The Bayesian Approach to Inverse Problems

Andrew M. Stuart
Mathematics Institute
Warwick University
CV4 7AL, UK
a.m.stuart@warwick.ac.uk

Abstract: These -ecture notes provide an introduction to the forthco ing boo DHS 3

Contents

1. Introduction

1.1. Bayesian Inversion $\mathbb{R}^n$

Consider the problem of finding $\mathbf{u} \in \mathbb{R}^n$ from $\mathbf{y} \in \mathbb{R}^J$ where $\mathbf{u}$ and $\mathbf{y}$ are related by the equation

$$\mathbf{y} = \mathbf{G}(\mathbf{u}).$$

We refer to $\mathbf{y}$ as **observed data** and to $\mathbf{u}$ as the **unknown**. This problem may be difficult for a number of reasons. We highlight two of these, both particularly relevant to our future developments.

1. The first difficulty, which may be illustrated in the case where $\mathbf{n} = \mathbf{J}$, concerns the fact that often the equation is perturbed by noise and so we should really consider the equation

   $$\mathbf{y} = \mathbf{G}(\mathbf{u}) + \boldsymbol{\eta}, \tag{1.1}$$

   where $\boldsymbol{\eta} \in \mathbb{R}^J$ represents the **observational noise** which enters the observed data. It may then be the case that, because of the noise, $\mathbf{y}$

- Let μ^y be measure on $\mathbb{R}^n$ with density ρ^y and μ_0 measure on $\mathbb{R}^n$ with density ρ_0. Then the conclusion of Theorem 1.1 may be written as:

$$\begin{aligned}\frac{d\mu^y}{d\mu_0}(u) &= \frac{1}{Z}\exp\Big(-\Phi(u;y)\Big),\\ Z &= \int_{\mathbb{R}^n}\exp\Big(-\Phi(u;y)\Big)\mu_0(du).\end{aligned} \tag{1.2}$$

  Thus the posterior is absolutely continuous with respect to the prior, and the Radon-Nikodym derivative is proportional to the likelihood. The expression for the Radon-Nikodym derivative is to be interpreted as the statement that, for all measurable $f : \mathbb{R}^n \to \mathbb{R}$,

$$\mathbb{E}^{\mu^y} f(u) = \mathbb{E}^{\mu_0}\Big(\frac{d\mu^y}{d\mu_0}(u) f(u)\Big).$$

  Alternatively we may write this in integral form as

$$\int_{\mathbb{R}^n} f(u)\mu^y(du) = \int_{\mathbb{R}^n}\Big(\frac{1}{Z}\exp\Big(-\Phi(u;y)\Big) f(u)\Big)\mu_0(du).$$

1.2. I e

We will be interested in the inverse problem of finding $\mathbf{u}$ from $\mathbf{y}$ where

$$\begin{aligned}\mathbf{y} &= \mathbf{v}(1)+\boldsymbol{\eta}\\ &= \mathbf{G}(\mathbf{u})+\boldsymbol{\eta}.\end{aligned}$$

Here $\boldsymbol{\eta} \in \mathbf{H}$ is noise and $\mathbf{G}(\mathbf{u}) := \mathbf{v}(1) = \mathbf{e}^{-A}\mathbf{u}$. Formally this looks like an infinite dimensional linear version of the inverse problem (1.1), extended from finite dimensions to a Hilbert space setting. However the operator $\mathbf{e}^{A} : \mathbf{H} \to \mathbf{H}$ is not continuous and so we need regularization to make sense of the problem. Thus, if the noise $\boldsymbol{\eta} \in \mathbf{H}$, it will not be possible to simply apply $\mathbf{G}^{-1}$ to $\mathbf{y}$, difficulty 1. from the preceding subsection. We will apply a Bayesian approach and hence will need to put probability measures on the Hilbert space $\mathbf{H}$; in particular we will want to study $\mathbb{P}(\mathbf{u})$, $\mathbb{P}(\mathbf{y}|\mathbf{u})$ and $\mathbb{P}(\mathbf{u}|\mathbf{y})$, all probability measures on $\mathbf{H}$.

1.3. Elliptic Inverse Problem

One motivation for adopting the Bayesian approach to inverse problems is that prior modelling is a transparent approach to dealing with under-determined inverse problems; it forms a rational approach to dealing with the second difficulty, labelled 2. in the previous subsection. The elliptic inverse problem we now describe is a concrete example of an under-determined inverse problem.

As in Section 1.2, $\mathbf{D} \subset \mathbb{R}^d$ denotes a bounded open set, with smooth boundary $\boldsymbol{\partial}\mathbf{D}$. We define the Hilbert spaces (Gelfand triple) $\mathbf{V} \subset \mathbf{H} \subset \mathbf{V}^*$ as follows:

$$\begin{aligned}&\mathbf{H} = \left(\mathbf{L}^2(\mathbf{D}), \langle\cdot,\cdot\rangle, \|\cdot\|\right);\\ &\mathbf{V} = \mathbf{H}_0^1(\mathbf{D}) \text{ with norm } \|\cdot\|_V = \|\nabla\cdot\|;\\ &\mathbf{V}^* \text{ dual space;}\\ &\|\cdot\| \le \mathbf{C}_p\|\cdot\|_V \text{ (Poincaré inequality).}\end{aligned}$$

Let $\boldsymbol{\kappa} \in \mathbf{X} := \mathbf{L}^\infty(\mathbf{D})$ satisfy

$$\operatorname*{ess\,inf}_{x\in D} \boldsymbol{\kappa}(\mathbf{x}) = \boldsymbol{\kappa}_{\min} > 0. \tag{1.4}$$

Now consider the equation

$$\begin{aligned}-\nabla\cdot(\boldsymbol{\kappa}\nabla\mathbf{p}) &= \mathbf{f}, \quad \mathbf{x}\in\mathbf{D}, &(1.5a)\\ \mathbf{p} &= 0, \quad \mathbf{x}\in\boldsymbol{\partial}\mathbf{D}. &(1.5b)\end{aligned}$$

Lemma 1.5. **Assume that $\mathbf{f} \in \mathbf{V}^*$ and that $\boldsymbol{\kappa}$ satisfies (1.4). Then (1.5) has a unique weak solution $\mathbf{p} \in \mathbf{V}$. This solution satisfies**

$$\|\mathbf{p}\|_V \le \|\mathbf{f}\|_{V^*}/\boldsymbol{\kappa}_{\min}$$

and, if $\mathbf{f} \in \mathbf{H}$,

$$\|\mathbf{p}\|_V \le \mathbf{C}_p\|\mathbf{f}\|/\boldsymbol{\kappa}_{\min}.$$

We will be interested in the inverse problem of finding $\boldsymbol{\kappa}$ from $\mathbf{y}$ where

$$\mathbf{y}_j = \mathbf{l}_j(\mathbf{p}) + \boldsymbol{\eta}_j, \quad \mathbf{j} = 1,\cdots,\mathbf{J}. \tag{1.6}$$

Here $\mathbf{l}_j \in \mathbf{V}^*$ is a continuous linear functional on $\mathbf{V}$ and $\boldsymbol{\eta}_j$ is a noise.

Notice that the unknown, $\boldsymbol{\kappa} \in \mathbf{L}^\infty(\mathbf{D})$, is a function (infinite dimensional) whereas the data from which we wish to determine $\boldsymbol{\kappa}$ is finite dimensional: $\mathbf{y} \in \mathbb{R}^J$. The problem is severely under-determined, illustrating point 2. from the previous subsection. It is natural to treat such problems in the Bayesian framework, using prior modeling to fill-in missing information. We will take the unknown function to be $\mathbf{u}$ where either $\mathbf{u} = \boldsymbol{\kappa}$ or $\mathbf{u} = \log\boldsymbol{\kappa}$. In either case, we will define $\mathbf{G}_j(\mathbf{u}) = \mathbf{l}_j(\mathbf{p})$ and then (1.6) may be written as

$$\mathbf{y} = \mathbf{G}(\mathbf{u}) + \boldsymbol{\eta} \tag{1.7}$$

where $\mathbf{y}, \boldsymbol{\eta} \in \mathbb{R}^J$ and $\mathbf{G} : \mathbf{X}' \subseteq \mathbf{X} \to \mathbb{R}^J$. The set $\mathbf{X}'$ is introduced because $\mathbf{G}$ may not be defined on the whole of $\mathbf{X}$. In particular, the positivity constraint (1.4) is only satisfied on

$$\mathbf{X}' := \left\{ \mathbf{u} \in \mathbf{X} : \operatorname*{ess\,inf}_{x \in D} \mathbf{u}(\mathbf{x}) > 0 \right\} \subset \mathbf{X}$$

in the case where $\boldsymbol{\kappa} = \mathbf{u}$. On the other hand if $\boldsymbol{\kappa} = \exp(\mathbf{u})$ then the positivity constraint (1.4) is satisfied for any $\mathbf{u} \in \mathbf{X}$.

Notice that we again need probability measures on function space, here the Banach space $\mathbf{X} = \mathbf{L}^\infty(\mathbf{D})$. Furthermore, these probability measures should charge only positive functions, in view of the desired inequality (1.4).

2. Prior Modeling

2.1. General Setting

We let $\{\boldsymbol{\varphi}_j\}_{j=0}^\infty$ denote an infinite sequence in the Banach space $\left(\mathbf{X}, \|\cdot\|\right)$ of $\mathbb{R}-$valued functions defined on $\mathbf{D} \subset \mathbb{R}^d$, a bounded, open set with smooth boundary. (The extension to $\mathbb{R}^n-$valued functions is straightforward, but omitted for brevity). We normalize these functions so that $\|\boldsymbol{\varphi}_j\| = 1$ for $\mathbf{j} = 1, \cdots, \infty$; we do not assume that $\boldsymbol{\varphi}_0$ is normalized. Define the function $\mathbf{u}$ by

$$\mathbf{u} = \boldsymbol{\varphi}_0 + \sum_{j=1}^{\infty} \mathbf{u}_j \boldsymbol{\varphi}_j. \tag{2.1}$$

By randomizing $\mathsf{u} := \{\mathbf{u}_j\}_{j=1}^\infty$ we create random functions. To this end we define the deterministic sequence $\boldsymbol{\gamma} = \{\boldsymbol{\gamma}_j\}_{j=1}^\infty$ and the i.i.d. random sequence $\boldsymbol{\xi} = \{\boldsymbol{\xi}_j\}_{j=1}^\infty$, and set $\mathbf{u}_j = \boldsymbol{\gamma}_j \boldsymbol{\xi}_j$. We let $\left(\Omega, \mathcal{F}, \mathbb{P}\right)$ denote the probability space for the i.i.d. sequence $\boldsymbol{\xi} \in \Omega = \mathbb{R}^\infty$, with $\mathbb{E}$ denoting expectation. In the next three subsections we demonstrate how this general setting may be adapted to create a variety of useful prior measures on function space. On occasion we will find it useful to consider the truncated random functions

$$\mathbf{u}^N = \boldsymbol{\varphi}_0 + \sum_{j=1}^{N} \mathbf{u}_j \boldsymbol{\varphi}_j, \quad \mathbf{u}_j = \boldsymbol{\gamma}_j \boldsymbol{\xi}_j. \tag{2.2}$$

2.2. Uniform Priors

Choose $\mathbf{X} = \mathbf{L}^\infty(\mathbf{D})$. Assume $\mathbf{u}_j = \boldsymbol{\gamma}_j \boldsymbol{\xi}_j$ with $\boldsymbol{\xi} = \{\boldsymbol{\xi}_j\}_{j=1}^\infty$ an i.i.d. sequence with $\boldsymbol{\xi}_1 \sim \mathbf{U}[-1, 1]$ and $\boldsymbol{\gamma} = \{\boldsymbol{\gamma}_j\}_{j=1}^\infty \in \ell^1$. Assume further that there are finite, positive constants $\boldsymbol{\varphi}_{\min}, \boldsymbol{\varphi}_{\max}, \boldsymbol{\delta} > 0$ such that

$$\begin{aligned}
&\operatorname*{ess\,inf}_{x \in D} \boldsymbol{\varphi}_0(\mathbf{x}) \geq \boldsymbol{\varphi}_{\min}; \\
&\operatorname*{ess\,sup}_{x \in D} \boldsymbol{\varphi}_0(\mathbf{x}) \leq \boldsymbol{\varphi}_{\max}; \\
&\|\boldsymbol{\gamma}\|_{\ell^1} = \frac{\boldsymbol{\delta}}{1 + \boldsymbol{\delta}} \boldsymbol{\varphi}_{\min}.
\end{aligned}$$

Theorem 2.1. **The following holds $\mathbb{P}-$almost surely: the sequence of functions $\{\mathbf{u}^N\}_{N=1}^\infty$ given by (2.2) is Cauch in $\mathbf{X}$ and the limiting function $\mathbf{u}$ given by (2.1) satisfies**

$$\frac{1}{1 + \boldsymbol{\delta}} \boldsymbol{\varphi}_{\min} \leq \mathbf{u}(\mathbf{x}) \leq \boldsymbol{\varphi}_{\max} + \frac{\boldsymbol{\delta}}{1 + \boldsymbol{\delta}} \boldsymbol{\varphi}_{\min} \quad \text{a.e.} \quad \mathbf{x} \in \mathbf{D}.$$

Proof. Let $N > M$. Then, $\mathbb{P}$–a.s.,

$$
\begin{aligned}
\|u^N - u^M\|_\infty &= \Big\| \sum_{j=M+1}^{N} u_j \phi_j \Big\|_\infty \\
&\le \Big\| \sum_{j=M+1}^{N} \gamma_j \xi_j \phi_j \Big\|_\infty \\
&\le \sum_{j=M+1}^{\infty} |\gamma_j||\xi_j|\|\phi_j\|_\infty \\
&\le \sum_{j=M+1}^{\infty} |\gamma_j|.
\end{aligned}
$$

The right hand side tends to zero as $M \to \infty$ by the dominated convergence theorem and hence the sequence is Cauchy in X.

