P + K_R \times s_R \times s_{RR}) = 1.92$ Million evaluations of $\mathcal{G}(\cdot)$ and $N \times K_R \times s_R = 15,200$ evaluations of $\mathcal{R}(\cdot)$ (Table 5.2). For the propagation, the step size of the pCN proposals (Eq. 5.4) is adapted such that the ‘posterior’ steps have an acceptance rate of about 30 %. In Figure 5.14, we showcase three particles from the final posterior subset where $\mathcal{R}(\boldsymbol{\theta}) \leq 60$ days, along with their corresponding contamination fields. Figure 5.13c shows the mean of the particles lying in the last posterior subset. Upon executing the PostRisk-SMC method ten times, we compute an average of 4.56×10^{-4} and observe a COV of 0.21 for $\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \leq 60 \text{ days})$ (Table 5.2).

For a fair comparison with Monte Carlo estimation based on MH samples, we run ten chains with 1.92 Million steps and evaluate $\mathcal{R}(\cdot)$ for only 15,000 samples (per chain) that are obtained by thinning. We employ pCN proposals with an adjusted step size aiming for an acceptance rate of 30 %. We obtain a mean rare event probability estimate of 5.64×10^{-4} and a COV of 0.49 (Table 5.2). Using the first three chains, convergence with respect to the \hat{R} -statistics would be declared after 350,000 iterations. The corresponding merged 1,500 thinned samples per chain would specify the hazard occurrence probability as zero.

Similar to the one-dimensional flow example, we can observe that incorporating measurements leads to a shift in our estimation of the hazard occurrence probability. In the context of this two-dimensional transport example, the incorporation of local measurements and

pumping data increases the estimated probability of hazard occurrence by a factor of about six compared with the estimate based on prior knowledge only. We observe that for the ten considered estimates, the range of the values for the prior and posterior can be clearly separated (for both PostRisk-SMC and Monte Carlo estimation).

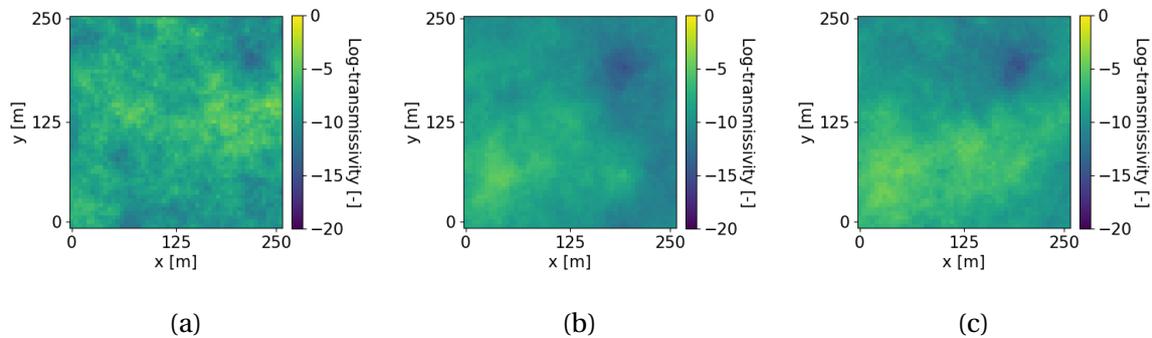


Figure 5.13: Results for the 2D transport example with the PostRisk-SMC method: Particle mean representing the log-hydraulic transmissivity field from the (a) prior subset where $\mathcal{R}(\boldsymbol{\theta}) \leq 60$ days, (b) posterior distribution and (c) posterior subset where $\mathcal{R}(\boldsymbol{\theta}) \leq 60$ days.

Table 5.2: Table summarizing the different trials of the PostRisk-SMC and MH method applied to the 2D transport test case under the prior and the posterior distribution. The second column shows the number of required simulations of the forward response $\mathcal{G}(\cdot)$ and quantity of interest $\mathcal{R}(\cdot)$ and mean, COV (coefficient of variation), min (minimum) and max (maximum) refer to the 10 estimates of the rare event probability.

Method		$\mathcal{G}(\cdot)/\mathcal{R}(\cdot)$ [$\times 10^4$]	Mean [$\times 10^{-4}$]	COV	Min [$\times 10^{-4}$]	Max [$\times 10^{-4}$]	N	$\frac{CESS^*}{N}$	s_P	s_R	ss_R	K_R
Prior	PostRisk-SMC	- / 1.5	0.71	0.37	0.24	1.09	40	-	-	10	-	38
	Monte Carlo	- / 1.5	0.87	0.60	0	1.33	-	-	-	-	-	-
Posterior	PostRisk-SMC	192 / 1.5	4.56	0.21	3.55	6.64	40	0.99	100	10	100	38
	MH	192 / 1.5	5.64	0.49	2.01	12.75	-	-	-	-	-	-

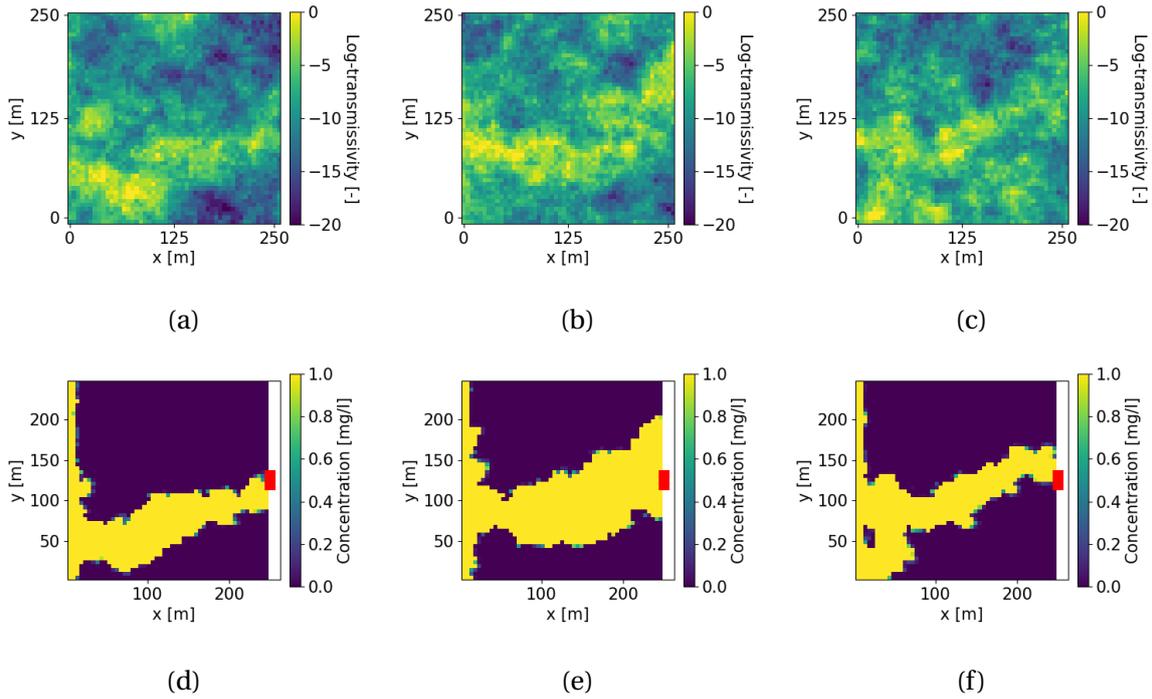


Figure 5.14: Rare event estimation for the 2D transport example with the PostRisk-SMC method (with inversion): (a-c) log-hydraulic transmissivity field examples from the final subset with $\mathcal{R}(\theta) \leq 60$ days and (d-f) their corresponding contamination breakthrough fronts.

5.5 Discussion

Sustainable groundwater management and assessment of associated hazards are pressing needs that are being accentuated under global change (e.g., Siebert et al. 2010, Famiglietti 2014, Gorelick and Zheng 2015). With the Posterior Risk Sequential Monte Carlo (PostRisk-SMC) method, we present an approach that combines Bayesian inversion and rare event probability estimation under uncertainty. It first generates a particle approximation of the posterior which is then propagated to provide an accurate estimation of the rare hazard probability. Thereby, the method relies on ‘subset sampling’ and aims to estimate a small probability as a product of larger conditional probabilities. In addition to probability estimation, the method also generates realizations of the rare event (as illustrated in Figs. 5.7 and 5.14), providing tangible representations of how the subsurface property field leading to the hazard could look like in practice.

In the first phase of the PostRisk-SMC method, we employ adaptive SMC proposed by Zhou et al. (2016), relying on an adaptive sequence of the exponent of power posteriors giving increasingly more weight to the likelihood. Generally, as the number of particles increases, the approximation becomes more accurate. Moreover, a slower increase of the exponent and a higher number of MH steps per iteration also contribute to improving the accuracy of the approximation. All three factors come at the expense of computational resources.

However, unlike many MCMC methods, the SMC method is particularly well-suited for parallel computation, as the particles can be distributed across multiple computing nodes. The configuration and computational power required ultimately depends on the complexity of the posterior distribution, which is influenced by various factors such as the dimension of the parameter space, the underlying physics, and the characteristics of the measurement setup (Amaya et al. 2021).

In the second phase of the PostRisk-SMC method, we rely on subset sampling to estimate the rare event probabilities. The selection of intermediate thresholds is a crucial aspect that involves a trade-off between the intermediate conditional probabilities and the number of particles (Au and Beck 2001). If the threshold increases slowly, the conditional probabilities are large and a small number of particles is needed to ensure accurate estimation. On the other hand, more intermediate thresholds are needed until the target threshold is reached. If the thresholds increase faster, more particles are needed for an accurate estimation, which also increases the total number of simulations. Cérou et al. (2012) propose an adaptive sequence of thresholds based on quantiles to increase the efficiency of their algorithm. The negative aspect of introducing adaptive thresholds is a positive bias in the rare event probability estimate, which diminishes with an increasing number of particles (Cérou et al. 2012). Cérou et al. (2012) propose a correction factor for the bias, however, their analytical study assumes that the particles are independent, which is hard to guarantee in practice due to the resampling and the finite number of MH steps s .

In the one-dimensional flow example (Section 5.3), the bias resulting from the adaptive thresholds turns out to be far from negligible, especially when using a relatively small number of $N = 40$ particles and targeting a rare event with probability of one in a billion (Fig. 5.8). To avoid this bias, and the computational burden associated with re-running or increasing the number of particles, we employ instead a fixed sequence of thresholds (Eq. 5.16). However, this leads to the possibility of the particle system “dying” during the rare event estimation process if no particles exceed the current threshold. We did not specifically consider this scenario, but one possible approach to address this issue is discussed by LeGland and Oudjane (2006). Their idea involves continuing to generate new particles until a specified count of particles has reached the given threshold. Both increasing the number of particles N and the number of subsets K_R decreases the risk of a dying particle system. While increasing the number of particles N seems to be a general recipe to decrease the rare event estimator’s variance, increasing K_R leads to fewer particles discarded in each step, which reduces the variance of the estimator. However, it potentially results in high conditional probabilities that are hard to estimate with a small number of particles, which could be a factor mitigating the benefit. Furthermore, in this first simplistic test case with a rather smooth posterior, a slower increase of the exponents (higher $CESS^*$) only had a limited influence on the variance of the rare event estimate. This happens as a small number s of MH steps can prevent the particles from collapsing, even after resampling steps.

In the context of the two-dimensional transport example (Section 5.4), posterior exploration presents a challenge as strong non-uniqueness and underdetermination enable a wide range of solutions to accurately explain the observed data (Soueid Ahmed et al. 2014 Cotter et al.

2013). Hence, the number of resampling steps and the propagation through the MH steps play a crucial role in preventing particle collapse. This latter aspect gains even greater significance during the phase of rare event estimation, as resampling cannot be avoided. For this reason, we implement a slight adaptation of the PostRisk-SMC method outlined in Figure 5.5. Rather than simulating both the forward response and the quantity of interest at each iteration of the rare event phase, we first perform a sequence of $s s_R$ posterior steps during each of the s_R MH steps. We then consider the last state as proposal from the posterior distribution and decide to accept or reject it depending on whether it lies within the current subset. In scenarios involving contamination simulations, where the computational cost of the contamination simulation typically surpasses that of the data simulation flow model, this strategy enhances particle propagation efficiency while simultaneously decreasing computational demands.

In both test examples, we investigate the significance of using the posterior instead of the prior PDF to determine the probability of hazard occurrence. In the context of the one-dimensional flow example, we showcase how the introduction of pumping test measurements in this scenario alters a rather likely event into a highly unlikely one. Indeed, the initial occurrence probability of roughly a quarter is after considering the data turned into a probability of one in a billion for T^{**} . In the case of the two-dimensional transport example, the situation is reversed: the inclusion of local measurements and pumping data helps in quantifying the probability of hazard occurrence as being six times higher than with prior knowledge alone. The integration of posterior inference serves as a clear demonstration of why it is crucial to design appropriate data acquisition strategies within the realm of risk assessment. Designing appropriate experimental designs for such tasks is a relatively open research area that we leave for the future.

We compare the performance of the PostRisk-SMC method with a conventional Monte Carlo approach relying on prior or posterior samples obtained by the MH algorithm. In the one-dimensional flow example (Table 5.1), the estimates obtained with PostRisk-SMC align with those of the traditional method for the less rare event. For the more rare event with occurrence probability approaching one in a billion, the Monte Carlo approach fails in simulating the hazardous scenario. The PostRisk-SMC method, on the other hand, is able to specify the occurrence probability with a coefficient of variation of 0.35. In the two-dimensional transport example (Table 5.2), the PostRisk-SMC method successfully reduces the coefficient of variation by more than 50 % compared to Monte Carlo estimation based on MH samples (for the inversion setting). This comparison is established within a scenario where Monte Carlo estimation remains feasible. For rarer events, we anticipate complete failure of Monte Carlo estimation, as showcased by the one-dimensional flow example (Table 5.1).

It is worth noting that the two phases of the PostRisk-SMC method exhibit different dynamics. While in our 1D flow example, the adaptive procedure for the exponents defining the power posteriors leads to an exponential increase, the sequence of thresholds follows a logarithmic progression. In Section 5.4, we take an initial step in addressing this distinct difference in dynamics by using different numbers of MH steps for the two phases of the method. However, there is considerable potential for further exploration and refinement in this regard. So far, we only dealt with rare sets $A = \{\boldsymbol{\theta} \in \mathbb{R}^P : \mathcal{R}(\boldsymbol{\theta}) \in \mathcal{T}\}$ with $\mathcal{T} = [T, \infty)$ or $\mathcal{T} = (-\infty, T]$ for some

real number T . If we would consider $\mathcal{T} = [T^*, T^{**}]$, one could gradually shrink the interval from both sides. Looking ahead, it could be interesting to incorporate surrogate modeling within the PostRisk-SMC method to tackle more complex and realistic problems. Surrogates (e.g. Razavi et al. 2012) in this context can serve as simplified models or approximations of the underlying system, allowing for faster evaluations and reducing the computational burden.

5.6 Conclusions

The combination of Bayesian inversion and rare event estimation is very helpful for understanding groundwater hazards and their implications for humans and ecosystems. To overcome the challenges of rare event estimation in an inversion setting, we present a two-stage formulation of Sequential Monte Carlo, denoted as the PostRisk-SMC method. First, particles are generated to approximate the posterior distribution by adaptively increasing the exponent of the likelihood function. Second, subset sampling is employed to evaluate the probability of the rare event of interest. To showcase the efficacy and accuracy of the PostRisk-SMC method, we present a one-dimensional flow example and a two-dimensional flow- and transport example. The one-dimensional example demonstrates that the PostRisk-SMC method allows us to estimate rare event probabilities as low as one in a billion. In the two-dimensional example, we showcase the method's capability for rare event probability estimation in a more realistic and complex setting. In both examples, the PostRisk-SMC method successfully reduces the coefficient of variation of the rare event probability estimate compared to Monte Carlo estimation based on posterior samples. In both cases, the addition of the measurement data lead to a distinctly different assessment of the occurrence probability than relying on the prior only. This highlights the need for experimental design in this context; a subject that we leave for future research. Future work will also consider inclusion of surrogate modeling to speed up computations and applications to actual field settings.

Chapter 6

An energy-based model approach to rare event probability estimation

Lea Friedli, David Ginsbourger, Arnaud Doucet and Niklas Linde

Submitted to SIAM/ASA Journal on Uncertainty Quantification

Abstract

The estimation of rare event probabilities plays a pivotal role in diverse fields. Our aim is to determine the probability of a hazard or system failure occurring when a quantity of interest exceeds a critical value. In our approach, the distribution of the quantity of interest is represented by an energy density, characterized by a free energy function. To efficiently estimate the free energy, a bias potential is introduced. Using concepts from energy-based models (EBM), this bias potential is optimized such that the corresponding probability density function approximates a pre-defined distribution targeting the failure region of interest. Given the optimal bias potential, the free energy function and the rare event probability of interest can be determined. The approach is applicable not just in traditional rare event settings where the variable upon which the quantity of interest relies has a known distribution, but also in inversion settings where the variable follows a posterior distribution. By combining the EBM approach with a Stein discrepancy-based stopping criterion, we aim for a balanced accuracy-efficiency trade-off. Furthermore, we explore both parametric and non-parametric approaches for the bias potential, with the latter eliminating the need for choosing a particular parameterization, but depending strongly on the accuracy of the kernel density estimate used in the optimization process. Through three illustrative test cases encompassing both traditional and inversion settings, we show that the proposed EBM approach, when properly configured, (i) allows stable and efficient estimation of rare event probabilities and (ii) compares favorably against subset sampling approaches.

6.1 Introduction

Estimating rare event probabilities is a fundamental challenge in various fields, including finance, engineering and environmental sciences. Rare events are characterized by their low occurrence rates, but they become crucial when their outcomes have substantial consequences (Juneja and Shahabuddin 2006). In the field of uncertainty quantification, rare events are frequently related to the failure of engineering systems that are designed to be highly reliable (Beck and Zuev 2015), examples of which include hydroelectric dams, airplanes, or nuclear reactors. In these cases, the accurate estimation of rare event probabilities is of utmost importance as it enables informed decision-making, effective risk management, and the design of robust systems.

In practical applications, there are commonly no analytical formulas for estimating rare event probabilities. To address this issue, asymptotic approximation methods such as the first-order reliability method (FORM; Hasofer and Lind 1974) have been proposed. However, relying solely on FORM results without understanding the characteristics of the linearized domains, particularly in higher dimensions, is not recommended (Straub et al. 2016). On the other hand, conventional Monte Carlo simulation methods are often computationally inefficient. Consequently, a considerable amount of research has been dedicated to developing more efficient stochastic simulation techniques for rare event probability estimations (e.g., Bucklew 2004, Rubino and Tuffin 2009). Beck and Zuev (2015) present two principal stochastic simulation approaches: importance sampling and subset simulation. Importance

sampling is a variance reduction technique, aiming to increase sampling frequency in the region of interest, which in the present context corresponds to the region containing the rare event and its vicinity (e.g., Au and Beck 1999). Subset simulation (Au and Beck 2001) represents a rare event probability as a product of larger probabilities, effectively breaking down the rare event into less rare conditional events (e.g., Au and Wang 2014). While subset sampling explores a wide range of parameter combinations, the more directed line sampling (e.g., Hohenbichler and Rackwitz 1988, Koutsourelakis et al. 2004) generates samples on a hyperplane that is orthogonal to a significant direction pointing to the rare event region.

