



Universiteit  
Leiden  
The Netherlands

## Bayesian inference for Gaussian models: Inverse problems and evolution equations

Yan, D.

### Citation

Yan, D. (2020, March 3). *Bayesian inference for Gaussian models: Inverse problems and evolution equations*. Retrieved from <https://hdl.handle.net/1887/86070>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/86070>

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/86070> holds various files of this Leiden University dissertation.

**Author:** Yan, D.

**Title:** Bayesian inference for Gaussian models: Inverse problems and evolution equations

**Issue Date:** 2020-03-03

## Chapter 7

# Inverse Problems with Discrete Observations: Gaussian Conjugacy

### 7.1 Introduction

Linear inverse problems have been studied since long in the statistical and numerical analysis literature; see, e.g., [3, 7, 15, 16, 19, 24, 56, 57, 101], and references therein. Emphasis in these works has been on the signal-in-white noise model,

$$Y = Af + \varepsilon W, \quad (7.1)$$

where the parameter of interest  $f$  lies in some infinite-dimensional function space,  $A$  is a linear operator with values in a possibly different space,  $W$  is white noise, and  $\varepsilon$  is the noise level. Applications of linear inverse problems include, e.g., computerized tomography, see [72], partial differential equations, see [54], and scattering theory, see [20].

Arguably, in practice one does not have access to a full record of observations on the unknown function  $f$  as in the idealised model (7.1), but rather one indirectly observes it at a finite number of points. This statistical setting can be conveniently formalised as follows: let the signal of interest  $f$  be an element in a Hilbert space  $H_1$  of functions defined on a compact interval  $[0, 1]$ . The forward operator  $A$  maps  $f$  to another Hilbert space  $H_2$ . We assume that  $H_1, H_2$  are subspaces of  $L^2([0, 1])$ , typically collections of functions of certain smoothness as specified in the later sections, and that the design points are chosen deterministically,

$$\left\{ x_i = \frac{i}{n} \right\}_{i=1, \dots, n}. \quad (7.2)$$

Assuming continuity of  $Af$  and defining

$$Y_i = Af(x_i) + \xi_i, \quad i = 1, \dots, n, \quad (7.3)$$

with  $\xi_i$  i.i.d. standard Gaussian random variables, our observations are the pairs  $(x_i, Y_i)_{i \leq n}$ , and we are interested in estimating  $f$ . A prototype example we think of is the case when  $A$  is the solution operator in the Dirichlet problem for the heat equation acting on the initial condition  $f$ ; see Example 7.8 below for details.

Model (7.3) is related to the inverse regression model studied e.g. in [5] and [6]. Although the setting we consider is somewhat special, our contribution is arguably the first one to study from a theoretical point of view a nonparametric Bayesian approach to estimation of  $f$  in the inverse problem setting with partial observations (see [35] for a monographic treatment of modern Bayesian nonparametrics). In the context of the signal-in-white noise model (7.1), a nonparametric Bayesian approach has been studied thoroughly in [59] and [60], and techniques from these works will turn out to be useful in our context as well. Our results will deal with derivation of posterior contraction rates and study of asymptotic frequentist coverage of Bayesian credible sets. A posterior contraction rate can be thought of as a Bayesian analogue of a convergence rate of a frequentist estimator, cf. [33] and [35]. Specifically, we will show that as the sample size  $n \rightarrow \infty$ , the posterior distribution concentrates around the ‘true’ parameter value, under which data have been generated, and hence our Bayesian approach is consistent and asymptotically recovers the unknown ‘true’  $f$ . The rate at which this occurs will depend on the smoothness of the true parameter and the prior and the ill-posedness degree of the problem. Correct combinations of these values lead to optimal posterior contraction rates (up to logarithmic factors). Furthermore, a Bayesian approach automatically provides uncertainty quantification in parameter estimation through the spread of the posterior distribution, specifically by means of posterior credible sets. We will give an asymptotic frequentist interpretation of these sets in our context. In particular, we will see that the frequentist coverage will depend on a combination of smoothness of the true parameter and the prior, and the ill-posedness of the problem. Oversmoothing priors lead to zero coverage, while undersmoothing priors produce highly conservative results.

The chapter is organized as follows: in Section 7.2, we give a detailed description of the problem, introduce the singular value decomposition and convert the model (7.3) into an equivalent truncated sequence model that is better amenable to our theoretical analysis. We show how a Gaussian prior in this sequence model leads to a Gaussian posterior and give an explicit characterisation of the latter. Our main results on posterior contraction rates and Bayesian credible sets are given in Section 7.3, followed by simulation examples in Section 7.4 that illustrate our theoretical results. Section 7.5 contains the proofs of the main theorems, while the technical lemmas used in the proofs are collected in Section 7.6.

### 7.1.1 Notation

The notational conventions we use in this work are the following: definitions are marked by the  $:=$  symbol;  $|\cdot|$  denotes the absolute value and  $\|\cdot\|_H$  indicates the norm related to the space  $H$ ;  $\langle \cdot, \cdot \rangle_H$  is understood as the canonical inner product in the inner product space  $H$ ; subscripts are omitted when there is no danger of confusion;  $\mathcal{N}(\mu, \Sigma)$  denotes the Gaussian distribution with mean  $\mu$  and covariance operator  $\Sigma$ ; subscripts  $\mathcal{N}_n$  and  $\mathcal{N}_H$  may be used to emphasize the fact that the distribution is defined on the space  $\mathbb{R}^n$  or on the abstract space  $H$ ;  $\text{Cov}(\cdot, \cdot)$  denotes the covariance or the covariance operator, depending on the context

## 7.2 Sequence model

### 7.2.1 Singular value decomposition

We impose a common assumption on the forward operator  $A$  from the literature on inverse problems, see, e.g., [3], [7] and [15].

**Assumption 7.1.** Operator  $A$  is injective and compact.

It follows that  $A^*A$  is also compact and in addition self-adjoint. Hence, by the spectral theorem for self-adjoint compact operators, see [21], we have a representation  $A^*Af = \sum_{k \in \mathbb{N}} a_k^2 f_k \varphi_k$ , where  $\{\varphi_k\}$  and  $\{a_k\}$  are the eigenbasis on  $H_1$  and eigenvalues, respectively, (corresponding to the operator  $A^*A$ ), and  $f_k = \langle f, \varphi_k \rangle$  are the Fourier coefficients of  $f$ . This decomposition of  $A^*A$  is known as the singular value decomposition (SVD), and  $\{a_k\}$  are also called singular values.

It is easy to show that the conjugate basis  $\psi_k := A\varphi_k/a_k$  of the orthonormal basis  $\{\varphi_k\}_k$  is again an orthonormal system in  $H_2$  and gives a convenient basis for  $\text{Ran } A$ , the range of  $A$  in  $H_2$ . Furthermore, the following relations hold (see [3]),

$$A\varphi_k = a_k\psi_k, \quad A^*\psi_k = a_k\varphi_k. \quad (7.4)$$

Recall a standard result (see, e.g., [44]): a Hilbert space  $H$  is isometric to  $\ell^2$ , and Parseval's identity  $\|f\|_{\ell^2}^2 := \sum_k |f_k|^2 = \|f\|_H^2$  holds; here  $f_k$  are the Fourier coefficients with respect to some known and fixed orthonormal basis.

We will employ the eigenbasis  $\{\varphi_k\}$  of  $A^*A$  to define the Sobolev space of functions. This will define the space in which the unknown function  $f$  resides.

**Definition 7.2.** We say  $f$  is in the Sobolev space  $S^\beta$  with smoothness parameter  $\beta \geq 0$ , if it can be written as  $f = \sum_{k=1}^{\infty} f_k \varphi_k$  with  $f_k = \langle f, \varphi_k \rangle$ , and if its norm  $\|f\|_\beta := (\sum_{k=1}^{\infty} f_k^2 k^{2\beta})^{1/2}$  is finite.

**Remark 7.3.** The above definition agrees with the classical definition of the Sobolev space if the eigenbasis is the trigonometric basis, see, e.g., [95]. With a fixed basis, which is always the case in this article, one can identify the function  $f$  and its Fourier coefficients  $\{f_k\}$ . Thus, we use  $S^\beta$  to denote both the function space and the sequence space. For example, it is easy to verify that  $S^0 = \ell^2$  (correspondingly  $S^0 = L^2$ ),  $S^\beta \subset \ell^2$  for any nonnegative  $\beta$ , and  $S^\beta \subset \ell^1$  when  $\beta > 1/2$ .

