



Citation for published version:

Korolev, Y, Latz, J & Schönlieb, C-B 2022 'Gaussian random fields on non-separable Banach spaces'.

Publication date:
2022

[Link to publication](#)

University of Bath

Alternative formats

If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Gaussian random fields on non-separable Banach spaces

Yury Korolev* Jonas Latz† Carola-Bibiane Schönlieb*

Abstract

We study Gaussian random fields on certain Banach spaces and investigate conditions for their existence. Our results apply inter alia to spaces of Radon measures and Hölder functions. In the former case, we are able to define Gaussian white noise on the space of measures directly, avoiding, e.g., an embedding into a negative-order Sobolev space. In the latter case, we demonstrate how Hölder regularity of the samples is controlled by that of the covariance kernel and, thus, show a connection to the Theorem of Kolmogorov-Chentsov.

Keywords: Gaussian measures, sample regularity, Radon measures, Hölder spaces, Besov spaces, tensor products of Banach spaces

MSC2020: 60G15, 46N30, 46B26.

1 Introduction

Function-valued Gaussian random variables play a fundamental role in various fields of mathematics, e.g., non-parametric statistics [28], stochastic partial differential equations [11], and function approximation [29]. We distinguish two kinds of function-valued Gaussian random variables: *Gaussian processes*, which are families of Gaussian random variables with a (general) index set, and *Gaussian random fields*, which are Gaussian random variables defined on certain structured function spaces, equipped with a space of continuous linear functionals. We give rigorous definitions below. Which concept is used depends very much on the field of study: Gaussian processes are well-understood in terms of classical regularity of samples and popular in certain applications, e.g., data science. Gaussian random fields allow one to study random functions from a functional analytic perspective, simplifying, e.g., the investigation of conditional distributions and stochastic partial differential equations. So far, the Gaussian random field theory is mainly developed on separable Hilbert spaces that often do not allow to study classical regularity of samples. In this work, we aim at closing this gap between the two concepts through investigating Gaussian random fields on certain Banach spaces. We prove their existence under assumptions on the covariance operators and discuss their construction. In particular, we consider Hölder spaces and spaces of Radon measures. In the former, we are able to investigate classical regularity. From the latter, we obtain a simple theory for Gaussian white noise.

Background. Let $(Y, \mathcal{F}, \mathbb{P})$ be the probability space on which we define random variables throughout this work and $\Omega \subset \mathbb{R}^n$ be some compact set. θ is a function-valued Gaussian random variable, say θ is a randomised function of type $f : \Omega \rightarrow \mathbb{R}$. We now give two different definitions, or really frameworks, of such Gaussian random variables that are common in the literature.

We commence with the *(Gaussian/stochastic) process viewpoint*. Here, $\theta := (\theta(x))_{x \in \Omega}$ is a collection of scalar random variables. That means, θ is a random element in $\mathbb{R}^\Omega := \{f : \Omega \rightarrow \mathbb{R}\}$ equipped with the cylindrical σ -algebra.

*Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, CB3 0WA, Cambridge, UK. email: {yk362, cbs31}@cam.ac.uk

†Maxwell Institute for Mathematical Sciences & School of Mathematical and Computer Sciences, Heriot-Watt University, EH14 4AS, Edinburgh, UK. email: j.latz@hw.ac.uk

Definition 1.1. Let $m : \Omega \rightarrow \mathbb{R}$ be a function and $c : \Omega \times \Omega \rightarrow \mathbb{R}$ be a continuous, symmetric, positive semi-definite function. We refer to θ as a Gaussian process with mean m and covariance c , if for any $k \in \mathbb{N} := \{1, 2, \dots\}$ and any set of points $x_1, \dots, x_k \in \Omega$, we have

$$\begin{pmatrix} \theta(x_1) \\ \vdots \\ \theta(x_k) \end{pmatrix} \sim \mathbb{N} \left(\begin{pmatrix} m(x_1) \\ \vdots \\ m(x_k) \end{pmatrix}, \begin{pmatrix} c(x_1, x_1) & \cdots & c(x_1, x_k) \\ \vdots & \ddots & \vdots \\ c(x_k, x_1) & \cdots & c(x_k, x_k) \end{pmatrix} \right),$$

where for appropriate $m' \in \mathbb{R}^k, C' \in \mathbb{R}^{k \times k}$, we use $\mathbb{N}(m', C')$ to denote multivariate Gaussian distributions on $(\mathbb{R}^k, \mathcal{B}\mathbb{R}^k)$.

One can show existence of such processes through the Kolmogorov extension theorem, see, e.g., Theorem 14.16 in [14]. Important properties of Gaussian processes are their regularity, especially continuity and Hölder continuity of the samples. To this end, first note that another stochastic process $\tilde{\theta} := (\tilde{\theta}(x))_{x \in \Omega}$ is a *modification* of θ , if $\mathbb{P}(\theta(x) = \tilde{\theta}(x)) = 1$ for all $x \in \Omega$. The Theorem of Kolmogorov-Chentsov (e.g., Theorem 4.23 in [13]) discusses the existence of regular modifications of a stochastic processes:

Theorem 1.2 (Kolmogorov–Chentsov). Let $a, b, c > 0$ be some constants such that

$$\mathbb{E}[\|\theta(x) - \theta(y)\|^a] \leq c\|x - y\|^{n+b}.$$

Then, there is a modification $\tilde{\theta}$ of θ that is α -Hölder continuous with $\alpha \in (0, b/a)$.

Bogachev [4] gives the following definition of a Gaussian measure, which we refer to as the (Gaussian) random field view point.

Definition 1.3. Let E be a locally convex space and F be a space containing continuous linear functionals on E . Moreover, let $\langle \cdot, \cdot \rangle_{F,E}$ be the bilinear pairing of E and F . A random variable $\theta : Y \rightarrow E$ is a Gaussian random field, if for any $k \in \mathbb{N}$ and $\ell_1, \dots, \ell_k \in F$ there are $m' \in \mathbb{R}^k$ and $C' \in \mathbb{R}^{k \times k}$ such that

$$\begin{pmatrix} \langle \ell_1, \theta \rangle_{F,E} \\ \vdots \\ \langle \ell_k, \theta \rangle_{F,E} \end{pmatrix} \sim \mathbb{N}(a, C'). \quad (1.1)$$

Remark 1.4. The Gaussian process viewpoint is essentially contained in the random field view point, where $E := \mathbb{R}^\Omega$ and F containing point evaluations of functions in E . To simplify the discussion, we still distinguish the two

As mentioned before, the Kolmogorov extension theorem gives us a simple way to show existence of a Gaussian process with a certain covariance function. Showing existence of Gaussian random fields is more involved: Indeed, Bogachev [4] discusses existence only in the case where E is a separable Hilbert space and $F \cong E$ is its (isomorphic) dual; see also Section 2.

We have already mentioned advantages and disadvantages of the different view points. In this article, we extend the treatment of Definition 1.3 to the case where E is a Banach space, possibly non-separable, that has a predual with a basis, and $F \subseteq E^*$ is contained in the dual of E . Our results apply inter alia to spaces of Radon measures and Hölder continuous functions. By studying Gaussian random fields on Hölder spaces we hope to close the gap in between the theories of Gaussian random fields and Gaussian processes by allowing us to study classical regularity of function-valued Gaussian random variables on a structured space without the necessity of continuous versions. Gaussian random fields on Radon spaces gives us a very simple and natural path to the definition of Gaussian white noise.