In this study, we are particularly interested in rare event probability estimation in the context of an underlying inverse problem. We employ non-linear Bayesian inversion aiming to infer unknown properties θ given measurements \mathbf{y} . Instead of being interested in the posterior distribution itself, we target the distribution of a real-valued quantity of interest that depends on the unknown properties through a non-linear relationship $\theta \mapsto \mathcal{R}(\theta)$. More particularly, we would like to estimate the probability of this quantity exceeding a critical threshold, $\mathbb{P}(\mathcal{R}(\theta) \geq T | \mathbf{y})$, which is related to the problem of evaluating the probability of failure of a system. As the underlying relationships are non-linear, analytical formulas for the distribution of $\mathcal{R}(\theta)$ conditioned on the data \mathbf{y} are typically not available. In structural reliability engineering, similar problems have been targeted with data first being used to update θ and then apply these updated distributions to the prediction of rare events (e.g., Papadimitriou et al. 2001, Jensen et al. 2013, Sundar and Manohar 2013, Hadjidoukas et al. 2015). Specialized estimation methods are essential since conventional Monte Carlo approaches relying on posterior samples would typically demand overwhelmingly large sample sizes.

In practice, most structural reliability methods encounter difficulties when starting with a sample approximation of the posterior, but subset sampling is an exception (Straub et al. 2016). In an inversion context, it was employed to estimate the so-called “updated robust failure probability” by Jensen et al. (2013) and Hadjidoukas et al. (2015). Also in the engineering literature, Straub (2011) introduced Bayesian Updating with Structural reliability methods for posterior inference (BUS; e.g. Straub and Papaioannou 2015). This method can be considered as an extension of rejection sampling that aims to overcome its inefficiency by treating the acceptance event as a rare event using structural reliability methods (e.g., FORM, subset simulation and line sampling; see Straub et al. 2016). It offers a framework, in which the rare event probability under the posterior can be estimated directly within the framework of structural reliability analysis. A recent advancement to the BUS framework involves the utilization of cross entropy-based importance sampling, which enables efficient sampling from the critical region of the posterior failure domain (Kanjilal et al. 2023).

Free energy is a fundamental concept in materials science and physical chemistry (Stecher et al. 2014). Performing the free energy estimation on crude Monte Carlo draws is rarely practical as it would need an excessively large number of samples to sufficiently cover the whole range of the states, particularly in a rare event setting as ours. To address this challenge, one solution is to enhance the sampling in specific regions of interest by introducing a bias term. This approach, known as umbrella or non-Boltzmann sampling, can be traced back

to the work of Torrie and Valleau (1977), who demonstrated how to recover the unbiased probability distribution from the biased samples. Applications for free energy functions are numerous (e.g., Stecher et al. 2014, Kästner and Thiel 2005). More recently, Shirts and Ferguson (2020) introduced a Bayesian formalism to estimate the free energy function by minimizing the Kullback–Leibler divergence between a continuous trial function and the empirical samples generated by biased sampling. Constructing a bias potential is a non-trivial task, therefore, Valsson and Parrinello (2014) introduced a variational approach that adaptively combines exploration and reconstruction by iteratively improving and refining the bias potential and the free energy function. The approach by Valsson and Parrinello (2014) not only utilizes the bias potential to improve the sampling technique employed for optimizing the free energy function (as with umbrella sampling), but also depends on it to directly estimate the free energy function. This variationally-enhanced sampling method has been applied in the context of coarse-graining methods by Invernizzi et al. (2017) and combined with machine-learning techniques by Bonati et al. (2019).

By employing concepts from energy-based models (EBMs), we propose a new formulation and approach to solve rare event probability estimation problems. We write the marginal posterior distribution of the quantity of interest $R = \mathcal{R}(\boldsymbol{\theta})$ as an energy density function with free energy $r \mapsto F(r)$. To estimate $F(r)$ efficiently in the region of interest, a bias potential $r \mapsto V(r)$ and a corresponding probability density function (PDF) $r \mapsto p_V(r) \propto \exp(-(F(r) + V(r)))$, are introduced. That is, the energy density function of R (knowing \mathbf{y}) is $p_0(r)$, where 0 refers to a zero-valued bias potential. The considered EBM (Goodfellow et al. 2016) approach relies on optimizing the bias potential $V(r)$ such that $p_V(r)$ approximates a pre-defined PDF $p_{\text{ref}}(r)$, which is selected such that it has high probability mass in the targeted region. Given the optimal bias potential $V(r)$, it is straightforward to derive the free energy $F(r)$ with an accurate estimation in the region where $p_{\text{ref}}(r)$ has most of its mass being emphasized. Our approach is related to the variational method by Valsson and Parrinello (2014), but we adopt the fundamental concepts of their method within a different formulation and apply it to a novel context, that is, rare event probability estimation. Practically, in EBM methods, the potential $V(r)$ is parameterized using methods such as neural networks, splines or radial basis functions. In Valsson and Parrinello (2014), the optimal $V(r)$ is approximated by minimizing a loss function which is related to the Kullback–Leibler divergence between $p_{\text{ref}}(r)$ and $p_V(r)$ using stochastic optimization methods. We adapt this approach and add a stopping criterion based on the Stein discrepancy (Gorham and Mackey 2015) to achieve a balance between computational efficiency and satisfactory model performance. Furthermore, we introduce a non-parametric approach for the bias potential, thereby alleviating the challenge of selecting an appropriate parameterization.

The proposed EBM approach reduces the potentially high-dimensional problem of estimating the posterior distribution of $\boldsymbol{\theta}$ and subsequently $\mathcal{R}(\boldsymbol{\theta})$ to an optimization of a one-dimensional function $V : \mathbb{R} \rightarrow \mathbb{R}$. Importantly, the EBM approach is not limited to inversion settings and can also be effectively applied to traditional rare event probability estimation. We examine the performance of our proposed method using three illustrative test examples. The first example targets the probability of high contamination values endangering organism living in the soil. The presence of an analytical solution allows us to explore the EBM

method's sensitivity to implementation variables. Next, we consider the two-dimensional four-branch function problem, commonly used as a benchmark in reliability analysis. This example does not involve inversion and demonstrates the method's suitability for general rare event estimation scenarios. Last, we use a simple load and capacity example from Straub et al. (2016) to compare the EBM method's performance against their BUS approach. The manuscript is organized as follows: Section 6.2 introduces the problem setting, outlines our energy-based model approach, and details the methods employed for comparison. In Section 6.3, the test examples are presented. Subsequently, Section 6.4 discusses the results, leading to the conclusions exposed in Section 6.5.

6.2 Methodology

6.2.1 Problem setting

In the considered rare event setting, we target a quantity of interest $R = \mathcal{R}(\boldsymbol{\theta})$ derived from the random vector $\boldsymbol{\theta}$ on \mathbb{R}^d by some non-linear function $\mathcal{R} : \mathbb{R}^d \rightarrow \mathbb{R}$. We consider a rare set $A = \{\boldsymbol{\theta} \in \mathbb{R}^d : \mathcal{R}(\boldsymbol{\theta}) \geq T\}$ with $T \in \mathbb{R}$ and want to estimate $\mathbb{P}(\boldsymbol{\theta} \in A)$. In a traditional rare event setting, we consider $\boldsymbol{\theta}$ distributed according to a prior PDF $p_{\boldsymbol{\theta}}(\boldsymbol{\theta})$, which is absolutely continuous with respect to a dominating measure, typically the Lebesgue measure on \mathbb{R}^d . The rare event probability can then be expressed as,

$$\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq T) = \mathbb{P}(\boldsymbol{\theta} \in A) = \int_A p_{\boldsymbol{\theta}}(\boldsymbol{\theta}) d\boldsymbol{\theta}. \quad (6.1)$$

Bayesian probabilistic inversion methods target generally unknown properties $\boldsymbol{\theta}$ and seek to infer their posterior PDF given the measurements $\mathbf{y} \in \mathbb{R}^m$. In most applications, the random data vector is given by $\mathbf{Y} = \mathcal{G}(\boldsymbol{\theta}) + \boldsymbol{\varepsilon}_{\mathcal{O}}$, with $\mathcal{G} : \mathbb{R}^d \rightarrow \mathbb{R}^m$ referring to the forward operator and $\boldsymbol{\varepsilon}_{\mathcal{O}}$ to the observational noise. In Bayes' theorem, the posterior PDF of the target parameters $\boldsymbol{\theta}$ given measurements \mathbf{y} is given by,

$$p_{\boldsymbol{\theta}|\mathbf{Y}}(\boldsymbol{\theta}|\mathbf{y}) = \frac{p_{\boldsymbol{\theta}}(\boldsymbol{\theta})p_{\mathbf{Y}|\boldsymbol{\theta}}(\mathbf{y}|\boldsymbol{\theta})}{p_{\mathbf{Y}}(\mathbf{y})}, \quad (6.2)$$

with the prior PDF $p_{\boldsymbol{\theta}}(\boldsymbol{\theta})$, the likelihood function $p_{\mathbf{Y}|\boldsymbol{\theta}}(\mathbf{y}|\boldsymbol{\theta})$ and the evidence $p_{\mathbf{Y}}(\mathbf{y})$ (assumed non-zero). Assuming independent Gaussian observational errors, the likelihood is expressed as $p_{\mathbf{Y}|\boldsymbol{\theta}}(\mathbf{y}|\boldsymbol{\theta}) = \varphi_m(\mathbf{y}; \mathcal{G}(\boldsymbol{\theta}), \boldsymbol{\Sigma}_{\mathbf{Y}})$, with $\varphi_m(\cdot; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denoting the PDF of a m -variate normal distribution with mean $\boldsymbol{\mu}$ and diagonal covariance matrix $\boldsymbol{\Sigma}$ with the variance of the observational error on its diagonal. For $\boldsymbol{\theta}$ distributed according to a Bayesian posterior PDF $\boldsymbol{\theta} \mapsto p_{\boldsymbol{\theta}|\mathbf{Y}}(\boldsymbol{\theta}|\mathbf{y})$, we aim to estimate,

$$\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq T|\mathbf{y}) = \mathbb{P}(\boldsymbol{\theta} \in A|\mathbf{y}) = \int_A p_{\boldsymbol{\theta}|\mathbf{Y}}(\boldsymbol{\theta}|\mathbf{y}) d\boldsymbol{\theta}. \quad (6.3)$$

6.2.2 Energy-based model approach

While our methodology is formulated for an inversion setting, it can be readily adjusted to suit a conventional rare event estimation scenario with $\boldsymbol{\theta}$ being distributed according to a prior PDF.

We write the posterior PDF of $\boldsymbol{\theta}$ as an energy density function,

$$p_{\boldsymbol{\theta}|\mathbf{y}}(\boldsymbol{\theta}|\mathbf{y}) = \frac{\exp(-U(\boldsymbol{\theta}))}{Z}, \quad Z := \int \exp(-U(\zeta))d\zeta, \quad (6.4)$$

with free energy $U(\boldsymbol{\theta}) = -\log p_{\mathbf{Y}|\boldsymbol{\theta}}(\mathbf{y}|\boldsymbol{\theta}) - \log p_{\boldsymbol{\theta}}(\boldsymbol{\theta})$. The marginal posterior distribution of the quantity of interest $R = \mathcal{R}(\boldsymbol{\theta})$ is thus given by,

$$p_{R|\mathbf{Y}}(r|\mathbf{y}) = \int \frac{\exp(-U(\boldsymbol{\theta}))}{Z} \delta(r - \mathcal{R}(\boldsymbol{\theta}))d\boldsymbol{\theta}. \quad (6.5)$$

More details on the definition of such integrals can be found in Lelièvre et al. (2010) (Lemma 3.2; co-area formula). We write $p_{R|\mathbf{Y}}(r|\mathbf{y})$ as an energy density function with (unknown) free energy $r \mapsto F(r)$,

$$p_{R|\mathbf{Y}}(r|\mathbf{y}) = \exp(-F(r)). \quad (6.6)$$

In order to estimate $F(r)$ and $p_{R|\mathbf{Y}}(r|\mathbf{y})$, we have the option of utilizing posterior samples obtained from $p_{\boldsymbol{\theta}|\mathbf{Y}}(\boldsymbol{\theta}|\mathbf{y})$ and applying a transformation $\boldsymbol{\theta} \mapsto \mathcal{R}(\boldsymbol{\theta})$. Since this follows a typical Monte Carlo approach, it would need an impractically large number of samples to include the low probability regions of $p_{R|\mathbf{Y}}(r|\mathbf{y})$ that are of interest. To enhance sampling in the region of interest, we seek to sample in accordance to a pre-defined $p_{\text{ref}}(r)$ with most mass on the set $\{R \geq T\}$. To achieve this, we introduce the bias potential $r \mapsto V(r)$ and the PDF $r \mapsto p_V(r)$,

$$p_V(r) = \frac{\exp(-(F(r) + V(r)))}{\int \exp(-(F(s) + V(s)))ds}. \quad (6.7)$$

We seek $p_V(r) = p_{\text{ref}}(r)$, which holds for $V_{\text{opt}}(r) = -F(r) - \log(p_{\text{ref}}(r))$ (for $p_{\text{ref}}(r) > 0$, ignoring an irrelevant constant). Given $V_{\text{opt}}(r)$, it is straightforward to obtain $F(r)$ as $F(r) = -\log(p_{\text{ref}}(r)) - V_{\text{opt}}(r)$, for $p_{\text{ref}}(r) > 0$ and hence $p_{R|\mathbf{Y}}(r|\mathbf{y})$. Eventually, we can express the probability of the rare event as,

$$\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq T|\mathbf{y}) = \int_T^\infty p_{R|\mathbf{Y}}(r|\mathbf{y})dr = \int_T^\infty \exp(-F(r))dr. \quad (6.8)$$

A basic visualization of the EBM approach is given in Figure 6.1.

An important choice is the definition of $p_{\text{ref}}(r)$. When the PDF of $p_{R|\mathbf{Y}}(r|\mathbf{y})$ is supported by a compact interval with length Ω_p , one possible and natural choice is to set $p_{\text{ref}}(r) = 1/\Omega_p$ (Valsson and Parrinello 2014). This results in uniform sampling and is commonly employed in other enhanced sampling approaches (Wang and Landau 2001). If the support is unbounded, then $p_{\text{ref}}(r)$ can be selected such that it shifts mass to the range of R which is of interest (here

$R \geq T$). To enable accurate rare event probability estimation, the right tails of the selected $p_{\text{ref}}(r)$ have to be at least as heavy as the tails of the target PDF $p_{R|Y}(r|\mathbf{y})$.

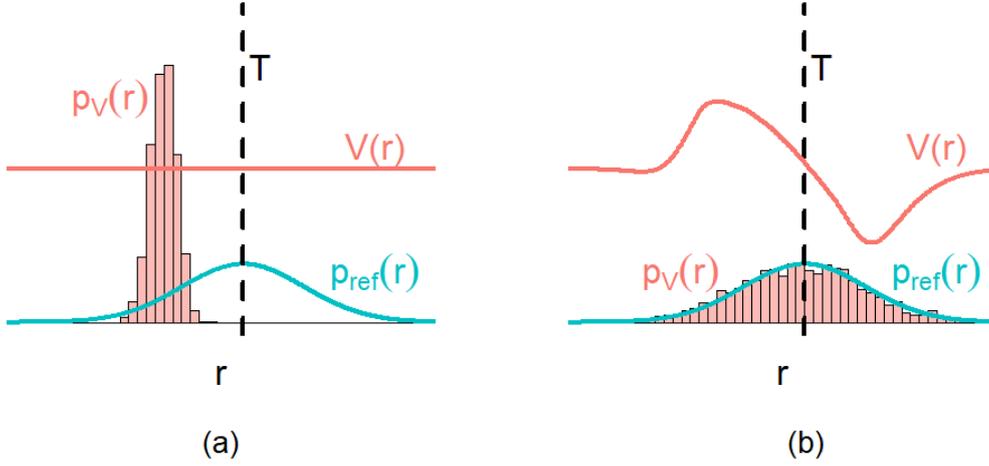


Figure 6.1: Illustration of the energy-based model approach. (a) For $V(r) = 0$, $p_V(r) = p_{R|Y}(r|\mathbf{y})$ has low probability mass on the set $\{R \geq T\}$. (b) For $V(r) \approx V_{\text{opt}}(r)$, $p_V(r) \approx p_{\text{ref}}(r)$ yielding by design a high probability mass on the set $\{R \geq T\}$.

Parameterization and optimization of $V(r)$

We parameterize the bias potential $V(r)$ using a flexible model $V_\psi(r)$ with free parameters ψ . We employ here radial basis functions (RBFs) with squared exponential kernels,

$$V_\psi(r) = \sum_{j=1}^B w_j \phi(r - b_j), \text{ where } \phi(r) = \exp(-(\kappa_j r)^2). \quad (6.9)$$

For a number B of RBFs, we obtain $\psi = (\mathbf{w}, \mathbf{b}, \kappa) \in \mathbb{R}^B \times \mathbb{R}^B \times \mathbb{R}^B$. Then, we seek the optimal $V_\psi(r)$ by minimizing the Kullback–Leibler divergence between $p_{\text{ref}}(r)$ and $p_{V_\psi}(r)$ with respect to ψ . We achieve this by employing stochastic gradient descent with momentum (SGDM; e.g., Liu et al. 2020),

$$m_n = \beta m_{n-1} + (1 - \beta) \frac{\partial \text{KL}(p_{\text{ref}} \| p_{V_\psi})}{\partial \psi}, \quad (6.10)$$

$$\psi_{n+1} = \psi_n - \gamma m_n,$$

using initial momentum $m_0 = 0$, momentum weight β , learning rate γ and a stochastic gradient of $\text{KL}(p_{\text{ref}} \| p_{V_\psi})$ at ψ_n . For a momentum weight of $\beta = 0$, SDGM reduces to traditional stochastic gradient descent.

It holds that

$$\begin{aligned} \text{KL}(p_{\text{ref}}||p_{V_\psi}) &= \int p_{\text{ref}}(r) \log p_{\text{ref}}(r) dr - \int p_{\text{ref}}(r) \log p_{V_\psi}(r) dr \\ &= \text{const.} + \int p_{\text{ref}}(r) V_\psi(r) dr + \log \int \exp(-(F(s) + V_\psi(s))) ds, \end{aligned} \quad (6.11)$$

and (assuming standard regularity assumptions that allow swapping the order of differentiation and integration),

$$\begin{aligned} \frac{\partial \text{KL}(p_{\text{ref}}||p_{V_\psi})}{\partial \psi} &= \frac{\partial}{\partial \psi} \int p_{\text{ref}}(r) V_\psi(r) dr + \frac{\partial}{\partial \psi} \log \int \exp(-(F(s) + V_\psi(s))) ds \\ &= \int p_{\text{ref}}(r) \frac{\partial}{\partial \psi} V_\psi(r) dr - \int p_{V_\psi}(s) \frac{\partial}{\partial \psi} V_\psi(s) ds. \end{aligned} \quad (6.12)$$

This allows us to approximate,

$$\frac{\partial \text{KL}(p_{\text{ref}}||p_{V_\psi})}{\partial \psi} \approx \frac{1}{n} \sum_{i=1}^n \frac{\partial}{\partial \psi} V_\psi(r_i) - \frac{1}{n} \sum_{i=1}^n \frac{\partial}{\partial \psi} V_\psi(s_i), \quad \text{with } r_i \sim p_{\text{ref}}(\cdot), s_i \sim p_{V_\psi}(\cdot). \quad (6.13)$$

This requires samples from $p_{V_\psi}(\cdot)$ and $p_{\text{ref}}(\cdot)$, with the latter being straightforward as the PDF $p_{\text{ref}}(\cdot)$ is pre-defined. Sampling from $p_{V_\psi}(\cdot)$ directly is not possible as it would require that $F(r)$ was known. Thus, we use samples of the original space of θ by introducing,

$$p_{V_\psi}(\theta) = \frac{\exp(-(U(\theta) + V_\psi(\mathcal{R}(\theta))))}{\int \exp(-(U(\zeta) + V_\psi(\mathcal{R}(\zeta)))) d\zeta}, \quad (6.14)$$

and then transform them using $\mathcal{R}(\cdot)$. This is correct as one can easily check that the following identity $\int p_{V_\psi}(\theta) \delta(r - \mathcal{R}(\theta)) d\theta = p_{V_\psi}(r)$ holds. In our implementation, the samples of $p_{V_\psi}(\theta)$ are obtained using a Metropolis–Hastings algorithm (MH; Metropolis et al. 1953; Hastings 1970).

