Bayesian approach to inverse problems for functions with a variable-index Besov prior

2016 Inverse Problems 32 085006

(http://iopscience.iop.org/0266-5611/32/8/085006)

IOP Publishing Inverse Problems

Inverse Problems **32** (2016) 085006 (32pp) doi:10.1088/0266-5611/32/8/085006

Bayesian approach to inverse problems for functions with a variable-index Besov prior

Junxiong Jia[1], **Jigen Peng**[1] **and Jinghuai Gao**[2]

[1] Department of Mathematics, Xi'an Jiaotong University, Xi'an 710049, People's Republic of China; Beijing Center for Mathematics and Information Interdisciplinary Sciences (BCMIIS)

[2] School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, People's Republic of China; Beijing Center for Mathematics and Information Interdisciplinary Sciences (BCMIIS)

E-mail: jjx323@xjtu.edu.cn, jgpeng@xjtu.edu.cn and jhgao@xjtu.edu.cn

Received 16 August 2015, revised 1 March 2016
Accepted for publication 31 May 2016
Published 23 June 2016

Abstract

The Bayesian approach has been adopted to solve inverse problems that reconstruct a function from noisy observations. Prior measures play a key role in the Bayesian method. Hence, many probability measures have been proposed, among which total variation (TV) is a well-known prior measure that can preserve sharp edges. However, it has two drawbacks, the staircasing effect and a lack of the discretization-invariant property. The variable-index TV prior has been proposed and analyzed in the area of image analysis for the former, and the Besov prior has been employed recently for the latter. To overcome both issues together, in this paper, we present a variable-index Besov prior measure, which is a non-Gaussian measure. Some useful properties of this new prior measure have been proven for functions defined on a torus. We have also generalized Bayesian inverse theory in infinite dimensions for our new setting. Finally, this theory has been applied to integer- and fractional-order backward diffusion problems. To the best of our knowledge, this is the first time that the Bayesian approach has been used for the fractional-order backward diffusion problem, which provides an opportunity to quantify its uncertainties.

Keywords: Bayesian inverse problems, fractional-order backward diffusion, variable Besov prior

1. Introduction

Inverse problems are defined, as the term itself indicates, as the inverse of direct or forward problems. For forward problems, partial differential equations are useful tools for modeling real-world physical systems. The outcome of some measurements can be predicted by specifying the coefficients in partial differential equations (e.g., sound velocity in the acoustic wave equation). Inverse problems aim to reconstruct the coefficients from some measurements of the solutions of partial differential equations. The forward model described using partial differential equations usually has a unique solution, while the inverse problem does not [40]. In order to overcome the issue of non-uniqueness, some form of regularization is required to ameliorate ill-posed behavior.

Regularization techniques are useful tools that produce a reasonable estimate of quantities of interest based on the data available. Studies on regularization techniques have a long history, dating back to A. N. Tikhonov in 1963 [42]. For a function u, Tikhonov regularization usually has the form $\|\nabla u\|_{L^2}^2$ or $\|u\|_{L^2}^2$, which penalizes the L^2 norm of the function or the gradient of the function. It is well known that Tikhonov regularization always leads to over-smoothing, and therefore, total variation (TV) regularization has been proposed in [37]. The TV penalty has gained increasing popularity because it can preserve important details such as the edges of the image. In 1997, Blomgren, Chan, Mulet and Wong [3] noticed that TV restoration typically exhibits 'blockiness', or a 'staircasing' effect, where the restored image comprises piecewise flat regions. Therefore, they proposed a regularization term as follows:

$$\int_{\Omega} |\nabla u|^{p(\nabla u)} \mathrm{d}x, \tag{1.1}$$

where p monotonically decreases from 2, when $|\nabla u| = 0$, to 1, as $|\nabla u| \nearrow \infty$. However, it is difficult to study (1.1) mathematically because the lower semi-continuity of (1.1) is not readily evident. Later in 2006, Chen, Levine and Rao [4] proposed another kind of variable-index TV norm, using which they also constructed complete mathematical theories for the variational model. In 2014, Tiirola [41] used a variable-index TV norm and variable-index Besov regularization terms in image decomposition problems.

However, regularization techniques cannot be used for uncertainty analysis. Statistical inversion theory reformulates inverse problems as problems of statistical inference using Bayesian statistics. Dating back to 1970, Franklin [14] formulated PDE inverse problems in terms of Bayes' formula on some Hilbert space X. Franklin derived a regularization using the Bayesian approach and obtained the relation between regularization techniques and the Bayesian approach. Recently, Lasanen [26–29] developed a fully nonlinear theory. Cotter, Dashti Robinson, Stuart, Law and Voss [5, 9, 39] established a mathematical framework for a range of inverse problems for functions, given noisy observations. They revealed the relationship between regularization techniques and the Bayesian framework, and estimated the error of finite-dimensional approximate solutions. Based on this framework, Cotter, Roberts, Stuart and White [6] developed faster MCMC algorithms.

We now state the relationship between TV regularization and the Bayesian approach. Solving an optimization problem with regularization terms (e.g., Tikhonov regularization, TV regularization) could sometimes be seen as acquiring the maximum point of the posterior probability measure [5, 21]. Under the Bayesian framework, except for the maximum point, Bayesian conditional mean estimates can provide us valuable information. In [31], Lassas and Siltanen found that TV regularization does not have the discretization-invariant property. More specifically, Bayesian conditional mean estimates for the TV prior distribution are not edge-preserving with very fine discretizations of the model space. In order to overcome this deficiency, Lassas, Saksman and S. Siltanen [30] proposed the Besov prior $B_{1,1}^1$ which has the discretization-invariant property. Dashti, Harris and Stuart [8] studied the Besov prior under the mathematical framework established in [5]. Under this framework, the Besov prior naturally has the discretization-invariant property because the framework is originally built on infinite-dimensional space.

Considering the Besov and TV regularization techniques and Bayes' inverse theory, variable-index Besov prior theory does not seem to be available. As mentioned before, variable-index TV and Besov regularization terms have been used in image analysis and yielded good performance [41]. In this paper, we attempt to build a variable-index Besov prior and generalize Bayes' inverse theory by using this new prior probability measure. In section 5, we apply our theory to integer-order backward diffusion problems and fractional-order backward diffusion problems.

The main contributions of this paper are as follows:

1. We construct a variable-index Besov prior using wavelet characterization of the variable-index Besov space and prove a Fernique-like result [7] for the variable-index Besov prior.
2. Based on the variable-index Besov prior, we generalize the results in [39] to build Bayes' inverse theory. Under the same conditions for the forward operator, we also prove the convergence of variational problems using the variable-index Besov regularization term.
3. Although there are many studies on inverse problems for fractional diffusion equations [44, 45], the number of studies on fractional-order backward diffusion problems under the Bayes' inverse framework is limited. Using our theory, we prove that a posterior measure exists as well as the continuity of the posterior measure with respect to the data for integer- and fractional-order backward diffusion problems.

The contents of this paper are organized as follows. In section 2, some basic knowledge on the variable-index space is presented, and the wavelet characterization of the variable-index Besov space on a periodic domain is proved. In section 3, we first construct the variable-index Besov prior and prove a Fernique-like theorem. Second, we generalize Bayesian inverse theory to our variable-index Besov prior setting. In section 4, under the same conditions as in section 3 for the forward problem, we prove that the variational problem with the variable-index Besov regularization term converges. In section 5, our theory is applied to integer- and fractional-order backward diffusion problems. In the last section, we provide some technical lemmas, and for the reader's convenience, we list some of the useful theorems and lemmas used in our paper.

2. Variable-order space and wavelet characterization

In this section, we provide a short introduction to space of variable smoothness and integrability on a periodic domain and then prove a wavelet characterization of the variable-index Besov space on the periodic domain.

2.1. Modular spaces

Definition 2.1. [11] Let X be a vector space over $\mathbb{R}$ or $\mathbb{C}$. A function $\rho : X \to [0, \infty]$ is semimodular on X if the following properties hold:

1. $\rho(0) = 0$.
2. $\rho(\lambda f) = \rho(f)$ for all $f \in X$ and $|\lambda| = 1$.
3. $\rho(\lambda f) = 0$ for all $\lambda > 0$ implies $f = 0$.
4. $\lambda \mapsto \rho(\lambda f)$ is left continuous on $[0, \infty)$ for every $f \in X$.

A semimodular ρ is modular if

4. $\rho(f) = 0$ implies $f = 0$.

   A semimodular ρ is continuous if

5. for every $f \in X$, the mapping $\lambda \mapsto \rho(\lambda f)$ is continuous on $[0, \infty)$.

   A semimodular ρ can be additionally qualified by the term convex, which means, as usual, that

$$\rho(\theta f + (1 - \theta)g) \leqslant \theta\rho(f) + (1 - \theta)\rho(g),$$

for all $f, g \in X$.

Once we have a semimodular in place, we obtain a normed space in the standard way:

Definition 2.2. [11] If ρ is (semi)modular on X, then

$$X_\rho \coloneqq \{x \in X: \exists\ \lambda > 0, \rho(\lambda x) < \infty\}$$

is called a (semi)modular space.

Theorem 2.3. *[11] Let ρ be a convex semimodular on X. Then, X_ρ is a normed space with the Luxemburg norm given by*

$$\|x\|_\rho \coloneqq \inf\left\{\lambda > 0: \ \rho\left(\frac{1}{\lambda}x\right) \leqslant 1\right\}.$$

2.2. Spaces with variable integrability

The variable exponents that we consider are always measurable functions on an n-dimensional torus $\mathbb{T}^n$ with the range $[1, \infty)$. We denote the set of such functions by $\mathcal{P}$. We denote $p^+ = \text{esssup}_{x\in\mathbb{T}^n} p(x)$ and $p^- = \text{essinf}_{x\in\mathbb{T}^n} p(x)$. The function φ_p is defined as follows:

$$\varphi_p(t) = \begin{cases} t^p, & \text{if } p \in (0, \infty), \\ 0, & \text{if } p = \infty \text{ and } t \leqslant 1, \\ \infty, & \text{if } p = \infty \text{ and } t > 1. \end{cases}$$

The convention $1^\infty = 0$ is adopted so that φ_p is left continuous. We now write t^p instead of $\varphi_p(t)$. The variable-exponent modular is defined by

$$\rho_{L^{p(\cdot)}}(f) \coloneqq \int_{\mathbb{T}^n} |f(x)|^{p(x)}\,\mathrm{d}x.$$

The variable-exponent Lebesgue space $L^{p(\cdot)}$ and its norm $\|f\|_{p(\cdot)}$ are defined by the modular, as explained in the previous subsection.

We say that $g: \mathbb{T}^n \to \mathbb{R}$ is locally log-Hölder continuous, abbreviated as $g \in C^{\log}_{\text{loc}}(\mathbb{T}^n)$, if there exists $c > 0$ such that

$$|g(x) - g(y)| \leqslant \frac{c}{\log(e + 1/|x - y|)}$$

for all $x, y \in \mathbb{T}^n$. We say that g is globally log-Hölder continuous, abbreviated as $g \in C^{\log}$, if it is locally log-Hölder continuous and there exists $g_\infty \in \mathbb{R}$ such that

$$|g(x) - g_\infty| \leqslant \frac{c}{\log(e + |x|)}$$

for all $x \in \mathbb{T}^n$. The notation $\mathcal{P}^{\log}$ is used for the variable exponents $p \in \mathcal{P}$ with $\frac{1}{p} \in C^{\log}$; that is to say, $1 \leqslant p^- \leqslant p(x) \leqslant p^+ < \infty$ and $\frac{1}{p}$ is globally log-Hölder continuous.

2.3. Variable-index Besov space

Before we introduce the variable-index Besov space, we require the following definition of a mixed Lebesgue-sequence space.

Definition 2.4. [1] Let $p, q \in \mathcal{P}$. The mixed Lebesgue-sequence space $\ell^{q(\cdot)}(L^{p(\cdot)})$ is defined on sequences of $L^{p(\cdot)}$-functions by the modular

$$\rho_{\ell^{q(\cdot)}(L^{p(\cdot)})}(\{f_v\}_v) \coloneqq \sum_v \inf\left\{\lambda_v > 0 \mid \rho_{L^{p(\cdot)}}(f_v/\lambda_v^{\frac{1}{q(\cdot)}}) \leqslant 1\right\}.$$

As usual, we denote the Fourier transform of a distribution or a function f as $\mathcal{F}(f)$ or $\hat{f}$. We denote the inverse Fourier transform of a distribution or a function f as $f^\vee$. Similar to the constant-index case, we require the following definition of admissible functions.

Definition 2.5. [1] We say that a pair (φ, Φ) is admissible if $\varphi, \Phi \in \mathcal{S}$ satisfy

- supp $\hat{\varphi} \subset \{\xi \in \mathbb{R}^n: \ 1/2 \leqslant |\xi| \leqslant 2\}$ and $|\hat{\varphi}(\xi)| \geqslant c > 0$ when $3/5 \leqslant |\xi| \leqslant 5/3$,
- supp $\hat{\Phi} \subset \{\xi \in \mathbb{R}^n: \ |\xi| \leqslant 2\}$ and $|\hat{\Phi}(\xi)| \geqslant c > 0$ when $|\xi| \leqslant 5/3$.

We set $\varphi_v(x) \coloneqq 2^{vn}\varphi(2^v x)$ for $v \in \mathbb{N}$ and $\varphi_0(x) \coloneqq \Phi(x)$. Denote as $\mathcal{S}$ the Schwartz function space and as $\mathcal{S}'$ the tempered distribution that is the dual space of $\mathcal{S}$. Then, the variable-index Besov space in our setting can be defined as follows.

Definition 2.6. [1] Let φ_v be defined as in definition 2.5. For $\alpha: \mathbb{T}^n \to \mathbb{R}$ and $p, q \in \mathcal{P}$, the variable-index Besov space $B^{\alpha(\cdot)}_{p(\cdot),q(\cdot)}$ consists of all distributions $f \in \mathcal{S}'$ such that

$$\|f\|_{B^{\alpha(\cdot)}_{p(\cdot),q(\cdot)}} \coloneqq \|(2^{v\alpha(\cdot)}\varphi_v * f)_v\|_{\ell^{q(\cdot)}(L^{p(\cdot)})} < \infty.$$

In the above definition, $\varphi_v * f(x)$ can be simplified as

$$\sum_{m\in\mathbb{Z}^n} (2\pi)^{n/2} a_m \hat{\varphi}_v(2\pi m) \mathrm{e}^{\mathrm{i}2\pi m\cdot x}$$

with $f(x) = \sum_{m\in\mathbb{Z}^n} a_m \mathrm{e}^{\mathrm{i}2\pi m\cdot x}$. For detailed information on the periodic space, we refer to chapter 1 in [43].

In the case of $p = q$, we use the notation $B^{\alpha(\cdot)}_{q(\cdot)} \coloneqq B^{\alpha(\cdot)}_{p(\cdot),q(\cdot)}$. We can also associate the following modular with the Besov space:

$$\rho_{b^{\alpha(\cdot)}_{p(\cdot),q(\cdot)}} \coloneqq \rho_{\ell^{q(\cdot)}(L^{p(\cdot)})}((2^{v\alpha(\cdot)}\varphi_v * f)_v),$$

which can be used to define the norm. For the reader's convenience, we also list the definition of the variable-index Triebel–Lizorkin space $F^{\alpha(\cdot)}_{p(\cdot),q(\cdot)}$.

Definition 2.7. [12] Let φ_v, $v \in \mathbb{N} \cup \{0\}$, be defined as in definition 2.5. The Triebel–Lizorkin space $F^{\alpha(\cdot)}_{p(\cdot),q(\cdot)}$ is defined to be the space of all distributions $f \in \mathcal{S}'$ with $\|f\|_{F^{\alpha(\cdot)}_{p(\cdot),q(\cdot)}} < \infty$, where

$$\|f\|_{F^{\alpha(\cdot)}_{p(\cdot),q(\cdot)}} := \| \|2^{v\alpha(\cdot)}\varphi_v * f\|_{\ell^{q(\cdot)}} \|_{L^{p(\cdot)}}.$$

In the case of $p = q$, we use the notation $F^{\alpha(\cdot)}_{q(\cdot)} := F^{\alpha(\cdot)}_{q(\cdot),q(\cdot)}$. In the rest of this paper, we set $A \approx B$ equal to $cA \leqslant B \leqslant CA$ where c, C are two constants.