Recall that  $Af = \sum a_i f_i \psi_i$ . Then we have  $Af \in S^{\beta+p}$  if  $a_k \asymp k^{-p}$ , and  $Af \in S^\infty := \cap_{k \in \mathbb{N}} S^k$ , if  $a_k$  decays exponentially fast. Such a lifting property is beneficial in the forward problem, since it helps to obtain a smooth solution. However, in the context of inverse problems it leads to a difficulty in recovery of the original signal  $f$ , since information on it is washed out by smoothing. Hence, in the case of inverse problems one does not talk of the lifting property, but of ill-posedness, see [15].

**Definition 7.4.** An inverse problem is called mildly ill-posed, if  $a_k \asymp k^{-p}$  as  $k \rightarrow \infty$ , and extremely ill-posed, if  $a_k \asymp e^{-k^s p}$  with  $s \geq 1$  as  $k \rightarrow \infty$ , where  $p$  is strictly positive in both cases.

In the rest of the article, we will confine ourselves to the following setting.

**Assumption 7.5.** The unknown true signal  $f$  in (7.3) satisfies  $f \in S^\beta \subset H_1$  for  $\beta > 0$ . Furthermore, the ill-posedness is of one of the two types in Definition Definition 7.4.

**Remark 7.6.** As an immediate consequence of the lifting property, we have  $H_2 \subset H_1$ .

We conclude this section with two canonical examples of the operator  $A$ .

**Example 7.7** (mildly ill-posed case: Volterra operator [59]). The classical Volterra operator  $A : L^2[0, 1] \rightarrow L^2[0, 1]$  and its adjoint  $A^*$  are

$$Af(x) = \int_0^x f(s) ds, \quad A^*f(x) = \int_x^1 f(s) ds.$$

The eigenvalues, eigenfunctions of  $A^*A$  and the conjugate basis are given by

$$\begin{aligned} a_i^2 &= \frac{1}{(i - 1/2)^2 \pi^2}, \\ \varphi_i(x) &= \sqrt{2} \cos((i - 1/2)\pi x), \\ \psi_i(x) &= \sqrt{2} \sin((i - 1/2)\pi x), \end{aligned}$$

for  $i \geq 1$ .

**Example 7.8** (extremely ill-posed case: heat equation [60]). Consider the Dirichlet problem for the heat equation:

$$\begin{aligned} \frac{\partial}{\partial t} u(x, t) &= \frac{\partial^2}{\partial x^2} u(x, t), \quad u(x, 0) = f(x), \\ u(0, t) &= u(1, t) = 0, \quad t \in [0, T], \end{aligned} \tag{7.5}$$

where  $u(x, t)$  is defined on  $[0, 1] \times [0, T]$  and  $f(x) \in L^2[0, 1]$  satisfies  $f(0) = f(1) = 0$ . The solution of (7.5) is given by

$$u(x, t) = \sqrt{2} \sum_{k=1}^{\infty} f_k e^{-k^2 \pi^2 t} \sin(k\pi x) =: Af(x),$$

where  $\{f_k\}$  are the coordinates of  $f$  in the basis  $\{\sqrt{2} \sin(k\pi x)\}_{k \geq 1}$ .

For the solution map  $A$ , the eigenvalues of  $A^*A$  are  $e^{-k^2 \pi^2 t}$ , the eigenbasis and conjugate basis coincide and  $\varphi_k(x) = \psi_k(x) = \sqrt{2} \sin(k\pi x)$ .

### 7.2.2 Equivalent formulation

In this subsection we develop a sequence formulation of the model (7.3), which is very suitable for asymptotic Bayesian analysis. First, we briefly discuss the relevant results that provide motivation for our reformulation of the problem.

In Examples 7.7 and 7.8, the sine and cosine bases form the eigenbasis. In fact, the Fourier basis (trigonometric polynomials) frequently arises as an eigenbasis for various operators, e.g. in the case of differentiation, see [27], or circular deconvolution, see [16]. For simplicity, we will use Fourier basis as a primary example in the rest of the article. Possible generalization to other bases is discussed in Remark 7.10.

Restriction of our attention to the Fourier basis is motivated by its special property: discrete orthogonality. The next lemma illustrates this property for the sine basis (Example 7.8).

**Lemma 7.9** (discrete orthogonality). *Let  $\{\psi_k\}_{k \in \mathbb{N}}$  be the sine basis, i.e.*

$$\psi_k(x) = \sqrt{2} \sin(k\pi x), \quad k = 1, 2, 3, \dots$$

Then:

(i.) *Discrete orthogonality holds:*

$$\langle \psi_j, \psi_k \rangle_d := \frac{1}{n} \sum_{i=1}^n \psi_j(i/n) \psi_k(i/n) = \delta_{jk}, \quad j, k = 1, \dots, n-1. \quad (7.6)$$

Here  $\delta_{jk}$  is the Kronecker delta.

(ii.) *Fix  $l \in \mathbb{N}$ . For any fixed  $1 \leq k \leq n-1$  and all  $j \in \{ln, ln+1, \dots, (l+1)n-1\}$ , there exists only one  $\bar{k} \in \{1, 2, \dots, n-1\}$  depending only on the parity of  $l$ , such that for  $\tilde{j} = ln + \bar{k}$ , the equality*

$$|\langle \psi_{\tilde{j}}, \psi_k \rangle_d| = 1 \quad (7.7)$$

*holds, while  $\langle \psi_{\tilde{j}}, \psi_k \rangle_d = 0$  for all  $\tilde{j} = ln + \tilde{k}$  such that  $\tilde{k} \neq \bar{k}$ ,  $\tilde{k} \in \{1, 2, \dots, n-1\}$ .*

**Remark 7.10.** For other trigonometric bases, discrete orthogonality can also be attained. Thus, the conjugate eigenbasis in Example 7.7 is discretely orthogonal with design points  $\{(i-1/2)/n\}_{i=1, \dots, n}$ . We refer to [2] and references therein for details. With some changes in the arguments, our asymptotic statistical results still remain valid with such modifications of design points compared to (7.2). We would like to stress the fact that restricting attention to bases with discrete orthogonality property does constitute a loss of generality. However, there exist classical bases other than trigonometric bases that are discretely orthogonal (possibly after a suitable modification of design points). See, for instance, [78] for an example of Lagrange polynomials.

Motivated by the observations above, we introduce our central assumption on the basis functions.

**Assumption 7.11.** Given the design points  $\{x_i\}_{i=1, \dots, n}$  in (7.2), we assume the conjugate basis  $\{\psi_k\}_{k \in \mathbb{N}}$  of the operator  $A$  in (7.3) possesses the following properties:

(i.) for  $1 \leq j, k \leq n-1$ ,

$$\langle \psi_j(x), \psi_k(x) \rangle_d := \frac{1}{n} \sum_{i=1}^n \psi_j(x_i) \psi_k(x_i) = \delta_{jk}$$

(ii.) For  $1 \leq k \leq n-1$  and  $j \in \{ln, \dots, (l+1)n-1\}$  with fixed  $l \in \mathbb{N}$ , there exists only one  $\tilde{j} = ln + \bar{k}$ , such that  $0 < |\langle \psi_{\tilde{j}}, \psi_k \rangle_d| < M$ , where  $M$  is a fixed constant, and  $\bar{k}$  depends on the parity of  $l$  only. For other  $j \neq \tilde{j}$ ,  $|\langle \psi_j, \psi_k \rangle_d| = 0$ .