Selected previous work. Regularity of Gaussian random variables on function spaces has been an extensive field of study. For further reading, we refer to the works by Adler [1, 2] and Potthoff [20–22]. From an application point of view, we refer to the books by Sullivan [30] and Lord, Powell, and Shardlow [17]. Gaussian white noise has no spatial correlation and fits neither into the setting of Definition 1.1 nor the separable Hilbert space setting in Definition 1.3. It is treated as a ‘generalised random field’ and, e.g., discussed in the book by Kuo [16].

Our contributions. Our main contributions are as follows

- we define Gaussian random fields on dual Banach spaces, bridging the gap between the Gaussian process and Gaussian random field point of view;
- we study random fields on Hölder spaces and show how Hölder regularity of the samples is controlled by the Hölder regularity of the covariance kernel;
- we study random fields on the space of Radon measures and define white noise on the space of measures.

Structure of the paper. The paper is organised as follows. In Section 2, we outline the random field view point for Hilbert-space valued random variables following [4]. Then we present a generalisation of this framework to random variables with values in a dual Banach space. This is the topic of Section 3. In Section 4, we discuss some spaces where the theory can be applied. Section 4.1 is concerned with sampling Radon measures. Here we show, for example, how white noise can be defined on the space of measures. In Section 4.2, we consider sampling Hölder functions and obtain results on the regularity of samples generated by exponential covariance kernels. In Section 4.3, we briefly discuss how sampling continuous functions may fit into this framework. Necessary results about tensor products of Banach spaces are collected in Appendix A.

Notation. Generic Banach spaces will be denoted by E or F . The injective and projective tensor products will be denoted by $E \widehat{\otimes}_\varepsilon F$ and $E \widehat{\otimes}_\pi F$, respectively. Symmetric products will be denoted by $\widehat{\otimes}_{\varepsilon_s} E$ and $\widehat{\otimes}_{\pi_s} E$. We will use (X, d) or simply X for a metric space with metric d . All our metric spaces are assumed compact. If $0 < \alpha < 1$ and (X, d) is a metric space, we will denote by X^α the space (X, d^α) , which is, of course, also a metric space. We will use $\Omega \subset \mathbb{R}^n$ for a domain in \mathbb{R}^n , which will be assumed compact. The Borel σ -algebra on Ω will be denoted by $\mathcal{B}\Omega$. We will use $x, x' \in \Omega$ for spatial variables, $f, g: \Omega \rightarrow \mathbb{R}$ for generic functions and $\mu, \nu: \mathcal{B}\Omega \rightarrow \mathbb{R}$ or $\mathcal{B}\Omega \rightarrow E$ for generic measures. The space of Lipschitz functions on Ω will be denoted by $\text{Lip}(\Omega)$. The subspace of functions vanishing at a basepoint will be $\text{Lip}_0(\Omega)$. Its predual Arens-Eells space [32] will be denoted by $\mathcal{A}(\Omega)$. We will use $\mathcal{M}(\Omega)$ for the space of scalar-valued Radon measures, $\mathfrak{M}(\Omega, E)$ for the space of E -valued measures, and $\mathfrak{M}_1(\Omega, E)$ for the space of E -valued measures with the Radon-Nikodým property. The Banach space where we would like to sample will be called U . Its dual will be denoted by U^* . We will also assume that U has a predual, U° . We will assume that this predual has a basis. The elements of these spaces will be $\eta, \eta' \in U^\circ$, $u, u' \in U$, and $v, v' \in U^*$. By the Banach-Alaoglu theorem, the weak* topology on bounded sets in U is metrisable. We will use the symbol $\Delta_*(\cdot, \cdot)$ to denote this “weak* metric”. The completion of U with respect to this metric will be denoted by U_* . Samples will be denoted by θ . A covariance kernel will be denoted by c and we will assume that $c \in \widehat{\otimes}_{\pi_s} U$. This kernel corresponds to a nuclear covariance operator $\mathcal{C}: U^\circ \rightarrow U$. The standard normal distribution will be denoted by $N(0, 1)$ and a scalar sample from it by $\xi \sim N(0, 1)$.

2 Gaussian random fields on Hilbert spaces

We outline the basic Hilbert space setting following [4]. For illustrative purposes, we will restrict ourselves to the case $H = L^2(\Omega)$, where $\Omega \subset \mathbb{R}^n$ is a compact set. Let $c \in C(\Omega \times \Omega)$ be a covariance kernel, i.e. a continuous, symmetric, and positive semidefinite function $\Omega \times \Omega \rightarrow \mathbb{R}$. We consider

the following integral operator $\mathcal{C} : L^2(\Omega) \rightarrow L^2(\Omega)$

$$\mathcal{C}f := \int_{\Omega} f(x)c(x, \cdot) dx, \quad f \in L^2(\Omega), \quad (2.1)$$

which is by definition self-adjoint and positive semidefinite. It can also be shown that it is nuclear.

Definition 2.1 (Nuclear operators on Hilbert spaces). *Let E be a separable Hilbert space and $N : E \rightarrow E$ a compact positive semidefinite self-adjoint operator with eigenvalues $\{\lambda_i\}_{i \in \mathbb{N}} \subset \mathbb{R}_+$. The operator N is called nuclear if*

$$\|N\|_{\mathcal{N}} := \sum_{i=1}^{\infty} \lambda_i < \infty.$$

Theorem 2.2 (Mercer's theorem, [19]). *Let $\Omega \subset \mathbb{R}^d$ be compact and $c : \Omega \times \Omega \rightarrow \mathbb{R}$ be continuous, symmetric and positive-semidefinite. Then the operator defined in (2.1) admits the following eigendecomposition*

$$\mathcal{C}f = \sum_{i=1}^{\infty} \lambda_i \langle f, \varphi_i \rangle_{L^2} \varphi_i,$$

where $(\varphi_i)_{i=1}^{\infty}$ is an orthonormal basis of $L^2(\Omega)$, $\lambda_i \geq 0$ and $\langle \varphi, \varphi' \rangle_{L^2} := \int_{\Omega} \varphi(x)\varphi'(x) dx$ is the scalar product in $L^2(\Omega)$. Moreover, we have that

$$\sum_{i=1}^{\infty} \lambda_i < \infty,$$

hence, \mathcal{C} is a nuclear operator $L^2(\Omega) \rightarrow L^2(\Omega)$.

We can use the eigendecomposition of the covariance operator from Theorem 2.2 to obtain a (zero-mean) Gaussian random field on H by letting

$$\theta := \sum_{i=1}^{\infty} \lambda_i^{1/2} \xi_i \varphi_i, \quad (2.2)$$

where $\xi_1, \xi_2, \dots \sim N(0, 1)$ are independent and identically distributed. The following result shows that if \mathcal{C} is nuclear then the samples (2.2) are in H almost surely.

