Introduction to the Bayesian Approach to Inverse Problems - Part 2: Algorithms

Lecture 3

Claudia Schillings, Aretha Teckentrup[*,†]

[*]School of Mathematics, University of Edinburgh
[†]Alan Turing Institute, London

LMS Short Course - May 11, 2017

THE UNIVERSITY of EDINBURGH
School of Mathematics

The Alan Turing Institute

Contents

- The second half of this short course will focus on **algorithms in Bayesian inverse problems**, in particular algorithms for computing expectations with respect to the posterior distribution.
- The emphasis will be on **convergence properties** of the algorithms rather than implementation.
- The first lecture will focus on **standard Monte Carlo methods**: sampling methods based on independent and identically distributed (i.i.d.) samples.
- The second lecture will focus on **Markov chain Monte Carlo methods**: sampling methods based on correlated and approximate samples.

Outline of first lecture

1. Bayesian Inverse Problems
2. Standard Monte Carlo Method
3. Convergence of Standard Monte Carlo Method
4. Multilevel Monte Carlo Method
5. Convergence of Multilevel Monte Carlo Method

Bayesian Inverse Problems

Mathematical Formulation [Kaipio, Somersalo '04] [Stuart '10]

- We are interested in the following inverse problem: given observational data $y \in \mathbb{R}^J$, determine model parameter $u \in \mathbb{R}^n$ such that

$$y = \mathcal{G}(u) + \eta,$$

  where $\eta \sim N(0, \Gamma)$ represents observational noise.

- In the Bayesian approach, the solution to the inverse problem is the posterior distribution μ^y on $\mathbb{R}^n$, given by

$$\frac{d\mu^y}{d\mu_0}(u) = \frac{1}{Z} \exp\big(-\Phi(u;y)\big),$$

  where $Z = \mathbb{E}_{\mu_0}\Big(\exp\big(-\Phi(\cdot;y)\big)\Big)$ and $\Phi(u;y) = \frac{1}{2}\|y - \mathcal{G}(u)\|_\Gamma^2$.

Bayesian Inverse Problems

Computing expectations

- We will here focus on computing the expected value of a quantity of interest $\phi(u)$, $\phi : \mathbb{R}^n \to \mathbb{R}$, under the posterior distribution μ^y.

- In most cases, we do not have a closed form expression for the posterior distribution μ^y, since the normalising constant Z is not known explicitly.
  (Exception: forward map $\mathcal{G}$ linear and prior μ_0 Gaussian $\Rightarrow$ posterior μ^y also Gaussian.)

- However, the prior distribution is known in closed form, and furthermore often has a simple structure (e.g. multivariate Gaussian or independent uniform).

Bayesian Inverse Problems

Computing Expectations

Using Bayes' Theorem, we can write $\mathbb{E}_{\mu^y}[\phi]$ as

$$\begin{aligned}\mathbb{E}_{\mu^y}[\phi] &= \int_{\mathbb{R}^n} \phi(u)\, d\mu^y(u) \\ &= \int_{\mathbb{R}^n} \phi(u) \frac{d\mu^y}{d\mu_0}(u)\, d\mu_0(u) \\ &= \frac{1}{Z} \int_{\mathbb{R}^n} \phi(u) \exp[-\Phi(u; y)]\, d\mu_0(u) \\ &= \frac{\mathbb{E}_{\mu_0}[\phi \exp[-\Phi(\cdot; y)]]}{\mathbb{E}_{\mu_0}[\exp[-\Phi(\cdot; y)]]}.\end{aligned}$$

We have rewritten the posterior expectation as a ratio of two prior expectations.

We will now use Monte Carlo methods to estimate the two prior expectations.

Standard Monte Carlo Method

Sampling methods and random number generators

- The standard Monte Carlo method is a *sampling method*.
- To estimate $\mathbb{E}_{\mu_0}[f]$, for some $f : \mathbb{R}^n \to \mathbb{R}$, sampling methods use a sample average:

$$\mathbb{E}_{\mu_0}[f] = \int_{\mathbb{R}^n} f(u)\, d\mu_0(u) \approx \sum_{i=1}^{N} w_i\, f(u^{(i)}),$$

  where the choice of samples $\{u^{(i)}\}_{i=1}^N$ and weights $\{w_i\}_{i=1}^N$ determines the sampling method.
- In standard Monte Carlo, $w_i = \frac{1}{N}$ and $\{u^{(i)}\}_{i=1}^N$ is a sequence of independent and identically distributed (i.i.d.) random variables: $\{u^{(i)}\}_{i=1}^N$ are mutually independent and $u^{(i)} \sim \mu_0$, for all $1 \leq i \leq N$.
- Since μ_0 is fully known and simple, i.i.d. samples from μ_0 can be generated on a computer using a (pseudo-)random number generator. For more details, see [Robert, Casella '99], [L'Ecuyer '11].