This idea of optimizing the bias potential to estimate the free energy coincides with the variational approach of Valsson and Parrinello (2014), Invernizzi et al. (2017) and Bonati et al. (2019) introduced in metadynamics, working with the functional,

$$\Omega(V) = \log \left(\frac{\int \exp(-(F(r) + V(r))) dr}{\int \exp(-F(r)) dr} \right) + \int p_{\text{ref}}(r) V(r) dr, \quad (6.15)$$

which they show is convex and invariant under the addition of an arbitrary constant to $V(r)$. It is easy to show that minimizing $\Omega(V)$ is equivalent to minimizing $\text{KL}(p_{\text{ref}}||p_V)$ (e.g. Invernizzi et al. 2017). Furthermore, we can reformulate the minimization of the Kullback–Leibler divergence as a maximum likelihood estimation problem (Appendix 6.6.1), which offers multiple theoretical possibilities as it allows us to transfer known theoretical results on maximum likelihood estimation to our approach.

Non-parametric bias potential $V(r)$

In addition to the parametric bias potentials V_ψ considered so far, the functional $\Omega(V)$ of Eq. (6.15) can be applied to a broader class of bias potentials V . While Invernizzi et al. 2017 is not really giving details regarding the source space on which $\Omega(V)$ is defined, it is noticeable that for V integrable with respect to p_{ref} , $\Omega(V)$ can be seen as a convex functional (taking infinite values in case of non-integrable $\exp(-F - V)$). Furthermore, for suitable V and directions H , one can work out directional derivatives of Ω

$$d\Omega(V; H) = \int p_{\text{ref}}(r)H(r)dr - \int p_V(r)H(r)dr = \int (p_{\text{ref}}(r) - p_V(r))H(r)dr. \quad (6.16)$$

In particular, for bounded V and p_{ref} as considered herein, the directional derivative at V in the direction $p_{\text{ref}} - p_V$ is well defined and $p_V - p_{\text{ref}}$ delivers a descent direction. In sequential settings, the function V_n can be updated with SGDM (Eq. 6.10) using $p_{\text{ref}} - p_V$ in lieu of stochastic gradient at V_n . While for the PDF p_{ref} , we have access to its analytical form, we rely here in practice for p_{V_n} on a kernel density estimate derived from the MH samples.

Stopping criteria

Appropriate stopping criteria are crucial when employing stochastic gradient-based optimization in order to strike a balance between computational efficiency and achieving satisfactory model performance. One naïve approach is to stop the training after a pre-defined number of optimization steps. Since this approach does not ensure convergence if the selected number is too low or might waste computational resources if chosen to be too high, it makes sense to monitor the values of the loss function and terminate the optimization when it reaches a desired criterion or when it no longer improves significantly.

Here we consider a convergence measure based on Stein’s method (Stein 1972): the Stein discrepancy (Gorham and Mackey 2015). To circumvent the latter’s computational intractability, we use a specific version known as the kernel Stein discrepancy (KSD; Liu et al. 2016, Chwialkowski et al. 2016, Gorham and Mackey 2017),

$$KSD(p_{\text{ref}}||p_{V_\psi}) = \sqrt{\frac{1}{n^2} \sum_{i,j=1}^n k_{p_{\text{ref}}}(r_i, r_j)}, \quad r_i \sim p_{V_\psi}(\cdot), \text{ with,} \quad (6.17)$$

$$k_{p_{\text{ref}}}(r, s) = \nabla_r \nabla_s k(r, s) + \langle \nabla_r k(r, s), \nabla_s \log p_{\text{ref}}(s) \rangle + \langle \nabla_s k(r, s), \nabla_r \log p_{\text{ref}}(r) \rangle + k(r, s) \langle \nabla_r \log p_{\text{ref}}(r), \nabla_s \log p_{\text{ref}}(s) \rangle,$$

with kernel $k(r, s)$. The use of the KSD is particularly advantageous in our setting due to the analytical knowledge about the probability distribution $p_{\text{ref}}(r)$.

As stopping criterion, we employ a statistical test for goodness-of-fit based on the squared KSD (Liu et al. 2016, Chwialkowski et al. 2016). Employing bootstrap sampling for the estimation of the test statistic, the algorithm stops as soon as the null hypothesis of $p_{\text{ref}} = p_{V_\psi}$ cannot be rejected with a significance level of α anymore. To counteract the conservative nature of the test procedure when dealing with the correlated samples resulting from MH,

Chwialkowski et al. (2016) suggest an approach based on the wild bootstrap technique (e.g., Shao 2010). The proposed test statistics S takes into account the correlation structure in the samples by mimicking it with an auxiliary Markov chain taking values in $\{-1, 1\}$,

$$S = \frac{1}{n^2} \sum_{i,j=1}^n W_i W_j k_{p_{\text{ref}}}(r_i, r_j), \quad \text{using} \quad W_i = \mathbb{1}(U_i > a_{BS}) W_{i-1} - \mathbb{1}(U_i < a_{BS}) W_{i-1}, \quad (6.18)$$

where $W_1 = 1$ and $U_i \sim \text{Unif}(0, 1)$. Chwialkowski et al. (2016) suggest combining this method with thinning of the generated samples. They recommend to thin the chain such that $\text{Cor}(r_i, r_{i-1}) < 0.5$, set $a_{BS} = 0.1/q$ with $q < 10$ and use at least $n = \max(500q, 100)$ samples.

Related work in machine learning

Energy-based models are a popular class of probabilistic models in machine learning (Goodfellow et al. 2016). In this context, EBMs are used to model the distribution of available data $(r_i)_{i=1}^N$ (say images) distributed according to an unknown PDF $p_{\text{ref}}(r)$ via a model of the form

$$p_{\psi}(r) = \frac{\exp(-U_{\psi}(r))}{Z_{\psi}}, \quad (6.19)$$

the potential $U_{\psi}(r)$ being typically parameterized by a neural network. The parameter ψ can be learned by maximizing the normalized log-likelihood, that is,

$$\ell_N(\psi) = \frac{1}{N} \sum_{i=1}^N \log p_{\psi}(r_i), \quad (6.20)$$

which approximates up to an additive constant independent of ψ the Kullback–Leibler divergence $\text{KL}(p_{\text{ref}}||p_{\psi})$ as $N \rightarrow \infty$ (Appendix 6.6.1). Similarly to our scenario, the gradient of this discrepancy requires sampling from $p_{\psi}(r)$ using Markov chain Monte Carlo.

Compared to applications in machine learning, the rare event simulation context discussed in this paper enjoys several attractive properties. First, we can obtain as many samples from $p_{\text{ref}}(r)$ as we want. Second, because we know $p_{\text{ref}}(r)$ analytically, we can assess how good our estimate of $p_{\text{ref}}(r)$ is via $p_{\psi}(r)$ as discussed in the previous section. Third, we only need to estimate a one-dimensional function $V_{\psi}(r)$ instead of a complex high-dimensional potential $U_{\psi}(r)$. We note finally that numerous alternatives to Kullback–Leibler minimization have been developed to train EBMs and that some of them might be applicable to the rare event simulation context, see Song and Kingma (2021) for a recent review.

6.2.3 Alternative rare event probability estimation methods

Subset sampling

Subset sampling for rare event probability estimation was introduced by Au and Beck (2001) and is based on the realization that the probability of a rare event can be expressed as a multiplication of higher conditional probabilities involving intermediate failure events. To estimate $\mathbb{P}(\boldsymbol{\theta} \in A|\mathbf{y}) = \mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq T|\mathbf{y})$, subset sampling employs a sequence of thresholds $\{T_0, \dots, T_K\}$ with $T_0 = -\infty$ and $T_K = T$ and corresponding nested sets $A_k = \{\boldsymbol{\theta} \in \mathbb{R}^d : \mathcal{R}(\boldsymbol{\theta}) \geq T_k\}$. It holds that,

$$\mathbb{P}(\boldsymbol{\theta} \in A|\mathbf{y}) = \prod_{k=1}^K \mathbb{P}(\boldsymbol{\theta} \in A_k | \boldsymbol{\theta} \in A_{k-1}, \mathbf{y}). \quad (6.21)$$

Au and Beck (2001) propose to select the sequence of thresholds adaptively such that a ratio of $(1 - \gamma)$ of the samples survive. Such an adaptive procedure leads to positively biased estimates, however, this bias decreases with an increasing number of samples per subset (C erou et al. 2012). To use subset sampling in our inversion setting, we initialize the first set of samples with a draw from the posterior using the states sampled by a MH algorithm after burn-in. In this context, the thinning factor is critical as the initial set of samples should be as independent as possible. We then propagate the set of samples to estimate the conditional probabilities by sampling from the subsets while accounting for the posterior. A similar approach was applied by Jensen et al. (2013), who rely on transitional MCMC to generate the first set of posterior samples.

Bayesian updating of rare event probabilities

In the article by Straub and Papaioannou (2015), building upon their previous work presented in Straub (2011), a methodological framework called Bayesian Updating with Structural reliability methods (BUS) is introduced. This method serves as an extension of the conventional rejection sampling technique in Bayesian analysis. In a basic version of rejection sampling, samples are generated from the prior distribution and subsequently accepted with a probability of $p_{Y|\boldsymbol{\theta}}(\mathbf{y}|\boldsymbol{\theta})/\bar{p}$, where \bar{p} is an upper bound on $\sup_{\boldsymbol{\theta}} p_{Y|\boldsymbol{\theta}}(\mathbf{y}|\boldsymbol{\theta})$. The issue with the basic rejection sampling algorithm is its inefficiency, particularly when dealing with high-dimensional or complex posterior distributions. The core concept behind BUS is to address the challenge posed by the small acceptance probabilities in rejection sampling with structural reliability methods.

We consider BUS with subset sampling, employing the principles of subset sampling to sample from the acceptance region of the extended rejection sampling algorithm, while enabling an adaptive estimation of the constant \bar{p} (Betz et al. 2014). In the study of Straub et al. (2016), this method strikes an appropriate balance between accuracy and the number of model evaluations and has the advantage of being more robust compared to line sampling that they also consider. To apply BUS for rare event probability estimation in an inversion setting, Straub et al. (2016) combine two subset sampling runs, where the first generates samples from the posterior and the second propagates these samples towards the targeted

rare event. This approach is comparable to the subset sampling method described in the previous section with the difference that the posterior sampling part is replaced with BUS.

6.3 Illustrative test examples

We now explore the performance of our EBM method in different illustrative test examples. We start with exploring the method's sensitivity to the choice of configuration using a hypothetical contamination problem for which an analytical solution is available. Next, we consider the two-dimensional four-branch function problem, a frequently utilized benchmark in reliability analysis. This example does not require inversion and effectively demonstrates the method's suitability for more traditional rare event estimation scenarios. For both of these first examples, we compare the EBM method against a standard implementation of subset sampling. Finally, we consider a simple load and capacity example of Straub et al. (2016) to compare the EBM method's performance against the BUS approach with subset sampling.

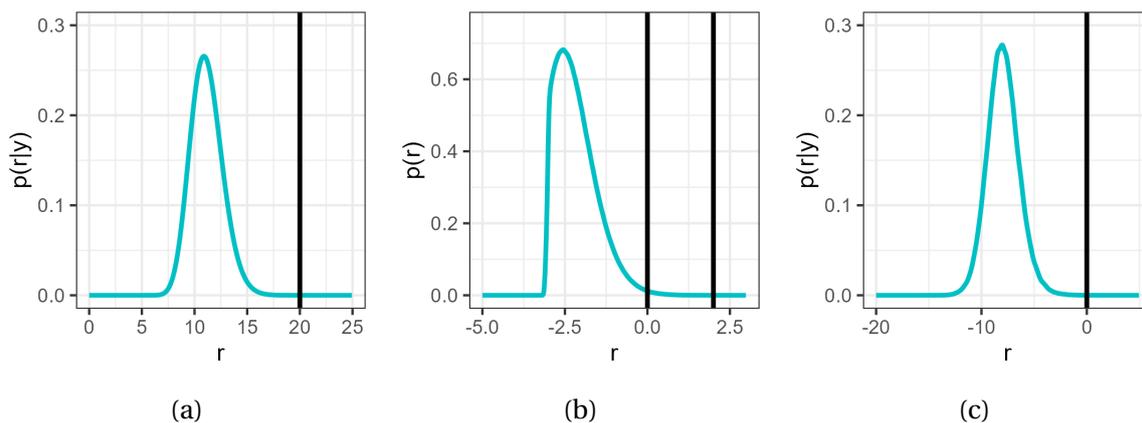


Figure 6.2: PDFs of the quantity of interest R in the three illustrative examples: (a) The dose response in Section 6.3.1, (b) the four-branch function in Section 6.3.2 and (c) the difference between load and capacity in Section 6.3.3 ($n_C = 10$). The vertical lines show the different critical thresholds T targeted within $\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq T | \mathbf{y})$.

6.3.1 Analytical contamination example

In our first test example, we seek to infer a hypothetical contamination field $\boldsymbol{\theta}$ using cell measurements \mathbf{y} and aim at predicting the probability of a resulting dose response $\mathcal{R}(\boldsymbol{\theta})$ being critically high for organisms living in the soil. We consider a setting with $M = 9$ contamination cells. The “true” values of the contamination level are sampled using independent Gaussian PDFs $\theta_m \sim \mathcal{N}(1, 0.3^2)$. For the data \mathbf{y} , we measure the concentration of the contamination in three cells and assume independent Gaussian observation errors with a standard deviation of 0.05. We target the quantity of interest $\mathcal{R}(\boldsymbol{\theta}) = \sum_{m=1}^M \theta_m^2$, which is related to the linear quadratic dose response (e.g., McMahan 2018). We assume that a dose response that is greater or equal than 20 is critical for the organisms in the soil and aim to estimate $\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq 20.0 | \mathbf{y})$. As the prior and the likelihood are Gaussian and the measurements depend linearly on the property field $\boldsymbol{\theta}$, there exists an analytical solution for the Gaussian posterior $p_{\boldsymbol{\theta} | \mathbf{Y}}(\boldsymbol{\theta} | \mathbf{y})$. The PDF $p_{\mathcal{R} | \mathbf{Y}}(r | \mathbf{y})$ of the quantity of interest follows a generalized chi-square distribution (Fig. 6.2a). Even if this distribution does not have a simple closed-form expression, we can derive numerically the true probability of the considered rare event, which is $\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq 20.0 | \mathbf{y}) = 1.76 \times 10^{-6}$.

We parameterize the bias potential $V_\psi(r)$ using RBFs with squared exponential kernels (Eq. 6.9). We assume a constant shape parameter κ , use evenly distributed centers b_j from L_B to H_B and only optimize the weights w_j during the optimization of $V_\psi(r)$. We choose $L_B = -80$ and $H_B = 120$ to ensure that the optimization of the bias potential is not limited by a too narrow range. After testing different numbers of RBFs B and shape parameters κ , we choose a parameterization with $B = 500$ and $\kappa = 1$. We also explore our non-parametric approach for modeling the bias potential, which involves employing a kernel density estimate to approximate the probability density function $p_V(r)$. To perform this estimation, we utilize the R function *density* (R Core Team 2021) with the ‘nrd’ bandwidth configuration (Scott 2015). In practice, for numerical evaluations, we discretize the function by considering equidistant points within the range of -80 to 120, with a spacing of 0.1.

We introduce a base scenario for the algorithmic setting. For $p_{\text{ref}}(r)$, we use a Gaussian PDF $\mathcal{N}(20, 7^2)$ centered on the critical threshold. We iterate for a maximum of 500 SGDM optimization steps employing a constant learning rate γ of 1.2 and a momentum weight β of 0.5 for the RBF parameterization (Eq. 6.10). For the non-parametric form of the bias potential, we use the same momentum weight and a constant learning rate of 15. For the sampling of $s_i \sim p_{V_\psi}(\cdot)$ (Eq. 6.13), we use MH steps employing Gaussian proposals with a step width chosen by initial testing to obtain an acceptance rate close to 30 %. To account for the fact that the posterior is much better defined for the cells where measurements have been made, we adjust the step size for these specific cells. We utilize the KSD stopping criterion with a significance level of $\alpha = 95\%$ and a bootstrap parameter $a_{BS} = 0.4$. Following the recommendations of Chwialkowski et al. (2016), we use $n = 125$ final samples with a thinning factor of $th = 10$ and a burn-in of 100 steps.

First, we examine various choices of $p_{\text{ref}}(r)$ based on Gaussian distributions $\mathcal{N}(20, \sigma^2)$, where the standard deviation σ ranges from 5 to 8. Figure 6.3a depicts the range of rare event prob-

ability estimates obtained from 50 runs using the different choices of $p_{\text{ref}}(r)$ and both the RBF and the non-parametric estimation of the bias potential. For both forms, the stability and accuracy of the method becomes apparent from a standard deviation of 7 and above. The black points in Figure 6.3a depict the mean probability values of the 50 runs. We observe a positive bias that diminishes as the standard deviation of $p_{\text{ref}}(r)$ increases. This bias in the estimate is more pronounced for the RBF bias potential. This is a consequence of initializing $V(r)$ as a constant function. During the optimization, the weights of the RBF $V(r)$ are only updated up to a value where samples of $p_{\text{ref}}(r)$ are located. For r larger than a high quantile of $p_{\text{ref}}(r)$, the resulting optimized bias potential is biased upwards. Accordingly, $F(r) = -\log(p_{\text{ref}}(r)) - V_{\text{opt}}(r)$ is too small in that region. This results in $p_{R|Y}(r|\mathbf{y})$ and $\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq T|\mathbf{y})$ being too large. This effect is less pronounced with the non-parametric bias potential, as the stochastic gradient is based on the analytical PDF $p_{\text{ref}}(r)$. Figure 6.3b illustrates the minimum and maximum number of forward simulations used in the 50 runs. Despite the stability of the probability estimate's uncertainty from a standard deviation of 7, the number of required optimization steps continues to rise as the standard deviation increases. The RBF parameterization demands a greater number of simulations compared to the non-parametric approach and yields estimates with a lower variance.