Proposition 2.3. *Let $\xi_1, \xi_2, \dots \sim N(0, 1)$ be independent and identically distributed and $\{\lambda_i\}_{i \in \mathbb{N}} \subset \mathbb{R}_+$ such that $\sum_{i=1}^{\infty} \lambda_i < \infty$. Let θ be as defined in (2.2). Then almost surely $\theta \in H$.*

Proof. Since $\{\varphi_i\}_{i \in \mathbb{N}}$ are orthonormal, we have

$$\|\theta\|_H^2 = \sum_{i=1}^{\infty} \lambda_i \xi_i^2. \quad (2.3)$$

Note that we have $\sum_{i=1}^{\infty} \lambda_i < \infty$ and $\sum_{i=1}^{\infty} \lambda_i^2 < \infty$ by assumption. Then, we study

$$\sum_{i=1}^{\infty} \lambda_i (\xi_i^2 - 1).$$

We have $\mathbb{E}[\lambda_i (\xi_i^2 - 1)] = 0$ and $\text{Var}(\lambda_i (\xi_i^2 - 1)) = 2\lambda_i^2$, for $i \in \mathbb{N}$. Thus, by [3, Thm. 22.6], the series $\sum_{i=1}^{\infty} \lambda_i (\xi_i^2 - 1)$ converges almost surely. On the other hand, we can write the sum (2.3) as $\sum_{i=1}^{\infty} \lambda_i (\xi_i^2 - 1) + \sum_{i=1}^{\infty} \lambda_i$, which is now the sum of two almost surely finite series. Thus, (2.3) is almost surely finite itself. \square

The covariance operator \mathcal{C} can be identified with the following bilinear form.

Proposition 2.4 ([4]). *For any $f, g \in L^2(\Omega)$ one has*

$$\langle f, \mathcal{C}g \rangle = \mathbb{E}_{\theta}(\langle f, \theta \rangle_H \langle g, \theta \rangle_H). \quad (2.4)$$

Proof. A short computation gives

$$\begin{aligned} \mathbb{E}_{\theta}(\langle f, \theta \rangle_H \langle g, \theta \rangle_H) &= \mathbb{E}_{\theta} \left(\left\langle f, \sum_{i=1}^{\infty} \lambda_i^{1/2} \xi_i \varphi_i \right\rangle_H \left\langle g, \sum_{j=1}^{\infty} \lambda_j^{1/2} \xi_j \varphi_j \right\rangle_H \right) \\ &= \sum_{i=1}^{\infty} \mathbb{E}_{\theta} (\lambda_i \xi_i^2 \langle f, \varphi_i \rangle_H \langle g, \varphi_i \rangle_H) \\ &= \sum_{i=1}^{\infty} \lambda_i \langle f, \varphi_i \rangle_H \langle g, \varphi_i \rangle_H = \langle f, \mathcal{C}g \rangle. \end{aligned}$$

\square

Characterisation via tensor products. It is also possible to discuss Theorem 2.2 in the language of topological tensor products. Since all Hilbert spaces have the approximation property [23], using Theorem A.7 one can see that the covariance operator \mathcal{C} is nuclear if and only if the associated bilinear form

$$(f, g) \mapsto \int_{\Omega \times \Omega} f(x)g(y)c(x, y) \, dx \, dy$$

has a finite projective norm. Somewhat abusing notation, we will use the letter c both for the covariance kernel and this bilinear form. If the covariance kernel is symmetric, we have that

$$c \in \widehat{\otimes}_{\pi_s}^s L^2(\Omega). \quad (2.5)$$

is necessary and sufficient for the nuclearity of \mathcal{C} . Theorem 2.2 gives a sufficient condition for this.

It is also helpful to rewrite (2.1) in the language of the inner product on $L^2(\Omega)$

$$\mathcal{C}f(x) := \langle f, c(\cdot, x) \rangle, \quad f \in L^2(\Omega), \quad x \in \Omega, \quad (2.6)$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product on $L^2(\Omega)$ and the equality holds almost everywhere. In the sequel we will also use the notation $\langle \cdot, \cdot \rangle$ for the duality pairing between Banach spaces, sometimes using subscripts to specify these spaces.

Another useful equivalent way of writing (2.1) and (2.6) is using pairings with elements of L^2

$$\langle f, \mathcal{C}g \rangle := \langle f, \langle g, c(\cdot, x) \rangle \rangle = \langle c, f \otimes_s g \rangle, \quad f, g \in L^2(\Omega), \quad (2.7)$$

where the last pairing is between the bilinear form c and the tensor $f \otimes_s g$. Since c is symmetric, we use the symmetric product $f \otimes_s g$.

3 Gaussian random fields on dual Banach spaces

In this section, we will discuss how the framework of Section 2 can be extended to (possibly, non-separable) Banach spaces that have a predual which possesses a basis.

3.1 General theory

Let U be a Banach space and U° a predual of U . We will assume that U has the approximation property and U° has a basis. In order to retain the structure of (2.6) and also ensure that the image of the covariance operator \mathcal{C} is in U , the domain of \mathcal{C} should be either U^* or U° . We are particularly interested in non-reflexive spaces, in which case the predual usually has ‘nicer’ properties than the dual. Hence, we will consider $\mathcal{C}: U^\circ \rightarrow U$.

Let c be a bilinear form on U , which we will also refer to as the covariance kernel. We generalise (2.7) and define the covariance operator as follows

$$\mathcal{C}: \langle \eta', \mathcal{C}\eta \rangle := \langle c, \eta \otimes_s \eta' \rangle, \quad \eta, \eta' \in U^\circ, \quad (3.1)$$

where the second pairing is between the bilinear form c and the tensor $\eta \otimes_s \eta'$.

The following result is a consequence of Theorem A.13.

Proposition 3.1. *Suppose that U has the approximation property. Then the covariance operator $\mathcal{C}: U^\circ \rightarrow U$ as defined in (3.1) is symmetric and nuclear if and only if*

$$c \in \widehat{\otimes}_{\pi_s}^s U. \quad (3.2)$$

In this case, there exists a sequence $\{\varphi_i\}_{i \in \mathbb{N}} \subset U$ satisfying $\|\varphi_i\| = 1$ for all i such that

$$\mathcal{C}\eta = \sum_{i=1}^{\infty} \lambda_i \langle \eta, \varphi_i \rangle_{U^\circ, U} \varphi_i, \quad \eta \in U^\circ \quad (3.3)$$

and $\sum_{i=1}^{\infty} |\lambda_i| < \infty$. In addition, \mathcal{C} is positive semidefinite if $\lambda_i \geq 0$ for all i .

Now we can define a Gaussian random field over U analogously to (2.2)

$$\theta := \sum_{i=1}^{\infty} \lambda_i^{1/2} \xi_i \varphi_i, \quad (3.4)$$

where $\xi_i \sim N(0, 1)$ are i.i.d. It is important to note that, in general, we will not be able to show that $\theta \in U$ with probability 1, but only that the sum (3.4) converges weakly-* with probability 1.

Proposition 3.2. *Let $\xi_1, \xi_2, \dots \sim N(0, 1)$ be independent and identically distributed and $\{\lambda_i\}_{i \in \mathbb{N}} \subset \mathbb{R}_+$ such that $\sum_{i=1}^{\infty} \lambda_i < \infty$. Let θ be as defined in (3.4). Then, for any $\eta \in U^\circ$, we have*

$$\left| \langle \eta, \theta \rangle_{U^\circ, U} \right| < \infty$$

with probability 1.