Standard Monte Carlo Method

Definition of Monte Carlo Estimator

- In the Bayesian inverse problem, we want to compute

$$\mathbb{E}_{\mu^y}[\phi] = \frac{\mathbb{E}_{\mu_0}[\phi \exp[-\Phi(\cdot; y)]]}{\mathbb{E}_{\mu_0}[\exp[-\Phi(\cdot; y)]]}.$$

- Using Monte Carlo, we approximate this by

$$\mathbb{E}_{\mu_0}[\phi \exp[-\Phi(\cdot; y)]] \approx \frac{1}{N} \sum_{i=1}^{N} \phi(u^{(i)}) \exp[-\Phi(u^{(i)}; y)],$$

$$\mathbb{E}_{\mu_0}[\exp[-\Phi(\cdot; y)]] \approx \frac{1}{N} \sum_{i=1}^{N} \exp[-\Phi(u^{(i)}; y)],$$

where $\{u^{(i)}\}_{i=1}^N$ is an i.i.d. sequence distributed according to μ_0.

(It is also possible to use different samples in the two estimators.)

Standard Monte Carlo Method

Definition of Monte Carlo Estimator

- In applications, it is usually not possible to evaluate ϕ and Φ exactly, since this involves the solution of the forward problem.
  - In the groundwater flow example, it involves the solution of a PDE.
- Denote by ϕ_h and Φ_h numerical approximations to ϕ and Φ, respectively, where h is the step length of the numerical method.
- The computable Monte Carlo ratio estimator of $\mathbb{E}_{\mu^y}[\phi]$ is then

$$\mathbb{E}_{\mu^y}[\phi] \approx \frac{\frac{1}{N}\sum_{i=1}^{N} \phi_h(u^{(i)}) \exp[-\Phi_h(u^{(i)};y)]}{\frac{1}{N}\sum_{i=1}^{N} \exp[-\Phi_h(u^{(i)};y)]} := \frac{\widehat{Q}_{h,N}^{\mathrm{MC}}}{\widehat{Z}_{h,N}^{\mathrm{MC}}}.$$

- There are two sources of error in the Monte Carlo ratio estimator:
  - the sampling error due to using Monte Carlo,
  - the discretisation error due to the numerical approximation.

Convergence of Standard Monte Carlo Method

Expected Value and Variance [Billingsley '95]

Consider a general Monte Carlo estimator $\widehat{E}^{\mathrm{MC}}_{h,N} = \frac{1}{N}\sum_{i=1}^N f_h(u^{(i)})$, with $\{u^{(i)}\}_{i=1}^N$ an i.i.d. sequence distributed as μ_0.

Lemma (Expected Value and Variance)

$$\mathbb{E}[\widehat{E}^{\mathrm{MC}}_{h,N}] = \mathbb{E}_{\mu_0}[f_h], \qquad \mathbb{V}[\widehat{E}^{\mathrm{MC}}_{h,N}] = \frac{\mathbb{V}_{\mu_0}[f_h]}{N}.$$

Proof: Since $\{u^{(i)}\}_{i=1}^N$ is an i.i.d. sequence, we have

$$\mathbb{E}\Big[\frac{1}{N}\sum_{i=1}^N f_h(u^{(i)})\Big] = \frac{1}{N}\mathbb{E}\Big[\sum_{i=1}^N f_h(u^{(i)})\Big] = \frac{1}{N}\sum_{i=1}^N \mathbb{E}_{\mu_0}[f_h] = \mathbb{E}_{\mu_0}[f_h],$$

and

$$\mathbb{V}\Big[\frac{1}{N}\sum_{i=1}^N f_h(u^{(i)})\Big] = \frac{1}{N^2}\mathbb{V}\Big[\sum_{i=1}^N f_h(u^{(i)})\Big] = \frac{1}{N^2}\sum_{i=1}^N \mathbb{V}_{\mu_0}[f_h] = \frac{1}{N}\mathbb{V}_{\mu_0}[f_h].$$

Convergence of Standard Monte Carlo Method

Central Limit Theorem [Billingsley '95]