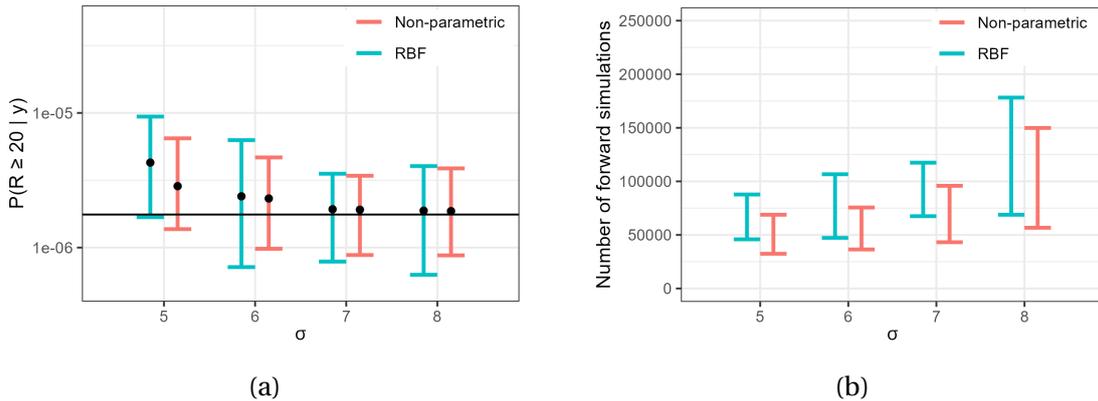


Figure 6.3: Analytical contamination example. (a) Range of rare event probability estimates obtained from 50 EBM runs using different standard deviations in defining $p_{\text{ref}}(r)$, as well as parametric and non-parametric forms of the bias potential, (b) the corresponding minimum and maximum number of forward simulations. The horizontal line in (a) indicates the true rare event probability and the dots the mean of the 50 estimates.

We now focus on the optimization scheme while maintaining $p_{\text{ref}}(r) = \mathcal{N}(20, 7^2)$ and the non-parametric form of the bias potential. Figures 6.4a - 6.4d show exemplary trajectories of the Kullback–Leibler divergence, KSD values and rare event probability estimates for different learning rates γ and momentum weights β . We perform 200 optimization steps without any stopping criterion and illustrate the step at which the KSD criteria would have halted the optimization process (indicated by vertical lines). We emphasize that while the Kullback–Leibler divergence exhibits a great amount of scatter, the KSD values demonstrate a much clearer

convergence (Figs. 6.4a and 6.4b). The mean, root-mean-square-error (RMSE) and coefficient of variation (COV) of the estimates using a learning rate $\gamma = 15$ and momentum weight $\beta = 0.5$ are summarized in Table 6.1. It is observed that whether a stopping criterion is applied or not, the accuracy of the estimate is similar. However, the computational budget can be reduced by almost a factor of 4 when considering 200 fixed optimization steps. Increasing the learning rate comes with a trade-off: while it reduces the computational requirement by reducing the number of optimization steps needed, there is a corresponding loss in accuracy (Fig. 6.4c). The momentum weight plays a crucial role in mitigating the impact of stochastic gradient fluctuations. However, when paired with a too high learning rate, there is a risk of overshooting, necessitating careful adjustment of the learning rate for optimal performance (Fig. 6.4d, further discussed in Section 6.3.3).

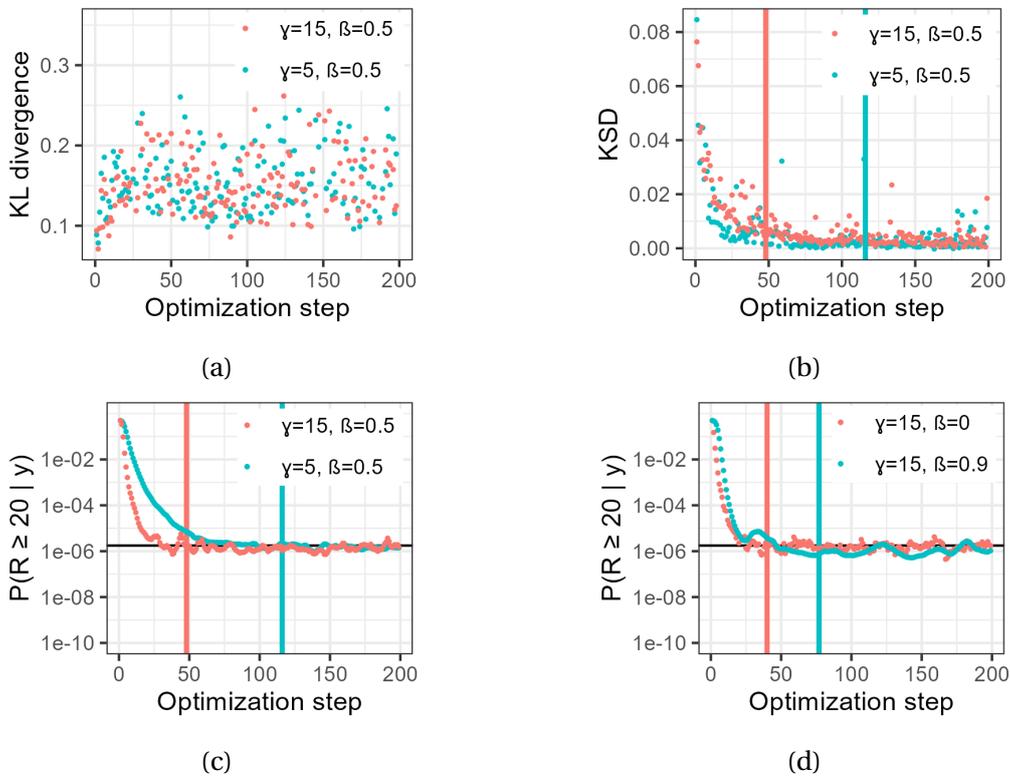


Figure 6.4: Analytical contamination example. Exemplary trajectories for the evolving (a) Kullback–Leibler divergence, (b) KSD and (c, d) rare event probability estimate using different learning rates γ and momentum weights β . The horizontal lines in (c, d) indicate the true rare event probability and the vertical lines in (b, c, d) the step at which the KSD criteria would have halted the optimization process.

We compare the results of the EBM method with subset sampling. To perform subset sampling, we establish a fixed burn-in of 100 steps during the initial posterior estimation process. Additionally, we utilize a thinning factor of 500 for the MH samples to promote an initial sample representation that is free from correlation. To prevent bias from an adaptive sequence of

thresholds and avoid re-running the algorithm, we use a fixed sequence of thresholds given by a logarithmic function increasing from 5.0 to T within nr_t steps. Then we assume the same total computational budget as for the EBM method (including the posterior sampling) and choose the number of samples per subset such as the number of MH steps per subset based on the lowest RMSE using 50 test runs. Here, we choose 60 samples per subset that are each propagated using 20 MH steps. The statistics of the resulting rare event probability estimates are summarized in Table 6.1. Both the lowest RMSE and COV values obtained with subset sampling are about 60 % higher than the ones of our EBM approach.

Table 6.1: Comparison of the EBM method and subset sampling for the analytical contamination example. Mean refers to the mean of the 50 estimates of $\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq T|\mathbf{y})$, RSME to their root-mean-square-error and COV to the coefficient of variation. The budget noted for EBM is equal to the mean budget.

Method	Budget	Mean	RMSE	COV
EBM (Stopping criterion)	72 k	1.74×10^{-6}	0.62×10^{-6}	0.36
EBM (200 iterations)	270 k	1.79×10^{-6}	0.62×10^{-6}	0.35
Subset sampling	72 k	1.73×10^{-6}	1.00×10^{-6}	0.58

6.3.2 Four-branch function

The four-branch function represents a widely used benchmark in structural reliability analysis, characterizing the failure of a system comprising four distinct component limit states (e.g., Schöbi et al. 2017). It is used here to demonstrate the applicability of our EBM method outside of an inversion setting. The four-branch function is defined as,

$$\mathcal{R}(\boldsymbol{\theta}) = \min \left\{ \begin{array}{l} 3 + 0.1(\theta_1 - \theta_2)^2 - \frac{\theta_1 + \theta_2}{\sqrt{2}} \\ 3 + 0.1(\theta_1 - \theta_2)^2 + \frac{\theta_1 + \theta_2}{\sqrt{2}} \\ (\theta_1 - \theta_2) + \frac{6}{\sqrt{2}} \\ (\theta_2 - \theta_1) + \frac{6}{\sqrt{2}} \end{array} \right\}, \quad (6.22)$$

with input variables $\boldsymbol{\theta} = (\theta_1, \theta_2)$, which are modeled by two independent standard normally-distributed variables. The traditional rare probability of interest is $\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \leq T^*) = \mathbb{P}(-\mathcal{R}(\boldsymbol{\theta}) \geq T^*)$, with $T^* = 0$. Given that with probability 4.46×10^{-3} (according to a MC estimation with 10^8 samples in Schöbi et al. 2017) this is not a particularly rare event, we also direct our focus towards $\mathbb{P}(-\mathcal{R}(\boldsymbol{\theta}) \geq T^{**})$ with $T^{**} = 2$. The distribution of the quantity of interest $R = -\mathcal{R}(\boldsymbol{\theta})$ is depicted in Figure 6.2b and the limit states $\boldsymbol{\theta}$ with $\mathcal{R}(\boldsymbol{\theta}) = T^*$ are depicted in Figure 6.5b-d as black outlines.

We use the non-parametric form of the bias potential with the same Gaussian kernel density estimate as in the previous section. As evaluation points, we employ equally-spaced intervals between -10 and 100 with a spacing of 0.1. Our objective is to investigate the performance of the EBM method under a restricted computational budget of 10,000 evaluations. To perform the sampling of $s_i \sim p_{V_\psi}(\cdot)$, we use MH steps with Gaussian proposals aiming for an acceptance rate close to 30%. We test a configuration using a stopping criterion employing a significance level of $\alpha = 0.99$ and $a_{BS} = 0.5$. With an appropriate thinning factor (4) and burn-in period (5 steps), this results in 100 samples obtained during 405 MH iterations. We run the EBM process with a constant learning rate of $\gamma = 6.5$ and momentum weight $\beta = 0.5$.

The optimized bias potential $V_{\text{opt}}(r)$ takes on different forms depending on the chosen PDF $p_{\text{ref}}(r)$, impacting not only $p_{V_{\text{opt}}}(r)$, but also $p_{V_{\text{opt}}}(\boldsymbol{\theta})$ (Eq. 6.14), which has to be explored by MH steps to obtain the samples $s_i \sim p_{V_\psi}(\cdot)$. This is particularly interesting in this test example, as the failure event is connected to four distinct regions of the parameter space. Figure 6.5 shows samples of the resulting joint PDFs $p_{V_{\text{opt}}}(\theta_1, \theta_2)$ for different choices of $p_{\text{ref}}(r)$ together with the limit states $\boldsymbol{\theta}$ with $\mathcal{R}(\boldsymbol{\theta}) = T^*$ (black outlines). The corresponding means and COVs for the rare event probability estimates are shown in Table 6.2. We start with a Gaussian distribution centred on the first threshold of interest, $\mathcal{N}(0, 2^2)$. The resulting joint PDF $p_{V_{\text{opt}}}(\theta_1, \theta_2)$ (Fig. 6.5b) appears to be relatively compact and only explores a small part of the area outside the limit states. This leads to a high mean estimate (Table 6.2). Merely increasing the standard deviation means that $p_{\text{ref}}(r)$ also puts high mass on the left tails of the distribution of R , a wasteful allocation of resources.

As an alternative, we consider an asymmetric and heavy-tailed Generalized Extreme Value (GEV) distribution as $p_{\text{ref}}(r)$. We start with $\text{GEV}(2, 2, 0.33)$ using a location parameter of 2, a scale parameter of 2, and a shape parameter of 0.33. This shift results in a PDF $p_{\text{ref}}(r)$ that has minimal mass in the high-probability region of R (Fig. 6.5a). The resulting joint PDF $p_{V_{\text{opt}}}(\theta_1, \theta_2)$ (Fig. 6.5c) displays an increased level of exploration outside the limit states but partly omits the central region that links the four branches. This creates challenges when transitioning between modes during the MH sampling process, leading to biased probability estimates with high COVs (Table 6.2). These findings suggest that a well-chosen $p_{\text{ref}}(r)$ should not only have heavy right tails, but also include the high probability region of R while excluding the left tails of its distribution. Therefore, we change to a type I extreme value distribution $\text{GEV}(2, 3, 0)$ (Fig. 6.5a), which expands the joint PDF $p_{V_{\text{opt}}}(\theta_1, \theta_2)$ while the central region is still incorporated (Fig. 6.5d). Consequently, our results yield a mean probability estimate for T^* that aligns with the one in Schöbi et al. (2017). Moreover, for T^{**} , we observe a significantly reduced mean compared to the alternative $p_{\text{ref}}(r)$, and for both thresholds, the COVs are at their lowest with the latter choice.

To ensure a connected $p_{V_{\text{opt}}}(\theta_1, \theta_2)$ and enhanced exploration of the failure regions, we consider $\text{GEV}(2, 3, 0)$ to be the best choice. In evaluating subset sampling for comparison purposes, we employ the average computational budget of the EBM approach. A satisfactory performance was achieved with 80 samples per subset and a propagation using 5 MH steps. For both thresholds, we obtain comparable mean estimates with both methods, yet higher COVs with subset sampling compared to EBM with $\text{GEV}(2, 3, 0)$. In fact, the COV for the

second threshold T^{**} is about 50 % higher.

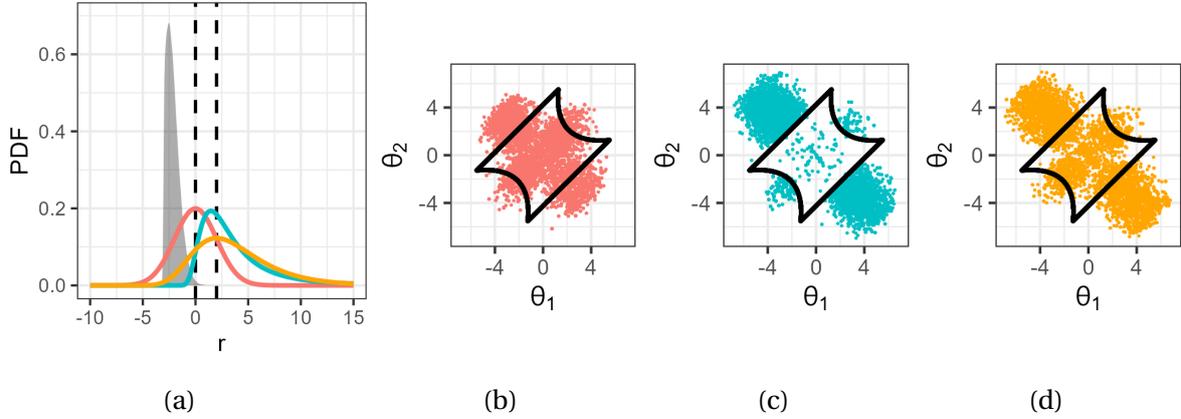


Figure 6.5: Impact of $p_{\text{ref}}(r)$ for the four-branch function example. (a) Distribution of the quantity of interest R (filled) with two considered thresholds (dashed vertical), along with the PDFs $\mathcal{N}(0, 2^2)$ (red), $\text{GEV}(2, 2, 0.33)$ (blue) and $\text{GEV}(2, 3, 0)$ (orange). A total of 10,000 samples of the resulting trained $p_V(\theta_1, \theta_2)$ obtained with (b) $\mathcal{N}(0, 2^2)$, (c) $\text{GEV}(2, 2, 0.33)$, and (d) $\text{GEV}(2, 3, 0)$. The black outlines in (b, c, d) depict the limit states θ with $\mathcal{R}(\theta) = T^*$.

Table 6.2: Comparison of the EBM method and subset sampling for the four-branch example. Mean refers to the mean of the 50 estimates of $\mathbb{P}(-\mathcal{R}(\theta) \geq T)$ for the two thresholds T^* and T^{**} and COV to the coefficient of variation. The budget noted for EBM is equal to the mean budget. The Monte Carlo results for T^* are copied from Schöbi et al. (2017).

Method	$p_{\text{ref}}(r)$	Budget	Mean T^* [$\times 10^{-3}$]	COV T^*	Mean T^{**} [$\times 10^{-5}$]	COV T^{**}
MC	-	10^5 k	4.46	0.0015	-	-
EBM	$\mathcal{N}(0, 2^2)$	5 k	5.79	0.47	45.69	0.63
EBM	$\text{GEV}(2, 2, 0.33)$	12 k	82.45	1.63	31.29	1.09
EBM	$\text{GEV}(2, 3, 0)$	10 k	4.97	0.31	1.56	0.53
Subset sampling	-	10 k	4.91	0.36	1.36	0.80

6.3.3 Load capacity example

As a third test case, we consider the load capacity example of Straub et al. (2016), enabling comparison of our EBM approach with their BUS method. The problem under consideration is a basic reliability problem concerned with the failure of an engineering system when the load θ^L exceeds the capacity θ^C . The corresponding rare event probability of interest is,

$$\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq 0) = \mathbb{P}(\theta^L - \theta^C \geq 0), \quad (6.23)$$

with $\boldsymbol{\theta} = (\theta^L, \theta^C)$. For the load θ^L , we assume a prior Gumbel distribution with mean 2 and standard deviation 1 and the capacity θ^C is lognormal distributed with mean 12 and standard deviation 2; the variables are assumed to be independent. To investigate the performance for a problem with increasing number of model parameters, Straub et al. (2016) express the capacity as a product of component capacities θ_i^C ($1 \leq i \leq n_C$). The component capacities θ_i^C follow a lognormal distribution such that θ^C still has mean 12 and standard deviation 2.

In this inversion setting, individual measurements y_i of the component capacities θ_i^C are considered. To ensure analytical solutions, multiplicative lognormal measurement errors,

$$p_{\mathbf{Y}|\boldsymbol{\theta}}(\mathbf{y}|\boldsymbol{\theta}) = p(\mathbf{y}|\theta_1^C, \theta_2^C, \dots, \theta_{n_C}^C) \propto \exp\left(-\frac{1}{2} \sum_{i=1}^{n_C} \left(\frac{\log y_i - \log \theta_i^C}{\sigma_y}\right)^2\right), \quad (6.24)$$

with $\sigma_y = 0.05$ and $y_i = 8^{1/n_C}$ are employed. Following Straub et al. (2016), we consider $n_C = 10$ and $n_C = 100$ with the resulting failure probabilities $\mathbb{P}(\theta^L - \theta^C \geq 0|\mathbf{y})$ being 6.8×10^{-5} and 2.1×10^{-5} , respectively. The posterior distribution of $R = \mathcal{R}(\boldsymbol{\theta}) = \theta^L - \theta^C$ for $n_C = 10$ with the critical threshold of $T = 0$ is depicted in Figure 6.2c.

The results of the BUS approach with subset sampling are reproduced in Table 6.3. As Straub et al. (2016), we transform the reliability problem to the standard normal space by marginal transformations of the load and capacity variables. To enable a direct comparison of the EBM with the BUS approach, we employ the same computational budgets. In order to accommodate a consistent and constrained computational budget in the context of EBM, we use a fixed number of optimizations steps (no stopping criterion). Following our analysis in the previous section, we employ as $p_{\text{ref}}(r)$ a type I extreme value distribution $\text{GEV}(0, 7, 0)$. To make the method dimensionally robust, we use prior-preserving preconditioned Crank-Nicolson proposals (pCN; e.g. Cotter et al. 2013) within the sampling of $s_i \sim p_{V_\psi}(\cdot)$. Then, we use a burn-in of 10 steps, a thinning factor of $th = 6$ and $a_{BS} = 0.5$. As we do not rely on a stopping criterion, we further divide the number of final samples by a factor of two (Chwialkowski et al. 2016), leading to $n = 50$ obtained by 310 MH steps. Then we fix the number of optimization steps in accordance with the computational budget (resulting in 25 and 27 steps for the two values of n_C).