We have $\mathbb{P}$–a.s. and for a.e. $x \in D$,

$$
\begin{aligned}
u(x) &\ge \phi_0(x) - \sum_{j=1}^{\infty} |u_j|\|\phi_j\|_\infty \\
&\ge \operatorname*{ess\,inf}_{x\in D} \phi_0(x) - \sum_{j=1}^{\infty} |\gamma_j| \\
&\ge \phi_{\min} - \|\gamma\|_{\ell^1} \\
&= \frac{1}{1+\delta}\phi_{\min}.
\end{aligned}
$$

Proof of the upper bound is similar. □

Example Consider the random function (2.1) as specified in this section. By Theorem 2.1 we have that, $\mathbb{P}$–a.s.,

$$
u(x) \ge \frac{1}{1+\delta}\phi_{\min} > 0, \quad \text{a.e.} \quad x \in D. \tag{2.3}
$$

Set $\kappa = u$ in the elliptic equation (1.4), so that the coefficient κ in the equation and the solution p are random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. Since (2.3) holds $\mathbb{P}$–a.s., Lemma 1.5 shows that, again $\mathbb{P}$–a.s.,

$$
\|p\|_V \le (1+\delta)\|f\|_{V^*}/\phi_{\min}.
$$

Since the r.h.s. is non-random we have that for all $r \in \mathbb{Z}^+$ the random variable $p \in L^r_{\mathbb{P}}(\Omega; V)$:

$$
\mathbb{E}\|p\|_V^r < \infty.
$$

In fact $\mathbb{E}\exp(\alpha\|p\|_V^r) < \infty$ for all $r \in \mathbb{Z}^+$ and $\alpha \in (0,\infty)$. □

2.3. Besov Priors

Now we set $\phi_0 = 0$ and let $\{\phi_j\}_{j=1}^{\infty}$ be an orthonormal basis for X. Let

$$
X := \dot{L}^2(\mathbb{T}^d) = \Big\{ u \Big| \int_{\mathbb{T}^d} |u(x)|^2 dx < \infty, \int_{\mathbb{T}^d} u(x)dx = 0 \Big\}
$$

for $d \le 3$ with inner-product and norm denoted by $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$ respectively. Then, for any $u \in X$, we have

$$
u(x) = \sum_{j=1}^{\infty} u_j \phi_j, \quad u_j = \langle u, \phi_j\rangle. \tag{2.4}
$$

Define $\zeta_j = j^{(t-s)q/d}|\xi_j|^q$. Using the fact that the ζ_j are non-negative and independent we deduce from Lemma 2.3 that

$$\sum_{j=1}^{\infty} \mathbb{E}\big(\zeta_j \wedge 1\big) = \sum_{j=1}^{\infty} \mathbb{E}\Big(j^{(t-s)q/d}|\xi_j|^q \wedge 1\Big) < \infty.$$

This implies that $t < s$. We note that then

$$\begin{aligned}
\mathbb{E}\zeta_j &= \mathbb{E}\Big(j^{-(s-t)q/d}|\xi_j|^q\Big)\\
&= \mathbb{E}\Big(j^{-(s-t)q/d}|\xi_j|^q \mathbb{I}_{\{|\xi_j|\le j^{(s-t)/d}\}}\Big) + \mathbb{E}\Big(j^{-(s-t)q/d}|\xi_j|^q \mathbb{I}_{\{|\xi_j|>j^{(s-t)/d}\}}\Big)\\
&\le \mathbb{E}\Big(\big(\zeta_j \wedge 1\big)\mathbb{I}_{\{|\xi_j|\le j^{(s-t)/d}\}}\Big) + I\\
&\le \mathbb{E}\big(\zeta_j \wedge 1\big) + I,
\end{aligned}$$

where

$$I \propto j^{-(s-t)q/d}\int_{j^{(s-t)/d}}^{\infty} x^q e^{-x^q/2}dx.$$

Noting that, since $q \ge 1$, the function $x \mapsto x^q e^{-x^q/2}$ is bounded, up to a constant of proportionality, by the function $x \mapsto e^{-\alpha x}$ for any $\alpha < \frac{1}{2}$, we see that there is a positive constant K such that

$$\begin{aligned}
I &\le K j^{-(s-t)q/d}\int_{j^{(s-t)/d}}^{\infty} e^{-\alpha x}dx\\
&= \frac{1}{\alpha} K j^{-(s-t)q/d}\exp\Big(-\alpha j^{(s-t)/d}\Big)\\
&:= \iota_j.
\end{aligned}$$

Thus we have shown that

$$\sum_{j=1}^{\infty} \mathbb{E}\Big(j^{-(s-t)q/d}|\xi_j|^q\Big) \le \sum_{j=1}^{\infty} \mathbb{E}\big(\zeta_j \wedge 1\big) + \sum_{j=1}^{\infty} \iota_j < \infty.$$

Since the ξ_j are i.i.d. this implies that

$$\sum_{j=1}^{\infty} j^{(t-s)q/d} < \infty,$$

from which it follows that $(s-t)q/d > 1$ and (iii) follows. □

Lemma 2.3. *Let $\{I_j\}_{j=1}^{\infty}$ be an independent sequence of $\mathbb{R}^+$−valued random variable. Then*

$$\sum_{j=1}^{\infty} I_j < \infty \quad \text{a.s.} \Leftrightarrow \sum_{j=1}^{\infty} \mathbb{E}(I_j \wedge 1) < \infty.$$

2.4. Gaussian Priors

Let X be a Hilbert space $\mathcal{H}$ with inner-product and norm denoted by $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$ respectively. Assume that $\{\varphi_j\}_{j=1}^{\infty}$ is an orthonormal basis for $\mathcal{H}$. Define

$$\mathcal{H}^t = \Big\{u \Big| \sum_{j=1}^{\infty} j^{\frac{2t}{d}}|u_j|^2 < \infty, \quad u_j = \langle u, \varphi_j\rangle\Big\}. \tag{2.5}$$

As in the Section 2.3, we consider the setting in which $\varphi_0 = 0$ so that function u is given by (2.4). We choose $\xi_1 \sim \mathcal{N}(0,1)$ and $\gamma_j \asymp j^{-\frac{s}{d}}$. We are interested in convergence of the following series, found from (2.2) with $\varphi_0 = 0$:

$$u^N = \sum_{j=1}^{N} u_j \varphi_j, \quad u_j = \gamma_j \xi_j. \tag{2.6}$$

To understand this sequence of functions, indexed by N, it is useful to introduce the following function space:

$$L^2_{\mathbb{P}}(\Omega; \mathcal{H}^t) := \left\{ v : \Omega \times D \to \mathbb{R} \,\middle|\, \mathbb{E}\|v\|^2_{\mathcal{H}^t} < \infty \right\}.$$

This is in fact a Hilbert space.

Theorem 2.4. *The sequence of functions $\{u^N\}_{N=1}^{\infty}$ is Cauchy in the Hilbert space $L^2_{\mathbb{P}}(\Omega; \mathcal{H}^t)$, $t < s - \frac{d}{2}$. Thus the infinite series*

$$u(x) = \sum_{j=1}^{\infty} u_j \varphi_j(x), \quad u_j = \gamma_j \xi_j \tag{2.7}$$

exists as an L^2-limit and takes values in $\mathcal{H}^t$ for $t < s - \frac{d}{2}$.

Proof. For $N > M$,

$$\begin{aligned} \mathbb{E}\|u^N - u^M\|^2_{\mathcal{H}^t} &= \mathbb{E} \sum_{j=M+1}^{N} j^{\frac{2t}{d}} |u_j|^2 \\ &\asymp \sum_{j=M+1}^{N} j^{\frac{2(t-s)}{d}} \leq \sum_{j=M+1}^{\infty} j^{\frac{2(t-s)}{d}}. \end{aligned}$$

The sum on the right hand side tends to 0 as $M \to \infty$, provided $\frac{2(t-s)}{d} < -1$, by the dominated convergence theorem. This completes the proof. □

Remarks 2.5. *We make the following remarks concerning the Gaussian random functions constructed in the preceding theorem.*

- *The preceding theorem shows that the sum (2.6) has an $L^2_{\mathbb{P}}$ limit in $\mathcal{H}^t$ when $t < s - d/2$. The same methods used to prove Theorem 2.2 show that the sum also has an almost sure limit in $\mathcal{H}^t$ when $t < s - d/2$. Indeed, for $t < s - \frac{d}{2}$,*

$$\begin{aligned} \mathbb{E}\|u\|^2_{\mathcal{H}^t} &= \sum_{j=1}^{\infty} j^{\frac{2t}{d}} \mathbb{E}(\gamma_j^2 \xi_j^2) \\ &= \sum_{j=1}^{\infty} j^{\frac{2t}{d}} \gamma_j^2 \\ &\asymp \sum_{j=1}^{\infty} j^{\frac{2(t-s)}{d}} < \infty. \end{aligned}$$

  Thus $u \in \mathcal{H}^t$ a.s., $t < s - \frac{d}{2}$.
- *From the preceding theorem we see that, provided $s > \frac{d}{2}$, the random function in (2.7) generates a mean zero Gaussian measure on $\mathcal{H}$. The expression (2.7) is known as the* Karhunen-Loéve expansion*, and the eigenfunctions $\{\varphi_j\}_{j=1}^{\infty}$ as the* Karhunen-Loéve basis*.*

- The following formal calculation gives an expression for the covariance operator:

$$\begin{aligned}
\mathcal{C} &= \mathbb{E}\mathbf{u}(\mathbf{x}) \otimes \mathbf{u}(\mathbf{x}) \\
&= \mathbb{E}\left(\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \gamma_j\gamma_k\xi_j\xi_k\varphi_j(\mathbf{x}) \otimes \varphi_k(\mathbf{x})\right) \\
&= \left(\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \gamma_j\gamma_k\mathbb{E}(\xi_j\xi_k)\varphi_j(\mathbf{x}) \otimes \varphi_k(\mathbf{x})\right) \\
&= \left(\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \gamma_j\gamma_k\delta_{jk}\varphi_j(\mathbf{x}) \otimes \varphi_k(\mathbf{x})\right) \\
&= \sum_{j=1}^{\infty} \gamma_j^2\varphi_j(\mathbf{x}) \otimes \varphi_j(\mathbf{x}).
\end{aligned}$$

  From this expression for the covariance, we may find eigenpairs explicitly:

$$\begin{aligned}
\mathcal{C}\varphi_k &= \left(\sum_{j=1}^{\infty} \gamma_j^2\varphi_j(\mathbf{x}) \otimes \varphi_j(\mathbf{x})\right)\varphi_k \\
&= \sum_{j=1}^{\infty} \gamma_j^2\langle\varphi_j, \varphi_k\rangle\varphi_j = \sum_{j=1}^{\infty} \gamma_j^2\delta_{jk}\varphi_k = \gamma_k^2\varphi_k.
\end{aligned}$$

- The Gaussian measure is denoted $\mathcal{N}(0, \mathcal{C})$ and the eigenfunctions of $\mathcal{C}$, $\{\varphi_j\}_{j=1}^{\infty}$, are the Karhunen-Loève basis for measure μ_0. The γ_j^2 are the eigenvalues associated with this eigenbasis, and thus γ_j is the standard deviation of the Gaussian measure in the direction φ_j.

Example In the case where $\mathcal{H} = \dot{\mathbf{L}}^2(\mathbb{T}^d)$ we are in the setting of Section 2.3. Furthermore, we now assume that the $\{\varphi_j\}_{j=1}^{\infty}$ constitute the Fourier basis. It then follows that $\mathcal{H}^t = \dot{\mathbf{H}}^t(\mathbb{T}^d)$, the Sobolev space of periodic functions on $[0, 1)^d$ with mean zero and $\mathbf{t}$ (possibly negative or fractional) square integrable derivatives, denoted by $\dot{\mathbf{H}}^t$. Thus $\mathbf{u} \in \dot{\mathbf{H}}^t$ a.s., $\mathbf{t} < \mathbf{s} - \frac{d}{2}$.

A commonly arising choice of prior covariance operator is $\mathcal{C} = (\mathbf{A})^{-\alpha}$ with $\mathbf{A} = -\triangle$, $\mathbf{D}(\mathbf{A}) = \dot{\mathbf{H}}^2(\mathbb{T}^d)$. It then follows, analogously to the result of Lemma 1.3 in the case of Dirichlet boundary conditions, that $\gamma_j^2 \asymp \mathbf{j}^{-\frac{2\alpha}{d}}$. Thus $\mathbf{s} = \alpha$ and $\mathbf{u} \in \dot{\mathbf{H}}^t$, $\mathbf{t} < \alpha - \frac{d}{2}$. As a result, for any $\mathbf{t} < \alpha - \frac{d}{2}$, it is possible to view the resulting Gaussian measure as defined on the Hilbert space $\dot{\mathbf{H}}^t$. In fact, by use of the Kolmogorov continuity theorem, the Gaussian measures may also be defined on Hölder spaces $\mathbf{C}^{0,t}$, for $\mathbf{t} < \alpha - \frac{d}{2}$, if $\alpha - \frac{d}{2} \in (0, 1)$ and $\mathbf{C}^{r,\epsilon}$ with $\mathbf{r} = \lfloor \alpha - \frac{d}{2} \rfloor$, $\epsilon = \alpha - \frac{d}{2} - \mathbf{r} \in (0, 1)$. □

The previous example illustrates the fact that, although we have constructed Gaussian measures in a Hilbert space setting, they may also be defined on Banach spaces, such as the space of Hölder continuous functions. The following theorem then applies.

Theorem 2.6. Fernique Tle

(1.5) so that the coefficient κ and the solution p are random variables on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then $\kappa_{\min}$ given in (1.3) satisfies

$$\kappa_{\min} \geq \exp\big(-\|u\|_\infty\big).$$

By Lemma 1.5 we obtain

$$\|p\|_V \leq \exp\big(\|u\|_\infty\big)\|f\|_{V^*}.$$

Since $C^{0,t} \subset L^\infty(\mathbb{T}^d)$, $t \in (0, 1)$, we deduce that,

$$\|u\|_{L^\infty} \leq K_1\|u\|_{C^t}.$$

Furthermore, for any $\epsilon > 0$, there is constant $K_2 = K_2(\epsilon)$ such that $\exp(K_1 r x) \leq K_2 \exp(\epsilon x^2)$ for all $x \geq 0$. Thus

$$\begin{aligned}\|p\|_V^r &\leq \exp\big(K_1 r\|u\|_{C^t}\big)\|f\|_{V^*}^r\\ &\leq K_2 \exp\big(\epsilon\|u\|_{C^t}^2\big)\|f\|_{V^*}^r.\end{aligned}$$

Hence, by Theorem 2.6, we deduce that

$$\mathbb{E}\|p\|_V^r < \infty, \quad \text{i.e.} \quad p \in L^r_{\mathbb{P}}(\Omega; V) \quad \forall\, r \in \mathbb{Z}^+.$$

Thus, when the coefficient of the elliptic PDE is **log-normal**, that is κ is the exponential of a Gaussian function, moments of all orders exist for the random variable p. However, unlike the case of the uniform prior, we cannot obtain exponential moments on $\mathbb{E}\exp(\alpha\|p\|_V^r)$ for any $(r, \alpha) \in \mathbb{Z}^+ \times (0, \infty)$. This is because the coefficient, whilst positive a.s., does not satisfy a uniform positive lower bound across the probability space. $\square$

2.5. Summary

In the preceding three subsections we have shown how to create random functions by randomizing the coefficients of a series of functions. We have also studied the regularity properties of the resulting functions. For the uniform prior we have shown that the random functions all live in a subset of $X = L^\infty$ characterized by the upper and lower bounds given in Theorem 2.1; denote this subset by X'. For the Besov priors we have shown in Theorem 2.2 that the random functions live in the Banach spaces $X^{t,q}$ for all $t < s - d/q$; denote any one of these Banach spaces by X'. And finally for the Gaussian priors we have shown in Theorem 2.4 that the random function exists as an L^2-limit in any of the Hilbert spaces $\mathcal{H}^t$ for $t < s - d/2$. Furthermore, we have indicated that, by use of the Kolmogorov test, we can also show that the Gaussian random functions lie in certain Hölder spaces; denote any of the Hlbert or Banach spaces where the Gaussian random function lies by X'. Thus, in all of these examples, we have created a probability measure μ_0 which is the pushforward of the measure $\mathbb{P}$ on the i.i.d. sequence ξ under the map which takes the sequence into the random function. This measure lives on X', and we will often write $\mu_0(X') = 1$ to denote this fact. This is shorthand for saying that functions drawn from μ are in X' almost surely.

3. Posterior Distribution

3.1. Conditioned Random Variables

Key to the development of Bayes's Theorem, and the posterior distribution, is the notion of conditional random variables. In this section we develop an important theorem concerning conditioning.

Let (X, A) and (Y, B) denote a pair of measurable spaces and let ν and π be probability measures on $X \times Y$. We assume that $\nu \ll \pi$. Thus there exists $\pi-$measurable $\varphi : X \times Y \to \mathbb{R}$ with $\varphi \in L^1_\pi$ and

$$\frac{d\nu}{d\pi}(x, y) = \varphi(x, y). \tag{3.1}$$

That is, for $(x, y) \in X \times Y$,

$$\mathbb{E}^{\nu} f(x, y) = \mathbb{E}^{\pi}\big(\varphi(x, y) f(x, y)\big),$$

or, equivalently,

$$\int_{X\times Y} f(x, y)\nu(dx, dy) = \int_{X\times Y} \varphi(x, y) f(x, y)\pi(dx, dy).$$

Theorem 3.1. **Assume that the conditional random variable $x|y$ exists under π with probability distribution denoted $\pi^y(dx)$. Then the conditional random variable $x|y$ under ν exists, with probability distribution denoted by $\nu^y(dx)$. Furthermore, $\nu^y \ll \pi^y$ and**

$$\frac{d\nu^y}{d\pi^y}(x) = \begin{cases} \frac{1}{c(y)}\varphi(x, y), & \text{if} \quad c(y) > 0 \\ 1, & \text{otherwise} \end{cases}$$

with $c(y) = \int_X \varphi(x, y) d\pi^y(x)$.