Theorem (Central Limit Theorem)

If $\mathbb{V}_{\mu_0}[f_h] \in (0, \infty)$*, then as* $N \to \infty$ *we have*

$$\widehat{E}_{h,N}^{\mathrm{MC}} \xrightarrow{D} \mathcal{N}(\mathbb{E}_{\mu_0}[f_h], \frac{\mathbb{V}_{\mu_0}[f_h]}{N}).$$

Here, $\xrightarrow{D}$ denotes convergence in distribution, i.e. point-wise convergence of the distribution function: with $X \sim \mathcal{N}(\mathbb{E}_{\mu_0}[f_h], \frac{\mathbb{V}_{\mu_0}[f_h]}{N})$,

$$\Pr[\, \widehat{E}_{h,N}^{\mathrm{MC}} \le x \,] \to \Pr[\, X \le x \,], \qquad \forall x \in \mathbb{R}.$$

The Central Limit Theorem crucially uses the fact that $\widehat{E}_{h,N}^{\mathrm{MC}}$ is based on i.i.d. samples. If this is not the case, we require stronger assumptions and/or obtain a different limiting distribution.

Convergence of Standard Monte Carlo Method

Strong Law of Large Numbers [Billingsley '95]

Theorem (Strong Law of Large Numbers)

If $\mathbb{E}_{\mu_0}[|f_h|] < \infty$, *then as* $N \to \infty$ *we have*

$$\widehat{E}^{\mathrm{MC}}_{h,N} \xrightarrow{a.s.} \mathbb{E}_{\mu_0}[f_h].$$

Here, $\xrightarrow{a.s.}$ denotes almost sure convergence, i.e. convergence with probability 1:

$$\mathrm{Pr}\big[\, \widehat{E}^{\mathrm{MC}}_{h,N} \to \mathbb{E}_{\mu_0}[f_h] \,\big] = 1.$$

The Strong Law of Large Numbers crucially uses the fact that $\widehat{E}^{\mathrm{MC}}_{h,N}$ is based on i.i.d. samples. If this is not the case, we may require stronger assumptions

Convergence of Standard Monte Carlo Method

Mean Square Error [Billingsley '95]

A measure of accuracy of $\widehat{E}_{h,N}^{\mathrm{MC}} = \frac{1}{N}\sum_{i=1}^{N} f_h(u^{(i)})$ as an estimator of $\mathbb{E}_{\mu_0}[f]$ is given by the mean square error (MSE):

$$e(\widehat{E}_{h,N}^{\mathrm{MC}})^2 := \mathbb{E}[(\widehat{E}_{h,N}^{\mathrm{MC}} - \mathbb{E}_{\mu_0}[f])^2].$$

Theorem (Mean Square Error)

$$e(\widehat{E}_{h,N}^{\mathrm{MC}})^2 = \underbrace{\frac{\mathbb{V}_{\mu_0}[f_h]}{N}}_{\textit{sampling error}} + \underbrace{(\mathbb{E}_{\mu_0}[f_h - f])^2}_{\textit{numerical error}}.$$

Proof: Since $\mathbb{E}[\widehat{E}_{h,N}^{\mathrm{MC}}] = \mathbb{E}_{\mu_0}[f_h]$ and $\mathbb{V}[\widehat{E}_{h,N}^{\mathrm{MC}}] = \frac{\mathbb{V}_{\mu_0}[f_h]}{N}$, we have

$$\begin{aligned} e(\widehat{E}_{h,N}^{\mathrm{MC}})^2 &= \mathbb{E}\Big[\big(\widehat{E}_{h,N}^{\mathrm{MC}} - \mathbb{E}_{\mu_0}[f_h] + \mathbb{E}_{\mu_0}[f_h] - \mathbb{E}_{\mu_0}[f]\big)^2\Big] \\ &= \mathbb{E}\Big[\big(\widehat{E}_{h,N}^{\mathrm{MC}} - \mathbb{E}_{\mu_0}[f_h]\big)^2\Big] + \mathbb{E}\Big[\big(\mathbb{E}_{\mu_0}[f_h - f]\big)^2\Big] \\ &= \frac{\mathbb{V}_{\mu_0}[f_h]}{N} + (\mathbb{E}_{\mu_0}[f_h - f])^2. \end{aligned}$$

Convergence of Standard Monte Carlo Method

Mean Square Error of Monte Carlo ratio estimator [Scheichl, Stuart, ALT '16]