We test both a parametric and non-parametric form for the bias potential. For the parametric approach, we employ RBFs with squared exponential kernels (Eq. 6.9), using $L_B = -100$ and $H_B = 100$, $B = 500$ and $\kappa = 0.5$. We run the EBM process with an exponentially decreasing

learning rate initiated as $\gamma = 0.5$ (factor -0.005) and momentum weight $\beta = 0.5$. For the non-parametric form, we use the same Gaussian kernel density estimate of $p_V(r)$ as in the previous sections and evaluation points between -100 and 100 with a spacing of 0.1. Here we employ an exponentially decreasing learning rate initiated as $\gamma = 19$ (factor -0.005) and momentum weight $\beta = 0.95$. The means and 95 % confidence intervals of 50 runs are shown in Table 6.3 and the corresponding estimates of $V(r)$ are depicted in Figure 6.6a for the RBFs and in Figure 6.6b for the non-parametric form. With a non-parametric $V(r)$, we sometimes obtain bias potentials characterized by minor-scale fluctuations, in contrast to the smoother bias potentials generated by RBFs, which exhibit more pronounced large-scale fluctuations. The small-scale fluctuations arise due to the constraints of our limited computational resources, compelling us to base the kernel density estimates on a relatively small set of samples, which may not be adequately thinned. To mitigate this issue and promote a smoother estimation, we have increased the momentum weight to $\beta = 0.95$. However, this adjustment also introduces challenges in fine-tuning the learning process, as it tends to cause overshooting during optimization. Due to the changed PDF $p_{R|Y}(r|\mathbf{y})$ and computational budget for $n_C = 100$, we have to adapt the learning rate schedule for an optimal performance and use $\gamma = 22$ and a decay factor of -0.003. On the other hand, the RBF parameterization with lower momentum weight is less sensitive to the choice of the learning rate schedule and we can use the same for both scenarios. For both n_C , both presented configurations of the EBM approach lead to a reduction of the 95 % confidence interval's length compared to the BUS approach. While the non-parametric form, when coupled with a higher momentum weight, yields noticeably increased accuracy, it also becomes significantly more sensitive to the learning rate schedule.

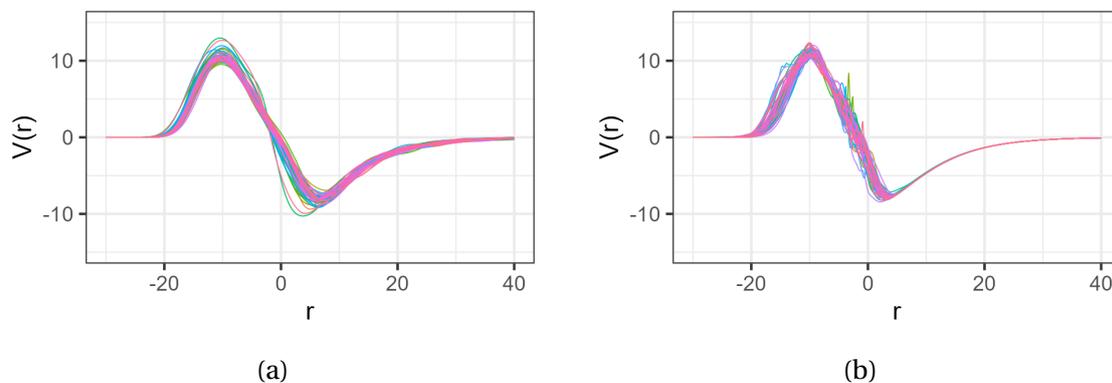


Figure 6.6: Twenty optimized bias potentials $V(r)$ for the load capacity example ($n_C = 100$) using (a) the RBF parameterization and (b) the non-parametric form.

Table 6.3: Table summarizing the comparison of EBM and BUS for the load capacity example. The 95 % confidence intervals of the BUS method with subset sampling are copied from Straub et al. (2016), the mean value is not provided.

Case	Method	Budget	Mean	95 % confidence interval
$n_C = 10$	Analytical	-	6.8×10^{-5}	-
	EBM (non-par.)	7.7 k	7.0×10^{-5}	$[2.7, 12.0] \times 10^{-5}$
	EBM (RBF)	7.7 k	7.1×10^{-5}	$[0.9, 19.5] \times 10^{-5}$
	BUS	7.7 k	-	$[2.1, 35.0] \times 10^{-5}$
$n_C = 100$	Analytical	-	2.1×10^{-5}	-
	EBM (non-par.)	12.6 k	2.7×10^{-5}	$[1.2, 4.7] \times 10^{-5}$
	EBM (RBF)	8.4 k	4.0×10^{-5}	$[0.3, 10.0] \times 10^{-5}$
	BUS	8.4 k	-	$[0.2, 12.0] \times 10^{-5}$

6.4 Discussion

This paper presents a novel energy-based model (EBM) approach for estimating probabilities of rare events that is applicable not only for conventional rare event estimation but also in inversion settings. The approach centers on representing the distribution of the quantity of interest as an energy density function. The estimation of the related free energy is achieved by optimizing a bias potential, aligning it with a predefined distribution $p_{\text{ref}}(r)$ that has most mass on the region of interest.

When parameterizing the bias potential $V_\psi(r)$, it is crucial to allow for an appropriate amount of flexibility. We recommend to compare results obtained by a chosen parameterization with those obtained with a more flexible parameterization. An alternative approach is to use a non-parametric form for the bias potential $V(r)$. Depending on the test case and the computational budget, we illustrate the effectiveness of the non-parametric bias potential in alleviating bias issues and reducing computational needs (Fig. 6.3). However, the non-parametric approach relies heavily on the quality of the kernel density estimate used during the optimization. In cases where computational resources are limited, and when dealing with correlated samples, it is possible to obtain bias potentials that exhibit fine-scale fluctuations resulting from individual optimization steps (Fig. 6.6b). A parameterization, on the other hand, does not face this problem as it can be chosen such that it enforces a smooth bias potential (Fig. 6.6a). However, when opting for a parameterization, we face the critical decision of choosing the appropriate form, which can significantly impact the final outcomes. When manually fine-tuning the bandwidth of the kernel density estimate within the non-parametric approach, we also retain the capability to shape the potential's form, but this comes at the cost of introducing additional configuration choices.

When selecting the distribution $p_{\text{ref}}(r)$, one needs to consider that too light right tails lead to a positive bias in the estimates of $\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \geq T | \mathbf{y})$. This effect is the most pronounced for the parametric approach, as the optimization relies on samples of $p_{\text{ref}}(r)$ and $p_V(r)$. Beyond a certain quantile, the probability of obtaining samples is extremely small, leading to an overestimation of $V(r)$ in this region (under the assumption that the initial potential is constant everywhere). The same principle would hold for $\mathbb{P}(\mathcal{R}(\boldsymbol{\theta}) \leq T | \mathbf{y})$ and the left tails of $p_{\text{ref}}(r)$. Since our primary focus is on failure probability estimates, a slight positive bias in the probability value is not too problematic. Additionally, the non-parametric form reduces this bias as the optimization relies on PDFs enabling an integration of the tails. The choice of $p_{\text{ref}}(r)$ also determines the form of the PDF $p_V(\boldsymbol{\theta})$ (Eq. 6.14), which needs to be explored by MH in each optimization step. In the second test example utilizing the four-branch function (Section 6.3.2), we showcase the benefit of using an asymmetric Generalized Extreme Value (GEV) distribution as $p_{\text{ref}}(r)$ (Fig. 6.5). The form of the GEV distribution enables a $p_{\text{ref}}(r)$ with heavy right tails that includes the high probability region of R . For this example with distinct regions contributing to the rare event, the incorporation of the region of high probability of the quantity of interest within $p_{\text{ref}}(r)$ facilitates the MH algorithm's ability to shift between the modes, leading to a lower coefficient of variation in the probability estimates.

We optimize the bias potential using stochastic gradient descent with momentum. Regarding the learning rate, a smaller value tends to lead to more stable estimates but also to a larger number of optimization steps (Fig. 6.4). Conversely, higher learning rates result in increased fluctuations but require fewer steps for convergence. The momentum has the advantage of dampening oscillations in the convergence process and making it more robust to noisy gradients, however, it requires careful tuning of the learning rate and momentum weight, to achieve optimal performance. Besides stochastic gradient descent with momentum, there exist numerous alternative implementations for the optimization. For instance, for the last test case requiring fine tuning of the optimization schedule (Section 6.3.3), exploring more advanced optimization algorithms like Adam (Kingma and Ba 2015) would be interesting. We combine the optimization with a stopping criterion based on a goodness-of-fit test employing the Kernelized Stein discrepancy. This is done as the Kullback–Leibler divergence used in the optimization of the bias potential exhibits a great amount of scatter compared with KSD that demonstrates a much clearer convergence. This characteristic of the KSD makes it a valuable diagnostic tool in determining the appropriate moment to stop the optimization process. This achieves two key benefits: First, it helps stabilize the rare event probability estimates by ensuring that the biased probability distribution $p_V(r)$ closely approximates the target probability distribution $p_{\text{ref}}(r)$. Second, it reduces computational effort by allowing the optimization process to halt as soon as the approximation becomes sufficiently accurate (Fig. 6.4). To account for the correlation in the samples, we use the specific bootstrap method proposed by Chwialkowski et al. (2016).

We compare the performance of the EBM approach against alternative methods (Section 6.2.3) using three illustrative test examples. In the first contamination test case involving inversion (Section 6.3.1), the EBM method outperforms subset sampling, reducing the root-mean-square-error (RMSE) and coefficient of variation (COV) by about 40 % (Table 6.1). In the second example involving the four-branch function (Section 6.3.2), subset sampling also

exhibits poorer performance than the EBM approach, with the latter reducing the COV by up to 35 % compared to subset sampling. In the third example in Section 6.3.3, we compare the EBM method’s performance against the BUS approach with subset sampling in a simplistic engineering test case from Straub et al. (2016). For both numbers of considered capacity components, the EBM method is able to narrow down the 95 % confidence interval compared to the BUS approach (Table 6.3).

Regarding different approaches for sampling $s_i \sim p_{V_\psi}(\cdot)$, we experiment with various bootstrap parameters a_{BS} (results not shown). We notice that there is an increase in the variance of the probability estimates with growing a_{BS} . However, the reduced values for small bootstrap parameters come at the expense of a higher number of forward simulations. Similar to the learning rate, there seems to be a trade-off between computational cost and accuracy. Besides testing other parameterizations for $V(r)$, distributions for $p_{\text{ref}}(r)$ and optimization schemes for updating $V(r)$, one could also consider replacing the sampler for $p_V(\theta)$. For instance, other proposal schemes, interacting MH chains or a particle method could be considered (e.g., Robert et al. 2018). In data rich settings with narrow posterior distributions, it is expected that $V(r)$ will vary over very large ranges. In such scenarios, it could be beneficial to know in advance the expected properties of $V(r)$ such as shape, smoothness class and range. This could potentially be achieved by using a simplistic free energy function estimate (e.g. histogram based) of some thousands of posterior samples giving together with $p_{\text{ref}}(r)$ to obtain a first approximation of $V_{\text{opt}}(r)$.

6.5 Conclusion

This paper introduces an energy-based model (EBM) approach to estimate rare event probabilities. The approach is based on formulating the distribution of the quantity of interest as an energy density function with a corresponding free energy function. By optimizing a bias potential such that the corresponding energy density approaches a pre-defined PDF $p_{\text{ref}}(r)$, the method estimates the free energy accurately in the region targeted by $p_{\text{ref}}(r)$. The presented approach is applicable both for traditional rare event estimation and in the context of inversion settings when one is not interested in the posterior itself, but rather in the distribution of a quantity that depends on the posterior. When employed in such a setting, this formulation reduces the potentially high-dimensional problem of first estimating the posterior and subsequently the quantity of interest to the optimization of a one-dimensional function. The optimization of the bias potential involves minimizing the Kullback–Leibler divergence and a stopping criterion based on the Kernelized Stein Discrepancy is introduced to terminate the optimization process. The stopping criterion not only enhances the stability and accuracy of the rare event probability estimation but also optimizes computational resources by terminating the optimization process when the approximation is deemed satisfactory. A non-parametric form of the bias potential is introduced, which eliminates the need to make a parameterization choice while simultaneously enabling efficient and accurate probability estimates. For the three presented exemplary test cases, a properly configured EBM approach ensures precise estimations of rare event probabilities and outperforms the examined variants of subset sampling methods.

6.6 Appendix

6.6.1 Relation to maximum likelihood estimation

We can reformulate the minimization of the Kullback-Leibler divergence as a maximum likelihood estimation problem. While we summarize the general concepts and implications of this relation, we refer to Van der Vaart (2000) for the theoretical details about maximum likelihood estimation.

Let R be a random variable, whose distribution P has a PDF $p_{V_\psi}(r)$ with a statistical model $\{p_{V_\psi}(\cdot) : \psi \in \Psi\}$. Furthermore, r_i for $1 \leq i \leq n$ are n independently generated realizations of R . The log-likelihood function of the model is given by,

$$l_n(\psi) = \log \prod_{i=1}^n p_{V_\psi}(r_i) = \sum_{i=1}^n \log p_{V_\psi}(r_i). \quad (6.25)$$

A maximum likelihood estimator (MLE) is defined as $\hat{\psi}_{MLE} \in \Psi$ with $l_n(\hat{\psi}_{MLE}) = \max_{\psi \in \Psi} l_n(\psi)$.

We define,

$$\ell(\psi) := \mathbb{E}_{\psi_0}(\log p_{V_\psi}(R)) = \int \log p_{V_\psi}(r) p_{V_{\psi_0}}(r) dr. \quad (6.26)$$

Under the assumption that the model is well specified with $p_{V_{\psi_0}}(r)$ being the PDF of P and $\mathbb{E}(|\log p_{V_\psi}(R)|) < \infty$, $\psi \mapsto \ell(\psi)$ is maximized at ψ_0 . The normalized log-likelihood $\psi \mapsto \frac{1}{n} l_n(\psi)$ is a sample approximation of $\psi \mapsto \ell(\psi)$, and under some regularity constraints, $\frac{1}{n} l_n(\psi) \xrightarrow{a.s.} \mathbb{E}_{\psi_0}(\log p_{V_\psi}(R))$ almost surely such that the MLE $\hat{\psi}_{MLE}$ is a consistent estimator for ψ_0 . Furthermore, under some additional conditions, the MLE follows asymptotic normality,

$$\sqrt{n}(\hat{\psi}_{MLE} - \psi_0) \xrightarrow{d} \mathcal{N}(0, I(\psi_0)^{-1}), \quad \text{for } n \rightarrow \infty, \quad (6.27)$$

with the Fisher information matrix $I(\psi_0) = \mathbb{E}_{\psi_0} \left(\nabla_{\psi} \log p_{V_\psi}(R) \nabla_{\psi} \log p_{V_\psi}(R)^T \right)$.

In our setting, we assume that $p_{\text{ref}}(r) = p_{V_{\psi_0}}(r)$ for $\psi_0 \in \Psi$. Furthermore, the Kullback-Leibler divergence between two distributions with PDFs $p_{V_\psi}(\cdot)$ and $p_{V_{\psi_0}}(\cdot)$ can be expressed as $\ell(\psi_0) - \ell(\psi)$:

$$\text{KL}(p_{V_{\psi_0}} \| p_{V_\psi}) = \int \log p_{V_{\psi_0}}(r) p_{V_{\psi_0}}(r) dr - \int \log p_{V_\psi}(r) p_{V_{\psi_0}}(r) dr. \quad (6.28)$$

The reformulation $\ell(\psi) = \ell(\psi_0) - \text{KL}(p_{V_{\psi_0}} \| p_{V_\psi})$ shows that maximizing the likelihood is equal to minimizing the Kullback-Leibler divergence.

Chapter 7

Conclusions and outlook

7.1 Conclusions

This thesis presents a compilation of methodological advancements leveraging traditional probabilistic Bayesian inversion approaches. Addressing geoscientific applications, their value for accurate posterior estimation and reliable uncertainty quantification is demonstrated. First, this thesis delves into the practical application of statistical estimation and approximation approaches within the context of scenarios involving intractable likelihood functions and latent variable models. Second, it introduces advanced statistical techniques that integrate inversion and risk assessment, providing specialized approaches for scenarios where the emphasis lies not on the posterior distribution, but rather on the probability of a rare event of interest. The presented methodological advancements improve the computational efficiency and realism of probabilistic Bayesian inversion approaches for targeted estimation of properties and attributes of environmental systems.

Chapters 2, 3, and 4 of this thesis center around inversion problems concerned with intractable likelihood functions, arising from latent variables, which obscure the relationship between the target model parameters and the observed measurements. Our findings demonstrate that by accurately estimating the intractable likelihood function, we can directly obtain reliable posterior estimates of the target parameters. In the context of geophysical inversion, we leverage the original formulation of lithological tomography by employing the correlated pseudo-marginal (CPM) method (Chapter 2). This approach estimates the intractable likelihood function by employing Monte Carlo averaging over samples of the latent variable. By doing so, it can directly target the (hydro)geological parameters of interest while accounting for the uncertainty in the petrophysical relationship. Through combination of the CPM method with importance sampling and prior-sampling MCMC proposals, we effectively address the limitations of the original lithological tomography formulation when confronted with high-dimensional problems and large datasets featuring high signal-to-noise ratios. Moreover, the incorporation of prior-sampling proposals enhances exploration capabilities in the context of high-dimensional problems. Subsequently, in Chapter 3, we compare the CPM method against a linearized Gaussian approximation (LinGau) method. We find that the LinGau method provides a computationally efficient approximation, which works well when the scatter of petrophysical prediction uncertainty is small compared to the observational noise. Nevertheless, the computationally more demanding CPM method offers an accurate

and more general approach that outperforms the LinGau method in scenarios with high petrophysical uncertainty. However, for the CPM method to perform effectively in practical applications, it requires an efficient importance sampling distribution. In Chapter 4, we employ the CPM method to infer geostatistical hyperparameters of the model domain from hydrogeological or geophysical data. By avoiding the bias and overconfidence that often arise from simplified model assumptions, the CPM method enables precise posterior estimates of the hyperparameters. Additionally, it eliminates the need to infer a large number of parameters related to the local properties of the model domain.

In Chapters 5 and 6, we demonstrate that combining inversion and reliability methods enables stable rare event probability estimates even in scenarios for which the probability of occurrence is less than one in a million. This demonstrates the benefit of specialized methods against traditional Monte Carlo estimation based on posterior sampling approaches. In Chapter 5, we introduce a two-stage application of the Sequential Monte Carlo method within an inversion framework to assess groundwater hazards. Through flow and transport examples, we aim to demonstrate the method's ability to accurately estimate the occurrence of extremely rare hazards. In Chapter 6, we present a novel approach based on energy-based models (EBM) to estimate probabilities of rare events. In inversion settings, this approach reduces the potentially high-dimensional problem of estimating the posterior and subsequently the quantity of interest to the problem of optimizing a one-dimensional function. We demonstrate the accuracy and effectiveness of the EBM method through test cases sourced from the geosciences and engineering literature. The results showcase the method's favorable performance compared to state-of-the-art reliability methods in traditional rare event estimation but also in inversion settings.