In the following parts of this paper, for a function $s(\cdot)$, we always use $s^+ = \sup_{x\in\mathbb{T}^n} s(x)$ and $s^- = \inf_{x\in\mathbb{T}^n} s(x)$.

2.4. Wavelet characterization

Now, we state some notations for wavelet theory and then prove a wavelet characterization of variable-index Besov and Triebel–Lizorkin space on the periodic domain.

Let ψ^M, ψ^F be the Meyer or Daubechies wavelets described in proposition A.1 in the appendix. Now, we define

$$G^0 = \{F, M\}^n \quad \text{and} \quad G^j = \{F, M\}^{n*} \text{ if } j \geqslant 1,$$

where $*$ indicates that at least one G_i of $G = (G_1, \cdots, G_n) \in \{F, M\}^{n*}$ must be an M. It is clear from the definition that the cardinal number of $\{F, M\}^{n*}$ is $2^n - 1$. Let $x \in \mathbb{R}^n$ be

$$\Psi^j_{Gm}(x) = 2^{j\frac{n}{2}} \prod_{r=1}^{n} \psi^{Gr}(2^j x_r - m_r), \tag{2.1}$$

where $G \in G^j$, $m \in \mathbb{Z}^n$ and $j \in \mathbb{N}_0$. Then, $\{\Psi^j_{Gm} : j \in \mathbb{N}_0, G \in G^j, m \in \mathbb{Z}^n\}$ is an orthonormal basis in $L^2(\mathbb{R}^n)$.

Define

$$\psi^M_{j,k}(x) := 2^{\frac{j}{2}}\psi^M(2^j x - k) \quad \psi^F_{j,k}(x) := 2^{\frac{j}{2}}\psi^F(2^j x - k)$$

where $k \in \mathbb{Z}$. Then, we define

$$\tilde{\psi}^M_{j,k}(x) := \sum_{\ell\in\mathbb{Z}} \psi^M_{j,k}(x + \ell) = 2^{j\frac{n}{2}} \sum_{\ell\in\mathbb{Z}} \psi^M(2^j(x + \ell) - k) \tag{2.2}$$

and

$$\tilde{\psi}^F_{j,k}(x) := \sum_{\ell\in\mathbb{Z}} \psi^F_{j,k}(x + \ell) = 2^{j\frac{n}{2}} \sum_{\ell\in\mathbb{Z}} \psi^F(2^j(x + \ell) - k). \tag{2.3}$$

Obviously, $\tilde{\psi}_{j,k}$, $\tilde{\phi}_{j,k}$ are 1-periodic functions belonging to $L^1([0, 1])$.

Define

$$\tilde{\Psi}^j_{Gm}(x) = 2^{j\frac{n}{2}} \prod_{r=1}^{n} \tilde{\psi}^{Gr}(2^j x_r - m_r). \tag{2.4}$$

By proposition A.1, we know that ψ^M, ψ^F is included in the functions with radial decreasing L^1-majorants; that is,

$$|\psi^M(x)| \leqslant R_1(|x|) \quad |\psi^F(x)| \leqslant R_2(|x|),$$

where R_1 and R_2 are bounded decreasing functions belonging to $L^1([0, \infty))$. We can use theorem 5.9 in [17] to show that $\{\tilde{\Psi}^j_{Gm} : j \in \mathbb{N} \cup \{0\}, G \in G^j, m \in \mathbb{M}_j\}$ with

$\mathbb{M}_j = \{m : m = 0, 1, 2, \cdots, 2^j - 1\}$ as an orthonormal basis in $L^2(\mathbb{T}^n)$. Considering corollary 5 in [23] and definition A.1 in the appendix, we easily obtain the following theorem for the wavelet characterization of the variable-index Besov and Triebel–Lizorkin space on a periodic domain.

Theorem 2.8. *Let $s(\cdot) \in L^\infty \cap C^{\log}_{\mathrm{loc}}(\mathbb{T}^n)$ and $p(\cdot) \in \mathcal{P}^{\log}(\mathbb{T}^n)$. The symbol A stands for B or F, and a symbolizes b or f.*

(i) Let $0 < q \leqslant \infty$ ($p^+ < \infty$ in the F-case) and

$$k > \max(\sigma_p - s^-, s^+) \quad (\sigma_{p,q} \text{ in the } F\text{-case}),$$

where $\sigma_p = n\left(\frac{1}{\min(1,p^-)} - 1\right)$ and $\sigma_{p,q} = n\left(\frac{1}{\min(1,p^-,q)} - 1\right)$. Then, $f \in \mathcal{S}'(\mathbb{T}^n)$ belongs to $A^{s(\cdot)}_{p(\cdot),q}$ if and only if it can be represented as

$$f = \sum_{j=0}^{\infty} \sum_{G \in G^j} \sum_{m \in \mathbb{M}_j} \lambda^j_{Gm} 2^{-j\frac{n}{2}} \tilde{\Psi}^j_{Gm} \quad \text{with } \lambda \in \tilde{a}^{s(\cdot)}_{p(\cdot),q}, \tag{2.5}$$

with $\mathbb{M}_j = \{m : m = 0, 1, 2, \cdots, 2^j - 1\}$ and where the series expansion (2.5) is unconditionally convergent in $\mathcal{S}'(\mathbb{R}^n)$ and in any space $A^{\sigma(\cdot)}_{p(\cdot),q}(\mathbb{T}^n)$, where $\sigma(x) < s(x)$ with $\inf(s(x) - \sigma(x)) > 0$ and $\sigma(x)/s(x) \to 0$ for $|x| \to \infty$. The representation (2.5) is unique, and we have

$$\lambda^j_{Gm} = \lambda^j_{Gm}(f) = 2^{j\frac{n}{2}} \langle f, \tilde{\Psi}^j_{Gm} \rangle$$

and

$$I : f \mapsto \{2^{j\frac{n}{2}} \langle f, \tilde{\Psi}^j_{Gm} \rangle\}$$

is an isomorphic map from $A^{s(\cdot)}_{p(\cdot),q}(\mathbb{T}^n)$ onto $\tilde{a}^{s(\cdot)}_{p(\cdot),q}$. If, in addition, $\max(p^+, q) < \infty$, then $\{\tilde{\Psi}^j_{Gm}\}_{j \in \mathbb{N}_0, G \in G^j, m \in \mathbb{M}_j}$ is an unconditional basis in $A^{s(\cdot)}_{p(\cdot),q}(\mathbb{T}^n)$.

(ii) Let $q(\cdot) \in \mathcal{P}^{\log}(\mathbb{T}^n)$ with $0 < p^- \leqslant p^+ < \infty$, $0 < q^- \leqslant q^+ \leqslant \infty$, and let

$$k > \max(\sigma_{p,q} - s^-, s^+)$$

with $\sigma_{p,q} = n\left(\frac{1}{\min(1,p^-,q^-)} - 1\right)$. Then, $f \in \mathcal{S}'(\mathbb{T}^n)$ belongs to $F^{s(\cdot)}_{p(\cdot),q(\cdot)}(\mathbb{T}^n)$ if and only if it can be represented as (2.5) with $\lambda \in \tilde{f}^{s(\cdot)}_{p(\cdot),q(\cdot)}(\mathbb{T}^n)$ and with unconditional convergence in $\mathcal{S}'(\mathbb{T}^n)$ and in $\lambda \in \tilde{f}^{s(\cdot)}_{p(\cdot),q(\cdot)}(\mathbb{T}^n)$. The representation (2.5) is unique, and we have

$$\lambda^j_{Gm} = \lambda^j_{Gm}(f) = 2^{j\frac{n}{2}} \langle f, \tilde{\Psi}^j_{Gm} \rangle$$

and

$$I : f \mapsto \{2^{j\frac{n}{2}} \langle f, \tilde{\Psi}^j_{Gm} \rangle\}$$

is an isomorphic map from $F^{s(\cdot)}_{p(\cdot),q(\cdot)}(\mathbb{T}^n)$ onto $\tilde{f}^{s(\cdot)}_{p(\cdot),q(\cdot)}(\mathbb{T}^n)$.

At the end of this section, we introduce the following notation, which is used frequently subsequently.

$$\rho_{B^{s(\cdot)}_{q(\cdot)}}(u) = \int_{\mathbb{T}^n} \sum_{j=0}^{\infty} \sum_{G \in G^j} \sum_{m \in \mathbb{M}_j} 2^{jq(x)s(2^{-j}m)} \, |\lambda^j_{Gm}|^{q(x)} \chi_{jm}(x) \mathrm{d}x, \tag{2.6}$$

where λ^j_{Gm} are defined as in theorem 2.8.

3. Bayesian approach with a variable-index Besov prior

In this section, we first explain the meaning of the variable-order Besov prior and prove some important properties about the variable-order Besov prior. Because the variable-order space is much more complex than the usual function space, this is the most difficult part of our work and involves non-trivial generalizations of the previous results. Second, based on the statements on the variable-order Besov prior, we may easily generalize the results in [5] and [8] to the variable-order Besov prior setting. As the generalizations are not difficult, only a sketch of the proof has been given in the appendix.

Before proceeding to the main part, we describe the general setting for the inverse problems under the Bayesian framework. Denote using X, Y separable Banach spaces, equipped with Borel σ-algebra, and let $\mathcal{G}: X \to Y$ be a measurable mapping. We wish to solve the inverse problem of finding u from y where

$$y = \mathcal{G}(u) + \eta \tag{3.1}$$

and $\eta \in Y$ denotes noise. Applying the Bayesian approach to this problem, we let $(u, y) \in X \times Y$ be a random variable and compute $u|y$. We specify the random variable (u, y) as follows:

- **Prior**: $u \sim \mu_0$ measure on X.
- **Noise**: $\eta \sim \mathbb{Q}_0$ measure on Y, and $\eta \perp u$.

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

$$\frac{\mathrm{d}\mathbb{Q}_u}{\mathrm{d}\mathbb{Q}_0}(y) = \exp(-\Phi(u; y)), \tag{3.2}$$

and

$$\int_Y \exp(-\Phi(u; y))\mathbb{Q}_0(\mathrm{d}y) = 1. \tag{3.3}$$

As in [5] and [8], we also make the following assumptions about the potential Φ.

Assumptions 1: Let X and Y be Banach spaces. The function Φ: $X \times Y \to \mathbb{R}$ satisfies the following:

(i) there is an $\alpha_1 > 0$ and, for every $r > 0$, an $M \in \mathbb{R}$, such that for all $u \in X$ and for all $y \in Y$ such that $\|y\|_Y < r$,

$$\Phi(u; y) \geqslant M - \alpha_1 \|u\|_X;$$

(ii) for every $r > 0$, there exists $K = K(r) > 0$ such that for all $u \in X$, $y \in Y$ with $\max\{\|u\|_X, \|y\|_Y\} < r$,

$$\Phi(u; y) \leqslant K;$$

(iii) for every $r > 0$, there exists $L = L(r) > 0$ such that for all $u_1, u_2 \in X$ and $y \in Y$ with $\max\{\|u_1\|_X, \|u_2\|_X, \|y\|_Y\} < r$,

$$|\Phi(u_1; y) - \Phi(u_2; y)| \leqslant L\, \|u_1 - u_2\|_X;$$

(iv) there is an $\alpha_2 > 0$ and, for every $r > 0$, a $C \in \mathbb{R}$ such that for all $y_1, y_2 \in Y$ with $\max\{\|y_1\|_Y, \|y_2\|_Y\} < r$ and for every $u \in X$,

$$|\Phi(u,\, y_1) - \Phi(u,\, y_2)| \leqslant \exp\left(\alpha_2 \, \|u\|_X \, + C\right) \, \|y_1 - y_2\|_Y.$$

Different choices of $\mathbb{Q}_0$ and μ_0 would lead to different Bayesian methods. How to choose $\mathbb{Q}_0$ and μ_0 depends on the problems. In the following parts of this paper, we choose $\mathbb{Q}_0$ to be Gaussian, and in order to give variable-order Besov regularization a meaningful explanation in the statistical world, we need to construct a new probability measure μ_0. The following subsection accomplishes this task by using the wavelet characterization of the variable-order Besov space.

3.1. Variable-order Besov prior

For the reader's convenience, we recall the general setting stated in [10] for our purpose. Denote using J an index set, and let $\{\phi_j\}_{j\in J}$ denote an infinite sequence in the Banach space X, with norm $\|\cdot\|$, of $\mathbb{R}$-valued functions defined on a domain D. For simplicity, we assume $D = \mathbb{T}^n$ to be the n-dimensional torus. We normalize these functions so that $\|\phi_j\| = 1$ for $j \in J$. We also introduce another element $m_0 \in X$, not necessarily normalized to 1. Define the function u by

$$u = m_0 + \sum_{j\in J} u_j \phi_j \tag{3.4}$$

By randomizing $u := \{u_j\}_{j\in J}$, we create real-valued random functions on D. (The extension to $\mathbb{R}^n$-valued random functions is straightforward, but omitted for brevity.) We now define the deterministic sequence $\gamma = \{\gamma_j\}_{j\in J}$ and the i.i.d. random sequence $\xi = \{\xi_j\}_{j\in J}$, and we set $u_j = \gamma_j \xi_j$. We assume that ξ is centered, i.e., that $\mathbb{E}(\xi_1) = 0$. Formally, we see that the average value of u is then m_0 so that this element of X should be thought of as the mean function.

In the following, we take X to be the Hilbert space

$$X := L^2(\mathbb{T}^n) = \left\{ u : \ \mathbb{T}^n \to \mathbb{R} : \ \int_{\mathbb{T}^n} |u(x)|^2 \mathrm{d}x < \infty \right\}$$

of real-valued periodic functions with dimension n with the inner product and norm denoted using $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$, respectively. We then set $m_0 = 0$ and let

$$J = \{G^j,\, \mathbb{M}_j\}_{\, j=0,1,\cdots}$$

with G^j and $\mathbb{M}_j$ defined as in theorem 2.8. Consequently, for any $u \in X$, we have

$$u(x) = \sum_{j=0}^{\infty} \sum_{G\in G^j} \sum_{m\in \mathbb{M}_j} u_{Gm}^j \tilde{\Psi}_{Gm}^j \quad \text{with } u_{Gm}^j = \langle u,\, \tilde{\Psi}_{Gm}^j \rangle, \tag{3.5}$$

where $\tilde{\Psi}_{Gm}^j$ is the wavelet basis stated in theorem 2.8. Given a function $u : \mathbb{T}^n \to \mathbb{R}$ and $\{u_{Gm}^j\}$ as defined in (3.5), we define the Banach space $B_{q(\cdot)}^{t(\cdot)}$ as

$$B_{q(\cdot)}^{t(\cdot)} = \{u : \mathbb{T}^n \to \mathbb{R} : \ \|u\|_{F^{t(\cdot)}_{q(\cdot),q(\cdot)}} < \infty\} \tag{3.6}$$

with

$$\|u\|_{F^{t(\cdot)}_{q(\cdot),q(\cdot)}} = \left\| \left(\sum_{k=0}^{\infty} |(\varphi_k \hat{u})^{\vee} \, 2^{kt(x)}|^{q(\cdot)} \right)^{1/q(\cdot)} \right\|_{L^{q(\cdot)}(\mathbb{T}^n)}, \tag{3.7}$$

where $\varphi_j(x) = \varphi(2^{-j}x)$ and $\varphi(\cdot)$ is a smooth decomposition of unity [2, 43]. By proposition 5.4 in [1], we can define the space $B^{s(\cdot)}_{q(\cdot),q(\cdot)}$ appropriately, which is equivalent to $F^{s(\cdot)}_{q(\cdot),q(\cdot)}$ if $s \in L^\infty$. We do not need to develop a full theory for the space $B^{s(\cdot)}_{q(\cdot),q(\cdot)}$, and we just understand it as $F^{s(\cdot)}_{q(\cdot),q(\cdot)}$. Hence, our space $B^{t(\cdot)}_{q(\cdot)}$ defined in (3.6) is just the usual variable-index Besov space with $p(\cdot) = q(\cdot)$. (Although the space defined in [1] is in the whole space $\mathbb{R}^n$, it can be adapted to the periodic case $\mathbb{T}^n$.)