- Recall: $\mathbb{E}_{\mu^y}[\phi] = \frac{\mathbb{E}_{\mu_0}[\phi \exp[-\Phi(\cdot;y)]]}{\mathbb{E}_{\mu_0}[\exp[-\Phi(\cdot;y)]]} =: \frac{Q}{Z} \approx \frac{\widehat{Q}^{\rm MC}_{h,N}}{\widehat{Z}^{\rm MC}_{h,N}}$.
- We know how to bound the MSEs of the individual estimators $\widehat{Q}^{\rm MC}_{h,N}$ and $\widehat{Z}^{\rm MC}_{h,N}$. Can we bound the MSE of $\widehat{Q}^{\rm MC}_{h,N}/\widehat{Z}^{\rm MC}_{h,N}$?
- Rearranging the MSE and applying the triangle inequality, we have

$$
\begin{aligned}
e\Big(\frac{\widehat{Q}^{\rm MC}_{h,N}}{\widehat{Z}^{\rm MC}_{h,N}}\Big)^2 &= \mathbb{E}\Big[\Big(\frac{Q}{Z} - \frac{\widehat{Q}^{\rm MC}_{h,N}}{\widehat{Z}^{\rm MC}_{h,N}}\Big)^2\Big] \\
&\leq \frac{2}{Z^2}\Big(\mathbb{E}\big[(Q - \widehat{Q}^{\rm MC}_{h,N})^2\big] + \mathbb{E}\big[(\widehat{Q}^{\rm MC}_{h,N}/\widehat{Z}^{\rm MC}_{h,N})^2(Z - \widehat{Z}^{\rm MC}_{h,N})^2\big]\Big).
\end{aligned}
$$

Convergence of Standard Monte Carlo Method

Mean Square Error of Monte Carlo ratio estimator [Scheichl, Stuart, ALT '16]

$$e\Big(\frac{\widehat{Q}^{\mathrm{MC}}_{h,N}}{\widehat{Z}^{\mathrm{MC}}_{h,N}}\Big)^2 \leq \frac{2}{Z^2}\Big(\mathbb{E}\big[(Q-\widehat{Q}^{\mathrm{MC}}_{h,N})^2\big] + \mathbb{E}\big[(\widehat{Q}^{\mathrm{MC}}_{h,N}/\widehat{Z}^{\mathrm{MC}}_{h,N})^2(Z-\widehat{Z}^{\mathrm{MC}}_{h,N})^2\big]\Big)$$

Theorem (Hölder's Inequality)

For any random variables X, Y *and* $p, q \in [1,\infty]$*, with* $p^{-1}+q^{-1}=1$*,*

$$\mathbb{E}[|XY|] \leq \mathbb{E}[|X|^p]^{1/p}\,\mathbb{E}[|Y|^q]^{1/q}.$$

Here, $\mathbb{E}[|X|^\infty]^{1/\infty} := \operatorname{ess\,sup} X$*.*

If $\operatorname{ess\,sup}_{\{u^{(i)}\}_{i=1}^N}(\widehat{Q}^{\mathrm{MC}}_{h,N}/\widehat{Z}^{\mathrm{MC}}_{h,N})^2 \leq C$, for a constant C independent of N and h, then the MSE of $\widehat{Q}^{\mathrm{MC}}_{h,N}/\widehat{Z}^{\mathrm{MC}}_{h,N}$ can be bounded in terms of the individual MSEs of $\widehat{Q}^{\mathrm{MC}}_{h,N}$ and $\widehat{Z}^{\mathrm{MC}}_{h,N}$.

For more details, see [Scheichl, Stuart, ALT '16].

Multilevel Monte Carlo Method

Motivation

The standard Monte Carlo estimator $\widehat{E}^{\mathrm{MC}}_{h,N} = \frac{1}{N}\sum_{i=1}^{N} f_h(u^{(i)})$ of $\mathbb{E}_{\mu_0}[f]$ has mean square error

$$e(\widehat{E}^{\mathrm{MC}}_{h,N})^2 = \frac{\mathbb{V}_{\mu_0}[f_h]}{N} + (\mathbb{E}_{\mu_0}[f_h - f])^2.$$

To make $e(\widehat{E}^{\mathrm{MC}}_{h,N})^2$, small we need to

- choose a large number of samples N,
- choose a small step length h in our numerical approximation.

Since the cost of sampling methods grows as

cost per sample $\times$ number of samples

the cost of standard Monte Carlo can be prohibitively large in applications.