7.2 Outlook

The findings of this thesis give rise to various potential directions for future research. We categorize these points into three overarching areas: General Bayesian inference, likelihood estimation and rare event estimation.

7.2.1 Bayesian inference

It is important to emphasize that this thesis has only considered the analysis of synthetic data. While synthetic data provide a controlled environment for methodological development and evaluation, it is essential to extend these investigations to field demonstrations in order to assess the practical applicability and performance of the proposed methods under real-world conditions. Furthermore, it is important to acknowledge that the entire body of our research presented in Chapters 2, 3 and 4 is built upon Gaussian assumptions for the target field, latent variables and observational noise. These assumptions have provided a useful framework for developing and analyzing the proposed methods. Future investigations could explore the broader applicability of the proposed methods in scenarios including non-Gaussian behavior.

In studies employing probabilistic inversion methods, there is not only a lack of consistent utilization of convergence diagnostics but also a notable absence of a universally applicable tool for assessing posterior estimates. While we provide a brief discussion on certain aspects related to this thesis, it is important to note that these are vast areas of research that offer numerous opportunities for further investigations. Convergence diagnostics play a crucial role in determining whether the posterior sampling approach has achieved a stationary distribution and if the obtained samples effectively represent the target distribution. In Chapter 2, we demonstrate that the widely used \hat{R} -statistics (Gelman and Rubin 1992) is an insufficient metrics to assess the convergence of an adaptive MCMC algorithm. Despite the availability of upgraded versions of the original \hat{R} -statistics (Vats and Knudson 2021), the limitation of this diagnostic to detect lack of convergence when the chains fail to explore regions of critical mass remains. Nevertheless, it is probably the most commonly utilized convergence diagnostics in probabilistic inversion studies.

Given that only synthetic data is utilized in this thesis, direct comparison between the posterior estimates and the true parameter values is always possible. We employ scoring rules (Gneiting and Raftery 2007) for this purpose, which assess both the statistical consistency between predictions and observations and the sharpness of the prediction. However, determining the most appropriate rule to emphasize a given objective in a geoscientific context remains an area for further investigation. In non-synthetic settings, where a true parameter value is unavailable for comparison, the posterior predictive distribution can be compared to the measurements (Krüger et al. 2021). The optimal way to achieve this is by reserving certain data for assessment purposes. Especially in the context of rare event estimation, exploring the relationship between the quality of the posterior predictive distribution and the accuracy of rare event probability estimates offers a fascinating avenue of investigation.

7.2.2 Likelihood estimation

Likelihood estimation is a crucial aspect when performing inversion in latent variable models, with applications in various contexts beyond those discussed in this thesis:

- **Petrophysical uncertainty**

Chapters 2 and 3 focus on accounting for the scatter in petrophysical relationships while assuming that the petrophysical model and parameters are known. It is important to note that uncertainty exists at these levels as well. Furthermore, our studies exclusively consider petrophysical uncertainty in latent variable models where the data are given by GPR first-arrival travel times. It is worth mentioning that more complex, non-linear problems such as electrical resistivity tomography or surface-based seismic refraction tomography may exhibit an even greater sensitivity to petrophysical uncertainty.

- **Curse of dimensionality**

In this thesis, we work in the full parameter space whenever possible such that any distortions in the posterior estimations due to model reductions can be avoided. However, the curse of dimensionality poses a general challenge for MCMC algorithms, as the convergence time increases with the number of target parameters (e.g., Robert et al. 2018). To address this issue while considering the full parameter space, one could infer some main features and account for the remaining components in the likelihood estimation (for instance with the CPM method).

- **Auxiliary parameters**

In most geoscientific inversion problems, the primary emphasis is on a specific property of interest. Nevertheless, the observed data can also be influenced by other factors (e.g., variations in porosity in tracer-test tomography). In such cases, the uncertainty of the auxiliary parameters can be accounted for by employing latent variables.

- **3D-effects**

This thesis exclusively focuses on model domains that are limited to two spatial dimensions. In real-world applications, the data are influenced by heterogeneities that exist beyond the two-dimensional plane of measurements (e.g., between boreholes) or by outer-space effects (Maurer and Friedel 2006). To further improve the estimation and uncertainty quantification in such setups, latent variables could be used to represent the heterogeneities in the out-of-plane dimension.

- **Multiple-point statistics with geophysical constraints**

Multiple-point statistics (MPS) methods use sequential simulations to generate geostatistical realizations that preserve the complex spatial patterns and relationships observed in so-called training images (Guardiano and Srivastava 1993). To effectively incorporate geophysical constraints into MPS methods, it is necessary to evaluate the likelihood of proposed candidate values. Thereby, the not yet informed model parameters can be considered as latent variables. This falls in the scope of the current work of Shiran Levy, where I am involved as co-author on one paper. Shiran Levy, together with Macarena Amaya, are the two PhD students in the same SNSF project under the guidance of Niklas Linde.

The CPM method provides a way to handle latent variable models by estimating the likelihood function. The effectiveness of the CPM method relies heavily on the variance of the log-likelihood ratio estimator. Particularly in scenarios with a large number of observations and a high signal-to-noise ratio, a well-working importance density is required for efficient sampling of the latent variables. While an inadequate importance sampling strategy can decrease efficiency, the posterior samples remain unbiased as the importance sampling is solely employed to reduce the variance of the likelihood estimator. A significant area for future research would involve the development and evaluation of importance sampling schemes that do not depend on Gaussian assumptions. Promising starting points for such research could include the exploration of efficient importance sampling techniques proposed by Richard and Zhang (2007), as well as the application of multiple importance sampling methods introduced by Veach and Guibas (1995) and popularised by Owen and Zhou (2000). In situations where it is challenging to determine a suitable importance sampling distribution, the applicability of the CPM method can be compromised. One potential solution in such cases is to infer additional key characteristics of the latent field alongside the target parameters and subsequently employ the CPM method to sample the remaining (smaller) variability associated with the latent variables.

In addition to CPM, we also employ the LinGau method, which utilizes a Gaussian approximation and local linearization for the likelihood estimation. The accuracy of the LinGau method relies on the quality of the first-order Taylor expansion used to approximate the forward solver, and its effectiveness diminishes as the variance in the latent variable increases. To address the approximation error inherent in the LinGau method, one potential approach is to treat it as a model error term (see e.g. Hansen et al. 2014).

7.2.3 Rare event estimation

Within this thesis, our focus lies in addressing rare event estimation in the presence of uncertain model parameters. This pertains not only to traditional rare event estimation scenarios but also extends to inversion frameworks. When it comes to evaluating risks and making well-informed decisions, a broader range of uncertainties may impact the outcome. It is important to recognize the potential influence of external factors on the system under study, such as climate change or natural disasters. With increasing diversity in the sources of uncertainty, optimizing the strategy for data acquisition becomes crucial. This optimization could involve integrating rare event estimation with joint inversion techniques or designing measurement collection in a manner that maximizes the information pertaining to rare event probability estimates.

In Chapter 5, we employ a sequential combination of Sequential Monte Carlo methods (PostRisk-SMC) for rare event estimation in inversion settings. In the first phase of obtaining a particle approximation of the posterior, we employ the adaptive Sequential Monte Carlo (SMC) method proposed by Zhou et al. (2016). This method has already been tested in geoscientific contexts by Amaya et al. (2021) and Amaya et al. (2021), where, in contrast

to state-of-the-art adaptive Markov Chain Monte Carlo methods, it proved to be effective for handling multi-modal distributions, well-suited for parallel computation, and relatively insensitive to the choice of model proposal schemes. Despite these benefits, the application of SMC methods in geoscientific inversion remains somewhat limited (Linde et al. 2017). In future research, one potential avenue is to explore its application in replacing Metropolis–Hastings in latent variable model scenarios, as examined in the initial part of this thesis. In the second phase, the PostRisk-SMC method relies on subset sampling to estimate the rare event probabilities. In this context, the selection of the intermediate thresholds is a critical aspect, requiring a trade-off between the number of conditional probabilities and the number of particles (Au and Beck 2001). To avoid the bias resulting from an adaptive sequence of thresholds or additional cost from re-running the algorithm with the previously optimized threshold schedule, we employ a fixed sequence of thresholds. This choice presents the possibility that the particle system may "die" at some stage during the rare event probability estimation process if no particles exceed the current threshold. One possible approach to address this issue is discussed by LeGland and Oudjane (2006) and entails generating additional particles until a specified particle count meets the given threshold. Moreover, the frequent resampling in the second phase of PostRisk-SMC necessitates to invest in well-working particle propagation with Metropolis–Hastings to ensure that a sufficient number of steps can prevent particle collapse. Finally, it is worth noting that the two phases of the PostRisk-SMC method exhibit different dynamics. It is conceivable to explore different configurations for the two phases, such as the number of particles and Metropolis-Hastings steps, to potentially optimize the method’s performance. In the two-dimensional transport example (Section 5.4), we initiate the process of addressing these differences by employing varying numbers of Metropolis-Hastings steps for each phase of the method.

The new energy-based model approach to rare event probability estimation (Chapter 6) relies heavily on the careful selection of the form of the bias potential, as well as the choice of the PDF $p_{\text{ref}}(r)$ determining the specific region that is to be explored. By assuming that the chosen parameterization exhibits sufficient flexibility, we can employ the bias potential to directly estimate the free energy associated with the posterior of the quantity of interest. In the context of this thesis, the parameterization is based on either Gaussian radial basis functions, with fixed centers and shapes, or a non-parametric form eliminating the need to make a parameterization choice. A comprehensive exploration of both the presented parameterization approaches and alternative methods could be considered for future investigation. In the process of selecting the probability distribution $p_{\text{ref}}(r)$, one has to recognize the impact of the distribution’s support and tail behavior on the rare event probability estimation. We optimize the bias potential using stochastic gradient descent with momentum combined with a stopping criteria based on a goodness-of-fit test employing the Kernelized Stein discrepancy. Besides the stochastic gradient descent with momentum, there exist more advanced optimization algorithms such as Adam (Kingma and Ba 2015). When considering future applications, it would be interesting to explore the potential benefits of coupling the energy-based model approach with more sophisticated optimization approaches. Furthermore, the method could be combined with more advanced samplers for $p_V(r)$, such as particle methods. Finally, it would be valuable to delve into the theoretical properties of the method.

Bibliography

- Ajami, N. K., Hornberger, G. M., and Sunding, D. L. (2008). Sustainable water resource management under hydrological uncertainty. *Water Resources Research*, 44(11). WR006736.
- Amaya, M., Linde, N., and Laloy, E. (2021). Adaptive sequential Monte Carlo for posterior inference and model selection among complex geological priors. *Geophysical Journal International*, 226(2):1220–1238.
- Amaya, M., Linde, N., and Laloy, E. (2022). Hydrogeological multiple-point statistics inversion by adaptive sequential Monte Carlo. *Advances in Water Resources*, 166:104252.
- Andrieu, C. and Roberts, G. O. (2009). The pseudo-marginal approach for efficient Monte Carlo computations. *The Annals of Statistics*, 37(2):697–725.
- Aster, R. C., Borchers, B., and Thurber, C. H. (2018). *Parameter Estimation and Inverse Problems*. Elsevier.
- Au, S. and Beck, J. L. (2001). Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics*, 16(4):263–277.
- Au, S. and Wang, Y. (2014). *Engineering Risk Assessment with Subset Simulation*. John Wiley & Sons.
- Au, S.-K. and Beck, J. L. (1999). A new adaptive importance sampling scheme for reliability calculations. *Structural Safety*, 21(2):135–158.
- Au, S.-K. and Beck, J. L. (2003a). Important sampling in high dimensions. *Structural Safety*, 25(2):139–163.
- Au, S.-K. and Beck, J. L. (2003b). Subset simulation and its application to seismic risk based on dynamic analysis. *Journal of Engineering Mechanics*, 129(8):901–917.
- Backus, G. (1970). Inference from inadequate and inaccurate data, III. *Proceedings of the National Academy of Sciences*, 67(1):282–289.
- Bakker, M., Post, V., Langevin, C. D., Hughes, J. D., White, J. T., Starn, J., and Fienen, M. N. (2016). Scripting MODFLOW model development using Python and FloPy. *Groundwater*, 54(5):733–739.
- Barbier, E. (2002). Geothermal energy technology and current status: an overview. *Renewable and Sustainable Energy Reviews*, 6(1-2):3–65.

- Bayes, T. (1763). LII. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, FRS communicated by Mr. Price, in a letter to John Canton, AMFR S. *Philosophical Transactions of the Royal Society of London*, (53):370–418.
- Beaumont, M. A. (2003). Estimation of population growth or decline in genetically monitored populations. *Genetics*, 164(3):1139–1160.
- Beck, J. L. and Zuev, K. M. (2015). Rare event simulation. In Ghanem, R., Higdon, D., and Owhadi, H., editors, *Handbook of Uncertainty Quantification*, pages 1–26. Springer, Cham.
- Benekos, I. D., Shoemaker, C. A., and Stedinger, J. R. (2007). Probabilistic risk and uncertainty analysis for bioremediation of four chlorinated ethenes in groundwater. *Stochastic Environmental Research and Risk Assessment*, 21(4):375–390.
- Beskos, A., Jasra, A., Kantas, N., and Thiery, A. (2016). On the convergence of adaptive sequential Monte Carlo methods. *The Annals of Applied Probability*, 26(2):1111–46.
- Betz, W., Papaioannou, I., and Straub, D. (2014). Adaptive variant of the BUS approach to Bayesian updating. In Cunha, A., Caetano, E., Ribeiro, P., and Müller, G., editors, *Proceedings of the 9th International Conference on Structural Dynamics, EURODYN 2014*, pages 3021–3028.
- Bikowski, J., Huisman, J. A., Vrugt, J. A., Vereecken, H., and van der Kruk, J. (2012). Integrated analysis of waveguide dispersed GPR pulses using deterministic and Bayesian inversion methods. *Near Surface Geophysics*, 10(6):641–652.
- Binley, A., Hubbard, S. S., Huisman, J. A., Revil, A., Robinson, D. A., Singha, K., and Slater, L. D. (2015). The emergence of hydrogeophysics for improved understanding of subsurface processes over multiple scales. *Water Resources Research*, 51(6):3837–3866.
- Bishop, C. M. and Nasrabadi, N. M. (2006). *Pattern Recognition and Machine Learning*, volume 4. Springer.
- Blöschl, G. and Sivapalan, M. (1995). Scale issues in hydrological modelling: a review. *Hydrological Processes*, 9(3-4):251–290.
- Bohling, G. C., Liu, G., Dietrich, P., and Butler Jr, J. J. (2016). Reassessing the MADE direct-push hydraulic conductivity data using a revised calibration procedure. *Water Resources Research*, 52(11):8970–8985.
- Bonati, L., Zhang, Y.-Y., and Parrinello, M. (2019). Neural networks-based variationally enhanced sampling. *Proceedings of the National Academy of Sciences*, 116(36):17641–17647.
- Bosch, M. (1999). Lithologic tomography: From plural geophysical data to lithology estimation. *Journal of Geophysical Research: Solid Earth*, 104(B1):749–766.
- Bosch, M. (2004). The optimization approach to lithological tomography: Combining seismic data and petrophysics for porosity prediction. *Geophysics*, 69(5):1272–1282.

- Bosch, M., Cara, L., Rodrigues, J., Navarro, A., and Díaz, M. (2007). A Monte Carlo approach to the joint estimation of reservoir and elastic parameters from seismic amplitudes. *Geophysics*, 72(6):O29–O39.
- Botev, Z. I. and Kroese, D. P. (2008). An efficient algorithm for rare-event probability estimation, combinatorial optimization, and counting. *Methodology and Computing in Applied Probability*, 10(4):471–505.
- Bourinet, J.-M., Deheeger, F., and Lemaire, M. (2011). Assessing small failure probabilities by combined subset simulation and support vector machines. *Structural Safety*, 33(6):343–353.
- Braak, C. J. T. (2006). A Markov chain Monte Carlo version of the genetic algorithm Differential Evolution: easy Bayesian computing for real parameter spaces. *Statistics and Computing*, 16:239–249.
- Bradford, J. H. and Deeds, J. C. (2006). Ground-penetrating radar theory and application of thin-bed offset-dependent reflectivity. *Geophysics*, 71(3):K47–K57.
- Brodin, E. and Klüppelberg, C. (2008). Extreme value theory in finance. In Everitt, B. and Melnick, E., editors, *Encyclopedia of Quantitative Risk Analysis and Assessment*. Wiley, Chichester.
- Brunetti, C. and Linde, N. (2018). Impact of petrophysical uncertainty on Bayesian hydrogeophysical inversion and model selection. *Advances in Water Resources*, 111:346–359.
- Brynjarsdóttir, J. and O’Hagan, A. (2014). Learning about physical parameters: The importance of model discrepancy. *Inverse problems*, 30(11):114007.
- Bucklew, J. A. (2004). *Introduction to Rare Event Simulation*, volume 5. Springer, New York.
- Butler, J. J. (2005). Hydrogeological methods for estimation of spatial variations in hydraulic conductivity. *Hydrogeophysics*, pages 23–58.
- Cadini, F., Avram, D., Pedroni, N., and Zio, E. (2012). Subset simulation of a reliability model for radioactive waste repository performance assessment. *Reliability Engineering & System Safety*, 100:75–83.
- Carrera, J., Alcolea, A., Medina, A., Hidalgo, J., and Slooten, L. J. (2005). Inverse problem in hydrogeology. *Hydrogeology journal*, 13:206–222.
- Cérou, F., Del Moral, P., Furon, T., and Guyader, A. (2012). Sequential Monte Carlo for rare event estimation. *Statistics and Computing*, 22(3):795–808.
- Chen, J., Hubbard, S., and Rubin, Y. (2001). Estimating the hydraulic conductivity at the South Oyster Site from geophysical tomographic data using Bayesian techniques based on the normal linear regression model. *Water Resources Research*, 37(6):1603–1613.
- Chen, V., Dunlop, M. M., Papaspiliopoulos, O., and Stuart, A. M. (2018). Dimension-robust MCMC in Bayesian inverse problems. *arXiv preprint arXiv:1803.03344*.