As in the general setting, we assume that $u^j_{Gm} = \gamma^j_{Gm}\xi^j_{Gm}$ where $\xi = \{\xi^j_{Gm}\}_{j=1,2,\cdots,\infty,G\in G^j,m\in\mathbb{M}_j}$ is an i.i.d. sequence and $\gamma = \{\gamma^j_{Gm}\}_{j=1,2,\cdots,\infty,G\in G^j,m\in\mathbb{M}_j}$ is deterministic. We assume that ξ^j_{Gm} is drawn from the measure centered on $\mathbb{R}$ with density proportional to $\exp\left(-\frac{1}{2}\int_{\mathbb{T}^n}|x|^{q(y)}\kappa(\mathrm{d}y)\right)$ where $1 \leqslant q^- \leqslant q(y) \leqslant q^+ < \infty$ and $\kappa(\cdot)$ is a probability measure. We refer to the measure with the above density as a generalized $q(\cdot)$-exponential distribution. Note that if q is constant, it is just a q-exponential distribution [10]. Hence, our generalized $q(\cdot)$-exponential distribution is a natural extension of the q-exponential distribution and includes the Gaussian and Laplace distributions as special cases. For $s(x) \geqslant s^- > 0$ and $\delta > 0$, we define

$$\gamma^j_{Gm} = 2^{-j(s(2^{-j}m)+n/2-n/q^+)}\left(\frac{1}{\delta}\right)^{1/q^+}. \tag{3.8}$$

We now prove the convergence of the series

$$u^N = \sum_{j=0}^{N}\sum_{G\in G^j}\sum_{m\in\mathbb{M}_j} u^j_{Gm}\tilde{\Psi}^j_{Gm}, \quad u^j_{Gm} = \gamma^j_{Gm}\xi^j_{Gm} \tag{3.9}$$

to the limit function

$$u = \sum_{j=0}^{\infty}\sum_{G\in G^j}\sum_{m\in\mathbb{M}_j} u^j_{Gm}\tilde{\Psi}^j_{Gm}, \quad u^j_{Gm} = \gamma^j_{Gm}\xi^j_{Gm}, \tag{3.10}$$

in an appropriate space. To understand the sequence of functions $\{u^N\}$ fully, we introduce the following function space:

$$L^{q(\cdot)}_{\mathbb{P}}(\Omega; B^{t(\cdot)}_{q(\cdot)}) \coloneqq \{u : \mathbb{T}^n \times \Omega \to \mathbb{R} :\ \exists\ \lambda > 0,\ \rho^E_{B^{t(\cdot)}_{q(\cdot)}}(\lambda u) < \infty\}$$

where

$$\begin{aligned}\rho^E_{B^{t(\cdot)}_{q(\cdot)}}(u) &= \sum_{k=0}^{\infty}\inf\left\{\lambda_k > 0 :\ \int_\Omega \rho_{L^{q(\cdot)}}(u_k\lambda_k^{-1/q(\cdot)}2^{kt(\cdot)})\mathbb{P}(\mathrm{d}\omega) \leqslant 1\right\}\\ &= \int_\Omega\int_{\mathbb{T}^n}\sum_{k=0}^{\infty}2^{kt(x)q(x)}\ |u_k(x,\omega)|^{q(x)}\,\mathrm{d}x\mathbb{P}(\mathrm{d}\omega),\end{aligned} \tag{3.11}$$

with

$$u_k = (\varphi_k\hat{u})^\vee$$

and φ_k defined as in (3.7). As mentioned in section 2.1, if $\rho^E_{B^{t(\cdot)}_{q(\cdot)}}$ is a convex semimodular on $L^{q(\cdot)}_{\mathbb{P}}(\Omega; B^{t(\cdot)}_{q(\cdot)})$, then $L^{q(\cdot)}_{\mathbb{P}}(\Omega; B^{t(\cdot)}_{q(\cdot)})$ is a normed space with the Luxemburg norm given by

$$\|u\|_{L^{q(\cdot)}_{\mathbb{P}}(\Omega;B^{t(\cdot)}_{q(\cdot)})} = \inf\left\{\mu > 0 :\ \rho^E_{B^{t(\cdot)}_{q(\cdot)}}\left(\left(\frac{1}{\mu}\right)u\right) \leqslant 1\right\}. \tag{3.12}$$

In order to preserve the fluency of our statement, we list the proof that $\rho^E_{B^{t(\cdot)}_{q(\cdot)}}$ is a convex semimodular in section A. We clarify the relation for our space $L^{q(\cdot)}_{\mathbb{P}}(\Omega; B^{t(\cdot)}_{q(\cdot)})$ with the usual constant q, t space $L^q_{\mathbb{P}}(\Omega; B^t_q)$ used in [8]. Setting q, t in (3.11) to be constants, we have

$$\begin{aligned}\rho^{E}_{B^t_q}(u) &= \int_{\Omega}\sum_{k=0}^{\infty}2^{ktq}\int_{\mathbb{T}^n}|u_k|^q\mathrm{d}x\mathbb{P}(\mathrm{d}\omega)\\ &= \mathbb{E}(\|u\|^q_{B^t_{q,q}}).\end{aligned}$$

Hence, our variable space is indeed a natural generalization of the usual space $L^q_{\mathbb{P}}(\Omega; B^t_q)$. Define

$$\tilde{b}^{E\,s(\cdot)}_{q(\cdot)} := \{\lambda = \{\lambda^j_{Gm}\}_{j\in\mathbb{N}_0, G\in G^j, m\in\mathbb{M}_j}:\ \|\lambda\|_{\tilde{b}^{E\,s(\cdot)}_{q(\cdot)}} < \infty\}, \tag{3.13}$$

where

$$\|\lambda\|_{\tilde{b}^{E\,s(\cdot)}_{q(\cdot)}} = \left\| \left(\sum_{j=0}^{\infty}\sum_{G\in G^j}\sum_{m\in\mathbb{M}_j}2^{jqs(2^{-j}m)}\mathbb{E}(|\lambda^j_{Gm}|^{q(\cdot)})\chi_{jm}(\cdot)\right)^{1/q(\cdot)}\right\|_{L^{q(\cdot)}(\mathbb{T}^n)},$$

and

$$\mathbb{E}(|\lambda^j_{Gm}|^{q(x)}) = \int_{\mathbb{T}^n}|\lambda^j_{Gm}(\omega)|^{q(x)}\mathbb{P}(\mathrm{d}\omega).$$

With these definitions, before presenting the main results in this section, we need the following lemma, which is proved in section A.

Lemma 3.1. *Let $s(\cdot) \in L^\infty \cap C^{\log}_{\mathrm{loc}}(\mathbb{T}^n)$ and $q(\cdot) \in \mathcal{P}^{\log}(\mathbb{T}^n)$. Let*

$$k > \max(\sigma_q - s^-, s^+),$$

where $\sigma_q = n\left(\frac{1}{\min(1, q^-)} - 1\right)$. Then, $f \in \mathcal{S}'(\mathbb{T}^n)$ belongs to $L^{q(\cdot)}_{\mathbb{P}}(\Omega; B^{s(\cdot)}_{q(\cdot)})$ if and only if it can be represented as

$$f = \sum_{j=0}^{\infty}\sum_{G\in G^j}\sum_{m\in\mathbb{M}_j}\lambda^j_{Gm}2^{-j\frac{n}{2}}\tilde{\Psi}^j_{Gm} \quad \text{with } \lambda \in \tilde{b}^{E\,s(\cdot)}_{q(\cdot)}, \tag{3.14}$$

with $\mathbb{M}_j = \{m : m = 0, 1, 2, \cdots, 2^j - 1\}$ and where the series expansion (3.14) is unconditionally convergent in $\mathcal{S}'(\mathbb{R}^n)$. The representation (3.14) is unique, and we have

$$\lambda^j_{Gm} = \lambda^j_{Gm}(f) = 2^{j\frac{n}{2}}\langle f, \tilde{\Psi}^j_{Gm}\rangle$$

and

$$I : f \mapsto \{2^{j\frac{n}{2}}\langle f, \tilde{\Psi}^j_{Gm}\rangle\}$$

is an isomorphic map from $L^{q(\cdot)}_{\mathbb{P}}(\Omega; B^{s(\cdot)}_{q(\cdot)})$ onto $\tilde{b}^{E\,s(\cdot)}_{q(\cdot)}$.

We can prove the following theorem, which provides a sufficient condition, on $t(\cdot)$, for the existence of the limiting random function.

Theorem 3.2. *For $t, s \in C^{\log}_{\mathrm{loc}}(\mathbb{T}^n) \cap L^\infty(\mathbb{T}^n)$, $q \in \mathcal{P}^{\log}(\mathbb{T}^n)$ and*

$$\sup_{x\in\mathbb{T}^n}\left(t(x) - s(x) + \frac{n}{q^+}\right) < 0,$$

the sequence of functions $\{u^N\}_{N=1}^{\infty}$ given by (3.9) and (3.8) with ξ^j_{Gm} drawn from a centered generalized $q(\cdot)$-exponential distribution is Cauchy in the Banach space $L^{q(\cdot)}_{\mathbb{P}}(\Omega; B^{t(\cdot)}_{q(\cdot)})$.

Thus, the infinite series (3.10) exists as a limit in the space $L_{\mathbb{P}}^{q(\cdot)}(\Omega; B_{q(\cdot)}^{t(\cdot)})$ *for all* $\sup_{x\in\mathbb{T}^n}\left(t(x) - s(x) + \frac{n}{q^+}\right) < 0$.

Proof. By (3.10) and lemma 3.1, we obtain

$$\|u\|_{L_{\mathbb{P}}^{q(\cdot)}(\Omega;B_{q(\cdot)}^{t(\cdot)})} \approx \|\{2^{j\frac{n}{2}}u_{Gm}^{j}\}\|_{\tilde{b}_{q(\cdot)}^{E\,s(\cdot)}} \tag{3.15}$$

For $M > N$, for every $\lambda > 0$, we have the following estimate

$$\begin{aligned}
&\int_{\mathbb{T}^n} \sum_{j=N+1}^{M} \sum_{G\in G^j} \sum_{m\in\mathbb{M}_j} \lambda^{q(x)} 2^{jq(x)(t(2^{-j}m)+n/2)}\, |\gamma_{Gm}^{j}|^{q(x)}\, \mathbb{E}(|\xi_{Gm}^{j}|^{q(x)})\chi_{jm}(x)\mathrm{d}x \\
&\leqslant C\max(\lambda^{q^+}, \lambda^{q^-})\delta^{-1}\int_{\mathbb{T}^n} \sum_{j=N+1}^{M} \sum_{G\in G^j} \sum_{m\in\mathbb{M}_j} 2^{jq(x)(t(2^{-j}m)-s(2^{-j}m)+n/q^+)}\chi_{jm}(x)\mathrm{d}x \qquad (3.16)\\
&\leqslant C\max(\lambda^{q^+}, \lambda^{q^-})\delta^{-1} \sum_{j=N+1}^{M} \sum_{m\in\mathbb{M}_j} 2^{jq^-(t(2^{-j}m)-s(2^{-j}m)+n/q^+)}2^{-jn}
\end{aligned}$$

where we use

$$\begin{aligned}
\mathbb{E}(|\xi_{Gm}^{j}|^{q(x)}) &\leqslant C\int_{\mathbb{R}\cap\{|\xi|>1\}} |\xi|^{q^+}\exp\left(-\frac{1}{2}\int_{\mathbb{T}^n}|\xi|^{q(x)}\kappa(\mathrm{d}x)\mathrm{d}\xi\right)\\
&\quad + C\int_{\mathbb{R}\cap\{|\xi|\leqslant1\}} |\xi|^{q^-}\exp\left(-\frac{1}{2}\int_{\mathbb{T}^n}|\xi|^{q(x)}\kappa(\mathrm{d}x)\mathrm{d}\xi\right)\\
&\leqslant C\int_{\mathbb{R}\cap\{|\xi|>1\}} |\xi|^{q^+}\exp\left(-\frac{1}{2}|\xi|^{q^-}\mathrm{d}\xi\right)\\
&\quad + C\int_{\mathbb{R}\cap\{|\xi|\leqslant1\}} |\xi|^{q^-}\exp\left(-\frac{1}{2}|\xi|^{q^+}\mathrm{d}\xi\right) < \infty.
\end{aligned}$$

The sum on the last line of (3.16) tends to 0 as $N \to \infty$, provided

$$\sup_{x\in\mathbb{T}^n}\left(t(x) - s(x) + \frac{n}{q^+}\right) < 0.$$

Hence, by lemma 2.1.9. in [11], we obtain

$$\lim_{N\to\infty} \|\{2^{j\frac{n}{2}}(u_{Gm}^{N\,j} - u_{Gm}^{M\,j})\}\|_{\tilde{b}_{q(\cdot)}^{E\,s(\cdot)}} = 0.$$

Finally, by (3.15), we complete the proof. □

Remark 3.3. We provide an intuitive meaning for the random series we defined in (3.10). We assume that the probability measure $\kappa(\cdot)$ in the centered generalized $q(\cdot)$-exponential distribution is a uniform measure that is $\kappa(\mathrm{d}x) = \mathrm{d}x$ in $\mathbb{T}^n$. Because $\tilde{\Psi}_{Gm}^{j}$ is an orthonormal basis and

$$u = \sum_{j=0}^{\infty} \sum_{G\in G^j} \sum_{m\in\mathbb{M}_j} u_{Gm}^{j}\tilde{\Psi}_{Gm}^{j} \tag{3.17}$$

with $u_{Gm}^{j}=2^{-j(s(2^{-j}m)+n/2-n/q^{+})}\left(\frac{1}{\delta}\right)^{1/q^{+}}\xi_{Gm}^{j}$, using $\lambda_{Gm}^{j}=2^{j\frac{n}{2}}u_{Gm}^{j}$, we have

$$\begin{aligned}
&\prod_{j=0}^{\infty}\prod_{G\in G^{j}}\prod_{m\in\mathbb{M}_{j}}\exp\left(-\frac{1}{2}\int_{\mathbb{T}^{n}}|\xi_{Gm}^{j}|^{q(x)}\mathrm{d}x\right)\\
&=\prod_{j=0}^{\infty}\prod_{G\in G^{j}}\prod_{m\in\mathbb{M}_{j}}\exp\left(-\frac{1}{2}\int_{\mathbb{T}^{n}}|\gamma_{Gm}^{j}|^{-q(x)}|u_{Gm}^{j}|^{q(x)}\mathrm{d}x\right)\\
&=\prod_{j=0}^{\infty}\prod_{G\in G^{j}}\prod_{m\in\mathbb{M}_{j}}\exp\left(-\frac{1}{2}\int_{\mathbb{T}^{n}}\delta^{\frac{q(x)}{q^{+}}}2^{jq(x)\left(s(2^{-j}m)-\frac{n}{q^{+}}\right)}|\lambda_{Gm}^{j}|^{q(x)}\mathrm{d}x\right)\\
&\leqslant\prod_{j=0}^{\infty}\prod_{G\in G^{j}}\prod_{m\in\mathbb{M}_{j}}\exp\left(-\frac{1}{2}\int_{\mathbb{T}^{n}}\delta^{\frac{q(x)}{q^{+}}}2^{jq(x)\left(s(2^{-j}m)-\frac{n}{q^{+}}\right)}|\lambda_{Gm}^{j}|^{q(x)}2^{jn}\chi_{jm}(x)\mathrm{d}x\right)\\
&\leqslant\exp\left(-\frac{1}{2}\min\left\{\delta^{\frac{q^{-}}{q^{+}}},\,\delta\right\}\rho_{B_{q(\cdot)}^{s(\cdot)}}(u)\right).
\end{aligned}$$

Thus, informally, the Lebesgue density of u can be controlled by a Lebesgue density proportional to $\exp\left(-\frac{1}{2}\min\left\{\delta^{\frac{q^{-}}{q^{+}}},\,\delta\right\}\rho_{B_{q(\cdot)}^{s(\cdot)}}(u)\right)$. Because $\rho_{B_{q(\cdot)}^{s(\cdot)}}(u)$ is related to the space $B_{q(\cdot)}^{s(\cdot)}(\mathbb{T}^{n})$ (theorem 2.8), and the space $B_{q(\cdot)}^{s(\cdot)}(\mathbb{T}^{n})$ is a generalization of a constant-index space, we guess that the Lebesgue density of u is related to a Lebesgue density similar to $\exp\left(-\frac{1}{2}\min\left\{\delta^{\frac{q^{-}}{q^{+}}},\,\delta\right\}\rho_{B_{q(\cdot)}^{s(\cdot)}}(u)\right)$. At least, if q is a function with a constant value, we have an equality that informally means that the Lebesgue density of u is proportional to $\exp\left(-\frac{1}{2}\delta\rho_{B_{q}^{s(\cdot)}}(u)\right)$. Hence, the probability measure we defined may be related to the space $B_{q(\cdot)}^{s(\cdot)}(\mathbb{T}^{n})$. Therefore, we may say that u is distributed according to a $B_{q(\cdot)}^{s(\cdot)}(\mathbb{T}^{n})$ measure with parameter δ, or, briefly, a $(\delta,\,B_{q(\cdot)}^{s(\cdot)}(\mathbb{T}^{n}))$ measure.