Multilevel Monte Carlo Method

Definition of Multilevel Monte Carlo Estimator [Giles, '08], [Heinrich '01]

The multilevel method works with a decreasing **sequence of step lengths** $\{h_\ell\}_{\ell=0}^{L}$, where h_L gives the most accurate numerical approximation.

Linearity of expectation gives us

$$\mathbb{E}_{\mu_0}\left[f_{h_L}\right] = \mathbb{E}_{\mu_0}\left[f_{h_0}\right] + \sum_{\ell=1}^{L} \mathbb{E}_{\mu_0}\left[f_{h_\ell} - f_{h_{\ell-1}}\right].$$

The **multilevel Monte Carlo (MLMC)** estimator

$$\widehat{E}^{\mathrm{ML}}_{\{M_\ell,N_\ell\}} = \frac{1}{N_0}\sum_{i=1}^{N_0} f_{h_0}(u^{(i,0)}) + \sum_{\ell=1}^{L} \frac{1}{N_\ell}\sum_{i=1}^{N_\ell} f_{h_\ell}(u^{(i,\ell)}) - f_{h_{\ell-1}}(u^{(i,\ell)}),$$

is a sum of $L+1$ independent MC estimators.

Convergence of Multilevel Monte Carlo Method

Expected Value and Variance [Giles, '08]

$$\widehat{E}^{\mathrm{ML}}_{\{M_\ell,N_\ell\}} = \frac{1}{N_0}\sum_{i=1}^{N_0} f_{h_0}(u^{(i,0)}) + \sum_{\ell=1}^{L}\frac{1}{N_\ell}\sum_{i=1}^{N_\ell} f_{h_\ell}(u^{(i,\ell)}) - f_{h_{\ell-1}}(u^{(i,\ell)})$$

Lemma (Expected Value and Variance)

$$\mathbb{E}[\widehat{E}^{\mathrm{ML}}_{\{M_\ell,N_\ell\}}] = \mathbb{E}_{\mu_0}[f_{h_L}], \quad \mathbb{V}[\widehat{E}^{\mathrm{ML}}_{\{M_\ell,N_\ell\}}] = \frac{\mathbb{V}[f_{h_0}]}{N_0} + \sum_{\ell=1}^{L}\frac{\mathbb{V}[f_{h_\ell} - f_{h_{\ell-1}}]}{N_\ell}.$$

Proof: Uses the linearity of expectation and the fact the $L+1$ estimators are independent, together with results for standard Monte Carlo.

Convergence of Multilevel Monte Carlo Method

Central Limit Theorem and Strong Law of Large Numbers [Billingsley '95]

Theorem (Central Limit Theorem)

If $\sigma_{\mathrm{ML}}^2 := \mathbb{V}[f_{h_0}]N_0^{-1} + \sum_{\ell=1}^{L} \mathbb{V}[f_{h_\ell} - f_{h_{\ell-1}}]N_\ell^{-1} \in (0,\infty)$ *and* $\{\mathbb{V}[f_{h_\ell} - f_{h_{\ell-1}}]\}_{\ell=1}^{L}$ *satisfies a Lindeberg condition, then as* $\{N_\ell\}_{\ell=0}^{L} \to \infty$ *we have*

$$\widehat{E}^{\mathrm{ML}}_{\{M_\ell,N_\ell\}} \xrightarrow{D} \mathcal{N}(\mathbb{E}_{\mu_0}[f_{h_L}], \sigma_{\mathrm{ML}}^2).$$

Proof: Requires Lindeberg condition to deal with sum of $L+1$ Monte Carlo estimators. For details, see [Collier et al '15] and [Billingsley '95].

Theorem (Strong Law of Large Numbers)

If $\mathbb{E}_{\mu_0}[|f_{h_\ell}|] < \infty$ *for* $0 \le \ell \le L$*, then as* $\{N_\ell\}_{\ell=0}^{L} \to \infty$ *we have*

$$\widehat{E}^{\mathrm{ML}}_{\{M_\ell,N_\ell\}} \xrightarrow{a.s.} \mathbb{E}_{\mu_0}[f_{h_L}].$$

Proof: Follows from the linearity of a.s. convergence, together with results for standard Monte Carlo.