Example Let $X = C\big([0, 1]; \mathbb{R}\big)$, $Y = \mathbb{R}$. Let π denote the measure on $X \times Y$ induced by the random variable $\big(w(\cdot), w(1)\big)$, where w is a draw from standard unit Wiener measure on $\mathbb{R}$, starting from $w(0) = z$.

Let π^y denote measure on X found by conditioning Brownian motion to satisfy $w(1) = y$, thus π^y is a Brownian bridge measure with $w(0) = z, w(1) = y$.

Assume that $\nu \ll \pi$ with

$$\frac{d\nu}{d\pi}(x, y) = \exp\big(-\Phi(x, y)\big).$$

(Such a formula arises from the Girsanov theorem, for example, in the theory of stochastic differential equations – SDEs.) Assume further that

$$\sup_{x\in S} \Phi(x, y) = \Phi^+(\eta), \quad \inf_{x\in S} \Phi(x, y) = \Phi^-(\eta)$$

and $\Phi^-, \Phi^+ \in (0, \infty)$ for every $y \in \mathbb{R}$. Then

$$c(y) = \int_{\mathbb{R}} \exp\big(-\Phi(x, y)\big) d\nu^y(x) > \exp\big(-\Phi^+(y)\big) > 0.$$

Thus $\nu^y(dx)$ exists and

$$\frac{d\nu^y}{d\pi^y}(x) = \frac{1}{c(y)} \exp\big(-\Phi(x, y)\big). \quad \square$$

The following lemma is useful for checking measurability.

Lemma 3.2. **Let (Z, C) be a measurable space and assume that $G \in C(Z; \mathbb{R})$ and that $\pi(Z) = 1$ for some probability measure π on Z. Then G is a $\pi-$measurable function.**

3.2. Bayes' Theorem for Inverse Problems

Let X, Y be separable Banach spaces, and G)cm[Ж)-23Y1762(e)a.89428(75-1.66393(a)-5.89115(s)-3.4847(u)0.931213(r)-6.486

The random variable $\mathbf{y}|\mathbf{u}$ is then distributed according to the measure $\mathbb{Q}_u$, the translate of $\mathbb{Q}_0$ by $\mathbf{G}(\mathbf{u})$. We **assume** throughout the following that $\mathbb{Q}_u \ll \mathbb{Q}_0$ for $\mathbf{u}$ μ_0- a.s. Thus, for some **potential** $\Phi : \mathbf{X} \times \mathbf{Y} \to \mathbb{R}$,

$$\frac{d\mathbb{Q}_u}{d\mathbb{Q}_0}(\mathbf{y}) = \exp\Big(-\Phi(\mathbf{u};\mathbf{y})\Big). \tag{3.3}$$

For given instance of the data $\mathbf{y}$, $\Phi(\mathbf{u};\mathbf{y})$ is the **negative log likelihood**. Define ν_0 to be the product measure defined by

$$\nu_0(d\mathbf{u}, d\mathbf{y}) = \mathbb{Q}_0(d\mathbf{y})\mu_0(d\mathbf{u}). \tag{3.4}$$

We also **assume** in what follows that $\Phi(\cdot,\cdot)$ is ν_0 measurable. Then the random variable $(\mathbf{u},\mathbf{y}) \in \mathbf{X} \times \mathbf{Y}$ is distributed according to measure $\nu(d\mathbf{u}, d\mathbf{y})$ where

$$\frac{d\nu}{d\nu_0}(\mathbf{u},\mathbf{y}) = \exp\Big(-\Phi(\mathbf{u};\mathbf{y})\Big).$$

We have the following infinite dimensional analogue of Theorem 1.1.

Theorem 3.3. Bayes Theorem Assume that $\Phi : \mathbf{X} \times \mathbf{Y} \to \mathbb{R}$ is ν_0 measurable and that, for $\mathbf{y}$ $\mathbb{Q}_0-$a.s.,

$$\mathbf{Z} := \int_X \exp\Big(-\Phi(\mathbf{u};\mathbf{y})\Big)\mu_0(d\mathbf{u}) > 0. \tag{3.5}$$

Then the conditional distribution of $\mathbf{u}|\mathbf{y}$ exists under ν, and is denoted μ^y. Furthermore $\mu^y \ll \mu_0$ and, for $\mathbf{y}$ $\nu-$a.s.,

$$\frac{d\mu^y}{d\mu_0}(\mathbf{u}) = \frac{1}{\mathbf{Z}}\exp\Big(-\Phi(\mathbf{u};\mathbf{y})\Big). \tag{3.6}$$

Proof. First note that the positivity of $\mathbf{Z}$ holds for $\mathbf{y}$ ν_0 almost surely, and hence by absolute continuity of ν with respect to ν_0, for $\mathbf{y}$ ν almost surely. The proof is an application of Theorem 3.1 with π replaced by ν_0, $\varphi(\mathbf{x},\mathbf{y}) = \exp\Big(-\Phi(\mathbf{x},\mathbf{y})\Big)$ and $(\mathbf{x},\mathbf{y}) \to (\mathbf{u},\mathbf{y})$. Since $\nu_0(d\mathbf{u}, d\mathbf{y})$ has product form, the conditional distribution of $\mathbf{u}|\mathbf{y}$ under ν_0 is simply μ_0. The result follows. $\square$

Remarks 3.4. In order to implement the derivation of Bayes' formula (3.6) four essential steps are required:

- Define a suitable prior measure μ_0 and noise measure $\mathbb{Q}_0$ whose independent product form the reference measure ν_0.
- Determine the potential Φ such that formula (3.3) holds.
- Show that Φ is ν_0 measurable.
- Show that the normalization constant $\mathbf{Z}$ given by (3.5) is positive almost surely with respect to $\mathbf{y} \sim \mathbb{Q}_0$.

Remark 3.5. In formula (3.6) we can shift $\Phi(\mathbf{u},\mathbf{y})$ by any constant $c(\mathbf{y})$, independent of $\mathbf{u}$, provided the constant is finite $\mathbb{Q}_0-$a.s. and hence $\nu-$a.s. Such a shift can be absorbed into a redefinition of the normalization constant $\mathbf{Z}$.

3.3. Heat Equation

We apply Bayesian inversion to the heat equation from Section 1.2. Recall that for $\mathbf{G}(\mathbf{u}) = \mathbf{e}^{-A}\mathbf{u}$, we have the relationship

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

which we wish to invert. Let $\mathbf{X} = \mathbf{H}$ and define $\mathcal{H}^t = \mathbf{D}(\mathbf{A}^{\frac{t}{2}})$. Then, for $\mathbf{u} = \sum_j \mathbf{u}_j \phi_j$,

$$\mathcal{H}^t = \Big\{\mathbf{u}\Big|\sum_j \alpha_j^t \mathbf{u}_j^2 < \infty, \quad \mathbf{u}_j = \langle \mathbf{u}, \phi_j\rangle\Big\}.$$

Recall from Lemma 1.3 that $\alpha_j \asymp \mathbf{j}^{\frac{2}{d}}$ so this agrees with $\mathcal{H}^t$ as defined in subsection 2.4. Furthermore, we observe that

$$\mathcal{H}^t = \mathbf{D}(\mathbf{A}^{t/2}) = \Big\{\mathbf{w}\Big|\mathbf{w} = \mathbf{A}^{-t/2}\mathbf{w}_0, \mathbf{w}_0 \in \mathbf{H}\Big\}.$$

We choose the prior $\mu_0 = N(0, A^{-\alpha})$, $\alpha > \frac{d}{2}$. Thus $\mu_0(X) = \mu_0(H) = 1$. Indeed the analysis in subsection 2.4 shows that $\mu_0(\mathcal{H}^t) = 1$, $t < \alpha - \frac{d}{2}$. For the likelihood we assume that $\eta \perp u$ with $\eta \sim \mathbb{Q}_0 = N(0, A^{-\beta})$, and $\beta \in \mathbb{R}$. This measure satisfies $\mathbb{Q}_0(\mathcal{H}^t) = 1$ for $t < \beta - \frac{d}{2}$ and we thus choose $Y = \mathcal{H}^{t'}$ for some $t' < \beta - \frac{d}{2}$. Notice that our analysis includes the case of white observational noise, for which $\beta = 0$. The Cameron-Martin Theorem, together with the fact that $e^{-\lambda A}$ commutes with arbitrary fractional powers of A, can be used to show that $y|u \sim \mathbb{Q}_u := N(G(u), A^{-\beta})$ where $\mathbb{Q}_u \ll \mathbb{Q}_0$ with

$$\frac{d\mathbb{Q}_u}{d\mathbb{Q}_0}(y) = \exp\Big(-\Phi(u; y)\Big),$$

$$\Phi(u; y) = \frac{1}{2}\|A^{\frac{\beta}{2}} e^{-A} u\|^2 - \langle A^{\frac{\beta}{2}} e^{-\frac{A}{2}} y, A^{\frac{\beta}{2}} e^{-\frac{A}{2}} u\rangle.$$

In the following we repeatedly use the fact that $A^\gamma e^{-\lambda A}$, $\lambda > 0$, is a bounded linear operator from $\mathcal{H}^a$ to $\mathcal{H}^b$, any $a, b, \gamma \in \mathbb{R}$. Recall that $\nu_0(du, dy) = \mu_0(du)\mathbb{Q}_0(dy)$. Note that $\nu_0(H \times \mathcal{H}^{t'}) = 1$. Using the boundedness of $A^\gamma e^{-\lambda A}$ it may be shown that

$$\Phi : H \times \mathcal{H}^{t'} \to \mathbb{R}$$

is continuous, and hence ν_0-measurable by Lemma 3.2.

Theorem 3.3 shows that the posterior is given by μ^y where

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

$$Z = \int_H \exp\Big(-\Phi(u; y)\Big)\mu_0(du),$$

provided that $Z > 0$ for y $\mathbb{Q}_0-$a.s. Since $y \in \mathcal{H}^t$ for any $t < \beta - \frac{d}{2}$, $\mathbb{Q}_0-$a.s., we have that $y = A^{-t'/2} w_0$ for some $w_0 \in H$ and $t' < \beta - \frac{d}{2}$. Thus we may write

$$\Phi(u; y) = \frac{1}{2}\|A^{\frac{\beta}{2}} e^{-A} u\|^2 - \langle A^{\frac{\beta - t'}{2}} e^{-\frac{A}{2}} w_0, A^{\frac{\beta}{2}} e^{-\frac{A}{2}} u\rangle. \tag{3.7}$$

Then, using the boundedness of $A^\gamma e^{-\lambda A}$, $\lambda > 0$, together with (3.7), we have

$$\Phi(u; y) \le C(\|u\|^2 + \|w_0\|^2)$$

where $\|w_0\|$ is finite $\mathbb{Q}_0-$a.s. Thus

$$Z \ge \int_{\|u\|^2 \le 1} \exp\Big(-C(1 + \|w_0\|^2)\Big)\mu_0(du)$$

and, since $\mu_0(\|u\|^2 \le 1) > 0$ (all balls have positive measure for Gaussians on a separable Banach space) the result follows.

3.4. Elliptic Inverse Problem

We consider the elliptic inverse problem from Section 1.3 from the Bayesian perspective. We consider the use of both uniform and Gaussian priors. Before studying the inverse problem, however, it is important to derive some continuiuty properties of the forward problem. Consider equation (1.5) and, define

$$X^+ = \Big\{v \in L^\infty(D) \Big| \operatorname*{ess\,inf}_{x \in D} v(x) > 0\Big\}$$

and define the map $\mathcal{R} : X^+ \to V$ by $\mathcal{R}(\kappa) = p$. This map is well-defined by Lemma 1.5.

Lemma 3.6. For $i = 1, 2$, let

$$\begin{aligned} -\nabla \cdot (\kappa_i \nabla p_i) &= f, \quad x \in D, \\ p_i &= 0, \quad x \in \partial D. \end{aligned}$$

Then

$$\|p_1 - p_2\|_V \le \frac{1}{\kappa_{\min}^2} \|f\|_{V^*} \|\kappa_1 - \kappa_2\|_{L^\infty}$$

where we assume that

$$\kappa_{\min} := \operatorname*{ess\,inf}_{x \in D} \kappa_1(x) \wedge \operatorname*{ess\,inf}_{x \in D} \kappa_2(x) > 0.$$

Thus the function $\mathcal{R} : X^+ \to V$ is locally Lipschitz.

Proof. Let $e = \kappa_1 - \kappa_2$, $d = p_1 - p_2$. Then

$$\begin{aligned} -\nabla \cdot (\kappa_1 \nabla d) &= \nabla \cdot \big((\kappa_1 - \kappa_2) \nabla p_2 \big), \quad x \in D \\ d &= 0, \quad x \in \partial D. \end{aligned}$$

By Lemma 1.5 (applied twice) and the Cauchy-Schwarz inequality on L^2 we have

$$\begin{aligned} \|d\|_V &\le \|(\kappa_2 - \kappa_1) \nabla p_2\| / \kappa_{\min} \\ &\le \|\kappa_2 - \kappa_1\|_{L^\infty} \|p_2\|_V / \kappa_{\min} \\ &\le \frac{1}{\kappa_{\min}^2} \|f\|_{V^*} \|e\|_{L^\infty}. \end{aligned}$$

□

We now study the inverse problem of finding κ from a finite set of continuous linear functionals $\{l_j\}_{j=1}^J$ on V, representing measurements of p; thus $l_j \in V^*$. We study both the use of uniform priors, and the use of Gaussian priors. We start with the uniform case, taking $\kappa = u$, and we define $G : X^+ \to \mathbb{R}^J$ by

$$G_j(u) = l_j \big(\mathcal{R}(u) \big), \quad j = 1, \dots, J.$$

Then $G(u) = \big(G_1(u), \cdots, G_J(u) \big)$. We set $X = L^\infty(D; \mathbb{R})$, $Y = \mathbb{R}^J$ and consider the inverse problem of finding u from y where

$$y = G(u) + \eta$$

and η is the noise.

Define $X' \subset X^+$ by

$$X' = \Big\{ v \in X \Big| \frac{1}{1+\delta} \varphi_{\min} \le v(x) \le \varphi_{\max} + \frac{\delta}{1+\delta} \varphi_{\min} \quad \text{a.e.} \quad x \in D \Big\}.$$

The measure μ_0 on functions from subsection 2.2, (found as the pushforward of the measure $\mathbb{P}$ on i.i.d. sequences, see subsection 2.5) is, by Theorem 2.1, a measure on X; furthermore $\mu_0(X') = 1$. We take μ_0 as the prior.

The likelihood is defined as follows. We assume $\eta \sim N(0, \Gamma)$, for positive symmetric $\Gamma \in \mathbb{R}^{J \times J}$. Thus $\mathbb{Q}_0 = N(0, \Gamma)$, $\mathbb{Q}_u = N\big(G(u), \Gamma\big)$ and

$$\frac{d\mathbb{Q}_u}{d\mathbb{Q}_0}(y) = \exp\big(-\Phi(u; y) \big),$$

$$\Phi(u; y) = \frac{1}{2} \big| \Gamma^{-\frac{1}{2}} (y - G(u)) \big|^2 - \frac{1}{2} \big| \Gamma^{-\frac{1}{2}} y \big|^2.$$

Recall that $\nu_0(dy, du) = \mathbb{Q}_0(dy)\mu_0(du)$. $G : X' \to \mathbb{R}^J$ is Lipschitz by Lemma 3.6 (in fact we only use that it is locally Lipschitz) and hence Lemma 3.2 implies that $\Phi : X' \times Y \to \mathbb{R}$ is ν_0-measurable. Thus Theorem 3.3 shows that $u|y \sim \mu^y$ where

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

provided $Z > 0$ for y $\mathbb{Q}_0$ almost surely. To see that $Z > 0$ note that

$$Z = \int_{X'} \exp\Big(-\Phi(u; y)\Big)\mu_0(du),$$

since $\mu_0(X') = 1.$ On X' we have that $\mathcal{R}(\cdot)$ is bounded in V, and hence G is bounded in $\mathbb{R}^J.$ Furthermore y is finite $\mathbb{Q}_0$ almost surely. Thus $\Phi(u; y)$ is bounded by $M = M(y) < \infty$ on X', $\mathbb{Q}_0$ almost surely. Hence

$$Z \geq \int_{X'} \exp(-M)\mu_0(du) = \exp(-M) > 0.$$

and the result is proved.