Using the shorthand notation

$$f = \sum_j f_j \varphi_j = \sum_{j=1}^{n-1} f_j \varphi_j + \sum_{j \geq n} f_j \varphi_j =: f^n + f^r,$$

we obtain for  $k = 1, \dots, n-1$  that

$$\begin{aligned} U_k &= \frac{1}{n} \sum_{i=1}^n Y_i \psi_k(x_i) = \langle Af^n, \psi_k \rangle_d + \langle Af^r, \psi_k \rangle_d + \frac{1}{n} \sum_{i=1}^n \xi_i \psi_k(x_i) \\ &= a_k f_k + R_k + \frac{1}{\sqrt{n}} \zeta_k, \end{aligned} \tag{7.8}$$

where

$$R_k := R_k(f) = \langle Af^r, \psi_k \rangle_d, \quad \zeta_k := \frac{1}{\sqrt{n}} \sum_{i=1}^n \xi_i \psi_k(x_i).$$

By Assumption 7.11, we have

$$|R_k| = |\langle Af^r, \psi_k \rangle_d| \leq \sum_{j \geq n} a_j |f_j| |\langle \psi_j, \psi_k \rangle_d| = \sum_{l=1}^{\infty} a_{ln+\bar{k}} |f_{ln+\bar{k}}|, \tag{7.9}$$

which leads to (via Cauchy-Schwarz)

$$R_k^2(f) \leq \left( \sum_{l=1}^{\infty} a_{ln+\bar{k}}^2 (ln + \bar{k})^{-2\beta} \right) \|f\|_{\beta}^2.$$

Hence, for a mildly ill-posed problem, i.e.  $a_k \asymp k^{-p}$ , the following bound holds, uniformly in the ellipsoid  $\{f : \|f\|_{\beta} \leq K\}$ ,

$$\begin{aligned} \sup_{f: \|f\|_{\beta} \leq K} R_k^2(f) &\lesssim \sum_{l=1}^{\infty} (ln)^{-2\beta-2p} = n^{-2(\beta+p)} \sum_{l=1}^{\infty} l^{-2(\beta+p)} \\ &\asymp n^{-2(\beta+p)} = o(1/n), \end{aligned} \tag{7.10}$$

for any  $1 \leq k \leq n-1$  when  $\beta + p > 1/2$ .

If the problem is extremely ill-posed, i.e.  $a_k \asymp e^{-k^s p}$ , we use the inequality

$$R_k^2(f) \leq \left( \sum_{j \geq n} a_j |f_j| \right)^2 \leq \left( \sum_{j \geq n} a_j^2 \right) \|f^r\|^2.$$

Since  $a_j \asymp \exp(-pj^s) \leq \exp(-pj)$ , it follows that  $\sum_{j \geq n} a_j^2$  is up to a constant bounded from above by  $\exp(-2pn)$ . Hence

$$\sup_{f: \|f\|_{\beta} \leq K} R_k^2(f) \lesssim \exp(-2pn) \ll 1/n. \quad (7.11)$$

In [59], [60], the Gaussian prior  $\Pi = \otimes_{i \in \mathbb{N}} \mathcal{N}(0, \lambda_i)$  is employed on the coordinates of the eigenbasis expansion of  $f$ . If  $\lambda_i = \rho_n^2 i^{-1-2\alpha}$ , the sum  $\sum_{i \in \mathbb{N}} \lambda_i = \rho_n^2 \sum_{i \in \mathbb{N}} i^{-1-2\alpha}$  is convergent, and hence this prior is the law of a Gaussian element in  $H_1$ .

In our case, we consider the same type of the prior with an additional constraint that only the first  $n-1$  components of the prior are non-degenerate, i.e.  $\Pi = (\otimes_{i < n} \mathcal{N}(0, \lambda_i)) \times (\otimes_{i \geq n} \mathcal{N}(0, 0))$ , where  $\lambda_i$  is as above. In addition, we assume the prior on  $f$  is independent of the noise  $\zeta_k$ ,  $k = 1, \dots, n-1$ , in (7.8). With these assumptions in force, we see  $\Pi(R_k = 0) = 1$ , for  $k = 1, \dots, n-1$ . Furthermore, the posterior can be obtained from the product structure of the model and the prior via the normal conjugacy,

$$\Pi(f|U^n) = \otimes_{k \in \mathbb{N}} \mathcal{N}(\hat{f}_k, \sigma_k^2), \quad (7.12)$$

$$\text{with } \hat{f}_k = \frac{na_k \lambda_k \mathbb{1}_{\{k < n\}}}{na_k^2 \lambda_k + 1} U_k, \quad \sigma_k^2 = \frac{\lambda_k \mathbb{1}_{\{k < n\}}}{na_k^2 \lambda_k + 1}.$$

We also introduce

$$\hat{f} = \mathbb{E}(f|U^n) = (\mathbb{E}(f_k|U_k)) = (\hat{f}_k)_{k \in \mathbb{N}} = (b_k U_k)_{k \in \mathbb{N}}, \quad (7.13)$$

where  $b_k = \frac{na_k \lambda_k \mathbb{1}_{\{k < n\}}}{na_k^2 \lambda_k + 1}$ . We conclude this section with a useful fact that will be applied in later sections:

$$\hat{f}_k = b_k U_k = b_k \left( a_k f_k + R_k + \frac{\zeta_k}{\sqrt{n}} \right) = \mathbb{E} \hat{f}_k + \tau_k \zeta_k, \quad (7.14)$$

where  $\mathbb{E} \hat{f}_k = a_k b_k f_k + b_k R_k$  and  $\tau_k = b_k / \sqrt{n}$ .

## 7.3 Main results

### 7.3.1 Contraction rates

In this section, we determine the rate at which the posterior distribution concentrates on shrinking neighbourhoods of the ‘true’ parameter  $f_0$  as the sample size  $n$  grows to infinity.

Assume the observations in (7.3) have been collected under the parameter value  $f_0 = \sum_{k \in \mathbb{N}} f_{0,k} \varphi_k$ . Thus our observations  $(U_k)_{k < n}$  given in (7.8) have the law  $\otimes_{k < n} \mathcal{N}(a_k f_{0,k} + R_k, 1/n)$ . We will use the notation  $\Pi_n(\cdot|U)$  to denote the posterior distribution given in (7.12).

**Theorem 7.12** (Posterior contraction: mildly ill-posed problem). *If the problem is mildly ill-posed as  $a_k \asymp k^{-p}$  with  $p > 0$ , the true parameter  $f_0 \in S^\beta$  with  $\beta > 0$ , and furthermore  $\beta + p > 1/2$ , by letting  $\lambda_k = \rho_n^2 k^{-1-2\alpha}$  with  $\alpha > 0$  and any positive  $\rho_n$  satisfying  $\rho_n^2 n \rightarrow \infty$ , we have, for any  $K > 0$  and  $M_n \rightarrow \infty$ ,*

$$\sup_{\|f_0\|_\beta \leq K} \mathbb{E}_{f_0} \Pi_n (f : \|f - f_0\|_{H_1} \geq M_n \varepsilon_n | U^n) \rightarrow 0,$$

where

$$\varepsilon_n = \varepsilon_{n,1} \vee \varepsilon_{n,2} = (\rho_n^2 n)^{-\beta/(2\alpha+2p+1) \wedge 1} \vee \rho_n (\rho_n^2 n)^{-\alpha/(2\alpha+2p+1)}. \quad (7.15)$$

In particular,

- (i.) if  $\rho_n = 1$ , then  $\varepsilon_n = n^{-(\alpha \wedge \beta)/(2\alpha+2p+1)}$ ;
- (ii.) if  $\beta \leq 2\alpha + 2p + 1$  and  $\rho_n \asymp n^{(\alpha-\beta)/(2\beta+2p+1)}$ , then  $\varepsilon_n = n^{-\beta/(2\beta+2p+1)}$ ;
- (iii.) if  $\beta > 2\alpha + 2p + 1$ , then for every scaling  $\rho_n$ ,  $\varepsilon_n \gg n^{-\beta/(2\beta+2p+1)}$ .

Thus we recover the same posterior contraction rates as obtained in [59], at the cost of an extra constraint  $\beta + p > 1/2$ . The frequentist minimax convergence rate for mildly ill-posed problems in the white noise setting with  $\varepsilon = n^{-1/2}$  is  $n^{-\beta/(2\beta+2p+1)}$ , see [15]. We will compare our result to this rate. Our theorem states that in case (i.) the posterior contraction rate reaches the frequentist optimal rate if the regularity of the prior matches the truth ( $\beta = \alpha$ ) and the scaling factor  $\rho_n$  is fixed. Alternatively, as in case (ii.), the optimal rate can also be attained by proper scaling, provided a sufficiently regular prior is used. In all other cases the contraction rate is slower than the minimax rate. Our results are similar to those in [59] in the white noise setting. The extra constraint  $\beta + p > 1/2$  that we have in comparison to that work demands an explanation. As (7.10) shows, the size of negligible terms  $R_k(f_0)$  in (7.8) decreases as the smoothness  $\beta + p$  of the transformed signal  $Af_0$  increases. In order to control  $R_k$ , a minimal smoothness of  $Af_0$  is required. The latter is guaranteed if  $p + \beta \geq 1/2$ , for it is known that in that case  $Af_0$  will be at least continuous, while it may fail to be so if  $p + \beta < 1/2$ , see [95].