Proof. We can write

$$\langle \eta, \theta \rangle_{U^\circ, U} = \left\langle \eta, \sum_{i=1}^{\infty} \lambda_i^{1/2} \xi_i \varphi_i \right\rangle_{U^\circ, U} = \sum_{i=1}^{\infty} \lambda_i^{1/2} \xi_i \langle \eta, \varphi_i \rangle_{U^\circ, U}.$$

Then we have $\mathbb{E}[\lambda_i^{1/2} \xi_i \langle \eta, \varphi_i \rangle_{U^\circ, U}] = 0$ and $\text{Var}(\lambda_i^{1/2} \xi_i \langle \eta, \varphi_i \rangle_{U^\circ, U}) = \lambda_i \langle \eta, \varphi_i \rangle_{U^\circ, U}^2 \leq \lambda_i \|\eta\|_{U^\circ}^2$. Moreover, $\sum_{i=1}^{\infty} \lambda_i \|\eta\|_{U^\circ}^2 < \infty$, by assumption. Hence, by [3, Thm. 22.6], we have $\left| \langle \eta, \theta \rangle_{U^\circ, U} \right| < \infty$ with probability 1. \square

Let $\{\eta_i\}_{i \in \mathbb{N}}$ be a countable dense system in U° . The weak* topology on the unit ball of U can be metrised by the following metric [7, Thm. V.5.1]

$$\Delta(u, u') = \sum_{i=1}^{\infty} \beta_i \frac{|\langle \eta_i, u - u' \rangle|}{1 + |\langle \eta_i, u - u' \rangle|},$$

where $\beta_i > 0$ are some coefficients such that $\sum_{i=1}^{\infty} \beta_i = 1$. If $\{\eta_i\}_{i \in \mathbb{N}}$ are normalised, we can use the following equivalent metric

$$\Delta_*(u, u') := \sum_{i=1}^{\infty} \beta_i |\langle \eta_i, u - u' \rangle|. \quad (3.5)$$

We denote by U_* be the completion of U with respect to the metric (3.5). It can be turned into a normed space by defining

$$\|u\|_{U_*} := \Delta_*(u, 0) = \sum_{i=1}^{\infty} \beta_i |\langle \eta_i, u \rangle|.$$

Theorem 3.3. *Let $\xi_1, \xi_2, \dots \sim N(0, 1)$ be independent and identically distributed and $\{\lambda_i\}_{i \in \mathbb{N}} \subset \mathbb{R}_+$ such that $\sum_{i=1}^{\infty} \lambda_i < \infty$. Let θ be as defined in (3.4). Then $\theta \in U_*$ with probability 1.*

Proof. Consider the unit ball B_{U° and the following probability measure on B_{U°

$$\mu := \sum_{i=1}^{\infty} \beta_i \delta_{\eta_i},$$

where $\{\eta_i\}_{i \in \mathbb{N}}$ is a countable dense and normalised system in U° and β_i 's are positive and sum up to 1. From Proposition 3.2, we know that for any $\eta \in U^\circ$, $\left| \langle \eta, \theta \rangle_{U^\circ, U} \right| < \infty$ with probability 1. Taking the expectation over μ , we get that

$$\|\theta\|_{U_*} = \sum_{i=1}^{\infty} \left| \langle \beta_i \eta_i, \theta \rangle_{U^\circ, U} \right| = \mathbb{E}_\mu \left| \langle \cdot, \theta \rangle_{U^\circ, U} \right| < \infty$$

with probability 1. \square

If the covariance operator is not only nuclear, but 1/2-nuclear, i.e. $\sum_{i=1}^{\infty} \lambda_i^{1/2} < \infty$, then we can even ensure that $\theta \in U$ with probability 1.

Proposition 3.4. *Let $\xi_1, \xi_2, \dots \sim N(0, 1)$ be independent and identically distributed and $\{\lambda_i\}_{i \in \mathbb{N}} \subset \mathbb{R}_+$ such that $\sum_{i=1}^{\infty} \lambda_i^{1/2} < \infty$. Let θ be as defined in (3.4). Then almost surely $\theta \in U$.*

Proof. Since $\{\varphi_i\}_{i \in \mathbb{N}}$ are normalised, the Cauchy-Schwarz inequality yields

$$\|\theta\|_U \leq \sum_{i=1}^{\infty} \lambda_i^{1/2} |\xi_i|. \quad (3.6)$$

We study the sum $\sum_{i=1}^{\infty} \lambda_i^{1/2} (|\xi_i| - \sqrt{2/\pi})$. Note that here

$$\mathbb{E} \left[\lambda_i^{1/2} (|\xi_i| - \sqrt{2/\pi}) \right] = 0, \quad \text{Var} \left(\lambda_i^{1/2} (|\xi_i| - \sqrt{2/\pi}) \right) = \lambda_i(\pi - 2)/\pi,$$

where by assumption, $\sum_{i=1}^{\infty} \lambda_i(\pi - 2)/\pi < \infty$. Thus, by [3, Thm. 22.6], $\sum_{i=1}^{\infty} \lambda_i^{1/2} (|\xi_i| - \sqrt{2/\pi})$ is finite with probability 1. Moreover, by assumption, we have $\sum_{i=1}^{\infty} \lambda_i^{1/2} \sqrt{2/\pi} < \infty$, giving us $\sum_{i=1}^{\infty} \lambda_i^{1/2} |\xi_i| < \infty$ with probability 1. \square

Finally, as in the Hilbert space setting, there is a natural bilinear form associated with the covariance operator.

Proposition 3.5. *The covariance operator \mathcal{C} can be identified with the following bilinear form*

$$\langle \eta', \mathcal{C}\eta \rangle_{U^\circ, U} = \mathbb{E}_\theta (\langle \eta, \theta \rangle_{U^\circ, U} \langle \eta', \theta \rangle_{U^\circ, U}), \quad \eta, \eta' \in U^\circ. \quad (3.7)$$

Proof. Let $\eta, \eta' \in U^\circ$. Then,

$$\begin{aligned} \mathbb{E}_\theta (\langle \eta, \theta \rangle_{U^\circ, U} \langle \eta', \theta \rangle_{U^\circ, U}) &= \mathbb{E}_\theta \left(\left\langle \eta, \sum_{i=1}^{\infty} \lambda_i^{1/2} \xi_i \varphi_i \right\rangle_{U^\circ, U} \left\langle \eta', \sum_{j=1}^{\infty} \lambda_j^{1/2} \xi_j \varphi_j \right\rangle_{U^\circ, U} \right) \\ &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \mathbb{E}_\xi (\lambda_i^{1/2} \lambda_j^{1/2} \xi_i \xi_j \langle \eta, \varphi_i \rangle_{U^\circ, U} \langle \eta', \varphi_j \rangle_{U^\circ, U}) \\ &= \sum_{i=1}^{\infty} \mathbb{E}_\xi (\lambda_i \langle \eta, \varphi_i \rangle_{U^\circ, U} \langle \eta', \varphi_i \rangle_{U^\circ, U}) \\ &= \left\langle \eta', \sum_{i=1}^{\infty} \lambda_i \langle \eta, \varphi_i \rangle_{U^\circ, U} \varphi_i \right\rangle_{U^\circ, U} = \langle \eta', \mathcal{C}\eta \rangle_{U^\circ, U}. \end{aligned}$$

\square

3.2 Finding the tensor decomposition

Let $\{\eta_i\}_{i \in \mathbb{N}}$ be a basis of U° and $\{u_i\}_{i \in \mathbb{N}} \subset U$ the corresponding coefficient functionals satisfying $\|\eta_i\| = \|u_i\| = 1$. Then, by Theorem A.15, the system $\{\eta_{i(k)} \otimes_s \eta_{j(k)}\}_{k \in \mathbb{N}}$, where the sequence of indices $\{i(k), j(k)\}_{k \in \mathbb{N}}$ corresponds to the ordering (A.7), is a basis in $\widehat{\otimes}_{\varepsilon_s}^s U^\circ$ and $\{u_{i(k)} \otimes_s u_{j(k)}\}_{k \in \mathbb{N}}$ are the corresponding coefficient functionals [10].