Remark 3.4. In this remark, a short verification has been provided for the non-Gaussian natural of our $(\delta,\,B_{q(\cdot)}^{s(\cdot)}(\mathbb{T}^{n}))$ measure. According to the definition of Gaussian measures, which can be found in [43], we know that a probability measure on a separable Banach space being Gaussian indicates that any one-dimensional projection is Gaussian. In our setting, the one-dimensional projection is $u_{Gm}^{j}\coloneqq\gamma_{Gm}^{j}\xi_{Gm}^{j}$ (j, G, m defined as in (3.5)). If u_{Gm}^{j} is drawn from a one-dimensional Gaussian measure, ξ_{Gm}^{j} should also be drawn from a one-dimensional Gaussian measure. However, ξ_{Gm}^{j} is drawn from the generalized $q(\cdot)$-exponential distribution, which is not Gaussian, except for $q(\cdot)$ that is a function with a constant value of 2. Therefore, the $(\delta,\,B_{q(\cdot)}^{s(\cdot)}(\mathbb{T}^{n}))$ measure is non-Gaussian in general, or the variable-index Besov prior is non-Gaussian.

Remark 3.5. We may need to demonstrate how to use our new prior in practical problems. Let the index function $q(\cdot)$ be monotonically decreasing with $q(0)=2$, $q(M)=1$ where M is a sufficiently large number. We consider that some information on the gradient of the target function u (the function we need to reconstruct) can be obtained. We can then specify the index function $q(\cdot)$ as $q(|\nabla u(\cdot)|)$, where ∇u is an estimated quantity of the target function u (using some quick algorithms such as the canny edge detection algorithm in the MATLAB toolbox). Then, our new prior may incorporate more information on the variation of u compared with the Gaussian prior and constant-index Besov prior case. For more information on the practical usage of the prior, we refer to [18]. Although this paper is on variable TV regularization, we do not perceive any conceptual difficulty in shifting to our variable-index Besov case.

Theorem 3.6. *Assume that* u *is given by (3.10) and that (3.8) with* ξ_{Gm}^{j} *for every* $\{j, G, m\}$ *is drawn from a centered generalized* $q(\cdot)$*-exponential distribution with* $\kappa(\mathrm{d}x) = \mathrm{d}x$*; that is to say,* $\kappa(\cdot)$ *is a uniform probability measure. In other words,* u *is distributed according to a* $(\delta, B_{q(\cdot)}^{s(\cdot)})$ *measure. In addition, we assume* $t, s \in C_{\mathrm{loc}}^{\log}(\mathbb{T}^n) \cap L^{\infty}(\mathbb{T}^n)$*,* $q \in \mathcal{P}^{\log}(\mathbb{T}^n)$ *and* $t(x) - s(x) + \frac{n}{q^+} \neq 0$ *for every* $x \in \mathbb{T}^n$*. Then, the following are equivalent:*

1. $\rho_{B_{q(\cdot)}^{t(\cdot)}}(u) < \infty \quad \mathbb{P}-$a.s.;
2. $\mathbb{E}(\exp(\alpha\rho_{B_{q(\cdot)}^{t(\cdot)}}(u))) < \infty$ *for any* $\alpha \in [0, \delta/2)$;
3. $\sup_{x\in\mathbb{T}^n}\left(t(x) - s(x) + \frac{n}{q^+}\right) < 0$.

Proof. (3) $\Rightarrow$ (2).

Because $\sup_{x\in\mathbb{T}^n}\left(t(x) - s(x) + \frac{n}{q^+}\right) < 0$, there exists a negative constant $\beta < 0$ such that $\sup_{x\in\mathbb{T}^n}\left(t(x) - s(x) + \frac{n}{q^+}\right) \leqslant \beta < 0$. Let K be a large enough positive constant; then, we have

$$\begin{aligned}
&\mathbb{E}(\exp(\alpha\rho_{B_{q(\cdot)}^{t(\cdot)}}(u))) \\
&= \prod_{j=0}^{\infty}\prod_{G\in G^j}\mathbb{E}\left(\exp(\alpha\delta^{-1}\sum_{m\in\mathbb{M}_j}\int_{\mathbb{T}^n} 2^{jq(x)(t(2^{-jm})-s(2^{-jm})+n/q^+)}\,|\xi_{Gm}^{j}|^{q(x)}\chi_{jm}(x)\mathrm{d}x)\right) \\
&\leqslant \prod_{j=0}^{\infty}\prod_{G\in G^j}\mathbb{E}\left(\exp\left(\alpha\delta^{-1}\sum_{m\in\mathbb{M}_j}\int_{\mathbb{T}^n} 2^{jq(x)\beta}\,|\xi_{Gm}^{j}|^{q(x)}\chi_{jm}(x)\mathrm{d}x\right)\right) \\
&\leqslant \prod_{j=0}^{\infty}\prod_{G\in G^j}\frac{\int_{\mathbb{R}}\exp\left(\alpha\delta^{-1}\int_{\mathbb{T}^n}2^{jq(x)\beta}\,|\xi|^{q(x)}\mathrm{d}x - \frac{1}{2}\int_{\mathbb{T}^n}|\xi|^{q(x)}\mathrm{d}x\right)\mathrm{d}\xi}{\int_{\mathbb{R}}\exp\left(-\frac{1}{2}\int_{\mathbb{T}^n}|\xi|^{q(x)}\mathrm{d}x\right)\mathrm{d}\xi} \\
&\leqslant \left(\prod_{j=0}^{\infty}\frac{\int_{\mathbb{R}}\exp\left(\alpha\delta^{-1}\int_{\mathbb{T}^n}2^{jq(x)\beta}\,|\xi|^{q(x)}\mathrm{d}x - \frac{1}{2}\int_{\mathbb{T}^n}|\xi|^{q(x)}\mathrm{d}x\right)\mathrm{d}\xi}{\int_{\mathbb{R}}\exp\left(-\frac{1}{2}\int_{\mathbb{T}^n}|\xi|^{q(x)}\mathrm{d}x\right)\mathrm{d}\xi}\right)^{K}.
\end{aligned}$$

If we want to prove that the above infinite product converges, we only need to prove that the following summation converges [25].

$$\begin{aligned}
&\sum_{j=0}^{\infty}\int_{\mathbb{R}}\left(\exp\left(\alpha\delta^{-1}\int_{\mathbb{T}^n}2^{jq(x)\beta}\,|\xi|^{q(x)}\mathrm{d}x\right) - 1\right)\exp\left(-\frac{1}{2}\int_{\mathbb{T}^n}|\xi|^{q(x)}\mathrm{d}x\right)\mathrm{d}\xi \\
&\quad\leqslant \int_{\mathbb{R}}\sum_{j=0}^{\infty}\sum_{k=1}^{\infty}\frac{(\alpha\delta^{-1})^k}{k!}\left(\int_{\mathbb{T}^n}2^{jq(x)\beta}\,|\xi|^{q(x)}\mathrm{d}x\right)^{k}\exp\left(-\frac{1}{2}\int_{\mathbb{T}^n}|\xi|^{q(x)}\mathrm{d}x\right)\mathrm{d}\xi. \qquad (3.18)
\end{aligned}$$

We now concentrate on the first summation term in the integral above.

$$\begin{aligned}
&\sum_{j=0}^{\infty}\sum_{k=1}^{\infty}\frac{(\alpha\delta^{-1})^k}{k!}\left(\int_{\mathbb{T}^n}2^{jq(x)\beta}\,|\xi|^{q(x)}\,\mathrm{d}x\right)^k\\
&=\sum_{k=1}^{\infty}\frac{(\alpha\delta^{-1})^k}{k!}\left(\sum_{j=0}^{\infty}\left(\int_{\mathbb{T}^n}2^{jq(x)\beta}\,|\xi|^{q(x)}\,\mathrm{d}x\right)^k\right)^{\frac{1}{k}k}\\
&\leqslant\sum_{k=1}^{\infty}\frac{(\alpha\delta^{-1})^k}{k!}\left(\int_{\mathbb{T}^n}\left(\sum_{j=0}^{\infty}2^{jq(x)\beta k}\right)^{1/k}\,|\xi|^{q(x)}\,\mathrm{d}x\right)^k\\
&\leqslant\frac{1}{1-2^{\beta q^-}}\sum_{k=1}^{\infty}\frac{(\alpha\delta^{-1})^k}{k!}\left(\int_{\mathbb{T}^n}|\xi|^{q(x)}\,\mathrm{d}x\right)^k\\
&\leqslant\frac{1}{1-2^{\beta q^-}}\exp\left(\alpha\delta^{-1}\int_{\mathbb{T}^n}|\xi|^{q(x)}\,\mathrm{d}x\right).
\end{aligned}$$

Substituting the above inequality into (3.18), we obtain

$$\begin{aligned}
&\sum_{j=0}^{\infty}\int_{\mathbb{R}}\left(\exp\left(\alpha\delta^{-1}\int_{\mathbb{T}^n}2^{jq(x)\beta}\,|\xi|^{q(x)}\,\mathrm{d}x\right)-1\right)\exp\left(-\frac{1}{2}\int_{\mathbb{T}^n}|\xi|^{q(x)}\,\mathrm{d}x\right)\mathrm{d}\xi\\
&\leqslant\frac{1}{1-2^{\beta q^-}}\int_{\mathbb{R}}\exp\left(\left(\alpha\delta^{-1}-\frac{1}{2}\right)\int_{\mathbb{T}^n}|\xi|^{q(x)}\,\mathrm{d}x\right)\mathrm{d}\xi<\infty,
\end{aligned}$$

where we used $\alpha\delta^{-1}<\frac{1}{2}$.

(2) $\Rightarrow$ (1).

If (1) does not hold, $\rho_{B^{t(\cdot)}_{q(\cdot)}}(u)$ is positive infinite on a set S with a positive measure. Then, because for $\alpha>0$, $\exp(\alpha\rho_{B^{t(\cdot)}_{q(\cdot)}}(u))=+\infty$ if $\rho_{B^{t(\cdot)}_{q(\cdot)}}(u)=+\infty$ and

$$\mathbb{E}(\exp(\alpha\rho_{B^{t(\cdot)}_{q(\cdot)}}(u)))\geqslant\mathbb{E}(1_S\exp(\alpha\rho_{B^{t(\cdot)}_{q(\cdot)}}(u))),$$

we get a contradiction.

(1) $\Rightarrow$ (3).

Because $\rho_{B^{t(\cdot)}_{q(\cdot)}}(u)<\infty$, we easily know that

$$\int_{\mathbb{T}^n}\sum_{j=0}^{\infty}\sum_{G\in G^j}\sum_{m\in\mathbb{M}_j}2^{jq(x)(t(2^{-j}m)-s(2^{-j}m)+n/q^+)}\,|\xi^j_{Gm}|^{q(x)}\,\chi_{jm}(x)\mathrm{d}x<+\infty.$$

Hence, for almost all $x\in\mathbb{T}^n$, the integrand in the above formula is finite. Choose $x\in\mathbb{T}^n$ such that

$$\sum_{j=0}^{\infty}\sum_{G\in G^j}\sum_{m\in\mathbb{M}_j}2^{jq(x)(t(2^{-j}m)-s(2^{-j}m)+n/q^+)}\,|\xi^j_{Gm}|^{q(x)}\,\chi_{jm}(x)<\infty.$$

Therefore, for every j, there is $m=m_{x,j}$ such that

$$\sum_{j=0}^{\infty}\sum_{G\in G^j}2^{jq(x)(t(2^{-j}m_{x,j})-s(2^{-j}m_{x,j})+n/q^+)}\,|\xi^j_{Gm_{x,j}}|^{q(x)}<\infty.$$

If $t(2^{-j}m_{x,j})-s(2^{-j}m_{x,j})+n/q^+\geqslant 0$, we have

$$\sum_{j=0}^{\infty}\sum_{G\in G^j}|\xi^j_{Gm_{x,j}}|^{q(x)}<\infty.$$

Because there exist c, C such that $0 < c \leqslant \mathbb{E}(|\xi_{Gm}^{j}|^{q(x)}) \leqslant C < \infty$ for every $\{j, G, m\}$, this contradicts the law of large numbers, which is described in chapter 1 in [38]. Therefore, we obtain $t(2^{-j}m_{x,j}) - s(2^{-j}m_{x,j}) + n/q^{+} < 0$ for infinite j. By the definition of $\chi_{jm}(\cdot)$, we know that $2^{-j}m_{x,j} \to x$; hence, we find that

$$t(x) - s(x) + \frac{n}{q^{+}} \leqslant 0.$$

In addition, by our assumption and the continuity of $t(\cdot)$ and $s(\cdot)$, we finally obtain

$$t(x) - s(x) + \frac{n}{q^{+}} < 0$$

for every $x \in \mathbb{T}^n$. □

Remark 3.7. Theorem 3.6 assumes that $t(x) - s(x) + \frac{n}{q^{+}} \neq 0$ and $\kappa(\cdot)$ is a uniform probability distribution, which seems technically sound; however, for the constant q, t, s case, these conditions are all satisfied naturally. Removing these conditions is left to future work.

In the previous two theorems, we proved basic properties for random variables constructed from infinite series (3.10). We now study a situation where the family $\tilde{\Psi}_{Gm}^{j}$ has a uniform Hölder exponent α and study the implications for the Hölder continuity of the random function u. We assume that there are C, a, $b > 0$ and $\alpha \in (0, 1]$ such that for all $j \geqslant 0$,

$$\begin{aligned} |\tilde{\Psi}_{Gm}^{j}(x)| &\leqslant C2^{jnb}, \quad x \in \mathbb{T}^n, \\ |\tilde{\Psi}_{Gm}^{j}(x) - \tilde{\Psi}_{Gm}^{j}(y)| &\leqslant C2^{jna}|x - y|^{\alpha}, \quad x, y \in \mathbb{T}^n. \end{aligned} \tag{3.19}$$

We also assume that $a > b$ as in [10].