**Remark 7.13.** The control on  $R_k(f_0)$  from (7.9) depends on the fact that the eigenbasis possesses the properties in Assumption 7.11. If instead of Assumption 7.11 (ii.) one only assumes  $|\langle \psi_j, \psi_k \rangle| \leq 1$  for any  $k \leq n - 1$  and  $j \geq n$ , the constraint on the smoothness of  $Af_0$  has to be strengthened to  $\beta + p \geq 1$  in order to obtain the same results as in Theorem 7.12, because the condition  $\beta + p \geq 1$  guarantees that the control on  $R_k(f_0)$  in (7.10) remains valid.

Now we consider the extremely ill-posed problem. The following result holds.

**Theorem 7.14** (Posterior contraction: extremely ill-posed problem). *Let the problem be extremely ill-posed as  $a_k \asymp e^{-pk^s}$  with  $s \geq 1$ , and let the true parameter  $f_0 \in S^\beta$  with  $\beta > 0$ . Let  $\lambda_k = \rho_n^2 k^{-1-2\alpha}$  with  $\alpha > 0$  and any positive  $\rho_n$  satisfying  $\rho_n^2 n \rightarrow \infty$ . Then*

$$\sup_{\|f_0\|_\beta \leq K} \mathbb{E}_{f_0} \Pi_n (f : \|f - f_0\|_{H_1} \geq M_n \varepsilon_n | U^n) \rightarrow 0,$$

for any  $K > 0$  and  $M_n \rightarrow \infty$ , where

$$\varepsilon_n = \varepsilon_{n,1} \vee \varepsilon_{n,2} = (\log(\rho_n^2 n))^{-\beta/s} \vee \rho_n (\log(\rho_n^2 n))^{-\alpha/s}. \quad (7.16)$$

In particular,

(i.) if  $\rho_n = 1$ , then  $\varepsilon_n = (\log n)^{-(\alpha \wedge \beta)/s}$ ,

(ii.) if  $n^{-1/2+\delta} \lesssim \rho_n \lesssim (\log n)^{(\alpha-\beta)/s}$  for some  $\delta > 0$ , then  $\varepsilon_n = (\log n)^{-\beta/s}$ .

Furthermore, if  $\lambda_k = \exp(-\alpha k^s)$  with  $\alpha > 0$ , the following contraction rate is obtained:  $\varepsilon_n = (\log n)^{-\beta/s}$ .

Since the frequentist minimax estimation rate in extremely ill-posed problems in the white noise setting is  $(\log n)^{-\beta/s}$  (see [15]), Theorem 7.14 shows that the optimal contraction rates can be reached by suitable choice of the regularity of the prior, or by using an appropriate scaling. In contrast to the mildly ill-posed case, we have no extra requirement on the smoothness of  $Af_0$ . The reason is obvious: because the signal is lifted to  $S^\infty$  by the forward operator  $A$ , the term (7.11) converges to zero exponentially fast, implying that  $R_k(f_0)$  in (7.8) is always negligible.

### 7.3.2 Credible sets

In the Bayesian paradigm, the spread of the posterior distribution is a common measure of uncertainty in parameter estimates. In this section we study the frequentist coverage of Bayesian credible sets in our problem.

When the posterior is Gaussian, it is customary to consider credible sets centered at the posterior mean, which is what we will also do. In addition, because in our case the covariance operator of the posterior distribution does not depend on the data, the radius of the credible ball is determined by the credibility level  $1 - \gamma$  and the sample size  $n$ . A credible ball centred at the posterior mean  $\hat{f}$  from (7.13) is given by

$$\hat{f} + B(r_{n,\gamma}) := \{f \in H_1 : \|f - \hat{f}\|_{H_1} \leq r_{n,\gamma}\}, \quad (7.17)$$

where the radius  $r_{n,\gamma}$  is determined by the requirement that

$$\Pi_n(\hat{f} + B(r_{n,\gamma}) | U^n) = 1 - \gamma. \quad (7.18)$$

By definition, the frequentist coverage or confidence of the set (7.17) is

$$\mathbb{P}_{f_0}(f_0 \in \hat{f} + B(r_{n,\gamma})), \quad (7.19)$$

where the probability measure is the one induced by the law of  $U^n$  given in (7.8) with  $f = f_0$ . We are interested in the asymptotic behaviour of the coverage (7.19) as  $n \rightarrow \infty$  for a fixed  $f_0$  uniformly in Sobolev balls, and also along a sequence  $f_0^n$  changing with  $n$ .

The following two theorems hold.

**Theorem 7.15** (Credible sets: mildly ill-posed problem). *Assume the same assumptions as in Theorem 7.12 hold, and let  $\tilde{\beta} = \beta \wedge (2\alpha + 2p + 1)$ . The asymptotic coverage of the credible set (7.17) is*

- (i.) 1, uniformly in  $\{f_0 : \|f_0\|_{\beta} \leq 1\}$ , if  $\rho_n \gg n^{(\alpha - \tilde{\beta})/(2\tilde{\beta} + 2p + 1)}$ ;
- (ii.) 1, for every fixed  $f_0 \in S^\beta$ , if  $\beta < 2\alpha + 2p + 1$  and  $\rho_n \asymp n^{(\alpha - \tilde{\beta})/(2\tilde{\beta} + 2p + 1)}$ ;  $c$ , along some  $f_0^n$  with  $\sup_n \|f_0^n\|_{\beta} < \infty$ , if  $\rho_n \asymp n^{(\alpha - \tilde{\beta})/(2\tilde{\beta} + 2p + 1)}$  (any  $c \in [0, 1)$ ).
- (iii.) 0, along some  $f_0^n$  with  $\sup_n \|f_0^n\|_{\beta} < \infty$ , if  $\rho_n \ll n^{(\alpha - \tilde{\beta})/(2\tilde{\beta} + 2p + 1)}$ .

**Theorem 7.16** (Credible sets: extremely ill-posed problem). *Assume the setup of Theorem 7.14. Then if  $\lambda_k = \rho_n^2 k^{-1 - 2\alpha}$  with  $\alpha > 0$  and any positive  $\rho_n$  satisfying  $\rho_n^2 n \rightarrow \infty$ , the asymptotic coverage of the credible set (7.17) is*

- (i.) 1, uniformly in  $\{f_0 : \|f_0\|_{S^\beta} \leq 1\}$ , if  $\rho_n \gg (\log n)^{(\alpha - \beta)/2}$ ;
- (ii.) 1, uniformly in  $f_0$  with  $\|f_0\|_{\beta} \leq r$  with  $r$  small enough;  
1, for any fixed  $f_0 \in S^\beta$ ,  
provided the condition  $\rho_n \asymp (\log n)^{(\alpha - \beta)/s}$  holds;
- (iii.) 0, along some  $f_0^n$  with  $\sup_n \|f_0^n\|_{\beta} < \infty$ , if  $\rho_n \lesssim (\log n)^{(\alpha - \beta)/s}$ .

Moreover, if  $\lambda_k = e^{-\alpha^s}$  with  $\alpha > 0$  and any positive  $\rho_n$  satisfying  $\rho_n^2 n \rightarrow \infty$ , the asymptotic coverage of the credible set (7.17) is

- (iv.) 0, for every  $f_0$  such that  $|f_{0,i}| \gtrsim e^{-ci^s/2}$  for some  $c < \alpha$ .

For the two theorems in this section, the most intuitive explanation is offered by the case  $\rho_n \equiv 1$ . The situations (i.), (ii.) and (iii.) correspond to  $\alpha < \beta$ ,  $\alpha = \beta$  and  $\alpha > \beta$ , respectively. The message is that the oversmoothing prior ((iii.) in Theorem 7.15 and (iii.), (iv.) in Theorem 7.16) leads to disastrous frequentist coverage of credible sets, while the undersmoothing prior ((i.) in both theorems) delivers very conservative frequentist results (coverage 1). With the right regularity of the prior (case (ii.)), the outcome depends on the norm of the true parameter  $f_0$ . Our results are thus similar to those obtained in the white noise setting in [59] and [60].