Convergence of Multilevel Monte Carlo Method

Mean Square Error of Multilevel Monte Carlo [Giles, '08]

Theorem (Mean Square Error)

$$e(\widehat{E}^{\mathrm{ML}}_{\{M_\ell,N_\ell\}})^2 = \underbrace{\frac{\mathbb{V}[f_{h_0}]}{N_0} + \sum_{\ell=1}^{L} \frac{\mathbb{V}[f_{h_\ell} - f_{h_{\ell-1}}]}{N_\ell}}_{\textit{sampling error}} + \underbrace{(\mathbb{E}_{\mu_0}[f_{h_L} - f])^2}_{\textit{numerical error}}.$$

Proof: The derivation is identical to the standard Monte Carlo case.

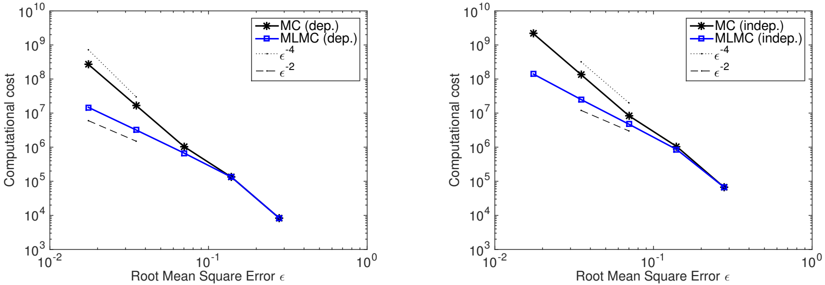
Thus,

- N_0 still needs to be large, **but** samples are much cheaper to obtain on coarse grid.
- N_ℓ $(\ell > 0)$ much smaller, **since** $\mathbb{V}[f_{h_\ell} - f_{h_{\ell-1}}] \to 0$ as $h_\ell \to 0$.

Convergence of Multilevel Monte Carlo Method

Numerical Comparison: Mean Square Error

- We compute $\mathbb{E}_{\mu^y}[\phi]$ for a typical model problem in groundwater flow, using a ratio of standard Monte Carlo and multilevel Monte Carlo estimators.
- Computational Cost is computed as number of FLOPS required.



[Scheichl, Stuart, ALT '16]

References I

- P. Billingsley, *Probability and measure*, John Wiley & Sons, 1995.
- N. Collier, A.-L. Haji-Ali, F. Nobile, E. von Schwerin, and R. Tempone, *A continuation multilevel Monte Carlo algorithm*, BIT Numerical Mathematics, 55 (2015), pp. 399–432.
- S. Cotter, M. Dashti, and A. Stuart, *Variational data assimilation using targetted random walks*, International Journal for Numerical Methods in Fluids, 68 (2012), pp. 403–421.
- T. Dodwell, C. Ketelsen, R. Scheichl, and A. Teckentrup, *A hierarchical multilevel Markov chain Monte Carlo algorithm with applications to uncertainty quantification in subsurface flow*, SIAM/ASA Journal on Uncertainty Quantification, 3 (2015), pp. 1075–1108.

References II

- Y. Efendiev, B. Jin, M. Presho, and X. Tan, *Multilevel Markov Chain Monte Carlo Method for High-Contrast Single-Phase Flow Problems*, Communications in Computational Physics, 17 (2015), pp. 259–286.
- M. Giles, *Multilevel Monte Carlo Path Simulation*, Operations Research, 56 (2008), pp. 607–617.
- S. Heinrich, *Multilevel Monte Carlo Methods*, in International Conference on Large-Scale Scientific Computing, Springer, 2001, pp. 58–67.
- V. Hoang, C. Schwab, and A. Stuart, *Complexity analysis of accelerated MCMC methods for Bayesian inversion*, Inverse Problems, 29 (2013), p. 085010.
- J. Kaipio and E. Somersalo, *Statistical and computational inverse problems*, Springer, 2004.

References III

P. L'Ecuyer, *Random number generation*, in Handbook of Computational Statistics, Springer, 2011, pp. 35–71.

C. Robert and G. Casella, *Monte Carlo Statistical Methods*, Springer, 1999.

D. Rudolf, *Explicit error bounds for Markov chain Monte Carlo*, arXiv preprint arXiv:1108.3201, (2011).

R. Scheichl, A. Stuart, and A. Teckentrup, *Quasi-Monte Carlo and Multilevel Monte Carlo Methods for Computing Posterior Expectations in Elliptic Inverse Problems*, arXiv preprint arXiv:1602.04704, (2016).

A. Stuart, *Inverse Problems: A Bayesian Perspective*, Acta Numerica, 19 (2010), pp. 451–559.