Theorem 3.8. *Assume that u is given by (3.10) and (3.8) with ξ_{Gm}^{j} drawn from a centered generalized $q(\cdot)$-exponential distribution. Suppose also that (3.19) holds and that $s \in C_{\text{loc}}^{\log}(\mathbb{T}^n) \cap L^{\infty}(\mathbb{T}^n)$, $q \in \mathcal{P}^{\log}(\mathbb{T}^n)$, $\inf_{x\in\mathbb{T}^n} s(x) > n\left(b + \frac{1}{q^{+}} + \frac{1}{2}\theta(a - b)\right)$ for some $\theta \in (0, 2)$. Then, $\mathbb{P}$-a.s., we have $u \in C^{\beta}(\mathbb{T}^n)$ for all $\beta < \frac{\alpha\theta}{2}$.*

Proof. We need to use theorem A.1 listed in the appendix, which is a variant of the Kolmogorov continuity theorem. Presented as in theorem A.1 but using our series (3.10), we obtain

$$\begin{aligned} S_1 &= \sum_{j=0}^{\infty}\sum_{G\in G^j}\sum_{m\in\mathbb{M}_j}|\gamma_{Gm}^{j}|^2\|\tilde{\Psi}_{Gm}^{j}\|_{L^\infty}^2 \\ &\leqslant C\sum_{j=0}^{\infty}\sum_{G\in G^j}\sum_{m\in\mathbb{M}_j}2^{-j\left(2s(2^{-jm})+n+\frac{2n}{q^{+}}\right)}2^{2jnb} \\ &\leqslant C\sum_{j=0}^{\infty}2^{jn}2^{-j\left(2s(2^{-jm})+n+\frac{2n}{q^{+}}\right)}2^{2jnb} = C\sum_{j=0}^{\infty}2^{-jnc_1} \end{aligned}$$

and

$$S_2 \leqslant C \sum_{j=0}^{\infty} \sum_{G \in G^j} \sum_{m \in \mathbb{M}_j} |\gamma_{Gm}^j|^{2-\theta} \|\tilde{\Psi}_{Gm}^j\|_{L^\infty}^{2-\theta} |\gamma_{Gm}^j|^{\theta} 2^{jna\theta}$$
$$\leqslant C \sum_{j=0}^{\infty} 2^{-j\left(s(2^{-jm}) - \frac{n}{q^+}\right)} 2^{j2nb} 2^{-j\theta(b-a)n} = C \sum_{j=0}^{\infty} 2^{-jnc_2},$$

where

$$c_1 = \frac{2s(2^{-jm})}{n} - \frac{2}{q^+} - 2b > 0,$$

and

$$c_2 = \frac{2s(2^{-jm})}{n} - \frac{2}{q^+} - 2b - \theta(a-b) > 0.$$

We require $c_1 > 0$ and $c_2 > 0$, and by our assumption $a > b$, we only need $c_2 > 0$. Our assumption $\inf_{x \in \mathbb{T}^n} s(x) > n\left(b + \frac{1}{q^+} + \frac{1}{2}\theta(a-b)\right)$ just ensures that $c_2 > 0$. Therefore, by theorem A.1, we can conclude our proof. □

Remark 3.9. If the mean function is nonzero and satisfies

$$|m_0(x)| \leqslant C, \quad x \in D,$$
$$|m_0(x) - m_0(y)| \leqslant C|x-y|^{\alpha}, \quad x, y \in D,$$

then the result of theorem 3.8 still holds.

Theorem 3.10. *Assume that u is given by (3.10) and (3.8) with ξ_{Gm}^j drawn from a centered generalized $q(\cdot)$-exponential distribution. Suppose also that $\tilde{\Psi}_{Gm}^j$, with $\{j = 0, 1, \cdots, \infty, G \in G^j, m \in \mathbb{M}_j\}$, forms a basis for $B_{q(\cdot)}^{t(\cdot)}$ with $t^- > 0$, $q(\cdot) \in \mathcal{P}^{\log}(\mathbb{T}^n)$ and $t \in C_{\mathrm{loc}}^{\log}(\mathbb{T}^n) \cap L^\infty(\mathbb{T}^n)$. Then, for any*

$$\sup_{x \in \mathbb{T}^n} \left(t(x) - s(x) + \frac{n}{q^+} \right) < 0,$$

we have $u \in C^{t(\cdot)}(\mathbb{T}^n)$ $\mathbb{P}$-a.s.

Proof. For any $k \geqslant 1$, using the definition of $\rho_{B_{q(\cdot)}^{t(\cdot)}}(u)$, we can write

$$\rho_{B_{kq(\cdot)}^{t(\cdot)}}(u) = C_{\delta,m} \int_{\mathbb{T}^n} \sum_{j=0}^{\infty} \sum_{G \in G^j} \sum_{m \in \mathbb{M}_j} 2^{jkq(x)\left(t(2^{-j}m) - s(2^{-j}m) + \frac{n}{q^+}\right)}$$
$$\times |\xi_{Gm}^j|^{kq(x)} \chi_{jm}(x) \mathrm{d}x.$$

For every $k \in \mathbb{N}$, there exist constants c_m, C_m such that $0 < c_m \leqslant \mathbb{E}(|\xi_{Gm}^j|^{q(x)}) \leqslant C_m < \infty$. Because each term of the above series is measurable, we can swap the sum and the integration and obtain

$$\mathbb{E}(\rho_{B_{kq(\cdot)}^{t(\cdot)}}(u)) \leqslant C_{\delta,m} \sum_{j=0}^{\infty} 2^{jq(x)q^-\left(t(2^{-j}m) - s(2^{-j}m) + \frac{n}{q^+}\right)} < \infty.$$

From the above inequality, we obtain $\rho_{B_{kq(\cdot)}^{t(\cdot)}}(u) < \infty$ $\mathbb{P}$-a.s. Therefore, we know that $\|u\|_{B_{kq(\cdot)}^{t(\cdot)}(\mathbb{T}^n)} < \infty$ $\mathbb{P}$-a.s. Because

$$\sup_{x\in\mathbb{T}^n}\left(t(x) - s(x) + \frac{n}{q^+}\right) < 0,$$

we can choose k to be large enough so that $\frac{n}{kq(x)} < s(x) - \frac{n}{q(x)} - t(x)$. Then, the embedding $B^{t_1(\cdot)}_{kq(\cdot)} \hookrightarrow C^{t(\cdot)}(\mathbb{T}^n)$ [1] for any t_1 satisfying $t(x) + \frac{n}{kq(x)} < t_1(x) < s(x) - \frac{n}{q(x)}$ implies that $\|u\|_{C^{t(\cdot)}(\mathbb{T}^n)} < \infty$ $\mathbb{P}$-a.s. It follows that $u \in C^{t(\cdot)}(\mathbb{T}^n)$ $\mathbb{P}$-a.s. □

Remark 3.11. If the mean function m_0 belongs to $C^{t(\cdot)}(\mathbb{T}^n)$, the result of the above theorem holds for a random series with a nonzero mean function as well.

3.2. The Bayesian approach to inverse problems

In the previous subsection, we constructed the probability measure μ_0, which is supported on a given variable-order Besov space $B^{t(\cdot)}_{q(\cdot)}$. We can now present the following theorem for the well-defined problem. The proof is very similar to the one proved in [8]; however, considering that we may need to use some properties of the variable-order Besov space, a sketch of the proof has been given in the appendix.

Theorem 3.12. *Let Φ satisfy (3.3) and Assumptions 1 (i)–(iii). Suppose that for some $q \in \mathcal{P}^{\log}(\mathbb{T}^n)$, $t \in C^{\log}_{\mathrm{loc}}(\mathbb{T}^n) \cap L^\infty(\mathbb{T}^n)$, $B^{t(\cdot)}_{q(\cdot)}$ is continuously embedded in X. There exists $\delta^* > 0$ such that if μ_0 is a $(\delta, B^{s(\cdot)}_{q(\cdot)})$ measure with*

$$\sup_{x\in\mathbb{T}^n}\left(t(x) - s(x) + \frac{n}{q^+}\right) < 0$$

and $\delta > \delta^*$, then μ^y is absolutely continuous with respect to μ_0 and satisfies

$$\frac{\mathrm{d}\mu^y}{\mathrm{d}\mu_0}(u) = \frac{1}{Z(y)}\exp(-\Phi(u; y)) \tag{3.20}$$

with the normalizing factor

$$Z(y) = \int_X \exp(-\Phi(u; y))\mu_0(\mathrm{d}u) < \infty. \tag{3.21}$$

The constant $\delta^* = 2\max\{c_{\mathrm{e}}^{q^-}, c_{\mathrm{e}}^{q^+}\}\alpha_1$, where c_{e} is the embedding constant satisfying $\|u\|_X \leqslant c_{\mathrm{e}}\|u\|_{B^{t(\cdot)}_{q(\cdot)}}$.

Now, we can show the well-posedness of the posterior measure μ^y with respect to the data y. Recall that the Hellinger metric d_{Hell} [15] is defined by

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

where ν is the reference measure with respect to which both μ and μ' are absolutely continuous.

The following theorem can be proved by using similar ideas to those used for theorem 3.3 of [8]. The minor difference is that we need to use $\rho_{B^{t(\cdot)}_{q(\cdot)}}(u)$ instead of $B^{t(\cdot)}_{q(\cdot)}$ when we need to bound $\|u\|_X$. The same situation appears in the proof of theorem 3.12, and therefore, we omit the details of the proof.

Theorem 3.13. *Let Φ satisfy (3.3) and Assumptions 1 (i)–(iv). Suppose that for some $q \in \mathcal{P}^{\log}(\mathbb{T}^n)$, $t \in C^{\log}_{\text{loc}}(\mathbb{T}^n) \cap L^{\infty}(\mathbb{T}^n)$, $B^{t(\cdot)}_{q(\cdot)}$ is continuously embedded in X. There exists $\delta^* > 0$ such that if μ_0 is a $(\delta, B^{s(\cdot)}_{q(\cdot)})$ measure with*

$$\sup_{x \in \mathbb{T}^n} \left(t(x) - s(x) + \frac{n}{q^+} \right) < 0$$

and $\delta > \delta^*$, then

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

where $C = C(r)$ with $\max\{\|y\|_Y, \|y'\|_Y\} \leqslant r$. The constant $\delta^* = 2\max\{c_{\text{e}}^{q^-}, c_{\text{e}}^{q^+}\}(\alpha_1 + 2\alpha_2)$, where c_{e} is the embedding constant satisfying $\|u\|_X \leqslant c_{\text{e}} \, \|u\|_{B^{t(\cdot)}_{q(\cdot)}}$.

For the approximation of the posterior, let Φ^N be an approximation of Φ. Define $\mu^{y,N}$ by

$$\frac{\mathrm{d}\mu^{y,N}}{\mathrm{d}\mu_0}(u) = \frac{1}{Z^N(y)} \exp(-\Phi^N(u)), \tag{3.22}$$

and

$$Z^N(y) = \int_X \exp(-\Phi^N(u)) \mu_0(\mathrm{d}u). \tag{3.23}$$

We do not use the dependence of Φ and Φ^N on y here, as it is considered fixed.

Theorem 3.14. *Assume that the measures μ and μ^N are both absolutely continuous with respect to μ_0 given by (3.20) and (3.22), respectively. Suppose that Φ and Φ^N satisfy Assumption 1 (i) and (ii), uniformly in N, and that there exist $\alpha_3 \geqslant 0$ and $C \in \mathbb{R}$ such that*

$$|\Phi(u) - \Phi^N(u)| \leqslant \exp(\alpha_3 \, \|u\|_X + C)\varphi(N),$$

where $\varphi(N) \to 0$ as $N \to \infty$. Suppose that for some $t \in C^{\log}_{\text{loc}}(\mathbb{T}^n) \cap L^{\infty}(\mathbb{T}^n)$, $q \in \mathcal{P}^{\log}(\mathbb{T}^n)$, $B^{t(\cdot)}_{q(\cdot)}$ is continuously embedded in X. Let μ_0 be a $(\delta, B^{s(\cdot)}_{q(\cdot)})$ measure with

$$\sup_{x \in \mathbb{T}^n} \left(t(x) - s(x) + \frac{n}{q^+} \right) < 0$$

and $\delta > 2\max\{c_{\text{e}}^{q^-}, c_{\text{e}}^{q^+}\}(\alpha_1 + 2\alpha_3)$ where c_{e} is the embedding constant satisfying $\|u\|_X \leqslant c_{\text{e}} \, \|u\|_{B^{t(\cdot)}_{q(\cdot)}}$. Then, there exists a constant independent of N such that

$$d_{\text{Hell}}(\mu, \mu^N) \leqslant C\varphi(N).$$

The proof of the above theorem is similar to the proof of theorem 3.3 of [39], and the differences and difficulties can be overcome by the same idea used in the proof of theorem 3.12.

4. Variational methods

The MAP estimator in Bayesian statistics literature [21] is an important concept. It specifies the relationship between the Bayesian approach and the classical regularization technique. It is well known that for a non-Gaussian prior, we can hardly obtain a rigorous relation between the prior measure and regularization term in infinite dimensions. Even for a simple constant-index Besov prior, theoretical study is not complete [8]. Tapio Helin and Martin Burger [16] addressed this issue in some sense only recently. Here, as stated in remark 3.3, we can get an upper bound for the probability density, and for the constant q case, we can get an equality. This is a more complex situation than the constant-index Besov prior case, and in this section, we provide only a partial illustration.

We define the following functional:

$$I(u) = \Phi(u) + \frac{1}{2}\rho_{B^{s(\cdot)}_{q(\cdot)}}(u). \tag{4.1}$$

Intuitively, the minimizers of the above functional or some variant of (4.1) may have the highest probability measure for a small ball centered on such minimizers. For more explanations, we refer to the Gaussian case [9]. For this functional, we have the following result, which provides an abstract theory for the existence of the MAP estimator, linked in a fundamental way to the natural assumption 1 in section 3, which implies that the posterior measure is well defined and well posed.

Theorem 4.1. *Let Assumptions 1 (i), (ii) hold,* $s \in C^{\log}_{\text{loc}}(\mathbb{T}^n) \cap L^\infty(\mathbb{T}^n)$, $q \in \mathcal{P}^{\log}(\mathbb{T}^n)$, $1 < q^- \leqslant q(x) \leqslant q^+ < \infty$ *and* $B^{s(\cdot)}_{q(\cdot)}$ *be compactly embedded in* X. *Then, there exists* $\bar{u} \in B^{s(\cdot)}_{q(\cdot)}$ *such that*

$$I(\bar{u}) = \bar{I} := \inf\{I(u) : u \in B^{s(\cdot)}_{q(\cdot)}\}.$$

Furthermore, if $\{u_n\}$ *is a minimizing sequence satisfying* $I(u_n) \to I(\bar{u})$, *then there is a subsequence* $\{u_{n'}\}$ *that converges strongly to* $\bar{u}$ *in* $B^{s(\cdot)}_{q(\cdot)}$.

Before proving this theorem, we need to prove the following lemma, which is of independent interest.