## 7.4 Simulation examples

In this section we carry out a small-scale simulation study illustrating our theoretical results. Examples we use to that end are those given in Subsection 7.2.1. These were also used in simulations in [59] and [60].

In the setting of Example 7.7, we use the following true signal,

$$f_0(x) = \sum_{i=1}^{\infty} f_{0,i} \varphi_i(x) \text{ with } f_{0,k} = k^{-3/2} \sin(k). \quad (7.20)$$

It is easy to check that  $f_0 \in S^1$ .

In the setup of Example 7.8, the initial condition is assumed to be

$$f_0(x) = 4x(x-1)(8x-5). \quad (7.21)$$

One can verify that in this case

$$f_{0,k} = \frac{8\sqrt{2}(13 + 11(-1)^k)}{\pi^3 k^3},$$

and  $f_0 \in S^\beta$  for any  $\beta < 5/2$ .

First, we generate noisy observations  $\{Y_i\}_{i=1,\dots,n}$  from our observation scheme (7.3) at design points  $x_i = \frac{i-1/2}{n}$  in the case of Volterra operator, and  $x_i = i/n$  in the case of the heat equation. Next, we apply the transform described in (7.8) and obtain transformed observations  $\{U_i\}_{i=1,\dots,n-1}$ . Then, by (7.12), the posterior of the coefficients with the eigenbasis  $\varphi_i$  is given by

$$f_k|U^n \sim \mathcal{N}\left(\frac{na_k\lambda_k\mathbb{1}_{\{k<n\}}}{na_k^2\lambda_k+1}U_k, \frac{\lambda_k\mathbb{1}_{\{k<n\}}}{na_k^2\lambda_k+1}\right).$$

Figures 7.1 and 7.2 display plots of (estimated) 95%  $L_2$ -credible bands for different sample sizes and different priors. For all priors we assume  $\rho_n \equiv 1$ , and use different smoothness degrees  $\alpha$ , as shown in the titles of the subplots. In addition, the columns from left to right corresponds to  $10^3, 10^4$  and  $10^5$  observations. The (estimated) credible bands are obtained by generating 1000 realizations from the posterior and retaining 95% of them that are closest in the  $L^2$ -distance to the posterior mean.

Two simulations reflect several similar facts. First, because of the difficulty due to the inverse nature of the problem, the recovery of the true signal is relatively slow, as the posteriors for the sample size  $10^3$  are still rather diffuse around the true parameter value. Second, it is evident that undersmoothing priors (the top rows in the figures) deliver conservative credible bands, but still capture the truth. On the other hand, oversmoothing priors lead to overconfident, narrow bands, failing to actually express the truth (bottom rows in the figures). As already anticipated due to a greater degree of ill-posedness, recovery of the initial condition in the heat equation case is more difficult than recovery of the true function in the case of the Volterra operator. Finally, we remark that qualitative behaviour of the posterior in our examples is similar to the one observed in [59] and [60]; for larger samples sizes  $n$ , discreteness of the observation scheme does not appear to have a noticeably adversary effect compared to the fully observed case in [59] and [60].

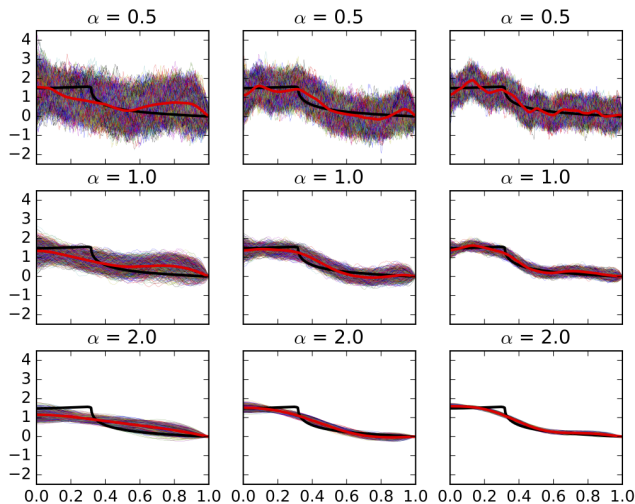


Figure 7.1: Realizations of the posterior mean (red) and 950 of 1000 draws from the posterior (colored thin lines) with smallest  $L^2$  distance to the posterior mean. From left to right columns, the posterior is computed based on sample size  $10^3$ ,  $10^4$  and  $10^5$  respectively. The true parameter (black) is of smoothness  $\beta = 1$  and given by coefficients  $f_{0,k} = k^{-3/2} \sin(k)$ .

## 7.5 Proofs

### 7.5.1 Proof of Lemma 7.9

This proof is a modification of the one of Lemma 1.7 in [95]. With the following temporary definitions  $a := e^{i\pi \frac{i}{n}}$  and  $b := e^{i\pi \frac{k}{n}}$ , using Euler's formula, we have

$$\begin{aligned}
 \langle \psi_j, \psi_k \rangle_d &= -\frac{1}{2n} \sum_{s=1}^n (a^s - a^{-s})(b^s - b^{-s}) \\
 &= -\frac{1}{2n} \sum_{s=1}^n [(ab)^s - (a/b)^s - (a/b)^{-s} + (ab)^{-s}], \\
 &= -\frac{1}{2n} \left[ \underbrace{\sum_{s=1}^n (ab)^s}_A - \underbrace{\sum_{s=1}^n (a/b)^s}_B - \underbrace{\sum_{s=1}^n (a/b)^{-s}}_C + \underbrace{\sum_{s=1}^n (ab)^{-s}}_D \right].
 \end{aligned} \tag{7.22}$$

Furthermore,

$$ab = e^{i\pi \frac{j+k}{n}}, \quad \frac{a}{b} = e^{i\pi \frac{j-k}{n}}.$$

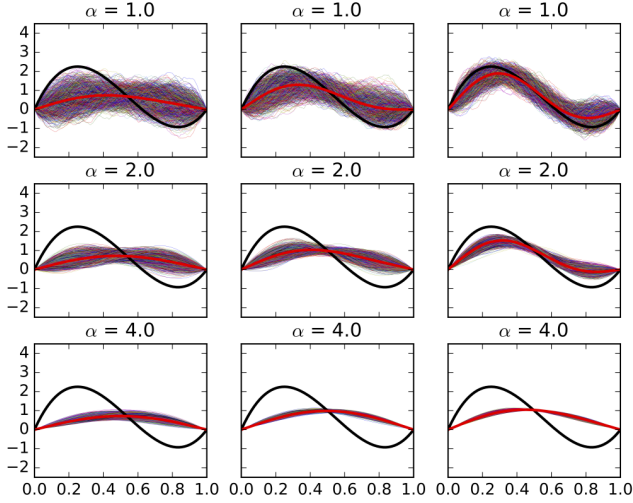


Figure 7.2: Realizations of the posterior mean (red) and 950 of 1000 draws from the posterior (colored thin lines) with smallest  $L^2$  distance to the posterior mean. From left to right columns, the posterior is computed based on sample size  $10^3$ ,  $10^4$  and  $10^5$  respectively. The true parameter (black) is of smoothness  $\beta$  for any  $\beta < 5/2$  and given by (7.21).

Observe that when  $ab \neq 1$ , we have

$$A = \frac{ab(1 - (ab)^n)}{1 - ab}, \quad D = \frac{1 - (ab)^{-n}}{ab - 1}, \quad A + D = \frac{ab(1 - (ab)^n) - (1 - (ab)^{-n})}{1 - ab}.$$

Similarly, if  $a/b \neq 1$ ,

$$B + C = \frac{(a/b)(1 - (a/b)^n) - (1 - (a/b)^{-n})}{1 - (a/b)}.$$

We fix  $1 \leq k \leq n - 1$  and discuss different situations depending on  $j$ .

(I.)  $1 \leq j \leq n - 1$  and  $j + k \neq n$ .