Since, by Theorem A.14, the projective product $\widehat{\otimes}_{\pi_s}^s U$ is isometrically embedded into $(\widehat{\otimes}_{\varepsilon_s}^s U^\circ)^*$ and $c \in \widehat{\otimes}_{\pi_s}^s U$, it can be expanded in the weak-* sense as follows

$$c = \sum_{k=1}^{\infty} \langle \eta_{i(k)} \otimes_s \eta_{j(k)}, c \rangle u_{i(k)} \otimes_s u_{j(k)} =: \sum_{k=1}^{\infty} c_{i(k)j(k)} u_{i(k)} \otimes_s u_{j(k)} \quad \text{weakly-}^*. \quad (3.8)$$

Treating the pairing (3.1) as the pairing between $\widehat{\otimes}_{\varepsilon_s}^s U^\circ$ and $(\widehat{\otimes}_{\varepsilon_s}^s U^\circ)^*$, we get

$$\begin{aligned} \langle \eta', \mathcal{C}\eta \rangle_{U^\circ, U} &= \sum_{k=1}^{\infty} c_{i(k)j(k)} \langle \eta \otimes_s \eta', u_{i(k)} \otimes_s u_{j(k)} \rangle \\ &= \frac{1}{2} \sum_{k=1}^{\infty} c_{i(k)j(k)} (\langle \eta, u_{j(k)} \rangle \langle \eta', u_{i(k)} \rangle + \langle \eta, u_{i(k)} \rangle \langle \eta', u_{j(k)} \rangle). \end{aligned}$$

This can be also written as follows

$$\mathcal{C}\eta = \sum_{k=1}^{\infty} c_{i(k)j(k)} \frac{\langle \eta, u_{j(k)} \rangle u_{i(k)} + \langle \eta, u_{i(k)} \rangle u_{j(k)}}{2} \quad \text{weakly-}^*. \quad (3.9)$$

To obtain a diagonal representation (3.3) of the covariance operator \mathcal{C} , we need to find a biorthogonal system $\{\tilde{\eta}_i, \mathcal{C}\tilde{\eta}_i\}_{i \in \mathbb{N}}$ satisfying $\|\tilde{\eta}_i\| = 1$ such that

$$\langle \tilde{\eta}_i, \mathcal{C}\tilde{\eta}_j \rangle_{U^\diamond, U} = \begin{cases} \|\mathcal{C}\tilde{\eta}_i\| & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

In other words, we need to biorthogonalise the basis $\{\eta_i\}_{i \in \mathbb{N}}$ with respect to the symmetric positive semidefinite operator \mathcal{C} , which can be done using Gram-Schmidt biorthogonalisation (e.g., [15]). Letting

$$\varphi_i := \frac{\mathcal{C}\tilde{\eta}_i}{\|\mathcal{C}\tilde{\eta}_i\|} \quad \text{and} \quad \lambda_i := \|\mathcal{C}\tilde{\eta}_i\|,$$

we obtain the desired representation (3.3).

4 Relevant spaces

In this section, we apply the above framework to two particular non-separable spaces, the space of Radon measures and the space of Hölder continuous functions.

4.1 Sampling Radon measures

We let $U = \mathcal{M}(\Omega)$ be the space of Radon measures on $(\Omega, \mathcal{B}\Omega)$, where Ω is compact, and $U^\diamond = C(\Omega)$, the space of continuous functions on Ω .

By [23, Thm. 5.25], the covariance operator \mathcal{C} can be written using a *representing measure* $c \in \mathfrak{M}(\Omega, \mathcal{M}(\Omega))$, where $\mathfrak{M}(\Omega, \mathcal{M}(\Omega))$ is the space of vector-valued with values in $\mathcal{M}(\Omega)$

$$\mathcal{C}f = \int_{\Omega} f \, dc \in \mathcal{M}(\Omega), \quad f \in C(\Omega), \quad (4.1)$$

The following result holds [23, Prop. 5.30].

Theorem 4.1. *The operator \mathcal{C} defined above is nuclear if and only if its representing measure c has the Radon-Nikodým property, i.e. $c \in \mathfrak{M}_1(\Omega, \mathcal{M}(\Omega))$. In this case*

$$\|\mathcal{C}\|_{\mathcal{N}} = \|c\|_{\mathfrak{M}}.$$

By Theorem A.10 we have that $\mathfrak{M}_1(\Omega, \mathcal{M}(\Omega)) = \mathcal{M}(\Omega) \widehat{\otimes}_{\pi} \mathcal{M}(\Omega)$, hence this is just another way of writing (3.2).

The Radon-Nikodým property can be ensured by construction. Let $\nu \in \mathcal{P}(\Omega)$ be a probability measure on Ω and $g \in L^1_{\nu}(D, \mathcal{M}(\Omega))$ a Bochner integrable function. Then

$$dc := g \, d\nu \in \mathfrak{M}_1(\Omega, \mathcal{M}(\Omega)) \quad (4.2)$$

has the Radon-Nikodým property and $\|\mathcal{C}\|_{\mathcal{N}} = \|g\|_{L^1}$.

There are many ways to construct a basis in $C(\Omega)$. For a cube $\Omega = [0, 1]^n$, we mention the basis of *Faber-Schauder functions* [25], see also [24]. Let

$$\psi(x) := \begin{cases} 1 - |x|, & x \in [-1, 1], \\ 0 & \text{otherwise} \end{cases} \quad (4.3)$$

be the distance function of the interval $[-1, 1]$. Consider the following dyadic system on the interval $[0, 1]$

$$D_0 := \{0, 1\}; \quad D_k := \{(2p-1)2^{-k}\}_{p=1, \dots, 2^{k-1}}; \quad D := \bigcup_{k=0}^{\infty} D_k,$$

and let

$$D_k^n := \{\tau = (\tau_1, \dots, \tau_n) \in D^n : \tau_i \in \bigcup_{j=0}^k D_j \text{ and } \exists i_0 \text{ s.t. } \tau_{i_0} \in D_k\}.$$

That is, k is the highest resolution in D_k^n . For any multi-index $\tau \in D_k^n$ define

$$f_\tau(x) := \prod_{i=1}^n \psi(2^k(x_i - \tau_i)), \quad x = (x_1, \dots, x_n) \in [0, 1]^n. \quad (4.4)$$

By [24, Prop. 7.1], the system $\{f_\tau\}_{\tau \in D^n}$ is a basis in $C([0, 1]^n)$. It is, in fact, a system of wavelets whose mother wavelet is the distance function (4.3). This basis will play an important role in Section 4.2, where we will consider sampling in Lipschitz and Hölder spaces.

The coefficient functionals corresponding to this basis are as follows [24, Prop. 7.1]. Let $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n) \in \{-1, 1\}^n$, $\tau = (\tau_1, \dots, \tau_n) \in D_k^n$ and $\tau^\varepsilon = (\tau_1^\varepsilon, \dots, \tau_n^\varepsilon)$, where

$$\tau_i^\varepsilon := \begin{cases} \tau + \varepsilon_i 2^{-k}, & \tau_i \in D_k, \\ \tau_i, & \tau_i \in \bigcup_{j=0}^{k-1} D_j. \end{cases}$$

The coefficient functionals $\{\mu_\tau\}_{\tau \in D^n} \subset \mathcal{M}(\Omega)$ corresponding to the basis functions (4.4) are

$$\mu_\tau := \begin{cases} \delta_\tau, & \tau \in D_0^n, \\ 2^{-n} \sum_{\varepsilon \in \{-1, 1\}^n} (\delta_\tau - \delta_{\tau^\varepsilon}), & \tau \in \bigcup_{j=1}^\infty D_j^n, \end{cases} \quad (4.5)$$

where δ_x is the Dirac measure at $x \in \Omega$. Hence, the coefficient functionals correspond to either point evaluations or differences of point evaluations.