Lemma 4.2. *Let* $s \in C^{\log}_{\text{loc}}(\mathbb{T}^n) \cap L^\infty(\mathbb{T}^n)$, $q \in \mathcal{P}^{\log}(\mathbb{T}^n)$, *and* $1 < q^- \leqslant q(x) \leqslant q^+ < \infty$; *then, the dual space of* $B^{s(\cdot)}_{q(\cdot)}$ *is* $B^{-s(\cdot)}_{q(\cdot)'}$, *where*

$$\frac{1}{q(x)} + \frac{1}{q(x)'} = 1.$$

Proof. Step 1. Let $s \in C^{\log}_{\text{loc}} \cap L^\infty$, $q \in \mathcal{P}^{\log}$. We prove in this step that $B^{-s(\cdot)}_{q(\cdot)}(\mathbb{R}^n) \subset B^{s(\cdot)}_{q(\cdot)}(\mathbb{R}^n)'$, where $B^{s(\cdot)}_{q(\cdot)}(\mathbb{R}^n)'$ stands for the dual space of $B^{-s(\cdot)}_{q(\cdot)}(\mathbb{R}^n)$. Let $f \in B^{-s(\cdot)}_{q'(\cdot)}(\mathbb{R}^n)$; define $f_k = \sum_{r=-1}^{1}(\varphi_{k+r}\hat{f})^\vee$, where φ_ℓ is defined as in (3.7) and φ is a smooth decomposition of unity [2, 43]. Then, we know that

$$f = \sum_{k=0}^{\infty}(\varphi_k\hat{f}_k)^\vee \quad \text{in } \mathcal{S}'(\mathbb{R}^n) \tag{4.2}$$

and

$$\begin{aligned}
&\|2^{-s(\cdot)}f_k\|_{L^{q'(\cdot)}(\mathbb{R}^n,\ell^{q'(\cdot)})}\\
&= \inf\left\{\lambda > 0 : \int_{\mathbb{R}^n}\sum_{j=0}^{\infty}\frac{2^{-js(x)q'(x)}}{\lambda^{q'(x)}}\,|(\varphi_j\hat{f}_k)^\vee(x)|^{q'(x)}\,\mathrm{d}x \leqslant 1\right\}\\
&\leqslant C\inf\left\{\frac{\lambda}{2} > 0 : \int_{\mathbb{R}^n}\sum_{j=0}^{\infty}\frac{2^{-js(x)q'(x)}}{\left(\frac{\lambda}{2}\right)^{q'(x)}}\,|(\varphi_k\hat{f})^\vee(x)|^{q'(x)}\,\mathrm{d}x \leqslant 1\right\}\\
&\leqslant C\,\|f\|_{B^{-s(\cdot)}_{q(\cdot)}}.
\end{aligned} \tag{4.3}$$

Take $\varphi \in \mathcal{S}(\mathbb{R}^n)$. We have the following:

$$
\begin{aligned}
|f(\varphi)| &= \left| \sum_{k=0}^{\infty} f(\mathcal{F}(\varphi_k \varphi^{\vee})) \right| \\
&= \left| \sum_{k=0}^{\infty} \sum_{r=-1}^{1} (\varphi_{k+r} \hat{f}_{k+r})^{\vee} (\mathcal{F}(\varphi_k \varphi^{\vee})) \right| \\
&= \left| \sum_{\ell=0}^{\infty} \sum_{r=-1}^{1} \int_{\mathbb{R}^n} 2^{-s(x)\ell} f_\ell(x) 2^{s(x)\ell} \mathcal{F}(\varphi_\ell \varphi_{\ell+r} \varphi^{\vee}) \mathrm{d}x \right| \\
&\leqslant C \, \|2^{-s(\cdot)k} f_k\|_{L^{q'(\cdot)}(\mathbb{R}^n, \ell^{q'(\cdot)})} \|2^{s(\cdot)k} \mathcal{F}(\varphi_k \varphi_{k+r} \varphi^{\vee})\|_{L^{q(\cdot)}(\mathbb{R}^n, \ell^{q(\cdot)})} \\
&\leqslant C \, \|f\|_{B_{q(\cdot)}^{-s(\cdot)}(\mathbb{R}^n)} \|\varphi\|_{B_{q(\cdot)}^{s(\cdot)}(\mathbb{R}^n)}.
\end{aligned}
$$

Hence, we have now proved that $B_{q(\cdot)}^{-s(\cdot)}(\mathbb{R}^n) \subset B_{q(\cdot)}^{s(\cdot)}(\mathbb{R}^n)'$.

Step 2. In this step, we need to prove $B_{q(\cdot)}^{s(\cdot)}(\mathbb{R}^n)' \subset B_{q(\cdot)}^{-s(\cdot)}(\mathbb{R}^n)$. Because $f \in B_{q(\cdot)}^{s(\cdot)}(\mathbb{R}^n) \to \{2^{s(\cdot)k}(\varphi_k \hat{f})^{\vee}\}_{k=0}^{\infty}$ is a one-to-one mapping from $B_{q(\cdot)}^{s(\cdot)}(\mathbb{R}^n)$ onto a subspace of $L^{q(\cdot)}(\mathbb{R}^n, \ell^{q(\cdot)})$, every functional $g \in (B_{q(\cdot)}^{s(\cdot)})'$ can be interpreted as a functional on that subspace. By the Hahn–Banach theorem, g can be extended to a continuous linear functional on $L^{q(\cdot)}(\mathbb{R}^n, \ell^{q(\cdot)})$, where the norm of g is preserved. If $\varphi \in \mathcal{S}(\mathbb{R}^n)$, considering corollary 1 of theorem 8 in Chapter 13 of [13], we have

$$
g(\varphi) = \int_{\mathbb{R}^n} \sum_{k=0}^{\infty} g_k(x) (\varphi_k \hat{\varphi})^{\vee}(x) \mathrm{d}x \tag{4.4}
$$

where $\|2^{-s(\cdot)k} g_k\|_{L^{q'(\cdot)}(\mathbb{R}^n, \ell^{q'(\cdot)})}$ is equivalent to the operator norm of g. (4.4) implies that

$$
g = \sum_{k=0}^{\infty} \mathcal{F}(\varphi_k g_k^{\vee}) \quad \text{in } \mathcal{S}'. \tag{4.5}
$$

Define $\eta_{km}(x) = \frac{2^{nk}}{(1+2^k|x|)^m}$ where m is a large enough constant. As φ can be chosen to be a radial smooth function with compact support, we know that it can be controlled as follows:

$$
2^{kn} \varphi^{\vee}(2^k x) \leqslant C \eta_{km}(x). \tag{4.6}
$$

Now, we have

$$
\begin{aligned}
&\|2^{-s(\cdot)k} (\varphi_k \hat{g})^{\vee}\|_{L^{q'(\cdot)}(\mathbb{R}^n, \ell^{q'(\cdot)})} \\
&= \inf\left\{ \lambda > 0 : \int_{\mathbb{R}^n} \frac{1}{\lambda^{q'(x)}} \sum_{k=0}^{\infty} 2^{-s(x)kq'(x)} |(\varphi_k \hat{g})^{\vee}(x)|^{q'(x)} \mathrm{d}x \leqslant 1 \right\} \\
&= \inf\left\{ \lambda > 0 : \int_{\mathbb{R}^n} \frac{1}{\lambda^{q'(x)}} \sum_{k=0}^{\infty} 2^{-s(x)kq'(x)} \left| \sum_{r=-1}^{1} (\varphi_k \varphi_{k+r} g_{k+r}^{\vee}(-\cdot))^{\vee}(x) \right|^{q'(x)} \mathrm{d}x \leqslant 1 \right\} \\
&\leqslant C \, \|2^{-s(\cdot)k} \varphi_k^{\vee} * g_k\|_{L^{q'(\cdot)}(\mathbb{R}^n, \ell^{q'(\cdot)})} \\
&\leqslant C \, \|2^{-s(\cdot)k} \eta_{km} * g_k\|_{L^{q'(\cdot)}(\mathbb{R}^n, \ell^{q'(\cdot)})} \\
&\leqslant C \, \|2^{-s(\cdot)k} g_k\|_{L^{q'(\cdot)}(\mathbb{R}^n, \ell^{q'(\cdot)})}
\end{aligned}
$$

where the last inequality follows from lemma 5.4 of [12]. (We need a small modification of lemma 5.4 of [12]; however, the modification is straightforward, and therefore, we omit it.) With the above estimates, the proof is completed. □

Although the above proof applies to the whole space $\mathbb{R}^n$, it is also valid for the periodic domain $\mathbb{T}^n$. We return to the proof of theorem 4.1.

Proof. For any $\delta > 0$, there is $N = N_1(\delta)$ such that

$$\bar{I} \leqslant I(u_n) \leqslant \bar{I} + \delta, \quad \forall n \geqslant N_1.$$

Thus,

$$\frac{1}{2}\rho_{B_{q(\cdot)}^{s(\cdot)}}(u_n) \leqslant \bar{u} + \delta \quad \forall n \geqslant N_1.$$

The sequence $\{u_n\}$ is bounded in $B_{q(\cdot)}^{s(\cdot)}$. By the above lemma 4.2, we know that $B_{q(\cdot)}^{s(\cdot)}$ is reflexive and that there exists $\bar{u} \in B_{q(\cdot)}^{s(\cdot)}$ such that $u_n \rightharpoonup \bar{u}$ in $B_{q(\cdot)}^{s(\cdot)}$. By the compact embedding of $B_{q(\cdot)}^{s(\cdot)}$ in X, we deduce that $u_n \to \bar{u}$, strongly in X. Notice that $\rho_{B_{q(\cdot)}^{s(\cdot)}}(u)$ is lower semi-continuous by theorem 2.2.8 of [11]. We can use similar ideas to those used for theorem 2.7 in [5] to complete the proof. □

5. Application to the backward diffusion problem

In this section, our theory is applied to backward diffusion problems for integer-order equations and fractional-order equations.

5.1. Integer-order diffusion equation

For simplicity, only the periodic domain $\mathbb{T}^n$ is considered. Define the operator A as follows:

$$H = (L^2(\mathbb{T}^n), \langle \cdot, \cdot \rangle, \|\cdot\|)$$
$$A = -\Delta, \quad \mathcal{D}(A) = H^2(\mathbb{T}^n).$$

Consider the diffusion equation on $\mathbb{T}^n$ with periodic boundary conditions as an ordinary differential equation in H:

$$\frac{\mathrm{d}}{\mathrm{d}t}v + Av = 0, \quad v(0) = u. \tag{5.1}$$

Define $G(u) = \mathrm{e}^{-A}u$, ℓ to be an operator defined as follows

$$\ell(G(u)) = (G(u)(x_1), G(u)(x_2), \cdots, G(u)(x_K))^T \tag{5.2}$$

where K is a fixed constant. Then, we have the relationship

$$y = \ell(G(u)) + \eta \tag{5.3}$$

where $\eta = \{\eta_j\}_{j=1}^K$ is a mean zero Gaussian with covariance Γ and $y = \{y_j\}_{j=1}^K$ are the data that we measured. We can show the well-definedness of the posterior measure and its continuity with respect to the data for the above inverse diffusion problem.

Theorem 5.1. *Consider the inverse problem for finding u from noisy observations of $G(u) = v(1, \cdot)$ in the form of (5.3). Let μ_0 be distributed as a variable-index Besov prior $(\delta, B_{q(\cdot)}^{s(\cdot)})$ with $s \in C_{\mathrm{loc}}^{\log}(\mathbb{T}^n) \cap L^\infty(\mathbb{T}^n)$, $q \in \mathcal{P}^{\log}(\mathbb{T}^n)$, $\inf_{x\in\mathbb{T}^n} s(x) > \frac{n}{q^+}$ and $\delta > 4$. Then, the measure $\mu^y(\mathrm{d}u)$ is absolutely continuous with respect to μ_0 with the Radon–Nikodym derivative satisfying*

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

where

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

and

$$Z(y) = \int_{B_{q(\cdot)}^{t(\cdot)}} \exp\left(-\frac{1}{2} |\Gamma^{-1/2}(y - \ell(G(u)))|^2 + \frac{1}{2} |\Gamma^{-1/2} y|^2\right) \mu_0(du)$$

with $\sup_{x\in\mathbb{T}^n}\left(t(x) - s(x) + \frac{n}{q^+}\right) < 0$. *Furthermore, the posterior measure is continuous in the Hellinger metric with respect to the data*

$$d_{\mathrm{Hell}}(\mu^y, \mu^{y'}) \leqslant C|y - y'|.$$

Proof. First, we prove two properties of the operator $\ell(G(\cdot))$.

Property 1: For large enough $\ell > 0$ and a small constant $\epsilon > 0$, by the Sobolev embedding theorem, we have

$$\begin{aligned} |\ell(G(u))| &\leqslant K\, \|G(u)\|_{L^\infty} = \|e^{-A}u\|_{L^\infty} \\ &\leqslant \|A^\ell e^{-A}u\|_{L^2} + \|e^{-A}u\|_{L^2} \\ &\leqslant C\, \|u\|_{B_{2,2}^{t^- - \frac{n}{q^-} + \frac{n}{2} - \epsilon}} \\ &\leqslant C\, \|u\|_{B_{q(\cdot)}^{t(\cdot)}} \end{aligned} \tag{5.4}$$

where we used the fact that $A^\gamma e^{-\lambda A}$, $\lambda > 0$, is a bounded linear operator from $B_{2,2}^a$ to $B_{2,2}^b$, with any $a, b, \gamma \in \mathbb{R}$. We also used embedding theorems for the variable-index Besov space [1].

Property 2: Let u_1, u_2 be two different initial data points for the diffusion equations. Similar to the proof of Property 1, we have

$$\begin{aligned} |\ell(G(u_1)) - \ell(G(u_2))| &\leqslant K\, \|G(u_1) - G(u_2)\|_{L^\infty} \\ &= K\, \|G(u_1 - u_2)\|_{L^\infty} \\ &\leqslant C\, \|u_1 - u_2\|_{B_{q(\cdot)}^{t(\cdot)}}. \end{aligned} \tag{5.5}$$

Let $X = B_{q(\cdot)}^{t(\cdot)}$. With Property 1 and Property 2, it is straightforward that $\Phi(u; y)$ satisfies Assumption 1 (i)–(iv) with $\alpha_1 = 0$ and $\alpha_2 = 1$. By theorem 3.12 and theorem 3.13, we immediately obtain our desired results. □

5.2. Fractional-order diffusion equation

There is a vast amount of literature on fractional-order diffusion equations. For the well-posedness theory, we refer to [19, 32–34]. We treat fractional diffusion equations on the periodic domain as follows:

$$\begin{aligned} &\partial_t^\alpha v(t, x) + (-\Delta)^\beta v(t, x) = 0, \quad t \geqslant 0, x \in \mathbb{T}^n, \\ &v(0) = u, \end{aligned} \tag{5.6}$$

where $0 < \alpha \leqslant 1$ and $0 < \beta \leqslant 1$ and ∂_t^α stands for the Caputo derivative of the α order, which can be defined as follows

$$\partial_t^\alpha f(t) := \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-s)^{-\alpha} f'(s) \mathrm{d}s,$$

where $\Gamma(\cdot)$ is the usual Gamma function. Define the operator A as follows:

$$H = (L^2(\mathbb{T}^n), \langle \cdot, \cdot \rangle, \|\cdot\|)$$
$$A = (-\Delta)^\beta, \quad \mathcal{D}(A) = H^{2\beta}(\mathbb{T}^n).$$

Consider the heat conduction equation on $\mathbb{T}^n$ with periodic boundary conditions as an ordinary differential equation in H:

$$\partial_t^\alpha v + Av = 0, \quad v(0) = u. \tag{5.7}$$

If A is a bounded operator, e.g., a positive number, then the solution of the above equation (5.7) has the following form:

$$v(t) = E_\alpha(-At^\alpha)u,$$

where $E_\alpha(\cdot)$ is the Mittag-Leffler function defined as

$$E_\alpha(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\alpha k + 1)}.$$

For more properties of the Mittag–Leffer function, we refer to [24, 36]. [35] proposed a fractional operator semigroup that characterizes the solution of the abstract fractional Cauchy problem (5.7). Because operator A in our setting can generate a fractional operator semigroup, we can define $G(u) = E_\alpha(-A)u$, ℓ to be an operator defined as follows:

$$\ell(G(u)) = (G(u)(x_1), G(u)(x_2), \cdots, G(u)(x_K))^T \tag{5.8}$$

where K is a fixed constant. Then, we have the relationship

$$y = \ell(G(u)) + \eta \tag{5.9}$$

where $\eta = \{\eta_j\}_{j=1}^K$ is a mean zero Gaussian with covariance Γ and $y = \{y_j\}_{j=1}^K$ are the data we measured. Reviewing the proof of theorem 5.1, the key points are the estimates of the operator $\ell(G(\cdot))$. In more complicated situations, fractional diffusion equations do not have the strong smoothing effects that normal diffusion equations have. For a more complete illustration, we refer to [20]. The Mittag–Leffer function appears naturally in fractional diffusion equations and only has a polynomial decay rate, which restricts the smoothing effect. More precisely, we list the following decay rate estimates.