Since  $n \neq j + k < 2n$ , we always have  $ab = e^{i\pi \frac{j+k}{n}} \neq 1$ , and the terms  $A$  and  $D$  can be calculated as above. Similarly, since  $-n < j - k < n$ ,  $a/b = 1$  only when  $j = k$ . Moreover,  $j + k$  and  $j - k$  have the same parity, and so  $j = k$  is only possible if  $j + k$  is even.

(i.)  $j + k$  is even.

In this case,  $(ab)^n = 1$ . This leads to  $A = D = 0$ .

Further, if  $j = k$ , we have  $a/b = b/a = 1$  and  $B = C = n$ . Otherwise, if  $j \neq k$ , we have  $a/b \neq 1$  and  $(a/b)^n = 1 = (b/a)^n$  (since  $j - k$  is even),

and so

$$B = \frac{a/b(1 - (a/b)^n)}{1 - a/b} = 0, \quad C = 0,$$

which implies (7.22) equals 1.

(ii.)  $j + k$  is odd. We have  $(ab)^n = (a/b)^n = -1$ , which results in  $A + D = B + C = -2$ , and so (7.22) equals 0.

(II.)  $1 \leq j < n$  and  $j + k = n$ . We have  $ab = -1$ . Arguing as above, if  $n$  is odd,  $A + D = -2$  and  $B + C = -2$ . If  $n$  is even,  $A = D = 0$  and  $B = C = n\delta_{jk}$ .

The remaining cases follow the same arguments, and hence we omit the (lengthy and elementary) calculations.

(III.)  $j = ln$  with  $l \in \mathbb{N}$ .

It can be shown that  $A + D = B + C$  always holds.

(IV.)  $j \in \{ln + 1, \dots, (l + 1)n - 1\}$ .

When  $l$  is even, one obtains  $\langle \psi_j, \psi_k \rangle_d = \delta_{\tilde{j}k}$ , where  $\tilde{j} = j - ln$ . Otherwise, for odd  $l$ ,  $\langle \psi_j, \psi_k \rangle_d = -\delta_{\tilde{j}k}$  where  $\tilde{j} = (l + 1)n - j$ .

### 7.5.2 Proof of Theorem 7.12

In this proof we use the notation  $\|\cdot\| = \|\cdot\|_{H_1} = \|\cdot\|_{\ell^2}$ . To show

$$\sup_{\|f_0\|_{\beta} \leq K} \mathbb{E}_{f_0} \Pi_n (f : \|f - f_0\| \geq M_n \varepsilon_n | U^n) \rightarrow 0,$$

we first apply Markov's inequality,

$$M_n^2 \varepsilon_n^2 \Pi_n (f : \|f - f_0\|^2 \geq M_n^2 \varepsilon_n^2 | U^n) \leq \int \|f - f_0\|^2 d\Pi_n(f | U^n).$$

From (7.12) and the bias-variance decomposition,

$$\int \|f - f_0\|^2 d\Pi_n(f | U^n) = \|\hat{f} - f_0\|^2 + \|\sigma\|^2,$$

where  $\sigma = (\sigma_k)_k$  is given in (7.12). Because  $\sigma$  is deterministic,

$$\mathbb{E}_{f_0} [\Pi_n (f : \|f - f_0\| \geq M_n \varepsilon_n | U^n)] \leq \frac{1}{M_n^2 \varepsilon_n^2} \left( \mathbb{E}_{f_0} \|\hat{f} - f_0\|^2 + \|\sigma\|^2 \right).$$

Since  $M_n \rightarrow \infty$  is assumed, it suffices to show that the terms in brackets are bounded by a constant multiple of  $\varepsilon_n^2$  uniformly in  $f_0$  in the Sobolev ellipsoid.

Using (7.14), we obtain

$$\mathbb{E}_{f_0} \|\hat{f} - f_0\|^2 = \|\mathbb{E}_{f_0} \hat{f} - f_0\|^2 + \|\tau\|^2 = \|\mathbb{E}_{f_0} \hat{f} - f_0^n\|^2 + \|f_0^r\|^2 + \|\tau\|^2,$$

where  $\tau = (\tau_k)_k$  given in (7.14) and

$$\begin{aligned} f_0^n &= (f_{0,1}, \dots, f_{0,n-1}, 0, \dots), \\ f_0^r &= (0, \dots, 0, f_{0,n}, f_{0,n+2}, \dots). \end{aligned}$$

We need to obtain a uniform upper bound over the ellipsoid  $\{f_0 : \|f_0\|_\beta \leq K\}$  for

$$\|\mathbb{E}_{f_0} \hat{f} - f_0^n\|^2 + \|f_0^r\|^2 + \|\tau\|^2 + \|\sigma\|^2. \quad (7.23)$$

We have

$$\begin{aligned} \|\mathbb{E}_{f_0} \hat{f} - f_0^n\|^2 &= \sum_{k=1}^{n-1} \left( \frac{na_k^2 \lambda_k}{na_k^2 \lambda_k + 1} f_{0,k} + \frac{na_k \lambda_k}{na_k^2 \lambda_k + 1} R_k - f_{0,k} \right)^2 \\ &\lesssim \underbrace{\sum_{k=1}^{n-1} \frac{1}{(na_k^2 \lambda_k + 1)^2} f_{0,k}^2}_{A_1} + n \sup_{k < n} R_k^2 \underbrace{\sum_{k=1}^{n-1} \frac{na_k^2 \lambda_k^2}{(na_k^2 \lambda_k + 1)^2}}_{A_2}, \end{aligned} \quad (7.24)$$

and

$$\|f_0^r\|^2 = \sum_{k \geq n} f_{0,k}^2, \quad \|\tau\|^2 = \sum_{k=1}^{n-1} \frac{na_k^2 \lambda_k^2}{(na_k^2 \lambda_k + 1)^2} = A_2, \quad \|\sigma\|^2 = \sum_{k=1}^{n-1} \frac{\lambda_k}{na_k^2 \lambda_k + 1}.$$

Recall that we write (7.15) as  $\varepsilon_n = \varepsilon_{n,1} \vee \varepsilon_{n,2}$ . The statements (i.)–(iii.) follow by elementary calculations. Specifically, in (ii.) the given  $\rho_n$  is the best scaling, as it gives the fastest rate. From [59] (see the argument below (7.3) on page 21),  $A_1$  is bounded by a fixed multiple of  $(\varepsilon_{n,1})^2$ , and  $\|\tau\|^2, \|\sigma\|^2$  are bounded by multiples of  $(\varepsilon_{n,2})^2$ . Hence, to show that the rate is indeed (7.15), it suffices to show that  $n \sup_{k < n} R_k^2 A_2$  and  $\|f_0^r\|^2$  can be bounded by a multiple of  $(\varepsilon_n)^2$  uniformly in the ellipsoid  $\{f_0 : \|f_0\|_\beta \leq K\}$ . Since  $A_2 = \|\tau\|^2$ , to that end it is sufficient to show that  $\sup_{k < n} n R_k^2 = O(1)$ , and that  $\|f_0^r\|^2 = O(\varepsilon_n)^2$ .

Since  $f_0 \in S^\beta$ , we have the following straightforward bound,

$$\|f_0^r\|^2 \leq n^{-2\beta} \sum_{k \geq n} f_{0,k}^2 k^{2\beta} \leq n^{-2\beta} \|f_0\|_\beta^2 \lesssim n^{-2\beta},$$

which is uniform in  $\{f_0 : \|f_0\|_\beta \leq K\}$ . By comparing to the rates in the statements (ii.)–(iii.), it is easy to see that  $n^{-2\beta}$  is always negligible with respect to  $\varepsilon_n^2$ .

Proving  $\sup_{k < n} n R_k^2 = O(1)$  is equivalent to showing  $\sup_{k < n} R_k^2 = O(1/n)$ ; but the latter has been already proved in (7.10). Notice that we actually obtained a sharper bound  $\sup_{k < n} n R_k^2 = o(1)$  than the one necessary for our purposes in this proof. However, this sharper bound will be used in the proof of Theorem 7.15. By taking supremum over  $f_0$ , we thus have

$$\sup_{\|f_0\|_{S^\beta} \leq K} \left( \|\mathbb{E}_{f_0} \hat{f} - f_0^n\|^2 + \|f_0^r\|^2 \right) \lesssim \varepsilon_n^2 + n^{-2\beta} \lesssim \varepsilon_n^2, \quad (7.25)$$

with which we conclude that up to a multiplicative constant, (7.23) is bounded by  $\varepsilon_n^2$  uniformly over the ellipsoid  $\sup_{\|f_0\|_\beta \leq K}$ . This completes the proof.