We may use Remark 3.5 to shift Φ by $\frac{1}{2}|\Gamma^{-\frac{1}{2}} y|^2$, since this is almost surely finite under $\mathbb{Q}_0$ and hence under $\nu(du, dy) = \mathbb{Q}_u(dy)\mu_0(du)$. We then obtain the equivalent form for the posterior distribution μ^y:

$$\frac{d\mu^y}{d\mu_0}(u) = \frac{1}{Z} \exp\Big(-\frac{1}{2}\Big|\Gamma^{-\frac{1}{2}}\big(y - G(u)\big)\Big|^2\Big), \tag{3.8a}$$
$$Z = \int_X \exp\Big(-\frac{1}{2}|\Gamma^{-\frac{1}{2}}\big(y - G(u)\big)\Big|^2\Big)\mu_0(du). \tag{3.8b}$$

We conclude this subsection by discussing the same inverse problem, but using Gaussian priors from subsection 2.4. We again set $X = L^\infty(D; \mathbb{R})$, $Y = \mathbb{R}^J$ and, for simplicity, take $D = [0, 1]^d$. We now take $\kappa = \exp(u)$, and define $G : X \to \mathbb{R}^J$ by

$$G_j(u) = l_j\Big(\mathcal{R}\big(\exp(u)\big)\Big), \quad j = 1, \dots, J.$$

We take as prior on u the measure $N(0, A^{-\alpha})$, from the example preceding the Fernique Theorem 2.6, with $\alpha > d/2.$ The measure μ_0 then satisfies $\mu(X') = 1$ with $X' = C(D; \mathbb{R})$. The likelihood is unchanged by the prior, since it concerns y given u, and is hence identical to that in the case of the uniform prior, although the mean shift from $\mathbb{Q}_0$ by $\mathbb{Q}_u$ by $G(u)$ now has a different interpretation. Thus we again obtain (3.8) for the posterior distribution (albeit with a different definition of $G(u)$) provided that we can establish that

$$Z = \int_X \exp\Big(-\frac{1}{2}\Big|\Gamma^{-\frac{1}{2}}\big(y - G(u)\big)\Big|^2\Big)\mu_0(du) > 0.$$

To this end we use the fact that the unit ball in X', denoted B, has positive measure, and that on this ball $\mathcal{R}\big(\exp(u)\big)$ is bounded in V by $e^{-1}\|f\|_{V^*}$, by Lemma 1.5, since the infimum of $\kappa = \exp(u)$ is e^{-1} on this ball B. Thus G is bounded on B and, noting that y is $\mathbb{Q}_0-$a.s. finite, we have for some $M = M(y) < \infty$,

$$\sup_{u \in B}\Big(\frac{1}{2}\Big|\Gamma^{-\frac{1}{2}}\big(y - G(u)\big)\Big|^2 - \frac{1}{2}\Big|\Gamma^{-\frac{1}{2}} y|^2\Big) < M.$$

Hence

$$Z \geq \int_B \exp(-R)\mu_0(du) = \exp(-R)\mu_0(B) > 0.$$

Thusc we again obtain (3.6) for the posterior measure, now with the new definition of G, and hence $\Phi.$

4. Common Structure

In this section we discuss various common features of the posterior distribution arising from the Bayesian approach to inverse problems. We start, in subsection 4.1, by studying the continuity properties of the posterior with respect to changes in data, proving a form of well-posedness; indeed we show that the posterior is Lipschitz in the data with respect to the Hellinger metric. In subsection 4.2 we use similar ideas to study the effect of approximation on the posterior distribution, showing that small changes in the potential Φ lead to small changes in the posterior distribution, again the Hellinger metric; this work may be used to translate error analysis pertaining to the forward problem into estimates on errors in the posterior distribution. In the remaining two subsections we work entirely in the case of Gaussian prior measure $\boldsymbol{\mu}_0$. Subsection 4.3 is concerned with derivation and study of a Langevin equation which is invariant with respect to the posterior $\boldsymbol{\mu}$, and subsection 4.4 concerns MCMC methods, also invariant with respect to $\boldsymbol{\mu}$, which exploit the structure of a target measure defined via density with respect to a Gaussian; in particular, the idea of using proposals which preserve the prior is introduced and benefits of doing so are explained.

4.1. Well-Posedness

In many classical inverse 4.2
so(t)2.56022i(o)-5.88993(n)0.929988(,)-475.393(a)-5.88993(n)0.929988(d)-324.278(s)-3.4847(o)-5.88993ime formof

In order to measure the effect of changes in y on the measure μ^y we need a metric on measures. We use the **Hellinger distance** defined as follows: given two measures μ and μ' on X, both absolutely continuous with respect to a common reference measure ν, the Hellinger distance is

$$d_{\text{Hell}}(\mu, \mu') = \sqrt{\left(\frac{1}{2}\int_X \left(\sqrt{\frac{d\mu}{d\nu}} - \sqrt{\frac{d\mu'}{d\nu}}\right)^2 d\nu\right)}.$$

In particular, if μ' is absolutely continuous with respect to μ then

$$d_{\text{Hell}}(\mu, \mu') = \sqrt{\left(\frac{1}{2}\int_X \left(1 - \sqrt{\frac{d\mu'}{d\mu}}\right)^2 d\mu\right)}.$$

Theorem 4.2. **Let Assumptions 4.1 hold. Assume that $\mu_0(X') = 1$ and that $\mu_0(X' \cap B) > 0$ for some bounded set B in X. Then, for every $y \in Y$, $Z(y)$ given by (4.1b) is positive and probability measure μ^y given by (4.1a) is well-defined.**

Proof. Since $u \sim \mu_0$ satisfies $u \in X'$ a.s., we have

$$Z(y) = \int_{X'} \exp\Big(-\Phi(u; y)\Big)\mu_0(du).$$

Note that $B' = X' \cap B$ is bounded in X. Define

$$R_1 := \sup_{u \in B'} \|u\|_X < \infty.$$

Since $\Phi : X' \times Y \to \mathbb{R}$ is continuous it is finite at every point in $B' \times \{y\}$. Thus, by the continuity of $\Phi(\cdot;\cdot)$ implied by Assumptions 4.1, we see that

$$\sup_{(u,y) \in B' \times B_Y(0,r)} \Phi(u; y) = R_2 < \infty.$$

Hence

$$Z(y) \geq \int_{B'} \exp(-R_2)\mu_0(du) = \exp(-R_2)\mu_0(B').$$

Since $\mu_0(B')$ is assumed positive and R_2 is finite we deduce that $Z(y) > 0$. □

Theorem 4.3. **Let Assumptions 4.1 hold. Assume that $\mu_0(X') = 1$ and that $\mu_0(X' \cap B) > 0$ for some bounded set B in X. Assume additionally that, for every fixed $r > 0$,**

$$\exp\Big(M_1(r, \|u\|_X)\Big) M_2^2(r, \|u\|_X) \in L^1_{\mu_0}(X; \mathbb{R}).$$

Then there is $C = C(r) > 0$ such that, for all $y, y' \in B_Y(0, r)$

$$d_{\text{Hell}}(\mu^y, \mu^{y'}) \leq C\|y - y'\|_Y.$$

Proof. Throughout this proof we use C to denote a constant independent of u, but possibly depending on the fixed value of r; it may change from occurence to occurence. We use the fact that, since $M_2(r, \cdot)$ is monotonic non-decreasing and since it is strictly positive on $[0, \infty)$, there is constant $C > 0$ such that

$$\exp\Big(M_1(r, \|u\|_X)\Big) M_2(r, \|u\|_X) \leq C \exp\Big(M_1(r, \|u\|_X)\Big) M_2(r, \|u\|_X)^2, \tag{4.2a}$$
$$\exp\Big(M_1(r, \|u\|_X)\Big) \leq C \exp\Big(M_1(r, \|u\|_X)\Big) M_2(r, \|u\|_X)^2. \tag{4.2b}$$

Let $Z = Z(y)$ and $Z' = Z(y')$ denote the normalization constants for μ^y and $\mu^{y'}$ so that, by Theorem 4.2,

$$Z = \int_{X'} \exp\Big(-\Phi(u; y)\Big)\mu_0(du) > 0,$$
$$Z' = \int_{X'} \exp\Big(-\Phi(u; y')\Big)\mu_0(du) > 0.$$

Then, using the local Lipschitz property of the exponential and the assumed Lipschitz continuity of $\Phi(\cdot;\mathbf{r})$, together with (4.2a), we have

$$\begin{aligned}
|\mathbf{Z}-\mathbf{Z}'| &\le \int_{X'} \big|\exp\big(-\Phi(\mathbf{u};\mathbf{y})\big)-\exp\big(-\Phi(\mathbf{u};\mathbf{y}')\big)\big|\boldsymbol{\mu}_0(\mathbf{du})\\
&\le \int_{X'} \exp\big(\mathbf{M}_1(\mathbf{r},\|\mathbf{u}\|_X)\big)|\Phi(\mathbf{u};\mathbf{y})-\Phi(\mathbf{u};\mathbf{y}')|\boldsymbol{\mu}_0(\mathbf{du})\\
&\le \Big(\int_{X'} \exp\big(\mathbf{M}_1(\mathbf{r},\|\mathbf{u}\|_X)\mathbf{M}_2(\mathbf{r},\|\mathbf{u}\|_X)\boldsymbol{\mu}_0(\mathbf{du})\Big)\|\mathbf{y}-\mathbf{y}'\|_Y\\
&\le \mathbf{C}\Big(\int_{X'} \exp\big(\mathbf{M}_1(\mathbf{r},\|\mathbf{u}\|_X)\mathbf{M}_2(\mathbf{r},\|\mathbf{u}\|_X)^2\boldsymbol{\mu}_0(\mathbf{du})\Big)\|\mathbf{y}-\mathbf{y}'\|_Y\\
&\le \mathbf{C}\|\mathbf{y}-\mathbf{y}'\|_Y.
\end{aligned}$$

The last line follows because the integrand is in $\mathbf{L}^1_{\mu_0}$ by assumption. From the definition of Hellinger distance we have

$$\big(\mathbf{d}_{\mathrm{Hell}}(\boldsymbol{\mu}^y,\boldsymbol{\mu}^{y'})\big)^2 \le \mathbf{I}_1+\mathbf{I}_2,$$

where

$$\begin{aligned}
\mathbf{I}_1 &= \frac{1}{\mathbf{Z}}\int_{X'}\Big(\exp\big(-\frac{1}{2}\Phi(\mathbf{u};\mathbf{y})\big)-\exp(-\frac{1}{2}\Phi(\mathbf{u};\mathbf{y}')\big)\Big)^2\boldsymbol{\mu}_0(\mathbf{du}),\\
\mathbf{I}_2 &= \big|\mathbf{Z}^{-\frac{1}{2}}-(\mathbf{Z}')^{-\frac{1}{2}}\big|^2\int_{X'}\exp(-\Phi(\mathbf{u};\mathbf{y}')\big)\boldsymbol{\mu}_0(\mathbf{du}).
\end{aligned}$$

Note that, again using similar Lipschitz calculations to those above, using the fact that $\mathbf{Z}>0$ and Assumptions 4.1,

$$\begin{aligned}
\mathbf{I}_1 &\le \frac{1}{\mathbf{Z}}\int_{X'}\exp\big(\mathbf{M}_1(\mathbf{r},\|\mathbf{u}\|_X)\big)|\Phi(\mathbf{u};\mathbf{y})-\Phi(\mathbf{u};\mathbf{y}')|^2\boldsymbol{\mu}_0(\mathbf{du})\\
&\le \frac{1}{\mathbf{Z}}\Big(\int_{X'}\exp\big(\mathbf{M}_1(\mathbf{r},\|\mathbf{u}\|_X)\big)\mathbf{M}_2(\mathbf{r},\|\mathbf{u}\|_X)^2\boldsymbol{\mu}_0(\mathbf{du})\Big)\|\mathbf{y}-\mathbf{y}'\|_Y^2\\
&\le \mathbf{C}\|\mathbf{y}-\mathbf{y}'\|_Y^2.
\end{aligned}$$

Also, using Assumptions 4.1, together with (4.2b),

$$\begin{aligned}
\int_{X'}\exp\big(-\Phi(\mathbf{u};\mathbf{y}')\big)\boldsymbol{\mu}_0(\mathbf{du}) &\le \int_{X'}\exp\big(\mathbf{M}_1(\mathbf{r},\|\mathbf{u}\|_X)\big)\boldsymbol{\mu}_0(\mathbf{du})\\
&\le \mathbf{C}\int_{X'}\exp\big(\mathbf{M}_1(\mathbf{r},\|\mathbf{u}\|_X)\big)\mathbf{M}_2(\mathbf{r},\|\mathbf{u}\|_X)^2\boldsymbol{\mu}_0(\mathbf{du})\\
&< \infty.
\end{aligned}$$

Hence

$$\mathbf{I}_2 \le \mathbf{C}\big(\mathbf{Z}^{-3}\vee(\mathbf{Z}')^{-3}\big)|\mathbf{Z}-\mathbf{Z}'|^2 \le \mathbf{C}\|\mathbf{y}-\mathbf{y}'\|_Y^2.$$

The result is complete. □

Remark 4.4. **The Hellinger metric has the very desirable property that it translates directly into bounds on expectations. For functions $\mathbf{f}$ which are in $\mathbf{L}^2_{\mu^y}(\mathbf{X};\mathbb{R})$ and $\mathbf{L}^2_{\mu^{y'}}(\mathbf{X};\mathbb{R})$ the closeness of the Hellinger metric implies closeness of expectations of $\mathbf{f}$. To be precise, for $\mathbf{y},\mathbf{y}'\in\mathbf{B}_Y(0,\mathbf{r})$ and $\mathbf{C}=\mathbf{C}(\mathbf{r})$, we have**

$$|\mathbb{E}^{\mu^y}\mathbf{f}(\mathbf{u})-\mathbb{E}^{\mu^{y'}}\mathbf{f}(\mathbf{u})|\le \mathbf{C}\mathbf{d}_{\mathrm{Hell}}(\boldsymbol{\mu}^y,\boldsymbol{\mu}^{y'})$$

so that then

$$|\mathbb{E}^{\mu^y}\mathbf{f}(\mathbf{u})-\mathbb{E}^{\mu^{y'}}\mathbf{f}(\mathbf{u})|\le \mathbf{C}\|\mathbf{y}-\mathbf{y}'\|.$$

4.2. Approximation

In this section we concentrate on continuity properties of the posterior measure with ximation of the potential Φ**.** The methods used are very similar to those in the previous sub establish a continuity property of the posterior distribution, in the Hellinger metric, with res anges in the potential Φ.