Lemma 5.2. *[36] If $0 < \alpha < 2$, μ is an arbitrary real number such that*

$$\frac{\pi\alpha}{2} < \mu < \min\{\pi, \pi\alpha\}.$$

Then, for an arbitrary integer $p \geqslant 1$, when $|z| \to \infty$, the following expansion holds:

$$E_\alpha(z) = \frac{1}{\alpha} \mathrm{e}^{z^{1/\alpha}} - \sum_{k=1}^{p} \frac{z^{-k}}{\Gamma(\beta - \alpha k)} + O(|z|^{-1-p}),$$

where $|\arg(z)| \leqslant \mu$.

Based on the above observation, we must restrict the fractional-order α, β to some appropriate interval to gain enough smoothing effects so that the forward operator is Lipschitz continuous. More precisely, we can obtain the following result.

Theorem 5.3. *Assume the dimension $n \leqslant 3$. Consider the inverse problem for determining u from noisy observations of $G(u) = v(1, \cdot)$ in the form of (5.9) with $0 < \alpha \leqslant 1$ and $\frac{n}{4} < \beta \leqslant 1$. Let $X = L^2(\mathbb{T}^n)$, μ_0 be distributed as a variable-index Besov prior $(\delta, B^{s(\cdot)}_{q(\cdot)})$ with $s \in C^{\log}_{\text{loc}}(\mathbb{T}^n) \cap L^\infty(\mathbb{T}^n)$, $q \in \mathcal{P}^{\log}(\mathbb{T}^n)$, $\inf_{x\in\mathbb{T}^n} s(x) > \frac{n}{q^+}$ and $\delta > 4$. Assume $t \in C^{\log}_{\text{loc}}(\mathbb{T}^n) \cap L^\infty(\mathbb{T}^n)$ and*

$$\frac{n}{q^-} - \frac{n}{2} < t^- \leqslant t^+ < s^- - \frac{n}{q^+}.$$

Then, the measure $\mu^y(\mathrm{d}u)$ is absolutely continuous with respect to μ_0 with the Radon–Nikodym derivative satisfying

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

where

$$\Phi(u; y) = \frac{1}{2}\,|\Gamma^{-1/2}(y - \ell(G(u)))|^2 - \frac{1}{2}\,|\Gamma^{-1/2}y|^2$$

and

$$Z(y) = \int_X \exp\left(-\frac{1}{2}\,|\Gamma^{-1/2}(y - \ell(G(u)))|^2 + \frac{1}{2}\,|\Gamma^{-1/2}y|^2\right)\mu_0(\mathrm{d}u).$$

Furthermore, the posterior measure is continuous in the Hellinger metric with respect to the data

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

Proof. In order to prove the above theorem, we first provide the following estimates. Let f_ℓ ($\ell \in Z^n$) be the Fourier coefficient of function f. Then, we have

$$\begin{aligned} \|AE_\alpha(-A)f\|^2_{L^2(\mathbb{T}^n)} &= \sum_{\ell\in Z^n} (|\ell|^{2\beta} E_\alpha(-|\ell|^{2\beta}))^2 |f_\ell|^2 \\ &\leqslant C\,\|f\|^2_{L^2(\mathbb{T}^n)}, \end{aligned} \tag{5.10}$$

where we used lemma 5.2. Considering $\frac{n}{q^-} - \frac{n}{2} < t^- \leqslant t^+ < s^- - \frac{n}{q^+}$, for an arbitrary small positive number $\epsilon > 0$, we know that

$$B^{t(\cdot)}_{q(\cdot)} \hookrightarrow B^{t^- - \frac{n}{q^-} + \frac{n}{2} - \epsilon}_{2,2} \hookrightarrow X.$$

Using (5.10), we easily have

$$\begin{aligned} |\ell(G(u))| &\leqslant K\,\|G(u)\|_{L^\infty} = \|E_\alpha(-A)u\|_{L^\infty} \\ &\leqslant \|AE_\alpha(-A)u\|_{L^2} + \|E_\alpha(-A)u\|_{L^2} \\ &\leqslant C\,\|u\|_{L^2}, \end{aligned} \tag{5.11}$$

where we used $\beta > \frac{n}{4}$ to obtain the first inequality. Similarly, we can obtain

$$|\ell(G(u_1)) - \ell(G(u_2))| \leqslant C\,\|u_1 - u_2\|_{L^2}.$$

At this stage, we can complete the proof easily, similar to the integer case. □

6. Conclusion

In this paper, the $(\delta, B_{q(\cdot)}^{s(\cdot)})$ measure has been constructed by using wavelet representations for the function space on a periodic domain. Roughly speaking, it can be seen as a counterpart of variable regularization terms. Using the new non-Gaussian measure as our prior measure, we establish the 'well-posedness' theory for inverse problems, similar to the work of [10]. The new non-Gaussian prior measure we provide may lead to better understanding of variable-order space regularization terms.

Our theory has been used for integer- and fractional-order backward diffusion problems. In particular, the 'well-posedness' theory for the fractional-order backward diffusion problem under a Bayesian inverse framework has been constructed, provided we restrict the time derivative to (0, 1] and the space derivative to $\left(\frac{n}{2}, 2\right]$ (n is the space dimension). Our study also reflects that fractional-order problems are not a straightforward generalization of integer-order problems. For fractional-order problems, fractional-order equations have completely different regularization properties compared to integer-order equations.

Acknowledgments

The authors would like to thank the anonymous referees for their comments and suggestions, which helped to improve the paper significantly. J Gao was supported partially by the National Natural Science Foundation of China under grant no. 41390454. J Jia was supported by the National Natural Science Foundation of China under grant no. 11501439 and the Postdoctoral Science Foundation Project of China under grant no. 2015M580826. J Peng was supported partially by the National Natural Science Foundation of China under grant no. 11131006, 41390454 and 91330204.

Appendix

A.1. Properties of $\rho^E_{B^{t(\cdot)}_{q(\cdot)}}$ appearing in section 3

Lemma A.1. *Let $1 \leqslant q^- \leqslant q(\cdot) \leqslant q^+ < \infty$ and $t(\cdot) \in C(\mathbb{T}^n)$. Then, $\rho^E_{B^{t(\cdot)}_{q(\cdot)}}$ is modular and continuous.*

Proof. Properties (1) and (2) in definition 2.1 are obviously satisfied. To prove (3), we suppose that

$$\rho^E_{B^{t\cdot}_{q(\cdot)}}(\lambda u) = 0$$

for all $\lambda > 0$. Clearly, for some k_0,

$$\int_\Omega \int_{\mathbb{T}^n} \lambda^{q(x)} 2^{k_0 t(x) q(x)} \, |u_{k_0}(x, \omega)|^{q(x)} \mathrm{d}x \mathbb{P}(\mathrm{d}\omega) \leqslant \rho^E_{B^{t\cdot}_{q(\cdot)}}(\lambda u) = 0$$

Because $1 \leqslant q^- \leqslant q(\cdot) \leqslant q^+ < \infty$ and $t(\cdot) \in C(\mathbb{T}^n)$, we easily obtain that $u_{k_0} = 0$. Hence, we obtain $u = 0$. Let $\mu \to 1$. We need to prove $\rho^E_{B^{t(\cdot)}_{q(\cdot)}}(\lambda u) \to \rho^E_{B^{t(\cdot)}_{q(\cdot)}}(u)$. Fix $\epsilon > 0$, choose $N > 0$, and let $\mu < 1$ and close enough to 1 such that

$$\rho^E_{B^{t(\cdot)}_{q(\cdot)}}(u) < \int_\Omega \int_{\mathbb{T}^n} \sum_{k=0}^{N} 2^{kt(x)q(x)} \, |u_k(x, \omega)|^{q(x)} \mathrm{d}x \mathbb{P}(\mathrm{d}\omega) + \epsilon$$
$$< \int_\Omega \int_{\mathbb{T}^n} \sum_{k=0}^{N} \lambda^{q(x)} 2^{kt(x)q(x)} \, |u_k(x, \omega)|^{q(x)} \mathrm{d}x \mathbb{P}(\mathrm{d}\omega) + 2\epsilon$$
$$< \rho^E_{B^{t(\cdot)}_{q(\cdot)}}(\lambda u) + 2\epsilon.$$

Hence, we find that $\rho^E_{B^{t(\cdot)}_{q(\cdot)}}(\lambda u)$ is left continuous with respect to λ. We can similarly show that it is right continuous. □

Lemma A.2. *Let $q \in \mathcal{P}$. Then, $\rho^E_{B^{t(\cdot)}_{q(\cdot)}}$ is convex.*

Proof. Let $\theta \in (0, 1)$. Then,

$$\rho^E_{B^{t(\cdot)}_{q(\cdot)}}(\theta f + (1-\theta)g) = \int_\Omega \int_{\mathbb{T}^n} \sum_{k=0}^{\infty} 2^{kt(x)q(x)} \, |\theta f_k(x, \omega) + (1-\theta) g_k(x, \omega)|^{q(x)} \mathrm{d}x \mathbb{P}(\mathrm{d}\omega)$$
$$\leqslant \rho^E_{B^{t(\cdot)}_{q(\cdot)}}(\theta f) + \rho^E_{B^{t(\cdot)}_{q(\cdot)}}((1-\theta)g)$$
$$\leqslant \theta \rho^E_{B^{t(\cdot)}_{q(\cdot)}}(f) + (1-\theta)\rho^E_{B^{t(\cdot)}_{q(\cdot)}}(g).$$

□

In our case, the parameter $p(\cdot)$ in [1] is equal to $q(\cdot)$; therefore, we only need $1 \leqslant q^-$ not $2 \leqslant q^-$ as indicated by theorem 3.6 in [1].

A.2. Proof of lemma 3.1

For the proof of lemma 3.1, we provide the following two important lemmas.

Lemma A.3. *Let $1 \leqslant q^- \leqslant q(\cdot) \leqslant q^+ < \infty$, $\delta > 0$. For any sequence $\{g_j\}_{j=0}^\infty$ of non-negative measurable functions on $\mathbb{T}^n$, assume the following:*

$$G_j(x, \omega) = \sum_{k=0} 2^{-|k-j|\delta} g_k(x, \omega).$$

Then,

$$\|\{G_j\}_{j=0}^\infty\|_{L^{q(\cdot)}(\ell_E^{q(\cdot)})} \leqslant C \, \|\{g_j\}_{j=0}^\infty\|_{L^{q(\cdot)}(\ell_E^{q(\cdot)})},$$

where

$$\|\{g_j\}_{j=0}^\infty\|_{L^{q(\cdot)}(\ell_E^{q(\cdot)})} = \left\| \left(\sum_{j=0}^{\infty} \mathbb{E}(g_j(\cdot)^{q(\cdot)}) \right)^{1/q(\cdot)} \right\|_{L^{q(\cdot)}(\mathbb{T}^n)},$$

and

$$\mathbb{E}(g_j(x)^{q(x)}) = \int_\Omega (g_j(x, \omega))^{q(x)} \mathbb{P}(\mathrm{d}\omega).$$

Proof. It is obvious that we only need to provide the following estimates:

$$\left(\sum_j \mathbb{E}(g_j(x)^{q(x)})\right)^{1/q(x)} \leqslant \left(\int_\Omega \left(\sum_{j=0}^{\infty} 2^{-|j|\delta}\right)^{q(x)} \sum_{j=0}^{\infty} |g_j(x,\,\omega)|^{q(x)}\,\mathbb{P}(\mathrm{d}\omega)\right)^{1/q(x)}$$
$$\leqslant C\left(\sum_{j=0}^{\infty}\int_\Omega |g_j(x,\,\omega)|^{q(x)}\,\mathbb{P}(\mathrm{d}\omega)\right)^{1/q(x)}$$
$$\leqslant C\left(\sum_{j=0}^{\infty}\mathbb{E}(g_j(x))^{q(x)}\right)^{1/q(x)}.$$

□

Lemma A.4. *Let $q(\cdot) \in C^{\log}(\mathbb{T}^n)$ with $1 < q^- \leqslant q(\cdot) \leqslant q^+ < \infty$. Then, the inequality*

$$\|\{\eta_{\nu,R} * f_j\}_{j\in\mathbb{N}_0}\|_{L^{q(\cdot)}(\ell_E^{q(\cdot)})} \leqslant C\,\|\{f_j\}_{j\in\mathbb{N}_0}\|_{L^{q(\cdot)}(\ell_E^{q(\cdot)})}$$

holds for every sequence $\{f_j(x,\,\omega)\}_{j\in\mathbb{N}_0}$ of L^1_{loc}-functions for variable x and $\mathbb{P}$-measurable functions for variable ω.

Proof. The proof of this lemma is similar to the proof of lemma 5.4 in [12]. Here, we only provide the difference. Let $\mathcal{D}_{\mathrm{i}}$ denote all dyadic cubes with side length $2^{-\mathrm{i}}$ and $\eta_{\nu m}(x) = \frac{2^{n\nu}}{(1+2^\nu|x|)^m}$. As in [12], we require the following estimate:

$$\int_{\mathbb{T}^n}\sum_{\nu=0}^{\infty}\int_\Omega |\eta_{\nu m} * f_\nu|^{q(x)}\,\mathbb{P}(\mathrm{d}\omega)\mathrm{d}x$$
$$\leqslant \int_{\mathbb{T}^n}\sum_{\nu=0}^{\infty}\left(\int_\Omega \sum_{j=0}^{\infty} 2^{-j(m-n)} \sum_{Q\in\mathcal{D}_{\nu-j}} \chi_{3Q}(x) M_Q f_\nu\,\mathbb{P}(\mathrm{d}\omega)\right)^{q(x)}\mathrm{d}x$$
$$\leqslant C\int_{\mathbb{T}^n}\sum_{\nu=0}^{\infty}\int_\Omega \sum_{j=0}^{\infty} 2^{-j(m-n)} \sum_{Q\in\mathcal{D}_{\nu-j}} \chi_{3Q}(x)(M_Q(|f_\nu|^{q(x)/q^-}))^{q^-}\mathbb{P}(\mathrm{d}\omega)\mathrm{d}x + II$$
$$= I + II,$$

where *II* is exactly the same as in the proof for lemma 5.4 in [12]. Next, we only provide an estimate for the term *I*.

$$I \leqslant C\int_{\mathbb{T}^n}\sum_{\nu=0}^{\infty}\int_\Omega (M(|f_\nu|^{q(x)/q^-}))^{q^-}\mathbb{P}(\mathrm{d}\omega)\sum_{j=0}^{\infty}2^{-j(m-n)}\sum_{Q\in\mathcal{D}_{\nu-j}}\chi_{3Q}(x)\mathrm{d}x$$
$$\leqslant C\int_{\mathbb{T}^n}\sum_{\nu=0}^{\infty}\mathbb{E}(|f_\nu(x)|^{q(x)})\mathrm{d}x < \infty.$$

With these estimates, it is easy to recover the whole proof. □

With the two lemmas, following the proofs of theorems 4.4 and 4.5, we can obtain the local mean characterizations of $L^{q(\cdot)}_{\mathbb{P}}(\Omega;\,B^{s(\cdot)}_{q(\cdot)})$ by using our lemma A.3 and lemma A.4 instead of lemmas 4.2 and 4.3 in [22]. Replacing lemmas 5 and 9 in [23] by our lemmas A.3 and A.4, we can use the proofs for corollarys 2 and 3 in [23] to provide the proof for lemma 3.1. As the proof is very long and does not involve any new component other than lemmas A.3 and A.4, we omit it here.

A.3. Proof of theorem 3.12

In the following, a short proof of theorem 3.12 is given.