We now consider some examples for Gaussian measures on the space of Radon measures.

Example 4.2 (Gaussian covariance). *A very widely used covariance operator for Gaussian measures is the so-called Gaussian (or square-exponential) covariance \mathcal{C} , which is usually defined by*

$$\mathcal{C}f(x') = \int_{\Omega} f(x) \exp\left(-\frac{1}{2}\|x - x'\|^2\right) dx.$$

In the terms of this section, we can write it in the following way. Let μ be the Lebesgue measure on $(D, \mathcal{B}D)$ and let $g \in L^1_\mu(D, \mathcal{M}(\Omega))$ be given by

$$D \ni z \mapsto k \cdot \mathbb{N}(z, \text{Id}_d)(\cdot \cap D),$$

for some constant $k > 0$. Hence, the measure-valued density g can be written as a Gaussian measure truncated on Ω multiplied by a prefactor with variable mean.

The samples of the Gaussian measure with Gaussian covariance are highly regular and, thus, may not be the best example for sampling a Radon measure. Instead, we can consider the following example, which gives a definition of Gaussian white noise on the space of Radon measures.

Example 4.3 (Gaussian white noise). *Let $\mu \in \mathcal{M}(\Omega)$ be a positive, finite measure on Ω . We consider the covariance kernel $c \in \mathfrak{M}_1(\Omega, \mathcal{M}(\Omega))$, where c is given by*

$$\mathcal{B}\Omega \times \mathcal{B}\Omega \ni (A, B) \mapsto \mu(A \cap B).$$

When testing the corresponding operator with predual functions $\eta_1, \eta_2 \in U^\diamond$, of course, we obtain

$$\langle \eta_2, \mathcal{C}\eta_1 \rangle_{U^\diamond, U} = \int_{\Omega} \eta_1 \eta_2 d\mu.$$

One can easily see that c satisfies the Radon-Nikodym property, by setting $\nu := \mu$ and

$$g := (\Omega \ni \omega \mapsto (\mathcal{B}\Omega \ni A \mapsto \delta_\omega(A))).$$

If either Ω is countable and μ is the counting measure or Ω contains an open set and μ is the n -dimensional Lebesgue measure, we refer to a random field with covariance kernel c as Gaussian white noise. Otherwise, we speak of spatially inhomogeneous Gaussian white noise. Gaussian white noise can also be defined on other function spaces, see [16].

4.2 Sampling Hölder functions

We start this section by recalling some facts from the theory of Lipschitz spaces. Let (X, d) or simply X be a compact metric space of diameter at most 2. Let $\text{Lip}(X) \subset C(X)$ be the space of Lipschitz continuous functions $X \rightarrow \mathbb{R}$ equipped with the following norm

$$\|f\|_{\text{Lip}} := \max\{\|f\|_{\infty}, L(f)\},$$

where $\|\cdot\|_{\infty}$ is the supremum norm and

$$L(f) := \sup_{x, x' \in X} \frac{|f(x) - f(x')|}{d(x, x')}$$

is the Lipschitz constant of f . If $0 < \alpha < 1$ then the space $\text{Lip}(X^{\alpha})$ is the space of α -Hölder continuous functions on X with respect to the original metric d (as well as the space of Lipschitz functions with respect to d^{α}).

If $e \in X$ is a distinguished *base point* (that is, X is a pointed metric space), then the subspace of all functions that vanish at e is denoted by

$$\text{Lip}_0(X) := \{f \in \text{Lip}(X) : f(e) = 0\}.$$

An equivalent norm on Lip_0 is given by the Lipschitz constant

$$\|f\|_{\text{Lip}_0} := L(f).$$

The next result shows that Lip spaces can be thought of as a certain special case of Lip_0 spaces.

Theorem 4.4 ([32, Prop. 2.13]). *Let (X, d) be a complete metric space whose diameter is at most 2 and let X^e be a pointed metric space consisting of X together with, as base point, a new element $\{e\}$*

$$X^e := X \cup \{e\}$$

and a new metric d' such that $d'(x, x') = d(x, x')$ for all $x, x' \in X$ and $d'(x, e) = 1$ for all $x \in X$. Then $\text{Lip}(X, d)$ can be naturally identified with $\text{Lip}_0(X^e, d')$.

Lipschitz spaces are always dual spaces, and in many cases the predual is unique. Our case where the underlying metric space has finite diameter is one of such cases [32, Sec. 3.4].

The predual of $\text{Lip}(X)$ is known as the Arens-Eells space or the Lipschitz-free space and can be seen as the completion of the space of zero-mean Radon measures $\mathcal{M}_0(X)$ with respect to the Kantorovich-Rubinstein norm

$$\|\mu\|_{KR} := \sup \left\{ \int_X u \, d\mu : u \in \text{Lip}(X), L(u) \leq 1, \|u\|_{\infty} \leq 1 \right\}.$$

Another expression for this norm and more details can be found in [32, Ch. 3].

Weak* convergence in $\text{Lip}(X)$ can be characterised as follows.

Theorem 4.5 ([32, Thm. 2.37 and Prop. 2.39]). *Let X be a pointed metric space of finite diameter. Then on bounded sets in $\text{Lip}(X)$ its weak* topology coincides with the topology of pointwise convergence. If X is compact, then it also coincides with the topology of uniform convergence.*

For $0 < \alpha < 1$, we denote by X^{α} the space (X, d^{α}) . It can be easily seen that $\text{Lip}(X^{\alpha})$ is the space of α -Hölder functions on X with respect to the original metric d . Since X has a finite diameter, $\text{Lip}(X^{\alpha}) \subset \text{Lip}(X^{\alpha'})$ for $0 < \alpha < \alpha' \leq 1$.

The space $\text{Lip}(X^{\alpha})$, $0 < \alpha < 1$, has a second predual known as the *little Lipschitz space* $\text{lip}(X^{\alpha})$ [32, Ch. 4], which consists of Lipschitz functions $f \in \text{Lip}(X^{\alpha})$ such that

$$\sup_{\substack{x, x' \in X \\ 0 < d(x, x') < \delta}} \frac{|f(x) - f(x')|}{d^{\alpha}(x, x')} \rightarrow 0 \quad \text{as } \delta \rightarrow 0. \quad (4.6)$$

Such functions are called *locally flat*. For $\alpha = 1$, constants are the only functions satisfying this condition. For $0 < \alpha < 1$ many functions satisfy this condition; for example, all piecewise linear functions on the interval $[0, 1]$ are locally flat. The norm in $\text{lip}(X^\alpha)$ coincides with that of $\text{Lip}(X^\alpha)$.

The following result holds.

Theorem 4.6 ([32, Thm. 8.49]). *Suppose that there exists a bi-Lipschitz embedding of the metric space X^α into \mathbb{R}^n for some n . Then the spaces $\text{lip}(X^\alpha)$, $\mathbb{A}(X^\alpha)$ and $\text{Lip}(X^\alpha)$ are linearly homeomorphic to the sequence spaces c_0 , ℓ^1 and ℓ^∞ , respectively.*

Remark 4.7. *In particular, Theorem 4.6 implies that both $\text{lip}(X^\alpha)$ and $\mathbb{A}(X^\alpha)$ have bases, and all three spaces have the metric approximation property (perhaps, upon switching to an equivalent norm).*

4.2.1 Hölder functions on a unit cube

If $X = [0, 1]^n$ equipped with the Euclidean metric, then by [24, Prop. 7.1] (see also [5]), Faber–Schauder functions (4.4) form a basis of $\text{lip}(X^\alpha)$, while (4.5) are the corresponding coefficient functionals. Note that as written in (4.4), Faber–Schauder functions are normalised in $C([0, 1]^n)$ but not in $\text{lip}(X^\alpha)$.