Because the data $\mathbf{y}$ plays no explicit role in this discussion, we drop explicit refer $\mathbf{X}$ be a Banach space and $\boldsymbol{\mu}_0$ a measure on $\mathbf{X}$. Assume that $\boldsymbol{\mu}$ and $\boldsymbol{\mu}^N$ are both absolutely co respect to $\boldsymbol{\mu}_0$ and given by

$$\frac{d\boldsymbol{\mu}}{d\boldsymbol{\mu}_0}(\mathbf{u}) = \frac{1}{\mathbf{Z}} \exp\big(-\Phi(\mathbf{u})\big), \qquad \text{(3a)}$$

$$\mathbf{Z} = \int_{X'} \exp\big(-\Phi(\mathbf{u})\big)\,\boldsymbol{\mu}_0(d\mathbf{u}\text{Band}$$

dN-0.249958632J ET Q q 75195f 1 0 Td Tf4 /IM true /W 1

$(\mathbf{u})$ 0 (duBesB (e)-3

Z N 0 $- Z^N$

Hence

$$Z \geq \int_{B'} \exp(-R_2)\mu_0(du) = \exp(-R_2)\mu_0(B').$$

Since $\mu_0(B')$ is assumed positive and R_2 is finite we deduce that $Z(y) > 0$. By Assumptions 4.5 we may choose N large enough so that

$$\sup_{u\in B'} |\Phi(u) - \Phi^N(u)| \leq R_2$$

so that

$$\sup_{u\in B'} \Phi^N(u) = 2R_2 < \infty.$$

Hence

$$Z^N \geq \int_{B'} \exp(-2R_2)\mu_0(du) = \exp(-2R_2)\mu_0(B').$$

Since $\mu_0(B')$ is assumed positive and R_2 is finite we deduce that $Z^N > 0$. Furthermore, the lower bound is independent of N, as required. □

Theorem 4.7. **Let Assumptions 4.1 hold. Assume that $\mu_0(X') = 1$ and that $\mu_0(X' \cap B) > 0$ for some bounded set B in X. Assume additionally that**

$$\exp\big(M_1(\|u\|_X)\big)M_2^2(\|u\|_X) \in L^1_{\mu_0}(X;\mathbb{R}).$$

Then there is $C > 0$ such that, for all N sufficiently large,

$$d_{\mathrm{Hell}}(\mu, \mu^N) \leq C\psi(N).$$

Proof. Throughout this proof we use C to denote a constant independent of u, and N; it may change from occurence to occurence. We use the fact that, since $M_2(\cdot)$ is monotonic non-decreasing and since it is strictly positive on $[0, \infty)$, there is constant $C > 0$ such that

$$\exp\big(M_1(\|u\|_X)\big)M_2(\|u\|_X) \leq C\exp\big(M_1(\|u\|_X)\big)M_2(\|u\|_X)^2, \tag{4.5a}$$
$$\exp\big(M_1(\|u\|_X)\big) \leq C\exp\big(M_1(\|u\|_X)\big)M_2(\|u\|_X)^2. \tag{4.5b}$$

Let Z and Z^N denote the normalization constants for μ and μ^N so that for all N sufficiently small, by Theorem 4.6,

$$Z = \int_{X'} \exp\big(-\Phi(u)\big)\mu_0(du) > 0,$$
$$Z^N = \int_{X'} \exp\big(-\Phi^N(u)\big)\mu_0(du) > 0,$$

with lower bounds independent of N. Then, using the local Lipschitz property of the exponential and the assumed Lipschitz continuity of $\Phi(\cdot)$, together with (4.5a), we have

$$\begin{aligned}
|Z - Z^N| &\leq \int_{X'} |\exp\big(-\Phi(u)\big) - \exp\big(-\Phi^N(u)\big)|\mu_0(du) \\
&\leq \int_{X'} \exp\big(M_1(\|u\|_X)\big)|\Phi(u) - \Phi^N(u)|\mu_0(du) \\
&\leq \Big(\int_{X'} \exp\big(M_1(\|u\|_X)\big)M_2(\|u\|_X)\mu_0(du)\Big)\psi(N) \\
&\leq C\Big(\int_{X'} \exp\big(M_1(\|u\|_X)\big)M_2(\|u\|_X)^2\mu_0(du)\Big)\psi(N) \\
&\leq C\psi(N).
\end{aligned}$$

The last line follows because the integrand is in $L^1_{\mu_0}$ by assumption. From the definition of Hellinger distance we have

$$\Big(d_{\mathrm{Hell}}(\mu^y, \mu^{y'})\Big)^2 \le I_1 + I_2,$$

where

$$I_1 = \frac{1}{Z}\int_{X'}\Big(\exp\Big(-\frac{1}{2}\Phi(u)\Big) - \exp(-\frac{1}{2}\Phi^N(u)\Big)^2\mu_0(du),$$

$$I_2 = \Big|Z^{-\frac{1}{2}} - (Z')^{-\frac{1}{2}}\Big|^2\int_{X'}\exp(-\Phi^N(u)\Big)\mu_0(du).$$

Note that, again using similar Lipschitz calculations to those above, using the fact (Theorem 4.6) that $Z, Z^N > 0$ uniformly in $N \to \infty$, and Assumptions 4.5,

$$\begin{aligned} I_1 &\le \frac{1}{Z}\int_{X'}\exp\Big(M_1(\|u\|_X\Big)|\Phi(u) - \Phi^N(u)|^2\mu_0(du) \\ &\le \frac{1}{Z}\Big(\int_{X'}\exp\Big(M_1(\|u\|_X)\Big)M_2(\|u\|_X)^2\mu_0(du)\Big)\psi(N)^2 \\ &\le C\psi(N)^2. \end{aligned}$$

Also, using Assumptions 4.5, together with (4.5b),

$$\begin{aligned} \int_{X'}\exp\Big(-\Phi^N(u)\Big)\mu_0(du) &\le \int_{X'}\exp\Big(M_1(\|u\|_X)\Big)\mu_0(du) \\ &\le C\int_{X'}\exp\Big(M_1(\|u\|_X)\Big)M_2(\|u\|_X)^2\mu_0(du) \\ &< \infty, \end{aligned}$$

and the upper bound is independent of N. Hence

$$I_2 \le C\Big(Z^{-3} \vee (Z^N)^{-3}\Big)|Z - Z^N|^2 \le C\psi(N)^2.$$

The result is complete. □

Remark 4.8. **Using the ideas underlying Remark 4.4, this result enables us to translate errors arising from approximation of the forward problem into errors in the Bayesian solution of the inverse problem. Furthermore, the errors in the forward and inverse problems scale the same way with respect to N. For functions f which are in L^2_μ and $L^2_{\mu^N}$, uniformly with respect to N, the closeness of the Hellinger metric implies closeness of expectations of f:**

$$|\mathbb{E}^\mu f(u) - \mathbb{E}^{\mu^N} f(u)| \le C\psi(N).$$

4.3. Measure Preserving Dynamics

The aim of this section is to exhibit a Hilbert space valued stochastic differential equation (SDE), which in many applicstions has an interpretation as a stochastic partial differential equation (SPDE), and which is invariant with respect to the posterior measure μ^y constructed in subsection 3.2. We restrict outselves to the case of Gaussian priors μ_0. The data y plays no role in what follows and indeed the theory applies to a wide range of measures μ which have density with respect to a Gaussian prior μ_0 including, but not limited to, Bayesian inverse problems; we work in this general setting.

Let $\mu_0 = N(0, \mathcal{C})$ be a Gaussian measure on Hilbert space $(\mathcal{H}, \langle\cdot,\cdot\rangle, \|\cdot\|)$. We assume that $\mu \ll \mu_0$ is given by

$$\frac{d\mu}{d\mu_0}(u) = \frac{1}{Z}\exp\Big(-\Phi(u)\Big), \tag{4.6a}$$

$$Z = \int_{\mathcal{H}}\exp\Big(-\Phi(u)\Big)\mu_0(du) \tag{4.6b}$$

where $Z \in (0, \infty)$. We assume that $\Phi : X \to \mathbb{R}$ where $X \subseteq \mathcal{H}$ satisfies $\mu_0(X) = 1$. We now specify X, thereby linking the properties of the reference measure μ_0 and the potential Φ.

We assume that $\mathcal{C}$ has eigendecomposition

$$\mathcal{C}\varphi_j = \gamma_j^2 \varphi_j \tag{4.7}$$

where $\{\varphi_j\}_{j=1}^{\infty}$ forms an orthonormal basis for $\mathcal{H}$, and where $\gamma_j \asymp j^{-s}$. Necessarily $s > \frac{1}{2}$ since $\mathcal{C}$ must be trace-class to be a covariance on $\mathcal{H}$. We define the following scale of Hilbert subspaces, defined for $r > 0$, by

$$\mathcal{X}^r = \{u \in \mathcal{H} \,\Big|\, \sum_{j=1}^{\infty} j^{2r} |\langle u, \varphi_j \rangle|^2 < \infty\}$$

and then extend to superspaces $r < 0$ by duality. We use $\|\cdot\|_r$ to denote the norm induced by the inner-product

$$\langle u, v \rangle_r = \sum_{j=1}^{\infty} j^{2r} u_j v_j$$

for $u_j = \langle u, \varphi_j \rangle$ and $v_j = \langle v, \varphi_j \rangle$. Application of Theorem 2.2 with $d = q = 1$ shows that $\mu_0(\mathcal{X}^r) = 1$ for all $r \in [0, s - \frac{1}{2})$. In what follows we will take $X = \mathcal{X}$ for some fixed $t \in [0, s - \frac{1}{2})$.

Notice that we have not assumed that the underlying Hilbert space is comprised of L^2 functions mapping $D \subset \mathbb{R}^d$ into $\mathbb{R}$, and hence we have not introduced the dimension d of an underlying physical space $\mathbb{R}^d$ into either the decay assumptions on the γ_j or the spaces $\mathcal{X}^r$. However, note that the spaces $\mathcal{H}^t$ introduced in subsection 2.4 are, in the case where $\mathcal{H} = L^2(D; \mathbb{R})$, the same as the spaces $\mathcal{X}^{t/d}$.

The aim of this section is to show that the equation

$$\frac{du}{dt} = -u - \mathcal{C} D\Phi(u) + \sqrt{2}\frac{dW}{dt}, \quad u(0) = u_0, \tag{4.8}$$

preserves the measure μ, where W is a $\mathcal{C}-$Wiener process, defined below. Precisely we show that $u_0 \sim \mu$ then $\mathbb{E}\phi\big(u(t)\big) = \mathbb{E}\phi(u_0)$ for all $t > 0$ for continuous bounded ϕ defined on an appropriately chosen subspace X of $\mathcal{H}$, under boundedness conditions on Φ and its derivatives.

In subsection 4.3.1 we introduce a family of Langevin equations which are invariant with respect to a given measure with smooth Lebesgue density. Using this, in subsection 4.3.2, we motivate equation (4.8) showing that, in finite dimensions, it corresponds to a particular choice of Langevin equation. In subsection 4.3.3 we describe the precise assumptions under which we will prove invariance of measure μ under the dynamics (4.8). Subsection 4.3.4 describes the elements of the finite dimensional approximation of (4.8) which will underly our proof of invariance. Finally, subsection 4.3.5 contains statement of the measure invariance result as Theorem 4.19, together with its' proof; this is preceded by Theorem 4.17 which establishes existence and uniqueness of a solution to (4.8), as well as continuous dependence of the solution on the initial condition and Brownian forcing. Theorems 4.11 and 4.9 are the finite dimensional analogues of Theorems 4.19 and 4.17 respectively and play a useful role in motivating the infinite dimensional theory.

4.3.1. Finite Dimensional Case

Before setting up the (rather involved) technical assumptions enquired for our proof of measure invariance, we give some finite-dimensional intuition. Recall that $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^n$ and we also use this notation for the induced matrix norm on $\mathbb{R}^n$. We assume that

$$I \in C^2(\mathbb{R}^n, \mathbb{R}^+), \quad \int_{\mathbb{R}^n} e^{-I(u)} du = 1.$$

Thus $\rho(u) = e^{-I(u)}$ is the Lebesgue density corresponding to a random variable on $\mathbb{R}^n$. Let μ be the corresponding measure.

Let $\mathbb{W}$ denote standard Wiener measure on $\mathbb{R}^n$. Thus $\mathbf{B} \sim \mathbb{W}$ is a standard Brownian motion in $\mathbf{C}([0,\infty);\mathbb{R}^n)$. Let $\mathbf{u} \in \mathbf{C}([0,\infty);\mathbb{R}^n)$ satisfy the SDE

$$\frac{\mathbf{du}}{\mathbf{dt}} = -\mathbf{A}\,\mathbf{D}\mathbf{I}(\mathbf{u}) + \sqrt{2\mathbf{A}}\frac{\mathbf{dB}}{\mathbf{dt}}, \ \mathbf{u}(0) = \mathbf{u}_0 \tag{4.9}$$

where $\mathbf{A} \in \mathbb{R}^{n\times n}$ is symmetric and strictly positive definite and $\mathbf{D}\mathbf{I} \in \mathbf{C}^1(\mathbb{R}^n,\mathbb{R}^n)$ is the gradient of $\mathbf{I}$. Assume that $\exists \mathbf{M} > 0 : \forall \mathbf{u} \in \mathbb{R}^n$, the Hessian of $\mathbf{I}$ satisfies

$$|\mathbf{D}^2\mathbf{I}(\mathbf{u})| \le \mathbf{M}.$$

We refer to equations of the form (4.9) as **Langevin equations**, and the matrix $\mathbf{A}$ as a **preconditioner**.

Theorem 4.9. **For every $\mathbf{u}_0 \in \mathbb{R}^n$ and $\mathbb{W}$−a.s., equation (4.9) has a unique global in time solution $\mathbf{u} \in \mathbf{C}([0,\infty);\mathbb{R}^n)$.**

Proof. A solution of the SDE is a solution of the integral equation

$$\mathbf{u}(\mathbf{t}) = \mathbf{u}_0 - \int_0^t \mathbf{A}\,\mathbf{D}\mathbf{I}\left(\mathbf{u}(\mathbf{s})\right)\mathbf{ds} + \sqrt{2\mathbf{A}}\mathbf{B}(\mathbf{t}$$

Theorem 4.11. **If $\mathbf{u}_0 \sim \boldsymbol{\mu}$ then $\mathbf{u}(\mathbf{t}) \sim \boldsymbol{\mu}$ for all $\mathbf{t} > 0$. More precisely, for all $\boldsymbol{\phi} : \mathbb{R}^n \to \mathbb{R}^+$ bounded and continuous, $\mathbf{u}_0 \sim \boldsymbol{\mu}$ implies**

$$\mathbb{E}\boldsymbol{\phi}\big(\mathbf{u}(\mathbf{t})\big) = \mathbb{E}\boldsymbol{\phi}(\mathbf{u}_0), \ \forall \mathbf{t} > 0.$$

Proof. Consider the additive noise SDE, for additive noise with strictly positive-definite diffusion matrix Σ,

$$\frac{\mathbf{du}}{\mathbf{dt}} = \mathbf{f}(\mathbf{u}) + \sqrt{2\Sigma}\frac{\mathbf{dW}}{\mathbf{dt}}, \ \mathbf{u}(0) = \mathbf{u}_0 \sim \boldsymbol{\nu}_0.$$

If $\boldsymbol{\nu}_0$ has pdf $\boldsymbol{\rho}_0$, then the Fokker-Planck equation for this SDE is

$$\begin{aligned} \frac{\partial \overline{\boldsymbol{\rho}}}{\partial \mathbf{t}} &= \nabla \cdot (-\mathbf{f}\overline{\boldsymbol{\rho}} + \Sigma\nabla\overline{\boldsymbol{\rho}}), \ (\mathbf{u}, \mathbf{t}) \in \mathbb{R}^n \times \mathbb{R}^+, \\ \overline{\boldsymbol{\rho}}|_{t=0} &= \boldsymbol{\rho}_0. \end{aligned}$$

At time $\mathbf{t} > 0$ the solution of the SDE is distributed according to measure $\boldsymbol{\nu}(\mathbf{t})$ with density $\overline{\boldsymbol{\rho}}(\mathbf{u}, \mathbf{t})$ solving the Fokker-Planck equation. Thus the initial measure $\boldsymbol{\nu}_0$ is preserved if

$$\nabla \cdot (-\mathbf{f}\boldsymbol{\rho}_0 + \Sigma\nabla\boldsymbol{\rho}_0) = 0$$

and then $\overline{\boldsymbol{\rho}}(\cdot, \mathbf{t}) = \boldsymbol{\rho}_0, \ \forall \mathbf{t} \geq 0$.