Proof. Define $\pi_0(\mathrm{d}u, \mathrm{d}y) = \mu_0(\mathrm{d}u) \otimes \mathbb{Q}_0(\mathrm{d}y)$ and $\pi(\mathrm{d}u, \mathrm{d}y) = \mu_0(\mathrm{d}u)\mathbb{Q}_u(\mathrm{d}y)$. Assumption 1 (iii) provides the continuity of Φ on X, and because $\mu_0(X) = 1$, we find that $\Phi : X \to \mathbb{R}$ is μ_0-measurable. Therefore, $\pi \ll \pi_0$ and π have a Radon–Nikodym derivative given by (3.2). Theorem 6.29 of [39] implies that $\mu^y(\mathrm{d}u)$ is absolutely continuous with respect to $\mu_0(\mathrm{d}u)$. This same lemma also gives (3.20) provided that the normalization constant (3.21) is positive, which we now establish. Because $\mu_0(B^{t(\cdot)}_{q(\cdot)}) = 1$, we note that all the integrals over X may be replaced by integrals over $B^{t(\cdot)}_{q(\cdot)}$ for any $\sup_{x\in\mathbb{T}^n}\left(t(x) - s(x) + \frac{n}{q^+}\right) < 0$. By Assumption 1 (i), we note that there is $M = M(y)$ such that

$$
\begin{aligned}
Z(y) &= \int_{B^{t(\cdot)}_{q(\cdot)}} \exp(-\Phi(u; y))\mu_0(\mathrm{d}u) \\
&\leqslant \int_{B^{t(\cdot)}_{q(\cdot)}} \exp(\alpha_1 \|u\|_X - M)\mu_0(\mathrm{d}u)
\end{aligned}
$$

By lemma 3.2.5 of [11], we know that

$$
\|u\|_X^{q^-} \leqslant c_{\mathrm{e}}^{q^-} \rho_{B^{t(\cdot)}_{q(\cdot)}}(u) \tag{A.1}
$$

when $\rho_{B^{t(\cdot)}_{q(\cdot)}}(u) \geqslant 1$ or

$$
\|u\|_X^{q^+} \leqslant c_{\mathrm{e}}^{q^+} \rho_{B^{t(\cdot)}_{q(\cdot)}}(u) \tag{A.2}
$$

when $\rho_{B^{t(\cdot)}_{q(\cdot)}}(u) < 1$. Hence, if $\|u\|_X \geqslant 1$, then we have

$$
\begin{aligned}
Z(y) &= \int_{B^{t(\cdot)}_{q(\cdot)}} \exp(-\Phi(u; y))\mu_0(\mathrm{d}u) \\
&\leqslant \int_{B^{t(\cdot)}_{q(\cdot)}} \exp(\alpha_1 \max\{c_{\mathrm{e}}^{q^-}, c_{\mathrm{e}}^{q^+}\}\rho_{B^{t(\cdot)}_{q(\cdot)}}(u) - M)\mu_0(\mathrm{d}u).
\end{aligned}
$$

This upper bound is finite by theorem 3.6 because $\delta > 2\alpha_1 \max\{c_{\mathrm{e}}^{q^-}, c_{\mathrm{e}}^{q^+}\}$. For $\|u\|_X \leqslant 1$, we have

$$
\begin{aligned}
Z(y) &= \int_{B^{t(\cdot)}_{q(\cdot)}} \exp(-\Phi(u; y))\mu_0(\mathrm{d}u) \\
&\leqslant \int_{B^{t(\cdot)}_{q(\cdot)}} \exp(\alpha_1 - M)\mu_0(\mathrm{d}u) < \infty.
\end{aligned}
$$

Now, we prove that the normalization constant does not vanish. Let $R = \mathbb{E}(\rho_{B^{t(\cdot)}_{q(\cdot)}}(u))$. We know that $R \in (0, \infty)$. As $\rho_{B^{t(\cdot)}_{q(\cdot)}}(u)$ is a non-negative random variable, we obtain that $\mu_0(\rho_{B^{t(\cdot)}_{q(\cdot)}}(u) < R) > 0$. With $r = \max\{\|y\|_Y, R\}$, Assumption 1 (ii) ensures that

$$
\begin{aligned}
Z(y) &= \int_{B^{t(\cdot)}_{q(\cdot)}} \exp(-\Phi(u; y))\mu_0(\mathrm{d}u) \\
&\geqslant \int_{\tilde{b}^{t(\cdot)}_{q(\cdot)}<R} \exp(-K)\mu_0(\mathrm{d}u) \\
&= \exp(-K)\mu_0(\rho_{B^{t(\cdot)}_{q(\cdot)}}(u) < R)
\end{aligned}
$$

which is positive. □

A.4. Some basic facts about wavelets

Proposition A.1. *[23] (i) There is a real scaling function $\varphi_F \in \mathcal{S}(\mathbb{R})$ and a real associated wavelet $\varphi_M \in \mathcal{S}(\mathbb{R})$ such that their Fourier transforms have compact support, $\hat{\varphi}_F(0) = 1$ and*

$$\operatorname{supp} \hat{\varphi}_M \subset \left[-\frac{8}{3}\pi, -\frac{2}{3}\pi\right] \bigcup \left[\frac{2}{3}\pi, \frac{8}{3}\pi\right].$$

(ii) For any $k \in \mathbb{N}$, there exist a real, compactly supported scaling function $\varphi_F \in C^k(\mathbb{R})$ and a real, compactly supported associated wavelet $\varphi_M \in C^k(\mathbb{R})$ such that $\hat{\varphi}_F(0) = 1$ and

$$\int_{\mathbb{R}} x^\ell \varphi_M(x)\mathrm{d}x = 0 \quad \text{for all } \ell \in \{0, 1, \cdots, k - 1\}.$$

In both cases, we observe that $\{\varphi_{vm} : v \in \mathbb{N} \cup 0, m \in \mathbb{Z}\}$ is an orthonormal basis in $L^2(\mathbb{R})$ where

$$\varphi_{vm}(t) := \begin{cases} \varphi_F(t - m), & \text{if } v = 0, m \in \mathbb{Z} \\ 2^{\frac{v-1}{2}} \varphi_M(2^{v-1}t - m), & \text{if } v \in \mathbb{N}, m \in \mathbb{Z} \end{cases}$$

and the functions φ_M, φ_F are according to (i) or (ii).

The wavelets in the first part of the above proposition are called Meyer wavelets. They do not have compact support but are fast-decaying functions, and φ_M has infinitely many moment conditions. The wavelets in the second part of the above proposition are called Daubechies wavelets. The functions φ_M, φ_F have compact support but only have limited smoothness.

Definition A.1. [23] Let $s(\cdot) \in L^\infty \cap C^{\log}_{\text{loc}}(\mathbb{T}^n)$, $0 < q \leqslant \infty$ and $p(\cdot) \in \mathcal{P}(\mathbb{T}^n)$ with $0 < p^- \leqslant p^+ \leqslant \infty$. Let $\mathbb{M}_j = \{m : m = 0, 1, 2, \cdots, 2^j - 1\}$.

(i) Then,

$$\tilde{b}^{s(\cdot)}_{p(\cdot),q} := \{\lambda = \{\lambda^j_{Gm}\}_{j\in\mathbb{N}_0, G\in G^j, m\in\mathbb{M}_j} : \|\lambda\|_{\tilde{b}^{s(\cdot)}_{p(\cdot),q}} < \infty\}$$

where

$$\|\lambda\|_{\tilde{b}^{s(\cdot)}_{p(\cdot),q}} = \left(\sum_{j=0}^{\infty} \sum_{G\in G^j} \left\| \sum_{m\in\mathbb{M}_j} 2^{js(2^{-j}m)} |\lambda_{Gm}|^j \chi_{jm}(\cdot) \right\|^q_{L^{p(\cdot)}(\mathbb{T}^n)}\right)^{1/q}.$$

(ii) For $p^+ < \infty$, we define

$$\tilde{f}^{s(\cdot)}_{p(\cdot),q(\cdot)} := \{\lambda = \{\lambda^j_{Gm}\}_{j\in\mathbb{N}_0, G\in G^j, m\in\mathbb{M}_j} : \|\lambda\|_{\tilde{f}^{s(\cdot)}_{p(\cdot),q(\cdot)}} < \infty\},$$

where

$$\|\lambda\|_{\tilde{f}^{s(\cdot)}_{p(\cdot),q(\cdot)}} = \left\| \left(\sum_{j=0}^{\infty} \sum_{G\in G^j} \sum_{m\in\mathbb{M}_j} 2^{jqs(2^{-j}m)} |\lambda^j_{Gm}|^{q(\cdot)} \chi_{jm}(\cdot)\right)^{1/q(\cdot)} \right\|_{L^{p(\cdot)}(\mathbb{T}^n)}$$

with $q(\cdot) \in \mathcal{P}(\mathbb{T}^n)$.

A.5. A useful corollary of Kolmogorov's continuity criterion

Let us consider a random function u given by the random series

$$u = \sum_{k \geqslant 0} \xi_k \psi_k \tag{A.3}$$

where $\{\xi_k\}_k$ is an i.i.d. sequence and the ψ_k are real- or complex-valued Hölder functions on the bounded open $D \subset \mathbb{R}^n$ satisfying, for some $\alpha \in (0, 1]$,

$$|\psi_k(x) - \psi_k(y)| \leqslant h(\alpha, \psi_k)|x - y|^\alpha \quad x, y \in D; \tag{A.4}$$

of course, if $\alpha = 1$, the functions are Lipschitz functions.

Theorem A.1. *[10] Let $\{\xi_k\}_{k\geqslant 0}$ be countably many centered i.i.d. random variables with bounded moments of all orders. Moreover, let $\{\psi_k\}_{k\geqslant 0}$ satisfy (A.4). Suppose there is some $\delta \in (0, 2)$ such that*

$$S_1 := \sum_{k \geqslant 0} \|\psi_k\|_{L^\infty}^2 < \infty, \tag{A.5}$$

and

$$S_2 := \sum_{k \geqslant 0} \|\psi_k\|_{L^\infty}^{2-\delta} h(\alpha, \psi_k)^\delta < \infty. \tag{A.6}$$

Then, u defined by (A.3) is a.s. finite for every $x \in D$, and u is Hölder continuous for every Hölder exponent smaller than $\alpha\delta/2$.

References

[1] Almeida A and Hästö P 2010 Besov spaces with variable smoothness and integrability *J. Funct. Anal.* **258** 1628–55

[2] Bahouri H, Chemin J-Y and Danchin R 2011 *Fourier Analysis and Nonlinear Partial Differential Equations* vol 343 (Springer Science & Business Media)

[3] Blomgren P, Chan T F, Mulet P and Wong C-K 1997 Total variation image restoration: numerical methods and extensions *icip* 384 IEEE

[4] Chen Y, Levine S and Rao M 2006 Variable exponent, linear growth functionals in image restoration *SIAM J. Appl. Math.* **66** 1383–406

[5] Cotter S L, Dashti M, Robinson J C and Stuart A M 2009 Bayesian inverse problems for functions and applications to fluid mechanics *Inverse Problems* **25** 115008

[6] Cotter S L *et al* 2013 Mcmc methods for functions: modifying old algorithms to make them faster *Stat. Sci.* **28** 424–46 3

[7] Da Prato G and Zabczyk J 2014 *Stochastic Equations in Infinite Dimensions* vol 152 (Cambridge University Press)

[8] Dashti M, Harris S and Stuart A 2012 Besov priors for bayesian inverse problems *Inverse Problems Imaging* **6** 183–200

[9] Dashti M, Law K J, Stuart A M and Voss J 2013 Map estimators and their consistency in bayesian nonparametric inverse problems *Inverse Problems* **29** 095017

[10] Dashti M and Stuart A M 2013 The bayesian approach to inverse problems arXiv preprint arXiv:1302.6989

[11] Diening L, Harjulehto P, Hästö P and Ruzicka M 2011 *Lebesgue and Sobolev Spaces with Variable Exponents* (Springer)

[12] Diening L, Hästö P and Roudenko S 2009 Function spaces of variable smoothness and integrability *J. Funct. Anal.* **256** 1731–68

[13] Diestel J and Uhl J J 1977 *Vector Measures* (American Mathematical Soc.) No. 15

[14] Franklin J N 1970 Well-posed stochastic extensions of ill-posed linear problems *J. Math. Anal. Appl.* **31** 682–716

[15] Gibbs A L and Su F E 2002 On choosing and bounding probability metrics *Int. Stat. Rev.* **70** 419–35

[16] Helin T and Burger M 2015 Maximum *a posteriori* probability estimates in infinite-dimensional bayesian inverse problems *Inverse Problems* **31** 085009

[17] Hernández E and Weiss G 1996 *A First Course on Wavelets* (CRC Press)

[18] Jia J, Peng J, Gao J and Li Y 2016 Variable total variation regularization for backward time-space fractional diffusion problem arXiv preprint arXiv:1601.04200

[19] Jia J, Peng J and Li K 2014 Well-posedness of abstract distributed-order fractional diffusion equations *Commun. Pure Appl. Anal.* **13** 2

[20] Jin B and Rundell W 2015 A tutorial on inverse problems for anomalous diffusion processes *Inverse Problems* **31** 035003

[21] Kaipio J and Somersalo E 2006 *Statistical and Computational Inverse Problems* vol 160 (Springer Science & Business Media)

[22] Kempka H 2009 2-microlocal besov and triebel-lizorkin spaces of variable integrability *Rev. Mat. Complutense* **22** 227–51

[23] Kempka H *et al* 2010 Atomic, molecular and wavelet decomposition of 2-microlocal besov and triebel-lizorkin spaces with variable integrability *Functiones et Approximatio Commentarii Mathematici* **43** 171–208

[24] Klages R, Radons G and Sokolov I M 2008 *Anomalous Transport: Foundations and Applications* (John Wiley & Sons)

[25] Knopp K 2013 *Theory and Application of Infinite Series* (Courier Corporation)

[26] Lasanen S 2007 Measurements and infinite-dimensional statistical inverse theory *Proc. Appl. Math. Mech.* **7** 1080101–2

[27] Lasanen S 2011 Posterior convergence for approximated unknowns in non-gaussian statistical inverse problems preprint arXiv:1112.0906

[28] Lasanen S 2012 Non-gaussian statistical inverse problems :I. Posterior distributions *Inverse Problems Imaging* **6** 2

[29] Lasanen S 2012 Non-gaussian statistical inverse problems II:. Posterior convergence for approximated unknowns *Inverse Problems Imaging* **6** 2

[30] Lassas M, Saksman E and Siltanen S 2009 Discretization-invariant bayesian inversion and besov space priors *Inverse Problems Imaging* **3**

[31] Lassas M and Siltanen S 2004 Can one use total variation prior for edge-preserving bayesian inversion? *Inverse Problems* **20** 1537

[32] Li K and Jia J 2011 Existence and uniqueness of mild solutions for abstract delay fractional differential equations *Comput. Math. Appl.* **62** 1398–404

[33] Li K, Peng J and Jia J 2012 Cauchy problems for fractional differential equations with riemann-liouville fractional derivatives *J. Funct. Anal.* **263** 476–510

[34] Mei Z, Peng J and Jia J 2014 A new characteristic property of mittag-leffler functions and fractional cosine functions *Stud. Math.* **220** 119–40

[35] Peng J and Li K 2012 A novel characteristic of solution operator for the fractional abstract cauchy problem *J. Math. Anal. Appl.* **385** 786–96

[36] Podlubny I 1998 *Fractional Differential Equations: An Introduction to Fractional Derivatives, Fractional Differential Equations, to Methods of their Solution and some of their Applications* vol 198 (Academic Press)

[37] Rudin L I, Osher S and Fatemi E 1992 Nonlinear total variation based noise removal algorithms *Physica D: Nonlinear Phenom.* **60** 259–68

[38] Stroock D W 2010 *Probability Theory: An Analytic View* (Cambridge university press)

[39] Stuart A M 2010 Inverse problems: a bayesian perspective *Acta Numer.* **19** 451–559

[40] Tarantola A 2005 *Inverse Problem Theory and Methods for Model Parameter Estimation* (SIAM)

[41] Tiirola J 2014 Image decompositions using spaces of variable smoothness and integrability *SIAM J. Imaging Sci.* **7** 1558–87

[42] Tikhonov A N *Regularization of Incorrectly Posed Problems* (Soviet Mathematics Doklady)

[43] Triebel H 2008 *Function Spaces and Wavelets on Domains* (European Mathematical Society) No. 7

[44] Wang L and Liu J 2013 Total variation regularization for a backward time-fractional diffusion problem *Inverse Problems* **29** 115013

[45] Zheng G H and Wei T 2010 Two regularization methods for solving a riesz-feller space-fractional backward diffusion problem *Inverse Problems* **26** 115017