Each function $t \mapsto \psi(2^k t)$ is linear on the intervals $[-2^{-k}, 0]$ and $[0, 2^{-k}]$ and zero outside these intervals, hence we have the following estimate for its norm in $\text{lip}([0, 2^{-k}]^\alpha)$, where $\alpha \in (0, 1)$,

$$\|\psi(2^k \cdot)\|_{\text{lip}([0, 2^{-k}]^\alpha)} = \sup_{t, t' \in [0, 2^{-k}]} \frac{2^k |t' - t|}{|t' - t|^\alpha} = 2^k \sup_{t, t' \in [0, 2^{-k}]} |t' - t|^{1-\alpha} = 2^{\alpha k}.$$

Due to the product structure of (4.4) we have that

$$\|f_\tau\|_{\text{lip}(X^\alpha)} = 2^{\alpha k}, \quad \tau \in D_k^n,$$

where k is the highest resolution in D_k^n . Thus, the renormalised Faber–Schauder functions are given by

$$\tilde{f}_\tau(x) := 2^{-\alpha k} \prod_{i=1}^n \psi(2^k(x_i - \tau_i)), \quad \tau \in D_k^n, \quad x = (x_1, \dots, x_n) \in [0, 1]^n. \quad (4.7)$$

and the renormalised coefficient functionals by

$$\tilde{\mu}_\tau := \begin{cases} \delta_\tau, & \tau \in D_0^n, \\ 2^{\alpha k - n} \sum_{\varepsilon \in \{-1, 1\}^n} (\delta_\tau - \delta_{\tau^\varepsilon}), & \tau \in \bigcup_{j=1}^\infty D_j^n, \end{cases} \quad \tau \in D_k^n. \quad (4.8)$$

We have the following

Proposition 4.8. *Let $f \in \text{Lip}(X^\alpha)$ and \tilde{f}_τ and $\tilde{\mu}_\tau$ as defined in (4.7) and (4.8), respectively. Then*

$$f = \sum_\tau \langle \tilde{\mu}_\tau, f \rangle \tilde{f}_\tau \quad \text{weakly-}^* \text{ in } \text{Lip}(X^\alpha), \quad (4.9)$$

which by Theorem 4.5 is equivalent to uniform convergence. If, in addition, $f \in \text{lip}(X^\alpha)$, the convergence is in the Lipschitz norm.

Proof. The second statement is trivial, since we already know that $\{\tilde{f}_\tau\}_\tau$ is a basis in $\text{lip}(X^\alpha)$ and $\{\tilde{\mu}_\tau\}_\tau$ are the corresponding coefficient functionals. By [12, Thm. 5.21], the system $\{\tilde{\mu}_\tau\}_\tau$ is a basis of $\overline{\text{span}}\{\tilde{\mu}_\tau\}_\tau$. It is easy to see that the system $\{\tilde{\mu}_\tau\}_\tau$ is dense in $\mathbb{A}(X^\alpha)$

$$\overline{\text{span}}\{\tilde{\mu}_\tau\}_\tau = \mathbb{A}(X^\alpha),$$

hence $\{\tilde{\mu}_\tau\}_\tau$ is a basis of $\mathbb{A}(X^\alpha)$, whose coefficient functionals are given by $\{\tilde{f}_\tau\}_\tau$. Therefore, for any $\eta \in \mathbb{A}(X)$ and any $f \in \text{Lip}(X^\alpha)$ we have

$$\langle \eta, f \rangle = \left\langle \sum_\tau \langle \eta, \tilde{f}_\tau \rangle \tilde{\mu}_\tau, f \right\rangle = \sum_\tau \langle \eta, \tilde{f}_\tau \rangle \langle \tilde{\mu}_\tau, f \rangle = \left\langle \eta, \sum_\tau \langle \tilde{\mu}_\tau, f \rangle \tilde{f}_\tau \right\rangle,$$

which proves (4.9). \square

We now turn to the projective tensor product $\text{Lip}(X^\alpha) \widehat{\otimes}_\pi \text{Lip}(X^\alpha)$. Using the canonical identification of $\text{Lip}(X^\alpha \times X^\alpha)$ with the space of vector-valued functions $\text{Lip}(X^\alpha; \text{Lip}(X^\alpha))$ as well as the identification $\text{Lip}(X^\alpha; \text{Lip}(X^\alpha)) \cong \mathcal{L}(\mathcal{A}(X^\alpha), \text{Lip}(X^\alpha))$ (e.g., [9]), we get that

$$\begin{aligned} \text{Lip}(X^\alpha) \widehat{\otimes}_\pi \text{Lip}(X^\alpha) &= \mathcal{N}(\mathcal{A}(X^\alpha), \text{Lip}(X^\alpha)) \\ &\subset \mathcal{L}(\mathcal{A}(X^\alpha), \text{Lip}(X^\alpha)) = \text{Lip}(X^\alpha; \text{Lip}(X^\alpha)) = \text{Lip}(X^\alpha \times X^\alpha). \end{aligned}$$

It is well known that Hölder spaces are isomorphic to certain Besov spaces [31]. More precisely,

$$\text{Lip}(X^\alpha \times X^\alpha) \cong B_{\infty, \infty}^\alpha(X \times X).$$

Therefore, $\text{Lip}(X^\alpha) \widehat{\otimes}_\pi \text{Lip}(X^\alpha) \subset B_{\infty, \infty}^\alpha(X \times X)$. A concrete description of the projective product does not seem to be known, but the following result provides a subspace.

Proposition 4.9. *Let $g \in B_{1,1}^\alpha(X \times X)$. Then $g \in \text{Lip}(X^\alpha) \widehat{\otimes}_\pi \text{Lip}(X^\alpha)$.*

Proof. Similarly to (4.9), we have the following expansion for any $g \in \text{Lip}(X^\alpha \times X^\alpha)$

$$g = \sum_{\tau, t} \langle \tilde{\mu}_\tau \otimes \tilde{\mu}_t, g \rangle \tilde{f}_\tau \otimes \tilde{f}_t \quad \text{weakly-* in } \text{Lip}(X^\alpha \times X^\alpha). \quad (4.10)$$

Since $\|\tilde{f}_\tau \otimes \tilde{f}_t\|_\pi = \|\tilde{f}_\tau\|_{\text{Lip}} \|\tilde{f}_t\|_{\text{Lip}} = 1$, we have

$$\|g\|_\pi \leq \sum_{\tau, t} |\langle \tilde{\mu}_\tau \otimes \tilde{\mu}_t, g \rangle| < \infty, \quad (4.11)$$

where we applied Hölder's inequality to (4.10).

Up to a multiplication by $(1 - \alpha)$, the functions \tilde{f}_τ are atoms of the Besov space $B_{1,1}^\alpha(X)$ in the sense of [31, Def. 2.17] (with $p = 1$ and $s = \sigma = \alpha$; see also Remark 2.14 in the same book). Therefore, (4.11) is equivalent to the condition $g \in B_{1,1}^\alpha(X \times X)$. \square

4.2.2 Exponential covariance kernels

We consider the following family of exponential kernels on the unit cube $X = [0, 1]^n$

$$c_\alpha(x, x') := e^{-\frac{d^{2\alpha}(x, x')}{2}} = e^{-\frac{\|x - x'\|^{2\alpha}}{2}}, \quad 0 < \alpha < 1, \quad (4.12)$$

where $d^\alpha := \|x - x'\|^\alpha$ defines a metric. The kernels (4.12) are Gaussian kernels on metric spaces $X^\alpha := (X, d^\alpha)$. The kernels (4.12) are symmetric and, by [26, Cor. 3], positive definite. Furthermore, the following result holds.