### 7.5.3 Proof of Theorem 7.14

We start by generalizing Theorem 3.1 in [60]. Following the same lines as in the proof of that theorem and using Lemma 7.17, 7.18, 7.19, 7.20 in Section 7.6 of the present paper instead of analogous technical results in [60], the statement of Theorem 3.1 in [60] can be extended from  $s = 2$  to a general  $s \geq 1$ , for which the posterior rate is given by (7.16), or  $\varepsilon_n = \varepsilon_{n,1} \vee \varepsilon_{n,2}$  in short.

In our model, we again obtain (7.23) and also that a fixed multiple of  $(\varepsilon_{n,1})^2$  is an upper bound of  $A_1$ , and that  $\|\tau\|^2, \|\sigma\|^2$  can be bounded from above by fixed multiples of  $(\varepsilon_{n,1})^2$ .

Now as in the proof of Theorem 7.12 in Section 7.5.2, we will show that  $\sup_{\|f_0\|_\beta \leq K} (\|\mathbb{E}_{f_0} \hat{f} - f_0^n\|^2 + \|f_0^r\|^2)$  can be bounded by a fixed multiple of  $(\varepsilon_n)^2$  by proving that  $\sup_{k \leq n} nR_k^2 = O(1)$ . By (7.11),  $n(R_k)^2 \lesssim \exp(-2pn)n$ , and the righthand side converges to zero. Therefore,

$$\sup_{\|f_0\|_\beta \leq K} \left( \|\mathbb{E}_{f_0} \hat{f} - f_0^n\|^2 + \|f_0^r\|^2 \right) \lesssim \varepsilon_n.$$

Parts (i.) and (ii.) of the statement of the theorem are obtained by direct substitutions, using the fact that  $\log n \ll n$ . Notice that if  $\rho_n \gtrsim (\log n)^{(\alpha-\beta)/s}$ , the rate  $\varepsilon_n$  deteriorates and is dominated by the second term in (7.16).

For the case  $\lambda_k = \exp(-\alpha k^s)$ , the argument follows the same lines as in Section 5.1 in [60], and our arguments above.

### 7.5.4 Proof of Theorem 7.15

The proof runs along the same lines as the proof of Theorem 4.2 in [59]. We will only show the main steps here.

In Section 7.2.2, we have shown that the posterior distribution is  $\otimes_{k \in \mathbb{N}} \mathcal{N}(\hat{f}_k, \sigma_k^2)$ , the radius  $r_{n,\gamma}$  in (7.17) satisfies  $\mathbb{P}_{X_n}(X_n < r_{n,\gamma}^2) = 1 - \gamma$ , where  $X_n$  is a random variable distributed as the square norm of an  $\otimes_{k \in \mathbb{N}} \mathcal{N}(\hat{f}_k, \sigma_k^2)$  variable. Let  $T = (\tau_k^2)_{k \in \mathbb{N}}$ . Under (7.8), the variable  $\hat{f}$  is distributed as  $\mathcal{N}_{H_1}(\mathbb{E}_{f_0} \hat{f}, T) := \otimes_{k \in \mathbb{N}} \mathcal{N}(\mathbb{E}_{f_0} \hat{f}_k, \tau_k^2)$ . Hence the coverage (7.19) can be rewritten as

$$\mathbb{P}_{W_n}(\|W_n + \mathbb{E}_{f_0} \hat{f} - f_0\|_{H_1} \leq r_{n,\gamma}), \quad (7.26)$$

where  $W_n \sim \mathcal{N}_{H_1}(0, T)$ . Denote  $V_n = \|W_n\|_{H_1}^2$  and observe that one has in distribution

$$X_n = \sum_{1 \leq i < n} \sigma_i^2 Z_i^2, \quad V_n = \sum_{1 \leq i < n} \tau_i^2 Z_i^2$$

for  $\{Z_i\}$  independent standard Gaussian random variables with

$$\sigma_i^2 = \frac{\lambda_i}{na_i^2 \lambda_i + 1}, \quad \tau_i^2 = \frac{na_i^2 \lambda_i^2}{(na_i^2 \lambda_i + 1)^2}.$$

By the same argument as in [59], one can show that the standard deviations of  $X_n$  and  $V_n$  are negligible with respect to their means,

$$\mathbb{E}X_n \asymp \rho_n^2 (\rho_n^2 n)^{-2\alpha/(2\alpha+2p+1)}, \quad \mathbb{E}V_n \asymp \rho_n^2 (\rho_n^2 n)^{-2\alpha/(2\alpha+2p+1)}, \quad (7.27)$$

and the difference of their means,

$$\mathbb{E}(X_n - V_n) \asymp \rho_n^2 (\rho_n^2 n)^{-2\alpha/(2\alpha+2p+1)}.$$

Since  $X_n \geq V_n$ , the distributions of  $X_n$  and  $V_n$  are asymptotically separated, i.e.  $\mathbb{P}(V_n \leq v_n \leq X_n) \rightarrow 1$  for some  $v_n$ , e.g.  $v_n = \mathbb{E}(V_n + X_n)/2$ . Since  $r_{n,\gamma}^2$  are  $1 - \gamma$  quantiles of  $X_n$ , we also have  $\mathbb{P}(V_n \leq r_{n,\gamma}^2(1 + o(1))) \rightarrow 1$ . In addition, by (7.27),

$$r_{n,\gamma}^2 \asymp \rho_n^2 (\rho_n^2 n)^{-2\alpha/(2\alpha+2p+1)}.$$

Introduce

$$B_n := \sup_{\|f_0\|_\beta \lesssim 1} \|\mathbb{E}_{f_0} \hat{f} - f_0\|_{H_1} = \sup_{\|f_0\|_\beta \lesssim 1} \left( \|\mathbb{E}_{f_0} \hat{f} - f_0^n\|_{H_1} + \|f_0^n\|_{H_1} \right). \quad (7.28)$$

It follows from the arguments for (7.10) in the proof of Theorem 7.12 that

$$B_n \lesssim \varepsilon_{n,1} \vee (\sqrt{n} R \varepsilon_{n,2}),$$

where  $R = \sup_{k < n} R_k \lesssim n^{-(p+\beta)}$ . Now apply the argument on the lower bound from Lemma 8.1 in [59] (with  $q = \beta, t = 0, u = 2\alpha + 2p + 1, v = 2, N = \rho_n^2 n$ ) to obtain that  $B_n \gtrsim \varepsilon_{n,1}$ . Thus we have

$$\varepsilon_{n,1} \lesssim B_n \lesssim \varepsilon_{n,1} \vee (\sqrt{n} R \varepsilon_{n,2}).$$

We consider separate cases. In case (i.), substituting the corresponding  $\rho_n$  into the expression of  $\varepsilon_{n,1}$  and  $\varepsilon_{n,2}$ , we have  $\varepsilon_{n,1} \ll \varepsilon_{n,2}$ . By (7.10),  $B_n \lesssim \varepsilon_{n,1} \vee (\sqrt{n} R \varepsilon_{n,2}) \ll \varepsilon_{n,2} \asymp r_{n,\gamma}$ . This leads to

$$\begin{aligned} \mathbb{P}(\|W_n + \mathbb{E}_{f_0} \hat{f} - f_0\|_{H_1} \leq r_{n,\gamma}) &\geq \mathbb{P}(\|W_n\|_{H_1} \leq r_{n,\gamma} - B_n) \\ &= \mathbb{P}(V_n \leq r_{n,\gamma}^2(1 + o(1))) \rightarrow 1 \end{aligned} \quad (7.29)$$

uniformly in the set  $\{f_0 : \|f_0\|_\beta \lesssim 1\}$ .

In case (iii.), the given  $\rho_n$  leads to  $\varepsilon_{n,1} \gg \varepsilon_{n,2}$  and consequently  $B_n \gg r_{n,\gamma}$ . Hence,

$$\mathbb{P}(\|W_n + \mathbb{E}_{f_0} \hat{f}^n - f_0^n\|_{H_1} \leq r_{n,\gamma}) \leq \mathbb{P}(\|W_n\|_{H_1} \geq B_n - r_{n,\gamma}) \rightarrow 0,$$

for any  $f_0^n$  (nearly) attaining the supremum.