We apply this Fokker-Planck equation to show that $\boldsymbol{\mu}$ is invariant for equation (4.10). We need to show that

$$\nabla \cdot \big(\mathbf{A}\mathbf{D}\mathbf{I}(\mathbf{u})\boldsymbol{\rho} + \mathbf{A}\nabla\boldsymbol{\rho}\big) = 0$$

if $\boldsymbol{\rho} = \mathbf{e}^{-I(u)}$. But then

$$\nabla\boldsymbol{\rho} = -\mathbf{D}\mathbf{I}(\mathbf{u})\mathbf{e}^{-I(u)} = -\mathbf{D}\mathbf{I}(\mathbf{u})\boldsymbol{\rho}.$$

Thus

$$\mathbf{A}\,\mathbf{D}\mathbf{I}(\mathbf{u})\boldsymbol{\rho} + \mathbf{A}\nabla\boldsymbol{\rho} = \mathbf{A}\,\mathbf{D}\mathbf{I}(\mathbf{u})\boldsymbol{\rho} - \mathbf{A}\,\mathbf{D}\mathbf{I}(\mathbf{u})\boldsymbol{\rho} = 0,$$

so that

$$\nabla \cdot \big(\mathbf{A}\,\mathbf{D}\mathbf{I}(\mathbf{u})\boldsymbol{\rho} + \mathbf{A}\nabla\boldsymbol{\rho}\big) = \nabla \cdot (0) = 0.$$

Hence the proof is complete. □

4.3.2. Motivation for Equation (4.8)

Using the preceding finite dimensional development, we now motivate the form of equation (4.8). For (4.6) we have, if $\mathcal{H}$ is $\mathbb{R}^n$,

$$\begin{aligned} \boldsymbol{\mu}(\mathbf{du}) &= \boldsymbol{\rho}(\mathbf{u})\mathbf{du}, \\ \boldsymbol{\rho}(\mathbf{u}) &= \exp\big(-\mathbf{I}(\mathbf{u})\big), \\ \mathbf{I}(\mathbf{u}) &= \frac{1}{2}|\mathcal{C}^{-\frac{1}{2}}\mathbf{u}|^2 + \Phi(\mathbf{u}) + \ln \mathbf{Z}. \end{aligned}$$

Thus

$$\mathbf{D}\mathbf{I}(\mathbf{u}) = \mathcal{C}^{-1}\mathbf{u} + \mathbf{D}\Phi(\mathbf{u})$$

and equation (4.9), which preserves $\boldsymbol{\mu}$, is

$$\frac{\mathbf{du}}{\mathbf{dt}} = -\mathbf{A}\big(\mathcal{C}^{-1}\mathbf{u} + \mathbf{D}\Phi(\mathbf{u})\big) + \sqrt{2\mathbf{A}}\frac{\mathbf{dB}}{\mathbf{dt}}.$$

Choosing the preconditioner $\mathbf{A} = \mathcal{C}$ gives

$$\frac{\mathbf{du}}{\mathbf{dt}} = -\mathbf{u} - \mathcal{C}\mathbf{D}\Phi(\mathbf{u}) + \sqrt{2\mathcal{C}}\frac{\mathbf{dB}}{\mathbf{dt}}.$$

This is exactly (4.8) provided $W = \sqrt{\mathcal{C}}B$, where B is a Brownian motion with covariance $\mathcal{I}$. Then W is a Brownian motion with covariance $\mathcal{C}$.

We provide further detail on the construction of W, using the discussion in Remark 4.10 to guide us. In the infinite dimensional case we define a **cylindrical Wiener process** by

$$B(t) = \sum_{j=1}^{\infty} \beta_j \varphi_j$$

where $\{\beta_j\}_{j=1}^{\infty}$ is an i.i.d. family of Brownian motions on $\mathbb{R}$ with $\beta_j \in C([0,\infty);\mathbb{R})$. Since $\sqrt{\mathcal{C}}\varphi_j = \gamma_j \varphi_j$, the $\mathcal{C}-$Wiener process $W = \sqrt{\mathcal{C}}B$ is then

$$W(t) = \sum_{j=1}^{\infty} \gamma_j \beta_j(t) \varphi_j. \tag{4.13}$$

The following formal calculation gives insight into the properties of W:

$$\begin{aligned}
\mathbb{E}W(t) \otimes W(s) &= \mathbb{E}\left(\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \gamma_j\gamma_k\beta_j(t)\beta_k(s)\varphi_j \otimes \varphi_k\right) \\
&= \left(\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \gamma_j\gamma_k\mathbb{E}\left(\beta_j(t)\beta_k(t)\right)\varphi_j \otimes \varphi_k\right) \\
&= \left(\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \gamma_j\gamma_k\delta_{jk}(t\wedge s)\varphi_j \otimes \varphi_k\right) \\
&= \sum_{j=1}^{\infty}\left(\gamma_j^2\varphi_j \otimes \varphi_j\right) t \wedge s \\
&= \mathcal{C}\,(t \wedge s).
\end{aligned}$$

Thus the process has the covaraince structure of Brownian motionin time, and covariance operator $\mathcal{C}$ in space. Hence the name $\mathcal{C}-$Wiener process.

In order to make sense of this infinite sum we follow an approach similar to that used in Theorem 2.4 to make sense of Gaussian random sums. To this end, consider the finite sum

$$W^N(t) = \sum_{j=1}^{N} \gamma_j \beta_j(t) \varphi_j.$$

Let $(\Omega, \mathcal{F}, \mathbb{P})$ denote the probability space underlying the i.i.d. sequence of unit Brownian motions used to construct W.

Theorem 4.12. **The sequence of functions $\{W^N\}_{N=1}^{\infty}$ is Cauchy in the Banach space $L^2_{\mathbb{P}}\left(\Omega; C([0,T];\mathcal{H}^t)\right)$, $t < s - \frac{d}{2}$. Thus the infinite series exists (4.13) as an L^2-limit and takes values in $C([0,T];\mathcal{H}^t)$ for $t < s - \frac{d}{2}$.**

We are now in a position to prove Theorems 4.17 and 4.19 which are the infinite dimensional analogues of Theorems 4.9 and 4.11.

4.3.3. Assumptions on Change of Measure

Recall that $\mu_0(\mathcal{X}^r) = 1$ for all $r \in [0, s - \frac{1}{2})$. The functional $\Phi(\cdot)$ is assumed to be defined on $\mathcal{X}^t$ for some $t \in [0, s - \frac{1}{2})$, and indeed we will assume appropriate bounds on the first and second derivatives, building on this assumption. These regularity assumptions on $\Phi(\cdot)$ that ensure that the probability distribution μ is not too different from μ_0, when projected into directions associated with φ_j for j large.

For each $\mathbf{x} \in \mathcal{X}^t$ the derivative $\mathbf{D}\Phi(\mathbf{u})$ is an element of the dual $(\mathcal{X}^t)^*$ of $\mathcal{X}^t$ comprising continuous linear functionals on $\mathcal{X}^t$. However, we may identify $(\mathcal{X}^t)^*$ with $\mathcal{X}^{-t}$ and view $\mathbf{D}\Phi(\mathbf{u})$ as an element of $\mathcal{X}^{-t}$ for each $\mathbf{x} \in \mathcal{X}^t$. With this identification, the following identity holds

$$\|\mathbf{D}\Phi(\mathbf{u})\|_{\mathcal{L}(\mathcal{X}^t,\mathbb{R})} = \|\mathbf{D}\Phi(\mathbf{u})\|_{-t}$$

and the second derivative $\mathbf{D}^2\Phi(\mathbf{u})$ can be identified as an element of $\mathcal{L}(\mathcal{X}^t, \mathcal{X}^{-t})$. To avoid technicalities we assume that $\Phi(\cdot)$ is quadratically bounded, with first derivative linearly bounded and second derivative globally bounded. Weaker assumptions could be dealt with by use of stopping time arguments.

Assumptions 4.13. **There exist constants $\mathbf{M}_i \in \mathbb{R}$, $\mathbf{i} \leq 4$ and $\mathbf{t} \in [0, \mathbf{s} - 1/2)$ such that, for all $\mathbf{u} \in \mathcal{X}^t$, the functional $\Phi : \mathcal{X}^t \to \mathbb{R}$ satisfies**

$$\begin{aligned} -\mathbf{M}_1 \leq \Phi(\mathbf{u}) &\leq \mathbf{M}_2 \left(1 + \|\mathbf{u}\|_t^2\right); \\ \|\mathbf{D}\Phi(\mathbf{u})\|_{-t} &\leq \mathbf{M}_3 \left(1 + \|\mathbf{u}\|_t\right); \\ \|\mathbf{D}^2\Phi(\mathbf{u})\|_{\mathcal{L}(\mathcal{X}^t,\mathcal{X}^{-t})} &\leq \mathbf{M}_4. \end{aligned}$$

Example The functional $\Phi(\mathbf{u}) = \frac{1}{2}\|\mathbf{u}\|_t^2$ satisfies Assumptions 4.13. It is defined on $\mathcal{X}^t$ and its derivative at $\mathbf{x} \in \mathcal{X}^t$ is given by $\mathbf{D}\Phi(\mathbf{u}) = \sum_{j\geq 0} \mathbf{j}^{2t}\mathbf{u}_j\boldsymbol{\varphi}_j \in \mathcal{X}^{-t}$ with $\|\mathbf{D}\Phi(\mathbf{u})\|_{-t} = \|\mathbf{u}\|_t$. The second derivative $\mathbf{D}^2\Phi(\mathbf{u}) \in \mathcal{L}(\mathcal{X}^t, \mathcal{X}^{-t})$ is the linear operator that maps $\mathbf{u} \in \mathcal{X}^t$ to $\sum_{j\geq 1} \mathbf{j}^{2t}\langle \mathbf{u}, \boldsymbol{\varphi}_j\rangle\boldsymbol{\varphi}_j \in \mathcal{X}^t$: its norm satisfies $\|\mathbf{D}^2\Phi(\mathbf{u})\|_{\mathcal{L}(\mathcal{X}^t,\mathcal{X}^{-t})} = 1$ for any $\mathbf{x} \in \mathcal{X}^t$. □

Since the eigenvalues $\boldsymbol{\gamma}_j^2$ of $\mathcal{C}$ decrease as $\boldsymbol{\gamma}_j \asymp \mathbf{j}^{-s}$, the operator $\mathcal{C}$ has a smoothing effect: $\mathcal{C}^\alpha\mathbf{h}$ gains $2\boldsymbol{\alpha}\mathbf{s}$ orders of regularity in the sense that the $\mathcal{X}^\beta$-norm of $\mathcal{C}^\alpha\mathbf{h}$ is controlled by the $\mathcal{X}^{\beta-2\alpha s}$-norm of $\mathbf{h} \in \mathcal{H}$. Indeed we have the following:

Lemma 4.14. **Under Assumptions 4.13, the following estimates hold:**

1. **The operator $\mathcal{C}$ satisfies**
$$\|\mathcal{C}^\alpha\mathbf{h}\|_\beta \asymp \|\mathbf{h}\|_{\beta-2\alpha s}.$$
2. **The function $\mathcal{C}\mathbf{D}\Phi : \mathcal{X}^t \to \mathcal{X}^t$ is globally Lipschitz on $\mathcal{X}^t$: there exists a constant $\mathbf{M}_5 > 0$ such that**
$$\|\mathcal{C}\mathbf{D}\Phi(\mathbf{u}) - \mathcal{C}\mathbf{D}\Phi(\mathbf{v})\|_t \leq \mathbf{M}_5\,\|\mathbf{u} - \mathbf{v}\|_t \qquad \forall \mathbf{u}, \mathbf{v} \in \mathcal{X}^t.$$
3. **The function $\mathbf{F} : \mathcal{X}^t \to \mathcal{X}^t$ defined by**
$$\mathbf{F}(\mathbf{u}) = -\mathbf{u} - \mathcal{C}\mathbf{D}\Phi(\mathbf{u}) \tag{4.14}$$
is globally Lipschitz on $\mathcal{X}^t$.
4. **The functional $\Phi(\cdot) : \mathcal{X}^t \to \mathbb{R}$ satisfies a second order Taylor formula[1]. There exists a constant $\mathbf{M}_6 > 0$ such that**
$$\Phi(\mathbf{v}) - \Big(\Phi(\mathbf{u}) + \langle \mathbf{D}\Phi(\mathbf{u}), \mathbf{v} - \mathbf{u}\rangle\Big) \leq \mathbf{M}_6\,\|\mathbf{u} - \mathbf{v}\|_t^2 \qquad \forall \mathbf{u}, \mathbf{v} \in \mathcal{X}^t. \tag{4.15}$$

4.3.4. Finite Dimensional Approximation

Our analysis now proceeds as follows. First we introduce an approximation of the measure $\boldsymbol{\mu}$, denoted by $\boldsymbol{\mu}^N$. To this end we let $\mathbf{P}^N$ denote orthogonal projection in $\mathcal{H}$ onto $\mathbf{X}^N := \mathrm{span}\{\boldsymbol{\varphi}_1, \cdots, \boldsymbol{\varphi}_N\}$ and denote by $\mathbf{Q}^N$ orthogonal projection in $\mathcal{H}$ onto $\mathbf{X}^\perp := \mathrm{span}\{\boldsymbol{\varphi}_{N+1}, \boldsymbol{\varphi}_{N+2}, \cdots\}$. Thus $\mathbf{Q}^N = \mathbf{I} - \mathbf{P}^N$. Then define the measure $\boldsymbol{\mu}^N$ by

$$\frac{\mathbf{d}\boldsymbol{\mu}^N}{\mathbf{d}\boldsymbol{\mu}_0}(\mathbf{u}) = \frac{1}{\mathbf{Z}^N}\exp\Big(-\Phi(\mathbf{P}^N\mathbf{u})\Big), \tag{4.16a}$$

$$\mathbf{Z}^N = \int_{X'} \exp\Big(-\Phi(\mathbf{P}^N\mathbf{u})\Big)\boldsymbol{\mu}_0(\mathbf{du}). \tag{4.16b}$$

[1] We extend $\langle\cdot,\cdot\rangle$ from an inner product on $\mathcal{X}$ to the dual pairing between $\mathcal{X}^{-t}$ and $\mathcal{X}^t$

This is a specific example of the approximating family in (4.4) if we define

$$\Phi^N = \Phi \circ \mathbf{P}^N. \tag{4.17}$$

Indeed if we take $\mathbf{X} = \mathcal{X}^\tau$ for any $\boldsymbol{\tau} \in (\mathbf{t}, \mathbf{s} - \frac{1}{2})$ we see that $\|\mathbf{P}^N\|_{\mathcal{L}(X,X)} = 1$ and that, for any $\mathbf{u} \in \mathbf{X}$,

$$\begin{aligned}\|\Phi(\mathbf{u}) - \Phi^N(\mathbf{u})\| &= \|\Phi(\mathbf{u}) - \Phi(\mathbf{P}^N\mathbf{u})\| \\ &\le \mathbf{M}_3(1 + \|\mathbf{u}\|_t)\|(\mathbf{I} - \mathbf{P}^N)\mathbf{u}\|_t \\ &\le \mathbf{C}\mathbf{M}_3(1 + \|\mathbf{u}\|_\tau)\|\mathbf{u}\|_\tau \mathbf{N}^{-(\tau - t)}.\end{aligned}$$

Since Φ, and hence Φ^N, are bounded below by $-\mathbf{M}_1$, and since the function $1 + \|\mathbf{u}\|_\tau^2$ is integrable by the Fernique Theorem 2.6, the approximation Theorem 4.7 applies. We deduce that the Hellinger distance between $\boldsymbol{\mu}$ and $\boldsymbol{\mu}^N$ is bounded above by $\mathcal{O}(\mathbf{N}^{-r})$ for any $\mathbf{r} < \mathbf{s} - \frac{1}{2} - \mathbf{t}$ since $\boldsymbol{\tau} - \mathbf{t} \in (0, \mathbf{s} - \frac{1}{2} - \mathbf{t})$.