- Chen, Y., Keyes, D., Law, K. J., and Ltaief, H. (2016). Accelerated dimension-independent adaptive Metropolis. *SIAM Journal on Scientific Computing*, 38(5):S539–S565.
- Chiles, J.-P. and Delfiner, P. (2012). *Geostatistics: Modeling Spatial Uncertainty*, volume 713. John Wiley & Sons.
- Ching, J. and Chen, Y.-C. (2007). Transitional Markov chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging. *Journal of Engineering Mechanics*, 133(7):816–832.
- Chwialkowski, K., Strathmann, H., and Gretton, A. (2016). A kernel test of goodness of fit. In Balcan, M. F. and Weinberger, K. Q., editors, *International Conference on Machine Learning*, PMLR, pages 2606–2615.
- Constable, S. C., Parker, R. L., and Constable, C. G. (1987). Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics*, 52(3):289–300.
- Cotter, S. L., Roberts, G. O., Stuart, A. M., and White, D. (2013). MCMC methods for functions: modifying old algorithms to make them faster. *Statistical Science*, 28(3):424 – 446.
- Dall'Alba, V., Neven, A., de Rooij, R., Filipponi, M., and Renard, P. (2023). Probabilistic estimation of tunnel inflow from a karstic conduit network. *Engineering Geology*, 312:106950.
- Davies, L., Ley-Cooper, A., Sutton, M., and Drovandi, C. (2023). Bayesian detectability of induced polarisation in airborne electromagnetic data. *Geophysical Journal International*. ggad073.
- Day-Lewis, F., Linde, N., Haggerty, R., Singha, K., and Briggs, M. A. (2017). Pore network modeling of the electrical signature of solute transport in dual-domain media. *Geophysical Research Letters*, 44(10):4908–4916.
- Day-Lewis, F. D., Singha, K., and Binley, A. M. (2005). Applying petrophysical models to radar travel time and electrical resistivity tomograms: Resolution-dependent limitations. *Journal of Geophysical Research: Solid Earth*, 110(B8). B08206.
- Del Moral, P., Doucet, A., and Jasra, A. (2006). Sequential Monte Carlo samplers. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(3):411–436.
- Del Moral, P., Doucet, A., and Jasra, A. (2007). Sequential Monte Carlo for Bayesian computation. *Bayesian Statistics*, 8(1):34.
- Deligiannidis, G., Doucet, A., and Pitt, M. (2018). The correlated pseudomarginal method. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, 80(5):839–870.
- Dentz, M., Le Borgne, T., Englert, A., and Bijeljic, B. (2011). Mixing, spreading and reaction in heterogeneous media: A brief review. *Journal of Contaminant Hydrology*, 120:1–17.
- Deparis, J. and Garambois, S. (2009). On the use of dispersive APVO GPR curves for thin-bed properties estimation: Theory and application to fracture characterization. *Geophysics*, 74(1):J1–J12.

- Diebold, F. X., Schuermann, T., and Stroughair, J. D. (2000). Pitfalls and opportunities in the use of extreme value theory in risk management. *The Journal of Risk Finance*, 1(2):30–35.
- Dietrich, C. R. and Newsam, G. N. (1997). Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix. *SIAM Journal on Scientific Computing*, 18(4):1088–1107.
- Domenico, P. A. and Schwartz, F. W. (1997). *Physical and Chemical Hydrogeology*. John Wiley & Sons.
- Dorn, C., Linde, N., Le Borgne, T., Bour, O., and de Dreuzy, J.-R. (2013). Conditioning of stochastic 3-D fracture networks to hydrological and geophysical data. *Advances in Water Resources*, 62:79–89.
- Douc, R. and Cappé, O. (2005). Comparison of resampling schemes for particle filtering. In *Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis, ISPA*, pages 64–69.
- Doucet, A., De Freitas, N., and Gordon, N. (2001). An introduction to sequential Monte Carlo methods. In Doucet, A., de Freitas, N., and Gordon, N., editors, *Sequential Monte Carlo methods in practice*, pages 3–14. Springer, New York.
- Doucet, A. and Johansen, A. M. (2009). A tutorial on particle filtering and smoothing: Fifteen years later. In Crisan, D. and Rozovskii, B., editors, *The Oxford Handbook of Nonlinear Filtering*, volume 12, pages 656–704. Oxford University Press, New York.
- Doucet, A., Pitt, M. K., Deligiannidis, G., and Kohn, R. (2015). Efficient implementation of Markov chain Monte Carlo when using an unbiased likelihood estimator. *Biometrika*, 102(2):295–313.
- Earl, D. J. and Deem, M. W. (2005). Parallel tempering: Theory, applications, and new perspectives. *Physical Chemistry Chemical Physics*, 7(23):3910–3916.
- Ellis, R. and Oldenburg, D. (1994). Applied geophysical inversion. *Geophysical Journal International*, 116(1):5–11.
- Enzenhofer, R., Nowak, W., and Helmig, R. (2012). Probabilistic exposure risk assessment with advective–dispersive well vulnerability criteria. *Advances in Water Resources*, 36:121–132.
- Famiglietti, J. S. (2014). The global groundwater crisis. *Nature Climate Change*, 4(11):945–948.
- Ferré, T., Bentley, L., Binley, A., Linde, N., Kemna, A., Singha, K., Holliger, K., Huisman, J. A., and Minsley, B. (2009). Critical steps for the continuing advancement of hydrogeophysics. *Eos, Transactions American Geophysical Union*, 90(23):200.
- Friedli, L. and Linde, N. (2023). Solving geophysical inversion problems with intractable likelihoods: Linearized Gaussian approximations versus the correlated pseudo-marginal method. *Mathematical Geosciences*, pages 1–21.

- Friedli, L., Linde, N., Ginsbourger, D., and Doucet, A. (2022). Lithological tomography with the correlated pseudo-marginal method. *Geophysical Journal International*, 228(2):839–856.
- Friedli, L., Linde, N., Ginsbourger, D., Visentini, A. F., and Doucet, A. (2023). Inference of geostatistical hyperparameters with the correlated pseudo-marginal method. *Advances in Water Resources*, 173:104402.
- Gelhar, L. W. and Axness, C. L. (1983). Three-dimensional stochastic analysis of macrodispersion in aquifers. *Water Resources Research*, 19(1):161–180.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian Data Analysis*. CRC press.
- Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457–472.
- Gelman, A. and Shirley, K. (2011). Inference from simulations and monitoring convergence. In Brooks, S., Gelman, A., Jones, G., and Meng, X.-L., editors, *Handbook of Markov chain Monte Carlo*, volume 6, pages 163–174. Chapman and Hall/CRC.
- Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378.
- González, E. F., Mukerji, T., and Mavko, G. (2008). Seismic inversion combining rock physics and multiple-point geostatistics. *Geophysics*, 73(1):R11–R21.
- Good, I. J. (1952). Rational decisions. *Journal of the Royal Statistical Society: Series B (Methodological)*, 14(1):107–114.
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT press.
- Gorelick, S. M. and Zheng, C. (2015). Global change and the groundwater management challenge. *Water Resources Research*, 51(5):3031–3051.
- Gorham, J. and Mackey, L. (2015). Measuring sample quality with Stein’s method. In Cortes, C., Lawrence, N., Lee, D., Sugiyama, M., and Garnett, R., editors, *Advances in Neural Information Processing Systems 28, NIPS*, volume 28.
- Gorham, J. and Mackey, L. (2017). Measuring sample quality with kernels. In Precup, D. and Teh, Y. W., editors, *International Conference on Machine Learning, PMLR*, pages 1292–1301.
- Grana, D. and Della Rossa, E. (2010). Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion. *Geophysics*, 75(3):O21–O37.
- Grana, D., Passos de Figueiredo, L., and Azevedo, L. (2019). Uncertainty quantification in Bayesian inverse problems with model and data dimension reduction. *Geophysics*, 84(6):M15–M24.
- Grégoire, C. and Hollender, F. (2004). Discontinuity characterization by the inversion of the spectral content of ground-penetrating radar (GPR) reflections—application of the Jonscher model. *Geophysics*, 69(6):1414–1424.

- Guardiano, F. B. and Srivastava, R. M. (1993). Multivariate geostatistics: Beyond bivariate moments. In A. Soares, editor, *Geostatistics Tróia '92*, volume 1, pages 133–144. Springer Netherlands.
- Haario, H., Saksman, E., and Tamminen, J. (2001). An adaptive Metropolis algorithm. *Bernoulli*, 7(2):223–242.
- Hadjidoukas, P. E., Angelikopoulos, P., Papadimitriou, C., and Koumoutsakos, P. (2015). $\pi 4u$: A high performance computing framework for bayesian uncertainty quantification of complex models. *Journal of Computational Physics*, 284:1–21.
- Hansen, T. M., Cordua, K. S., Jacobsen, B. H., and Mosegaard, K. (2014). Accounting for imperfect forward modeling in geophysical inverse problems—exemplified for crosshole tomography. *Geophysics*, 79(3):H1–H21.
- Hansen, T. M., Cordua, K. S., Looms, M. C., and Mosegaard, K. (2013a). SIPPI: A Matlab toolbox for sampling the solution to inverse problems with complex prior information: Part 1—methodology. *Computers & Geosciences*, 52:470–480.
- Hansen, T. M., Cordua, K. S., Looms, M. C., and Mosegaard, K. (2013b). SIPPI: A Matlab toolbox for sampling the solution to inverse problems with complex prior information: Part 2—application to crosshole GPR tomography. *Computers & Geosciences*, 52:481–492.
- Hansen, T. M., Cordua, K. S., and Mosegaard, K. (2012). Inverse problems with non-trivial priors: efficient solution through sequential Gibbs sampling. *Computational Geosciences*, 16(3):593–611.
- Hasofer, A. M. and Lind, N. C. (1974). Exact and invariant second-moment code format. *Journal of the Engineering Mechanics Division*, 100(1):111–121.
- Hastie, T., Tibshirani, R., and Friedman, J. H. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, volume 2. Springer.
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109.
- Helton, J. C. (1994). Treatment of uncertainty in performance assessments for complex systems. *Risk Analysis*, 14(4):483–511.
- Hermans, T., Oware, E., and Caers, J. (2016). Direct prediction of spatially and temporally varying physical properties from time-lapse electrical resistance data. *Water Resources Research*, 52(9):7262–7283.
- Hess, K. M., Wolf, S. H., and Celia, M. A. (1992). Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts: 3. Hydraulic conductivity variability and calculated macrodispersivities. *Water Resources Research*, 28(8):2011–2027.
- Hill, R. (1963). Elastic properties of reinforced solids: some theoretical principles. *Journal of the Mechanics and Physics of Solids*, 11(5):357–372.

- Hinnell, A., Ferré, T., Vrugt, J., Huisman, J., Moysey, S., Rings, J., and Kowalsky, M. (2010). Improved extraction of hydrologic information from geophysical data through coupled hydrogeophysical inversion. *Water Resources Research*, 46(4). WR007060.
- Hoege, M., Guthke, A., and Nowak, W. (2019). The hydrologist's guide to Bayesian model selection, averaging and combination. *Journal of Hydrology*, 572:96–107.
- Hoffman, F. O. and Hammonds, J. S. (1994). Propagation of uncertainty in risk assessments: the need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. *Risk analysis*, 14(5):707–712.
- Hohenbichler, M. and Rackwitz, R. (1988). Improvement of second-order reliability estimates by importance sampling. *Journal of Engineering Mechanics*, 114(12):2195–2199.
- Hubbard, S. S. and Rubin, Y. (2005). Introduction to hydrogeophysics. In Rubin, Y. and Hubbard, S. S., editors, *Hydrogeophysics*, pages 3–21. Springer.
- Hunziker, J., Laloy, E., and Linde, N. (2017). Inference of multi-Gaussian relative permittivity fields by probabilistic inversion of crosshole ground-penetrating radar data. *Geophysics*, 82(5):H25–H40.
- Invernizzi, M., Valsson, O., and Parrinello, M. (2017). Coarse graining from variationally enhanced sampling applied to the Ginzburg–Landau model. *Proceedings of the National Academy of Sciences*, 114(13):3370–3374.
- Jackson, D. D. and Matsu'Ura, M. (1985). A Bayesian approach to nonlinear inversion. *Journal of Geophysical Research: Solid Earth*, 90(B1):581–591.
- Jensen, H., Vergara, C., Papadimitriou, C., and Millas, E. (2013). The use of updated robust reliability measures in stochastic dynamical systems. *Computer Methods in Applied Mechanics and Engineering*, 267:293–317.
- Johansen, A. M., Del Moral, P., and Doucet, A. (2006). Sequential Monte Carlo samplers for rare events. In *6th International Workshop on Rare Event Simulation*, pages 256–267.
- Jougnot, D., Jiménez-Martínez, J., Legendre, R., Le Borgne, T., Méheust, Y., and Linde, N. (2018). Impact of small-scale saline tracer heterogeneity on electrical resistivity monitoring in fully and partially saturated porous media: Insights from geoelectrical milli-fluidic experiments. *Advances in Water Resources*, 113:295–309.
- Juneja, S. and Shahabuddin, P. (2006). Rare-event simulation techniques: an introduction and recent advances. *Handbooks in Operations Research and Management Science*, 13:291–350.
- Kahn, H. and Marshall, A. W. (1953). Methods of reducing sample size in Monte Carlo computations. *Journal of the Operations Research Society of America*, 1(5):263–278.
- Kanjilal, O., Papaioannou, I., and Straub, D. (2023). Bayesian updating of reliability by cross entropy-based importance sampling. *Structural Safety*, 102:102325.

- Kästner, J. and Thiel, W. (2005). Bridging the gap between thermodynamic integration and umbrella sampling provides a novel analysis method: "Umbrella integration". *The Journal of Chemical Physics*, 123(14):144104.
- Khadam, I. M. and Kaluarachchi, J. J. (2003). Multi-criteria decision analysis with probabilistic risk assessment for the management of contaminated ground water. *Environmental Impact Assessment Review*, 23(6):683–721.
- Kingma, D. P. and Ba, J. (2015). Adam: A method for stochastic optimization. In Bengio, Y. and LeCun, Y., editors, *3rd International Conference for Learning Representations, ICLR, Conference Track Proceedings*.
- Kitanidis, P. K. (1995). Quasi-linear geostatistical theory for inversing. *Water Resources Research*, 31(10):2411–2419.
- Klemeš, V. (1983). Conceptualization and scale in hydrology. *Journal of Hydrology*, 65(1-3):1–23.
- Kong, A., Liu, J. S., and Wong, W. H. (1994). Sequential imputations and Bayesian missing data problems. *Journal of the American Statistical Association*, 89(425):278–288.
- Koop, J. (1972). On the derivation of expected value and variance of ratios without the use of infinite series expansions. *Metrika*, 19(1):156–170.
- Koutsourelakis, P.-S., Pradlwarter, H. J., and Schueller, G. I. (2004). Reliability of structures in high dimensions, part I: algorithms and applications. *Probabilistic Engineering Mechanics*, 19(4):409–417.
- Kowalsky, M. B., Finsterle, S., Peterson, J., Hubbard, S., Rubin, Y., Majer, E., Ward, A., and Gee, G. (2005). Estimation of field-scale soil hydraulic and dielectric parameters through joint inversion of GPR and hydrological data. *Water Resources Research*, 41(11). WR004237.
- Krauskopf, K. (2013). *Radioactive Waste Disposal and Geology*, volume 1. Springer Science & Business Media.
- Krüger, F., Lerch, S., Thorarinsdottir, T., and Gneiting, T. (2021). Predictive inference based on Markov chain Monte Carlo output. *International Statistical Review*, 89(2):274–301.
- Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86.
- Lahkim, M. B. and Garcia, L. A. (1999). Stochastic modeling of exposure and risk in a contaminated heterogeneous aquifer. 1: Monte Carlo uncertainty analysis. *Environmental Engineering Science*, 16(5):315–328.
- Laloy, E., Linde, N., Jacques, D., and Vrugt, J. A. (2015). Probabilistic inference of multi-Gaussian fields from indirect hydrological data using circulant embedding and dimensionality reduction. *Water Resources Research*, 51(6):4224–4243.

- Laloy, E. and Vrugt, J. A. (2012). High-dimensional posterior exploration of hydrologic models using multiple-try DREAM (ZS) and high-performance computing. *Water Resources Research*, 48(1). WR010608.
- Langtangen, H. P. and Linge, S. (2017). Diffusion equations. In Langtangen, H. and Linge, S., editors, *Finite Difference Computing with PDEs: A Modern Software Approach*, pages 207–322. Springer, Cham.
- Leadbetter, M. R. (1991). On a basis for ‘peaks over threshold’ modeling. *Statistics & Probability Letters*, 12(4):357–362.
- LeGland, F. and Oudjane, N. (2006). A sequential particle algorithm that keeps the particle system alive. In Blom, H. and Lygeros, J., editors, *Stochastic Hybrid Systems: Theory and Safety Critical Applications*, pages 351–389. Springer, Berlin Heidelberg.
- Lelièvre, T., Stoltz, G., and Rousset, M. (2010). *Free Energy Computations: A Mathematical Perspective*. World Scientific.
- Li, J. and Xiu, D. (2010). Evaluation of failure probability via surrogate models. *Journal of Computational Physics*, 229(23):8966–8980.
- Linde, N. (2014). Falsification and corroboration of conceptual hydrological models using geophysical data. *Wiley Interdisciplinary Reviews: Water*, 1(2):151–171.
- Linde, N. and Doetsch, J. (2016). Joint inversion in hydrogeophysics and near-surface geophysics. In Moorkamp, M., Lelièvre, P. G., Linde, N., and Khan, A., editors, *Integrated imaging of the Earth: Theory and applications*, pages 117–135. Wiley Online Library.
- Linde, N., Ginsbourger, D., Irving, J., Nobile, F., and Doucet, A. (2017). On uncertainty quantification in hydrogeology and hydrogeophysics. *Advances in Water Resources*, 110:166–181.
- Linde, N., Renard, P., Mukerji, T., and Caers, J. (2015). Geological realism in hydrogeological and geophysical inverse modeling: A review. *Advances in Water Resources*, 86:86–101.
- Liu, Q., Lee, J., and Jordan, M. (2016). A kernelized Stein discrepancy for goodness-of-fit tests. In Balcan, M. F. and Weinberger, K. Q., editors, *International Conference on Machine Learning*, PMLR, pages 276–284.
- Liu, Y., Gao, Y., and Yin, W. (2020). An improved analysis of stochastic gradient descent with momentum. *Advances in Neural Information Processing Systems*, 33:18261–18271.
- Lochbühler, T., Breen, S. J., Detwiler, R. L., Vrugt, J. A., and Linde, N. (2014). Probabilistic electrical resistivity tomography of a CO₂ sequestration analog. *Journal of Applied Geophysics*, 107:80–92.
- Malinverno, A. and Briggs, V. A. (2004). Expanded uncertainty quantification in inverse problems: Hierarchical Bayes and empirical Bayes. *Geophysics*, 69(4):1005–1016.
- Marzouk, Y. M. and Najm, H. N. (2009). Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems. *Journal of Computational Physics*, 228(6):1862–1902.