In case (ii.), we have  $B_n \asymp r_{n,\gamma}$ . If  $\beta < 2\alpha + 2p + 1$ , by Lemma 8.1 in [59] the bias  $\mathbb{E}_{f_0} \hat{f} - f_0$  at a fixed  $f_0$  is of strictly smaller order than  $B_n$ . Following the argument of case (i.), the asymptotic coverage can be shown to converge to 1.

For existence of a sequence along which the coverage is  $c \in [0, 1)$ , we only give a sketch of the proof here; the details can be filled in as in [59].

The coverage (7.26) with  $f_0$  replaced by  $f_0^n$  tends to  $c$ , if for  $b_n = \mathbb{E}_{f_0} \hat{f}^n - f_0^n$  and  $z_c$  a standard normal quantile,

$$\frac{\|W_n + b_n\|_{H_1}^2 - \mathbb{E}\|W_n + b_n\|_{H_1}^2}{\text{sd}\|W_n + b_n\|_{H_1}^2} \rightsquigarrow \mathcal{N}(0, 1), \quad (7.30)$$

$$\frac{r_{n,\gamma}^2 - \mathbb{E}\|W_n + b_n\|_{H_1}^2}{\text{sd}\|W_n + b_n\|_{H_1}^2} \rightarrow z_c, \quad (7.31)$$

Since  $W_n$  is centred Gaussian  $\mathcal{N}_{H_1}(0, T)$ , (7.31) can be expressed as

$$\frac{r_{n,\gamma}^2 - \mathbb{E}V_n - \sum_{i=1}^{n-1} b_{n,i}^2}{\sqrt{\text{var } V_n + 4 \sum_{i=1}^{n-1} \tau_{i,n}^2 b_{n,i}^2}} \rightarrow z_c. \quad (7.32)$$

Here  $\{b_{n,i}\}$  has exactly one nonzero entry depending on the smoothness cases  $\beta \leq 2\alpha + 2p + 1$  and  $\beta > 2\alpha + 2p + 1$ . The nonzero entry, which we call  $b_{n,i_n}$ , has the following representation, with  $d_n$  to be yet determined,

$$b_{n,i_n}^2 = r_{n,\gamma}^2 - \mathbb{E}V_n - d_n \text{sd } V_n.$$

Since  $r_{n,\gamma}^2, \mathbb{E}V_n$  and  $r_{n,\gamma}^2 - \mathbb{E}V_n$  have the same order and  $\text{sd } V_n$  is of strictly smaller order, one can show that the lefthand side of (7.32) is equivalent to

$$\frac{d_n \text{sd } V_n}{\sqrt{\text{var } V_n + 4\tau_{i_n,n}^2 (r_{n,\gamma}^2 - \mathbb{E}V_n)(1 + o(1))}},$$

for bounded or slowly diverging  $d_n$ . Then (7.32) can be obtained by discussing different smoothness cases separately, by a suitable choice of  $i_n, d_n$ .

To prove the asymptotic normality in (7.30), the numerator can be written as

$$\|W_n + b_n\|_{H_1}^2 - \mathbb{E}\|W_n + b_n\|_{H_1}^2 = \sum_i \tau_{i,n}^2 (Z_i^2 - 1) + 2b_{n,i_n} \tau_{i_n,n} Z_{i_n}.$$

Next one applies the arguments as in [59].

### 7.5.5 Proof of Theorem 7.16

This proof is almost identical to the proof of Theorem 2.2 in [60]. We supply the main steps.

Following the same arguments as in the proof of Theorem 7.15, we obtain

$$\begin{aligned} \mathbb{E}X_n &\asymp \rho_n^2 (\log(\rho_n^2 n))^{-2\alpha/s} \gg \text{sd } X_n \asymp \rho_n^2 (\log(\rho_n^2 n))^{-1/(2s)-2\alpha/s}, \\ \mathbb{E}V_n &\asymp \rho_n^2 (\log(\rho_n^2 n))^{-1/s-2\alpha/s} \asymp \text{sd } V_n, \end{aligned}$$

as in the proof of Theorem 2.2 in [60]. This leads to

$$r_{n,\gamma}^2 \asymp \rho_n^2 (\log(\rho_n^2 n))^{-2\alpha/s},$$

and furthermore,

$$\mathbb{P}(V_n \leq \delta r_{n,\gamma}^2) = \mathbb{P}\left(\frac{V_n - \mathbb{E}V_n}{\text{sd } V_n} \leq \frac{\delta r_{n,\gamma}^2 - \mathbb{E}V_n}{\text{sd } V_n}\right) \rightarrow 1,$$

for every  $\delta > 0$ .

Similar to Theorem 7.15, the bounds on the square norm  $B_n$  (defined in (7.28)) of the bias are known: upper bound from the proof of Theorem 7.14, and lower bound from Lemma 7.17,

$$\varepsilon_{n,1} \lesssim B_n \lesssim \varepsilon_{n,1} \vee (\sqrt{n}R\varepsilon_{n,2}),$$

where  $\varepsilon_{n,1}, \varepsilon_{n,2}$  are given in (7.16), and  $\sqrt{n}R$  satisfies the bound (7.11).

In case (i.),  $B_n \ll r_{n,\gamma}$ , and hence (7.29) applies. The rest of the results can be obtained in a similar manner.

## 7.6 Auxiliary lemmas

The following lemmas are direct generalisations of the case  $s = 2$  in the Appendix of [60] to a general  $s$ . They can be easily proved by simple adjustments of the original proofs in [60], and we only state the results.

**Lemma 7.17** (Lemma 6.1 in [60]). *For  $q \in \mathbb{R}$ ,  $u \geq 0$ ,  $v > 0$ ,  $t + 2q \geq 0$ ,  $p > 0$ ,  $0 \leq r < pv$  and  $s \geq 1$ ,*

$$\sup_{\|f\|_{S^q} \leq 1} \sum_{i=1}^{\infty} \frac{f_i^2 i^{-t} e^{-ri^s}}{(1 + Ni^{-u} e^{-pi^s})^v} \asymp N^{-r/p} (\log N)^{-t/s - 2q/s + ru/ps},$$

as  $N \rightarrow \infty$ .

In addition, for any fixed  $f \in S^q$ ,

$$N^{r/p} (\log N)^{t/s + 2q/s - ru/ps} \sum_{i=1}^{\infty} \frac{f_i^2 i^{-t} e^{-ri^s}}{(1 + Ni^{-u} e^{-pi^s})^v} \rightarrow 0,$$

as  $N \rightarrow \infty$ .

**Lemma 7.18** (Lemma 6.2 in [60]). *For  $t, u \geq 0$ ,  $v > 0$ ,  $p > 0$ ,  $0 < r < vp$  and  $s \geq 1$ , as  $N \rightarrow \infty$ ,*

$$\sum_{i=1}^{\infty} \frac{i^{-t} e^{-ri^s}}{(1 + Ni^{-u} e^{-pi^s})^v} \asymp N^{-r/p} (\log N)^{-t/s + ru/ps}.$$

If  $r = 0$  and  $t > 1$ , while other assumptions remain unchanged,

$$\sum_{i=1}^{\infty} \frac{i^{-t} e^{-ri^s}}{(1 + Ni^{-u} e^{-pi^s})^v} \asymp (\log N)^{-(t+1)/s}.$$

**Lemma 7.19** (Lemma 6.4 in [60]). *Assume  $s \geq 1$ . Let  $I_N$  be the solution in  $i$  to  $Ni^{-u} e^{-pi^s} = 1$ , for  $u \geq 0$  and  $p > 0$ . Then*

$$I_N \sim \left( \frac{1}{p} \log N \right)^{1/s}$$

**Lemma 7.20** (Lemma 6.5 in [60]). *Let  $s \geq 1$ . As  $K \rightarrow \infty$ , we have*

(i.) for  $a > 0$  and  $b \in \mathbb{R}$ ,

$$\int_1^K e^{ax^s} x^b dx \sim \frac{1}{as} e^{aK^s} K^{b-s+1},$$

(ii.) for  $a, b, K > 0$ ,

$$\int_K^\infty e^{-ax^s} x^{-b} dx \leq \frac{1}{as} e^{-aK^s} K^{-b-s+1}.$$