We will not use this explicit convergence rate in what follows, but we will use the idea that $\boldsymbol{\mu}^N$ converges to $\boldsymbol{\mu}$ in order to prove invariance of the measure $\boldsymbol{\mu}$ under the SDE (4.8). The measure $\boldsymbol{\mu}^N$ has a product structure that we will exploit in the following. We note that any element $\mathbf{u} \in \mathcal{H}$ is uniquel-3.4em31213(c)-1.12(e)-1dΩ[(4(in)

4.3.5. Main Theorem and Proof

We define a solution of (4.8) to be a function $\mathbf{u} \in \mathbf{C}([0,\mathbf{T}];\mathcal{X}^t)$ satisfying the integral equation

$$\mathbf{u}(\tau) = \mathbf{u}_0 + \int_0^\tau \mathbf{F}\big(\mathbf{u}(\mathbf{s})\big)\mathbf{ds} + \sqrt{2}\,\mathbf{W}(\tau) \qquad \forall \tau \in [0,\mathbf{T}]. \tag{4.21}$$

The solution is said to be **global** if $\mathbf{T} > 0$ is arbitrary. Similarly a solution of (4.18) is a function $\mathbf{u} \in \mathbf{C}([0,\mathbf{T}];\mathcal{X}^t)$ satisfying the integral equation

$$\mathbf{u}^N(\tau) = \mathbf{u}_0 + \int_0^\tau \mathbf{F}^N\big(\mathbf{u}^N(\mathbf{s})\big)\mathbf{ds} + \sqrt{2}\,\mathbf{W}(\tau) \qquad \forall \mathbf{t} \in [0,\mathbf{T}]. \tag{4.22}$$

The following establishes basic existence, uniqueness, continuity and approximation properties of the solutions of (4.21) and (4.22).

Theorem 4.17. **For every $\mathbf{u}_0 \in \mathcal{X}^t$ and for almost every $\mathcal{C}-$Wiener process $\mathbf{W}$, equation (4.21) (respectively (4.22)) has a unique global solution. For any pair $(\mathbf{u}_0,\mathbf{W}) \in \mathcal{X}^t \times \mathbf{C}([0,\mathbf{T}];\mathcal{X}^t)$ we define the Itô map**

$$\Theta\colon \mathcal{X}^t \times \mathbf{C}([0,\mathbf{T}];\mathcal{X}^t) \to \mathbf{C}([0,\mathbf{T}];\mathcal{X}^t)$$

which maps $(\mathbf{u}_0,\mathbf{W})$ to the unique solution $\mathbf{u}$ (resp. $\mathbf{u}^N$ for (4.22)) of the integral equation (4.21) (resp. Θ^N for (4.22)). The map Θ (resp. Θ^N) is globally Lipschitz continuous. Finally we have that $\Theta^N(\mathbf{u}_0,\mathbf{W}) \to \Theta(\mathbf{u}_0,\mathbf{W})$ for every pair $(\mathbf{u}_0,\mathbf{W}) \in \mathcal{X}^t \times \mathbf{C}([0,\mathbf{T}];\mathcal{X}^t)$.

Proof. The existence and uniqueness of local solutions to the integral equation (4.21) is a simple application of the contraction mapping principle, following arguments similar to those employed when studying the Itó map below. Extension to a global solution may be achieved by repeating the local argument on succesive intervals.

Now let $\mathbf{u}^{(i)}$ solve

$$\mathbf{u}^{(i)} = \mathbf{u}_0^{(i)} + \int_0^\tau \mathbf{F}(\mathbf{u}^{(i)})(\mathbf{s})\mathbf{ds} + \sqrt{2}\mathbf{W}^{(i)}(\tau), \quad \tau \in [0,\mathbf{T}],$$

for $\mathbf{i} = 1,2$. Subtracting and using the Lipschitz property of $\mathbf{F}$ shows that $\mathbf{e} = \mathbf{u}^{(1)} - \mathbf{u}^{(2)}$ satisfies

$$\begin{aligned}\|\mathbf{e}(\tau)\|_t &\le \|\mathbf{u}_0^{(1)} - \mathbf{u}_0^{(2)}\|_t + \mathbf{L}\int_0^\tau \|\mathbf{e}(\mathbf{s})\|_t\mathbf{ds} + \sqrt{2}\|\mathbf{W}^{(1)}(\tau) - \mathbf{W}^{(2)}(\tau)\|_t \\ &\le \|\mathbf{u}_0^{(1)} - \mathbf{u}_0^{(2)}\|_t + \mathbf{L}\int_0^\tau \|\mathbf{e}(\mathbf{s})\|_t\mathbf{ds} + \sqrt{2}\sup_{0\le s\le T}\|\mathbf{W}^{(1)}(\mathbf{s}) - \mathbf{W}^{(2)}(\mathbf{s})\|_t.\end{aligned}$$

By application of the Gronwall inequality we find that

$$\sup_{0\le \tau\le T}\|\mathbf{e}(\tau)|_t \le \mathbf{C}(\mathbf{T})\Big(\|\mathbf{u}_0^{(1)} - \mathbf{u}_0^{(2)}\|_t + \sup_{0\le s\le T}\|\mathbf{W}^{(1)}(\mathbf{s}) - \mathbf{W}^{(2)}(\mathbf{s})\|_t\Big)$$

and the desired continuity is established.

Now we prove pointwise convergence of Θ^N to Θ. Let $\mathbf{e} = \mathbf{u} - \mathbf{u}^N$ where $\mathbf{u}$ and $\mathbf{u}^N$ solve (4.21), (4.22) respectively. The pointwise convergence of Θ^N to Θ is established by proving that $\mathbf{e} \to 0$ in $\mathbf{C}([0,\mathbf{T}];\mathcal{X}^t)$. Note that

$$\mathbf{F}(\mathbf{u}) - \mathbf{F}^N(\mathbf{u}^N) = \big(\mathbf{F}^N(\mathbf{u}) - \mathbf{F}^N(\mathbf{u}^N)\big) + \big(\mathbf{F}(\mathbf{u}) - \mathbf{F}^N(\mathbf{u})\big).$$

Also, by Lemma 4.16, $\|\mathbf{F}^N(\mathbf{u}) - \mathbf{F}^N(\mathbf{u}^N)\|_t \le \mathbf{L}\|\mathbf{e}\|_t$. Thus we have

$$\|\mathbf{e}\|_t \le \mathbf{L}\int_0^\tau \|\mathbf{e}(\mathbf{s})\|_t\mathbf{ds} + \int_0^\tau \|\mathbf{F}\big(\mathbf{u}(\mathbf{s})\big) - \mathbf{F}^N\big(\mathbf{u}(\mathbf{s})\big)\|_t\mathbf{ds}.$$

Thus, by Gronwall, it suffices to show that

$$\delta^N := \sup_{0\le s\le T}\|\mathbf{F}\big(\mathbf{u}(\mathbf{s})\big) - \mathbf{F}^N\big(\mathbf{u}(\mathbf{s})\big)\|_t$$

tends to zero as $\mathbf{N} \to \infty$. Note that

$$\begin{aligned}\mathbf{F}(\mathbf{u}) - \mathbf{F}^N(\mathbf{u}) &= \mathcal{C}\mathbf{D}\Phi(\mathbf{u}) - \mathcal{C}\mathbf{P}^N\mathbf{D}\Phi(\mathbf{P}^N\mathbf{u}) \\ &= (\mathbf{I} - \mathbf{P}^N)\mathcal{C}\mathbf{D}\Phi(\mathbf{u}) + \mathbf{P}^N\big(\mathcal{C}\mathbf{D}\Phi(\mathbf{u}) - \mathcal{C}\mathbf{D}\Phi(\mathbf{P}^N\mathbf{u})\big).\end{aligned}$$

Thus, since $\mathcal{C}\mathbf{D}\Phi$ is globally Lipschitz on $\mathcal{X}^t$, by Lemma 4.14, and $\mathbf{P}^N$ has norm one as a mapping from $\mathcal{X}^t$ into itself,

$$\|\mathbf{F}(\mathbf{u}) - \mathbf{F}^N(\mathbf{u})\|_t \leq \|(\mathbf{I} - \mathbf{P}^N)\mathcal{C}\mathbf{D}\Phi(\mathbf{u})\|_t + \mathbf{C}\|(\mathbf{I} - \mathbf{P}^N)\mathbf{u}\|_t.$$

By dominated convergence $\|(\mathbf{I} - \mathbf{P}_N)\mathbf{a}\|_t \to 0$ for any fixed element $\mathbf{a} \in \mathcal{X}^t$. Thus, because $\mathcal{C}\mathbf{D}\Phi$ is globally Lipschitz, by Lemma 4.14, because $\mathbf{u} \in \mathbf{C}([0, \mathbf{T}]; \mathcal{X}^t)$, we deduce that it suffices to bound $\sup_{0 \leq s \leq T} \|\mathbf{u}(\mathbf{s})\|_t$. But such a bound is a consequence of the existence Theorem 4.17. □

The following is a straightforward corollary of the preceding theorem:

Corollary 4.18. **For any pair $(\mathbf{u}_0, \mathbf{W}) \in \mathcal{X}^t \times \mathbf{C}([0, \mathbf{T}]; \mathcal{X}^t)$ we define the point Itô map**

$$\Theta_t \colon \mathcal{X}^t \times \mathbf{C}([0, \mathbf{T}]; \mathcal{X}^t) \to \mathcal{X}^t$$

which maps $(\mathbf{u}_0, \mathbf{W})$ to the unique solution $\mathbf{u}(\mathbf{t})$ of the integral equation (4.21) (resp. $\mathbf{u}^N(\mathbf{t})$ for (4.22)) at time t (resp. Θ_t^N for (4.22)). The map Θ_t (resp. Θ_t^N) is globally Lipschitz continuous. Finally we have that $\Theta_t^N(\mathbf{u}_0, \mathbf{W}) \to \Theta_t(\mathbf{u}_0, \mathbf{W})$ for every pair $(\mathbf{u}_0, \mathbf{W}) \in \mathcal{X}^t \times \mathbf{C}([0, \mathbf{T}]; \mathcal{X}^t)$.

Theorem 4.19. **Let Assumptions 4.13 hold. Then the measure $\boldsymbol{\mu}$ given by (4.3) is invariant for (4.8): for all continuous bounded functions $\boldsymbol{\phi} : \mathcal{X}^t \to \mathbb{R}$ it follows that, if $\mathbb{E}$ denotes expectation with respect to the product measure found from initial condition $\mathbf{u}_0 \sim \boldsymbol{\mu}$ and $\mathbf{W} \sim \mathbb{W}$, the $\mathcal{C}-$Wiener measure on $\mathcal{X}^t$, then $\mathbb{E}\boldsymbol{\phi}\big(\mathbf{u}(\mathbf{t})\big) = \mathbb{E}\boldsymbol{\phi}(\mathbf{u}_0)$.**

Proof. We have that

$$\mathbb{E}\boldsymbol{\phi}\big(\mathbf{u}(\mathbf{t})\big) = \int \boldsymbol{\phi}\big(\Theta_t(\mathbf{u}_0, \mathbf{W})\big)\boldsymbol{\mu}(\mathbf{du}_0)\mathbb{W}(\mathbf{dW}), \tag{4.23}$$

$$\mathbb{E}\boldsymbol{\phi}(\mathbf{u}_0) = \int \boldsymbol{\phi}(\mathbf{u}_0)\boldsymbol{\mu}(\mathbf{du}_0). \tag{4.24}$$

If we solve equation (4.18) with $\mathbf{u}_0 \sim \boldsymbol{\mu}^N$ then, using $\mathbb{E}^N$ with the obvious notation,

$$\mathbb{E}^N\boldsymbol{\phi}\big(\mathbf{u}^N(\mathbf{t})\big) = \int \boldsymbol{\phi}\big(\Theta_t^N(\mathbf{u}_0, \mathbf{W})\big)\boldsymbol{\mu}^N(\mathbf{du}_0)\mathbb{W}(\mathbf{dW}), \tag{4.25}$$

$$\mathbb{E}^N\boldsymbol{\phi}(\mathbf{u}_0) = \int \boldsymbol{\phi}(\mathbf{u}_0)\boldsymbol{\mu}^N(\mathbf{du}_0). \tag{4.26}$$

Lemma 4.20 below shows that, in fact,

$$\mathbb{E}^N\boldsymbol{\phi}\big(\mathbf{u}^N(\mathbf{t})\big) = \mathbb{E}^N\boldsymbol{\phi}(\mathbf{u}_0).$$

Thus it suffices to show that

$$\mathbb{E}^N\boldsymbol{\phi}\big(\mathbf{u}^N(\mathbf{t})\big) \to \mathbb{E}\boldsymbol{\phi}\big(\mathbf{u}(\mathbf{t})\big) \tag{4.27}$$

and

$$\mathbb{E}^N\boldsymbol{\phi}(\mathbf{u}_0) \to \mathbb{E}\boldsymbol{\phi}(\mathbf{u}_0). \tag{4.28}$$

Both of these facts follow from the dominated convergence theorem as we now show. First note that

$$\mathbb{E}^N\boldsymbol{\phi}(\mathbf{u}_0) = \int \boldsymbol{\phi}(\mathbf{u}_0)\mathbf{e}^{-\Phi(P^N u_0)}\boldsymbol{\mu}_0(\mathbf{du}_0).$$

Since $\boldsymbol{\phi}(\cdot)\mathbf{e}^{-\Phi\circ P^N}$ is bounded independently of $\mathbf{N}$, by $(\sup\boldsymbol{\phi})\mathbf{e}^{M_1}$, and since $(\Phi \circ \mathbf{P}^N)(\mathbf{u})$ converges pointwise to $\Phi(\mathbf{u})$ on $\mathcal{X}^t$, we deduce that

$$\mathbb{E}^N\boldsymbol{\phi}(\mathbf{u}_0) \to \int \boldsymbol{\phi}(\mathbf{u}_0)\mathbf{e}^{-\Phi(u_0)}\boldsymbol{\mu}_0(\mathbf{du}_0) = \mathbb{E}\boldsymbol{\phi}(\mathbf{u}_0)$$

so that (4.28) holds. The convergence in (4.27) holds by a similar argument. From (4.29) we have

$$\mathbb{E}^N \phi\big(u^N(t)\big) = \int \phi\big(\Theta_t^N(u_0, W)\big) e^{-\Phi(P^N u_0)} \mu_0(du_0) \mathbb{W}(dW). \tag{4.29}$$

The integrand is again dominated by $(\sup \phi) e^{M_1}$. Using the pointwise convergence of Θ_t^N to Θ_t on $\mathcal{X}^t \times C([0, T]; \mathcal{X}^t)$, as proved in Corollary 4.18, as well as the pointwise convergence of $(\Phi \circ P^N)(u)$ to $\Phi(u)$, the desired result follows from dominated convergence: we find that

$$\mathbb{E}^N \phi\big(u^N(t)\big) \to \int \phi\big(\Theta_t(u_0, W)\big) e^{-\Phi(u_0)} \mu_0(du_0) \mathbb{W}(dW) = \mathbb{E}\phi\big(u(t)\big).$$

The desired result follows. □

Lemma 4.20. **Let Assumptions 4.13 hold. Then the measure μ^N given by (4.16) is invariant for (4.18): for all continuous bounded functions $\phi : \mathcal{X}^t \to \mathbb{R}$ it follows that, if $\mathbb{E}^N$ denotes expectation with respect to the product measure found from initial condition $u_0 \sim \mu^N$ and $W \sim \mathbb{W}$, the $\mathcal{C}-$Wiener measure on $\mathcal{X}^t$, then $\mathbb{E}^N \phi\big(u^N(t)\big) = \mathbb{E}^N \phi(u_0)$.**

Proof. Recall from Lemma 4.15 that measure μ^N given by (4.16) factors as the independent product of two measures on μ_P on X^N and μ_Q on $X^\perp$. On $X^\perp$ the measure is simply the Gaussian $\mu_Q = \mathcal{N}(0, \mathcal{C}^\perp)$, whilst X^N the measure μ_P is finite dimensional with density proportional to

$$\exp\Big(-\Phi(p) - \frac{1}{2}\|(\mathcal{C}^N)^{-\frac{1}{2}} p\|^2\Big). \tag{4.30}$$

The equation (4.18) also decouples on the spaces X^N and $X^\perp$. On $X^\perp$ it is simply

$$\frac{dq}{dt} = -q + \sqrt{2} Q^N \frac{dW}{dt} \tag{4.31}$$

whilst on X^N it is

$$\frac{dp}{dt} = -p - \mathcal{C}^N D\Phi(p) + \sqrt{2} P^N \frac{dW}{dt}. \tag{4.32}$$

Measure μ_Q is preserved by (4.31), because (4.31) simply gives an Ornstein-Uhlenbeck process with desired Gaussian invariant measure. On the other hand, equation (4.32) is simply a Langevin equation for measure on $\mathbb{R}^N$ with density (4.30) and a calculation with the Fokker-Planck equation, as in Theorem 4.11, demonstrates the required invariance of μ_P. □

4.4. MCMC Methods

The perspective that we have described on inverse problems leads to new sampling methods which are specifically tailored to the infinite dimensional setting, and its approximation by finite dimensional measures. In particular it leads naturally to the design of algorithms which perform well under refinement of the finite dimensionalization. To illustrate this idea we consider the setting of Section 4.4 and study random walk type algorithms.