- Maurer, H. and Friedel, S. (2006). Outer-space sensitivities in geoelectrical tomography. *Geophysics*, 71(3):G93–G96.
- Mavko, G., Mukerji, T., and Dvorkin, J. (2020). *The Rock Physics Handbook*. Cambridge University Press.
- McCartney, J. S., Sánchez, M., and Tomac, I. (2016). Energy geotechnics: Advances in sub-surface energy recovery, storage, exchange, and waste management. *Computers and Geotechnics*, 75:244–256.
- McLaughlin, D. and Townley, L. R. (1996). A reassessment of the groundwater inverse problem. *Water Resources Research*, 32(5):1131–1161.
- McMahon, S. J. (2018). The linear quadratic model: usage, interpretation and challenges. *Physics in Medicine & Biology*, 64(1):01TR01.
- Menke, W. (2018). *Geophysical Data Analysis: Discrete Inverse Theory*. Academic Press.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092.
- Morrison, J. E. and Smith, J. A. (2002). Stochastic modeling of flood peaks using the generalized extreme value distribution. *Water Resources Research*, 38(12). WR000502.
- Mosegaard, K. and Tarantola, A. (1995). Monte Carlo sampling of solutions to inverse problems. *Journal of Geophysical Research: Solid Earth*, 100(B7):12431–12447.
- Mukerji, T., Avseth, P., Mavko, G., Takahashi, I., and González, E. F. (2001). Statistical rock physics: Combining rock physics, information theory, and geostatistics to reduce uncertainty in seismic reservoir characterization. *The Leading Edge*, 20(3):313–319.
- National Research Council (1996). *Rock Fractures and Fluid Flow: Contemporary Understanding and Applications*. National Academies Press.
- Neal, R. M. (2001). Annealed importance sampling. *Statistics and Computing*, 11(2):125–139.
- Neuman, S. P. and Di Federico, V. (2003). Multifaceted nature of hydrogeologic scaling and its interpretation. *Reviews of Geophysics*, 41(3). RG000130.
- Owen, A. and Zhou, Y. (2000). Safe and effective importance sampling. *Journal of the American Statistical Association*, 95(449):135–143.
- Papadimitriou, C., Beck, J. L., and Katafygiotis, L. S. (2001). Updating robust reliability using structural test data. *Probabilistic Engineering Mechanics*, 16(2):103–113.
- Parker, R. L. (1994). *Geophysical Inverse Theory*, volume 1. Princeton University Press.
- Podvin, P. and Lecomte, I. (1991). Finite difference computation of traveltimes in very contrasted velocity models: a massively parallel approach and its associated tools. *Geophysical Journal International*, 105(1):271–284.

- Press, F. (1968). Earth models obtained by Monte Carlo inversion. *Journal of Geophysical Research*, 73(16):5223–5234.
- R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Razavi, S., Tolson, B. A., and Burn, D. H. (2012). Review of surrogate modeling in water resources. *Water Resources Research*, 48(7). WR011527.
- Rehfeldt, K. R., Boggs, J. M., and Gelhar, L. W. (1992). Field study of dispersion in a heterogeneous aquifer: 3. Geostatistical analysis of hydraulic conductivity. *Water Resources Research*, 28(12):3309–3324.
- Renard, P. and De Marsily, G. (1997). Calculating equivalent permeability: a review. *Advances in Water Resources*, 20(5-6):253–278.
- Reynolds, J. M. (2011). *An Introduction to Applied and Environmental Geophysics*. John Wiley & Sons.
- Richard, J.-F. and Zhang, W. (2007). Efficient high-dimensional importance sampling. *Journal of Econometrics*, 141(2):1385–1411.
- Ripley, B. D. (2009a). *Stochastic Simulation*. John Wiley & Sons.
- Ripley, B. D. (2009b). *Stochastic Simulation*. John Wiley & Sons.
- Robert, C. P., Casella, G., and Casella, G. (1999). *Monte Carlo Statistical Methods*, volume 2. Springer.
- Robert, C. P., Elvira, V., Tawn, N., and Wu, C. (2018). Accelerating MCMC algorithms. *Wiley Interdisciplinary Reviews: Computational Statistics*, 10(5):e1435.
- Rosas-Carbajal, M., Linde, N., Kalscheuer, T., and Vrugt, J. A. (2014). Two-dimensional probabilistic inversion of plane-wave electromagnetic data: methodology, model constraints and joint inversion with electrical resistivity data. *Geophysical Journal International*, 196(3):1508–1524.
- Roth, K., Schulin, R., Flühler, H., and Attinger, W. (1990). Calibration of time domain reflectometry for water content measurement using a composite dielectric approach. *Water Resources Research*, 26(10):2267–2273.
- Rubin, Y., Chen, X., Murakami, H., and Hahn, M. (2010). A Bayesian approach for inverse modeling, data assimilation, and conditional simulation of spatial random fields. *Water Resources Research*, 46(10). WR008799.
- Rubin, Y. and Hubbard, S. (2005). Stochastic forward and inverse modeling: the “hydrogeophysical” challenge. In Rubin, Y. and Hubbard, S. S., editors, *Hydrogeophysics*, pages 487–511. Springer.
- Rubino, G. and Tuffin, B. (2009). *Rare Event Simulation Using Monte Carlo Methods*. John Wiley & Sons.

- Rudolf, D. and Sprungk, B. (2018). On a generalization of the preconditioned Crank–Nicolson Metropolis algorithm. *Foundations of Computational Mathematics*, 18:309–343.
- Ruggeri, P., Irving, J., and Holliger, K. (2015). Systematic evaluation of sequential geostatistical resampling within MCMC for posterior sampling of near-surface geophysical inverse problems. *Geophysical Journal International*, 202(2):961–975.
- Sanchez-Vila, X., Guadagnini, A., and Carrera, J. (2006). Representative hydraulic conductivities in saturated groundwater flow. *Reviews of Geophysics*, 44(3). RG000169.
- Scales, J. A. and Tenorio, L. (2001). Prior information and uncertainty in inverse problems. *Geophysics*, 66(2):389–397.
- Schöbi, R., Sudret, B., and Marelli, S. (2017). Rare event estimation using polynomial-chaos kriging. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 3(2):D4016002.
- Scott, D. W. (2015). *Multivariate Density Estimation: Theory, Practice, and Visualization*. John Wiley & Sons.
- Shahraeeni, M. S. and Curtis, A. (2011). Fast probabilistic nonlinear petrophysical inversion. *Geophysics*, 76(2):E45–E58.
- Shakas, A. and Linde, N. (2015). Effective modeling of ground penetrating radar in fractured media using analytic solutions for propagation, thin-bed interaction and dipolar scattering. *Journal of Applied Geophysics*, 116:206–214.
- Shakas, A., Linde, N., Le Borgne, T., and Bour, O. (2018). Probabilistic inference of fracture-scale flow paths and aperture distribution from hydrogeophysically-monitored tracer tests. *Journal of Hydrology*, 567:305–319.
- Shao, X. (2010). The dependent wild bootstrap. *Journal of the American Statistical Association*, 105(489):218–235.
- Shirts, M. R. and Ferguson, A. L. (2020). Statistically optimal continuous free energy surfaces from biased simulations and multistate reweighting. *Journal of Chemical Theory and Computation*, 16(7):4107–4125.
- Siebert, S., Burke, J., Faures, J., Frenken, K., Hoogeveen, J., Döll, P., and Portmann, F. T. (2010). Groundwater use for irrigation—a global inventory. *Hydrology and Earth System Sciences*, 14(10):1863–1880.
- Siirila, E. R., Navarre-Sitchler, A. K., Maxwell, R. M., and McCray, J. E. (2012). A quantitative methodology to assess the risks to human health from CO₂ leakage into groundwater. *Advances in Water Resources*, 36:146–164.
- Song, Y. and Kingma, D. P. (2021). How to train your energy-based models. *arXiv preprint arXiv:2101.03288*.

- Soueid Ahmed, A., Jardani, A., Revil, A., and Dupont, J.-P. (2014). Hydraulic conductivity field characterization from the joint inversion of hydraulic heads and self-potential data. *Water Resources Research*, 50(4):3502–3522.
- Stecher, T., Bernstein, N., and Csányi, G. (2014). Free energy surface reconstruction from umbrella samples using Gaussian process regression. *Journal of Chemical Theory and Computation*, 10(9):4079–4097.
- Stein, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Probability Theory*, pages 583–602.
- Stigler, S. M. (1986). *The History of Statistics: The Measurement of Uncertainty Before 1900*. Harvard University Press.
- Straub, D. (2011). Reliability updating with equality information. *Probabilistic Engineering Mechanics*, 26(2):254–258.
- Straub, D. and Papaioannou, I. (2015). Bayesian updating with structural reliability methods. *Journal of Engineering Mechanics*, 141(3):04014134.
- Straub, D., Papaioannou, I., and Betz, W. (2016). Bayesian analysis of rare events. *Journal of Computational Physics*, 314:538–556.
- Sundar, V. and Manohar, C. (2013). Updating reliability models of statically loaded instrumented structures. *Structural Safety*, 40:21–30.
- Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM.
- Tarantola, A. (2006). Popper, Bayes and the inverse problem. *Nature Physics*, 2(8):492–494.
- Tarantola, A. and Valette, B. (1982a). Generalized nonlinear inverse problems solved using the least squares criterion. *Reviews of Geophysics*, 20(2):219–232.
- Tarantola, A. and Valette, B. (1982b). Inverse problems= quest for information. *Journal of Geophysics*, 50(1):159–170.
- Tartakovsky, D. M. (2007). Probabilistic risk analysis in subsurface hydrology. *Geophysical Research Letters*, 34(5). GL029245.
- Tartakovsky, D. M. (2013). Assessment and management of risk in subsurface hydrology: A review and perspective. *Advances in Water Resources*, 51:247–260.
- Thibaut, R., Laloy, E., and Hermans, T. (2021). A new framework for experimental design using bayesian evidential learning: The case of wellhead protection area. *Journal of Hydrology*, 603:126903.
- Tikhonov, A. and Arsenin, V. (1977). *Solution of Ill-posed Problems*. Washington: Winston Sons.

- Torquato, S. and Haslach Jr, H. (2002). Random heterogeneous materials: microstructure and macroscopic properties. *Applied Mechanics Reviews*, 55(4):B62–B63.
- Torrie, G. M. and Valleau, J. P. (1977). Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics*, 23(2):187–199.
- Tsang, Y. (1992). Usage of “equivalent apertures” for rock fractures as derived from hydraulic and tracer tests. *Water Resources Research*, 28(5):1451–1455.
- Tsoflias, G. P., Perll, C., Baker, M., and Becker, M. W. (2015). Cross-polarized GPR imaging of fracture flow channeling. *Journal of Earth Science*, 26:776–784.
- Ullmann, E. and Papaioannou, I. (2015). Multilevel estimation of rare events. *SIAM/ASA Journal on Uncertainty Quantification*, 3(1):922–953.
- UNICEF (2023). Weltwassertag 2023: 10 fakten über wasser. <https://www.unicef.de/informieren/aktuelles/blog/-/weltwassertag-2023-zehn-fakten-ueber-wasser/275338>. Accessed: 2023-06-12.
- Valsson, O. and Parrinello, M. (2014). Variational approach to enhanced sampling and free energy calculations. *Physical Review Letters*, 113(9):090601.
- Van der Vaart, A. W. (2000). *Asymptotic Statistics*, volume 3. Cambridge University Press.
- Vats, D. and Knudson, C. (2021). Revisiting the Gelman–Rubin diagnostic. *Statistical Science*, 36(4):518–529.
- Veach, E. and Guibas, L. J. (1995). Optimally combining sampling techniques for Monte Carlo rendering. In Mair, S. G. and Cook, R., editors, *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques*, pages 419–428.
- Visentini, A. F., Linde, N., Le Borgne, T., and Dentz, M. (2020). Inferring geostatistical properties of hydraulic conductivity fields from saline tracer tests and equivalent electrical conductivity time-series. *Advances in Water Resources*, 146:103758.
- Vrugt, J. A. (2016). Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation. *Environmental Modelling & Software*, 75:273–316.
- Vrugt, J. A., Ter Braak, C., Diks, C., Robinson, B. A., Hyman, J. M., and Higdon, D. (2009). 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.
- Vrugt, J. A., Ter Braak, C. J., Clark, M. P., Hyman, J. M., and Robinson, B. A. (2008). Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation. *Water Resources Research*, 44(12). WR006720.
- Vrugt, J. A., ter Braak, C. J., Diks, C. G., and Schoups, G. (2013). Hydrologic data assimilation using particle Markov chain Monte Carlo simulation: Theory, concepts and applications. *Advances in Water Resources*, 51:457–478.

- Wang, F. and Landau, D. P. (2001). Efficient, multiple-range random walk algorithm to calculate the density of states. *Physical Review Letters*, 86(10):2050.
- Wang, L., Kitanidis, P. K., and Caers, J. (2022). Hierarchical Bayesian inversion of global variables and large-scale spatial fields. *Water Resources Research*, 58(5). WR031610.
- Ward, P. J., Blauhut, V., Bloemendaal, N., Daniell, J. E., de Ruiter, M. C., Duncan, M. J., Emberson, R., Jenkins, S. F., Kirschbaum, D., Kunz, M., et al. (2020). Natural hazard risk assessments at the global scale. *Natural Hazards and Earth System Sciences*, 20(4):1069–1096.
- Winter, C. L. and Tartakovsky, D. M. (2008). A reduced complexity model for probabilistic risk assessment of groundwater contamination. *Water Resources Research*, 44(6).
- Xiao, S., Xu, T., Reuschen, S., Nowak, W., and Hendricks Franssen, H.-J. (2021). Bayesian inversion of multi-Gaussian log-conductivity fields with uncertain hyperparameters: An extension of preconditioned Crank-Nicolson Markov chain Monte Carlo with parallel tempering. *Water Resources Research*, 57(9). WR030313.
- Yu, X. and Michael, H. A. (2022). Impacts of the scale of representation of heterogeneity on simulated salinity and saltwater circulation in coastal aquifers. *Water Resources Research*, 58(1). WR029523.
- Zhao, Y. and Luo, J. (2021). Bayesian inverse modeling of large-scale spatial fields on iteratively corrected principal components. *Advances in Water Resources*, 151:103913.
- Zhou, Y., Johansen, A. M., and Aston, J. A. D. (2016). Toward automatic model comparison: an adaptive sequential Monte Carlo approach. *Journal of Computational and Graphical Statistics*, 25(3):701–726.
- Zunino, A., Khan, A., Cupillard, P., and Mosegaard, K. (2016). Constitution and structure of earth's mantle: Insights from mineral physics and seismology. In Moorkamp, M., Lelièvre, P. G., Linde, N., and Khan, A., editors, *Integrated Imaging of the Earth: Theory and Applications*, pages 219–243. Wiley Online Library.

Acknowledgements

I am delighted to wholeheartedly express my gratitude to all those who have supported me throughout my PhD journey. Without you, the *likelihood* of those four years being a great experience and the associated *risk* would have been very different.

First, I want to thank **Niklas Linde** for being an exceptional supervisor. As a statistician navigating through the geosciences, there were times when I felt a bit lost, but your patience and support have been invaluable. I am deeply thankful for your unwavering reliability, which together with your extremely broad knowledge and curiosity have always been a source of support and motivation. I am sincerely grateful to my co-supervisor, **David Ginsbourger**, who has been by my side since my very first statistics lecture until the completion of my PhD. Your consistent guidance and unwavering trust have been invaluable, and I appreciate (most of the time) how you challenge me when I am tempted to take the easy path.

I want to extend my gratitude to **Arnaud Doucet**, who, despite a demanding schedule, consistently allocated time to provide feedback on drafts and engage in discussions. Your vision has left a deep impression on me and the time I spent in Oxford will be fondly remembered. Special thanks goes to the committee members **Klaus Holliger** and **Klaus Mosegaard** for the interesting and inspiring discussion during my private defense as well as to the president of the jury **Marie-Elodie Perga** for her availability and kindness. I would also like to thank the **Swiss National Science Foundation** for funding this work (project number 184574).

Next, I express my heartfelt gratitude to the one and only GEOcurls **Shiran** and **Macarena**, who have been my companions in this PhD journey, sharing the same boat through both calm and stormy times. I extend my thanks to **Jürg** and **Giovanni** for the enjoyable moments we shared in and around the office; your willingness to assist with any inquiries is greatly appreciated. Special thanks also goes to **Johanna** and everyone else in Géopolis who made the time I actually spent in Lausanne very memorable. Furthermore, I thank my statistics colleagues in Bern, especially **Sam**, **Cédric**, **Athénaïs** and **Claire**, who supported me across a range of topics from French translations to Banach spaces.

Infinite gratitude (so-called ‘unändlächi Dankbarkeit’) goes to my very best friend and flat-mate **Sara**, who is my home, my anchor and an incredibly patient, loyal and appreciative person. I am deeply grateful to all my friends in Bern for their unwavering support throughout the past four years. You make hard times so much easier and easy times simply fantastic. Especially, I would like to thank **Anna**, **Leandra**, **Eveline**, **Simon**, **Diana** and **Rahel**.

Finally, profound gratitude goes to my parents **Ursula** and **Markus** for their unwavering support and presence by my side. You are the harbor I can always rely on.

Bern, October 2023

Lea



Lea Friedli

As an applied statistician, I am deeply enthusiastic about interdisciplinary collaboration, derive great pleasure from gaining insights into new domains, and possess a strong passion for science communication.

📍 Bern, Switzerland

✉ lea.friedli@unil.ch

🌐 www.leafriedli.com

Educational background

- **Masters in Statistics (2016-2019)**
University of Bern, Switzerland
- **Bachelor in German Language and Literature (2013-2016)**
Minor Mathematics (60 ECTS)
University of Bern, Switzerland

Skills

Languages	German	Native
	English	Fluent
	French	Semi-fluent
IT	R	●●●●●○
	Python	●●●●○
	Matlab	●●●●○
	LaTeX	●●●●○

PhD in Environmental Science

Likelihood estimation and risk assessment for inversion problems in the geosciences

- Institute of Earth Sciences
University of Lausanne, Switzerland
- Since November 2019, defense in October 2023
- Supervisor: Prof. Dr. Niklas Linde
Co-Supervisor: Prof. Dr. David Ginsbourger
Scientific collaborator: Prof. Dr. Arnaud Doucet
- Teaching assistant for the Institute of Earth Sciences

Awards

- **Best student presentation award** during the MASCOT-NUM conference 2023 for the presentation with the title "An energy based model approach to the estimation of rare event probabilities"
- **C. Moser Award** by the institute of Mathematical Statistics and Actuarial Science of the University of Bern for the master thesis entitled "Area-covering post-processing of ensemble precipitation forecasts using topographical and seasonal conditions"

Science communication

- **Pint of Science 2023** (international science festival)
Event organization Bern
Presentation about hydrogeophysics
- **FameLab 2021** (science communication competition)
National finalist Switzerland

Sports instructor

Group fitness instructor Cardio and Toning, working for University Sports Bern since 2020.

Publications

- Friedli, L., & Linde, N. (2023). Solving Geophysical Inversion Problems with Intractable Likelihoods: Linearized Gaussian Approximations Versus the Correlated Pseudo-marginal Method. *Mathematical Geosciences*, pages 1–21.
- Friedli, L., Linde, N., Ginsbourger, D., Visentini, A. F., & Doucet, A. (2023). Inference of geostatistical hyperparameters with the correlated pseudo-marginal method. *Advances in Water Resources*, 173:104402.
- Friedli, L., Linde, N., Ginsbourger, D., & Doucet, A. (2022). Lithological tomography with the correlated pseudo-marginal method. *Geophysical Journal International*, 228(2):839–856.
- Friedli, L., Ginsbourger, D., & Bhend, J. (2021). Area-covering postprocessing of ensemble precipitation forecasts using topographical and seasonal conditions. *Stochastic Environmental Research and Risk Assessment*, 35(2):215–230.

Conferences

- **MASCOT-NUM** (Le Croisic, April 2023)
Presentation: An energy-based model approach to the estimation of rare event probabilities
- **Swiss Geoscience Meeting** (Lausanne, November 2022)
Poster: Sequential Monte Carlo for posterior risk assessment
- **IAMG** (Nancy, August 2022)
Presentation: Inference of geostatistical hyperparameters with the correlated pseudo-marginal method
- **GeoEnv** (Parma, June 2022)
Poster: Solving geophysical random effect models with intractable likelihoods: Linearized Gaussian approximations versus the correlated pseudomarginal method
- **EGU General Assembly** (Vienna, May 2022)
Presentation: Inference of (geostatistical) hyperparameters with the correlated pseudo-marginal method