First of all we describe the standard Random Walk Metropolis (RWM) algorithm, designed to sample a measure on $\mathbb{R}^N$. To this end we notice that the measure μ^N given by (4.16) factors as the product of two independent measures on X^N and $\mathcal{H} \backslash X^N$. The measure on $\mathcal{H} \backslash X^N$ is given by the prior and is easily sampled. Thus it remains to sample the measure on X^N. This space is isomorphic to $\mathbb{R}^N$. We define

$$I(u) = \Phi(u) + \frac{1}{2}\|\mathcal{C}^{-\frac{1}{2}} u\|^2. \tag{4.33}$$

Then, for $u \in X^N$, the measure of interest has Lebesgue density

$$\pi^N(u) \propto \exp\big(-I(u)\big).$$

This standard RWM algorithm defines a Markov chain $\{u^k\}$ on X^N as follows.

- Set $k = 0$ and Pick $u^{(0)} \in X^N$.
- Propose $v^{(k)} = u^{(k)} + \beta P^N \xi^{(k)}, \quad \xi^{(k)} \sim N(0, \mathcal{C})$.
- Set $u^{(k+1)} = v^{(k)}$ with proability $a(u^{(k)}, v^{(k)})$, independently of $u^{(k)}, \xi^{(k)}$.
- Set $u^{(k+1)} = u^{(k)}$ otherwise.
- $k \to k + 1$.

Here

$$a(u, v) = \min\{1, \exp\big(I(u) - I(v)\big)\}.$$

This Markov chain leaves the density π^N as defined above invariant. It is, however, badly behaved in the limit $N \to \infty$. This is because

$$\lim_{N\to\infty} I(P^N u) = \infty$$

almost surely for $u \sim \mu$.

To overcome this issue we introduce a new RWM algorithm which is defined on the whole of $\mathcal{H}$, not just on finite truncations. The algorithm is defined as follows, when applied on X^N:

- Set $k = 0$ and Pick $u^{(0)} \in X^N$.
- Propose $v^{(k)} = \sqrt{(1 - \beta^2)}u^{(k)} + \beta P^N \xi^{(k)}, \quad \xi^{(k)} \sim N(0, \mathcal{C})$.
- Set $u^{(k+1)} = v^{(k)}$ with proability $a(u^{(k)}, v^{(k)})$, independently of $u^{(k)}$ and $\xi^{(k)}$.
- Set $u^{(k+1)} = u^{(k)}$ otherwise.
- $k \to k + 1$.

Here

$$a(u, v) = \min\{1, \exp\big(\Phi(u) - \Phi(v)\big)\}.$$

Notice that the small change in proposal, when compared with the standard RWN, results in an acceptance probability defined via differences of Φ and not I. Because Φ is a.s. finite with respect to μ, whilst I is not, this leads to a considerably improved algorithm which has desirable $N-$independent properties when implemented on a sequence of approximating problems with $N \to \infty$.

To quantify this it is useful to introduce the concept of **spectral gap.** Define the spaces

$$L^2_\mu = \{f : X \to \mathbb{R} : \|f\|_2^2 := \mathbb{E}^\mu |f(u)|^2 < \infty\},$$
$$L^2_0 = \{f \in L^2_\mu : \mu(f) = 0.\}$$

Define the Markov kernel

$$(Pf)(u) = \mathbb{E}\Big(f\big(u^{(1)}\big)|u^{(0)} = u\Big).$$

Then set

$$\|P\|_{L^2_0 \to L^2_0} := \sup_{f \in L^2_0} \frac{\|Pf\|_2^2}{\|f\|_2^2}.$$

We have $L^2_\mu-$**spectral gap** γ if $\|P\|_{L^2_0\to L^2_0} < 1 - \gamma$. Clearly $\gamma \in (0, 1)$. Furthermore, the bigger γ the better the performance of the algorithm.

The following theorem quantifies the benefits of the new RWM algorithm over the standard one.

Theorem 4.21. *For the standard RWM algorithm:*

- *If $\beta = N^{-a}$ with $a \in [0, 1)$ then the spectral gap is bounded above by $C_p N^{-p}$ for any positive integer p.*
- *If $\beta = N^{-a}$ with $a \in [1, \infty)$ then the spectral gap is bounded above by $CN^{-\frac{a}{2}}$.*

Hence spectral gap is bounded above by $CN^{-\frac{1}{2}}$. For the new RWM algorithm the spectral gap is bounded below independently of N. Hence we have a central limit theorem and, for $u^{(0)} \sim \mu$ and C independent of N,

$$\mathbb{E}^\nu\Big(\frac{1}{K}\sum_{k=1}^{K} f(u^{(k)}) - \mathbb{E}^{\mu^N} f\Big)^2 \le CK^{-1}.$$

5. Bibliographical Notes

- Subsection 1.1. See [BS94] for a general overview of the Bayesian approach to statistics in the finite dimensional setting. The Bayesian approach to linear inverse problems with Gaussian noise and prior in finite dimensions is discussed in [Stu10, Chapters 2 and 6] and, with a more algorithmic flavour, in the book [KS05].
- Subsection 1.2, 1.3. See [Eva98] for theory relevant to both the heat equation and the elliptic equation. For more detail on the heat equation as an ODE in Hilbert space, see [Paz83, Lun95]. For further reading on severely ill-posed problems see [Stu10, Chapters 3 and 6], [KvdVvZ11b], [ASZ12]; for linear inverse problems in infinite dimensions see [Stu10, Chapters 3 and 6], [ALS12], [Man84], [LPS89], [KvDVvZ11a]; for the elliptic inverse problem – determining the permeability from the pressure in a Darcy model of flow in a porous medium and obtaining bounds on the solution using Lax-Milgram theorem [Ric81, DS11]; for the inverse heat equation, see [Kir96, EHN96].
- Subsection 2.1. For general discussion of the properties of random functions constructed via randomization of coefficients in a series expansion see [Kah85].
- Subsection 2.2. These uniform priors have been extensively studied in the context of the field of Uncertainty Quantification and the reader is directed to [CDS10, CDS12] for more details. Uncertainty Quantification in this context does not concern inverse problems, but rather studies the effect, on the solution of an equation, of randomizing the input data. Thus the interest is in the pushforward of a measure on input parameter space onto a measure on solution space, for a differential equation. Recently, however, these priors have been used to study the inverse problem; see [SS12].
- Subsection 2.3. Besov priors were introduced in the paper [LSS09] and Theorem 2.2 is taken from that paper. We notice that the theorem constitutes a special case of the Fernique Theorem in the Gaussian case $\mathsf{q} = 2$; it is restricted to a specific class of Hilbert spafe norms, however, whereas the Fernique Theorem in full generality applies in all norms on Banach spaces which have full Gaussian measure. A more general Fernique-like property of the Besov measures is proved in [DHS12] but it remains open to determine the appropriate complete generalization of the Fernique Theorem to Besov measures.
- Subsection 2.4. The general theory of Gaussian measures on Banach spaces is contained in [Lif95, Bog98]. The text [DZ92], concerning the theory of stochastic PDEs, also has a useful overview of the subject. The Karhunen-Loéve expansion (2.7) is contained in [Adl81]. The informal calculation concerning the covariance operator of the Gaussian measure which follows Theorem 2.4 may be proved using characteristic functions; see, for example, Proposition 2.18 in [DZ92]. All three texts include statement and proof of the Fernique Theorem in the generality given here. The Kolmogorov continuity theorem is discussed in [DZ92] and [Adl90]. Proof of Hölder regularity adapted to the case of the periodic setting may be found in [Hai09] and [Stu10, Chapter 6]. For further reading on Gaussian measures see [DP06].
- Subsection 3.1. Theorem 3.1 is taken from [HSVW05]where it is used to compute expressions for the meausure induced by various conditionings applied to SDEs. The Example following Theorem 3.1, concerning end-point conditioning of measures defined via a density with respect to Wiener measure, finds application to problems from molecular dynamics in [PS10, NST]. Further material concerning the equivalence of posterior with respect to the prior may be found in [Stu10, Chapters 3 and 6], [ALS12], [ASZ12]. The equivalence of Gaussian measures is studied via the Feldman-Hajek theorem; see [DPZ92] and [DZ92].
- Subsection 3.2. General development of Bayes' Theorems for inverse problems on function space, along the lines described here, may be found in [CDRS09, Stu10]. The reader is also directed to the papers [Las02, Las07] for earlier related material, and to [Las11, Las12a, Las12b] for recent developments.
- Subsection 3.3. The inverse problem for the heat equation was one of the first infinite dimensional inverse problems to receive Bayesian treatment; see [Fra70]. The problem is worked through in detail in [Stu10]. To fully understand the details the reader will need to study the Cameron-Martin theorem (concerning shifts in the mean of Gaussian measures) and the Feldman-Hajek theorem (concerning equivalence of Gaussian measures); both of these may be found in [DZ92, Lif95, Bog98] and are also discussed in [Stu10].
- Subsection 3.4. The elliptic inverse problem with the uniform prior is studied in [SS12]. A Gaussian

prior is adopted in [DS11], and a Besov prior in [DHS12].

- Subsection 4.1. Relationships between the Hellinger distance on probability measures, and the Total Variation distance and Kullback-Leibler divergence may be found in [GS02], [Pol].
- Subsection 4.2. The relationship between expectations and Hellinger distance, as used in Remark 4.8, is discussed in [Stu10].
- Subsection 4.3 concerns measure preserving continuous time dynamics. The finite dimensional aspects of this subsection, which we introduce for motivation, are covered in the texts [Oks03] and [Gar85]; the first of these books is an excellent introduction to the basisc existence and uniqueness theory, outlined in a simple case in Theorem 4.9, whilst the second provides an in depth treatment of the subject from the viewpoint of the Fokker-Planck equation, as used in Theorem 4.11. This subject has a long history which is overviewed in the paper [HSV07] where the idea is applied to fiding SPDEs which are invariant with respect to the measure generated by a conditioned diffusion process. This idea is generalized to certain conditioned hypoelliptic diffusions in [HSV11b]. It is also possible to study deterministic Hamiltonian dynamics which preserves the same measure. This idea is described in [BPSSS11] in the same set-up as employed here; that paper also contains references to the wider literature. Lemma 4.14 is proved in [MPS12]. Lemma 4.20 requires knowledge of the invariance of Ornstein-Uhlenbeck processes together with invariance of finite dimensional first order Langevin equations with the form of gradient dynamics subject to additive noise. The invariance of the Ornstein-Uhlenbeck process is covered in [DPZ96] and invariance of finite dimensional SDEs using the Fokker-Planck equation is discussed in [Gar85]. The $\mathcal{C}-$Wiener process, and its properties, are described in [DZ92].
- Subsection 4.4 concerns The standard RWM was introduced in [MRTT53] and led, via the paper [Has70], to the development of the more general class of Metropolis-Hastings methods. MCMC methods which are invariant with respect to the target measure **μ.** The paper [CRSW12] overviews this subject area, including the new RWM method. The specific idea of the new RWM is contained in the unpublished paper [Nea98], equation (15). The paper [Tie98] is a key reference which provides a framework for the study of Metropolis-Hastings methods on general state spaces, and may be used to establish that the new RWM method is well-defined on the Hilbert space $\mathcal{H}$. Theorem 4.21 is a summary of the results in the paper [HSV11a].

Acknowledgements AMS is grateful to Sergios Agapiou and to Yuan-Xiang Zhang for help in the preparation of these lecture notes. He is also grateful to EPSRC, ERC and ONR for financial support.

References

[Adl81] R.J. Adler. **The Geometry of Random Fields**, volume 62. SIAM, 1981.

[Adl90] R. J. Adler. **An Introduction to Continuity, Extrema, and Related Topics for General Gaussian Processes**, volume 12 of **Institute of Mathematical Statistics Lecture Notes—Monograph Series**. Institute of Mathematical Statistics, Hayward, CA, 1990.

[ALS12] S. Agapiou, S. Larsson, and A. M. Stuart. Posterior consistency of the Bayesian approach to linear ill-posed inverse problems, arxiv.org/abs/1203.5753. 2012.

[ASZ12] S. Agapiou, A.M. Stuart, and Y-X. Zhang. Bayesian posterior contraction rates for linear severely ill-posed inverse problems. 2012. Submitted, arxiv.org/abs/1210.1563.

[Bog98] Vladimir I. Bogachev. **Gaussian measures**, volume 62 of **Mathematical Surveys and Monographs**. American Mathematical Society, 1998.

[BPSSS11] A. Beskos, FJ Pinski, JM Sanz-Serna, and AM Stuart. Hybrid Monte Carlo on Hilbert spaces. **Stochastic Processes and their Applications**, 121:2201–2230, 2011.

[BS94] J.M. Bernardo and A.F.M. Smith. **Bayesian Theory**. Wiley, 1994.

[CDRS09] S.L. Cotter, M. Dashti, J.C. Robinson, and A.M. Stuart. Bayesian inverse problems for functions and applications to fluid mechanics. **Inverse Problems**, 25:doi:10.1088/0266–5611/25/11/115008, 2009.

[CDS10] A. Cohen, R. DeV24(e)-3262402Slo 024(,)-337.068(a)-5.8887(n)0.928763(S)-364.C38(T)-0.700861(h)0.931591(

[CDS12] A. Cohen, R. DeVore, and Ch. Schwab. Analytic regularity and polynomial approximation of parametric and stochastic elliptic pdes. **Anal. Appl.**, 2012.

inverse problems. **Arxiv preprint arXiv:1112.0906**, 2011.

[Las12a] S. Lasanen. Non-Gaussian statistical inverse problems. Part i: Posterior distributions. **Inverse Problems and Imaging**, 6(2):215–266, 2012.

[Las12b] S. Lasanen. Non-Gaussian statistical inverse problems. Part ii: Posterior distributions. **Inverse Problems and Imaging**, 6(2):267–287, 2012.

[Lif95] M.A. Lifshits. **Gaussian Random Functions**, volume 322 of **Mathematics and its Applications**. Kluwer, Dordrecht, 1995.

[LPS89] M. S. Lehtinen, L. Päivärinta, and E. Somersalo. Linear inverse problems for generalised random variables. **Inverse Problems**, 5(4):599–612, 1989.

[LSS09] M. Lassas, E. Saksman, and S. Siltanen. Discretization-invariant Bayesian inversion and Besov space priors. **Inverse Problems and Imaging**, 3:87–122, 2009.

[Lun95] A. Lunardi.
M9in2in261 1995.

[L¨S5]

[lȩJ.11297954]11187(l)26.0259334ssg[.bma7ndrn•51.52915(e)2.87417(o)5c38882(m)-1.5340]TJ /R549.96264Tf 124826310Td