Proposition 4.10. *Let $X = [0, 1]^n$ and $c_\alpha: X \times X \rightarrow \mathbb{R}$, $0 < \alpha < 1$, a family of functions as defined in (4.12). Then*

- (i) $c_\alpha \in \text{Lip}(X^\alpha \times X^\alpha)$ for all $0 < \alpha < 1$ and $c_\alpha \in \text{lip}(X^\gamma \times X^\gamma)$ for all $0 < \gamma < \alpha$;
- (ii) $c_\alpha \notin \text{Lip}(X^\beta \times X^\beta)$ for any $\alpha < \beta \leq 1$ and $c_\alpha \notin \text{lip}(X^\alpha \times X^\alpha)$.

Proof. A proof can be found in Appendix B. \square

Corollary 4.11. *Combining this result with Proposition 4.8, we conclude that the following series converges in the γ -Hölder norm for any $0 < \gamma < \alpha$*

$$c_\alpha = \sum_{\tau, t} \langle \tilde{\mu}_\tau \otimes \tilde{\mu}_t, c_\alpha \rangle \tilde{f}_\tau \otimes \tilde{f}_t \quad \text{strongly in } \text{lip}(X^\gamma \times X^\gamma).$$

Next we shall investigate whether the kernels (4.12) satisfy the assumptions of Proposition 4.9.

Proposition 4.12. *The kernels (4.12) satisfy*

$$c_\alpha \in B_{1,1}^\gamma(X \times X), \quad 0 < \gamma < \alpha,$$

and $c_\alpha \notin B_{1,1}^\alpha(X \times X)$.

Proof. We use the formula [31, Def. 9.12] (see also Remarks 9.13 and 9.27 in the same book). Since $c_\alpha \in \text{Lip}(X^\alpha \times X^\alpha)$, we obtain the following estimate for modulus of continuity of c_α

$$\begin{aligned}\omega(c_\alpha, t) &= \sup_{\|(h, h')\| \leq t} \int_{X \times X} |c_\alpha(x+h, x'+h') - c_\alpha(x, x')| dx dx' \\ &\leq \sup_{\|(h, h')\| \leq t} \int_{X \times X} C \|(h, h')\|^\alpha dx dx' = Ct^\alpha\end{aligned}$$

for some $C > 0$. Now we get

$$\begin{aligned}\|c_\alpha\|_{B_{1,1}^\gamma} &= \|c_\alpha\|_{L^1} + \int_0^1 t^{-\gamma} \omega(c_\alpha, t) \frac{dt}{t} \leq \|c_\alpha\|_{L^1} + C \int_0^1 t^{\alpha-1-\gamma} dt \\ &= \|c_\alpha\|_{L^1} + \frac{C}{\alpha-\gamma},\end{aligned}$$

which is finite for any $\gamma < \alpha$. If $\gamma = \alpha$, the integral diverges and, since $c_\alpha \notin \text{Lip}(X^\beta \times X^\beta)$ for any $\beta > \alpha$, we conclude that $\|c_\alpha\|_{B_{1,1}^\alpha} = \infty$. \square

Combining Proposition 4.12 and Proposition 4.9 with the results of Section 3.1, we conclude that exponential covariance kernels c_α defined in (4.12) produce samples that are γ -Hölder continuous for all $\gamma \in (0, \alpha)$.

Remark 4.13. *Let $\Omega \subseteq \mathbb{R}$. For $\alpha = 1/2$, the kernel (4.12) is the exponential covariance kernel $c_{1/2}(x, x') = e^{-\frac{\|x-x'\|}{2}}$, which describes the covariance of a stationary Ornstein–Uhlenbeck process. Using Theorem 1.2, one can show that a path of an Ornstein–Uhlenbeck process admits a modification that is γ -Hölder continuous for $\gamma \in (0, 1/2)$, which is in agreement with our result.*

4.3 A note on sampling continuous functions

In this section, we briefly discuss how sampling in the space of continuous functions fits into the framework presented above.

Let $U = C(\Omega)$, where $\Omega \subset \mathbb{R}^n$ is compact, and $c \in \widehat{\otimes}_{\pi_s}^s C(\Omega)$ a covariance kernel. Since $C(\Omega)$ does not have a predual, we need to choose U^* as the domain of the the covariance operator. By Theorem A.7 we have that $\mathcal{N}(U^*, U) \cong U^{**} \widehat{\otimes}_{\pi} U \subset U^{**} \widehat{\otimes}_{\pi} U^{**}$. To retain the symmetry of the covariance kernel c , we therefore need to consider it as an element of larger space, $c \in \widehat{\otimes}_{\pi_s}^s C^{**}(\Omega)$. As a result, the covariance operator will act as $\mathcal{C}: U^* \rightarrow U^{**}$. In accordance with this, we need to modify (3.1) as follows

$$\mathcal{C}: \langle v', \mathcal{C}v \rangle := \langle c, v \otimes_s v' \rangle, \quad v, v' \in U^*, \quad (4.13)$$

where the second pairing is between the bilinear form c and the tensor $v \otimes_s v'$.

From now on, we proceed similarly to Section 4.2. Since $c \in \widehat{\otimes}_{\pi_s}^s C(\Omega) \subset C(\Omega \times \Omega)$, we can expand the covariance kernel in the basis of Faber-Schauder functions (4.4)

$$c = \sum_{\tau, t} \langle c, \mu_\tau \otimes \mu_t \rangle f_\tau \otimes f_t \quad \text{strongly in } C(\Omega \times \Omega),$$

where $\mu_{\tau, t}$ are the coefficient functionals from (4.5). Proceeding as in Section 3.2, we can diagonalise this tensor representation and obtain Gaussian samples θ as in (3.4).

Remark 4.14. *We emphasise that samples obtained in this way will not necessarily lie in $C(\Omega)$. Indeed, by Theorem 3.3 the sum (3.4) converges only weakly- $*$ in $C^{**}(\Omega)$ (with probability 1). A representation theorem for $C^{**}([0, 1])$ can be found in [18].*

5 Outlook

We finish with a few open questions and possible directions for future research.

Second preduals. First, we go back to Section 4.2. Here, we have encountered a situation where the space $U = \text{Lip}(\Omega)$ has a second predual, the little Lipschitz space. The same happens in Section 4.3 with $U = C^{**}(\Omega)$. A natural question is now, whether it is possible in this case to strengthen Proposition 3.2 and Theorem 3.3 and show that $\theta \in U$ with probability 1.

Random fields with jumps. In this work we have focused on Gaussian measures and the generalisation to other probability distributions is interesting. Of special interest are random samples with jumps, e.g. piecewise continuous or piecewise constant samples where the subdomains on which the sampled functions are continuous/constant are also random. Chada et al. [6], for instance, discuss Cauchy random fields. It would be highly interesting to study such non-Gaussian random fields on, e.g., spaces of functions of bounded variation.

Conditioning. Gaussian processes are of particular interest in Bayesian statistics and data science as it is possible in linear settings to determine conditional mean and covariance in closed form. Scovel and Owhadi [27] have studied the conditioning of Gaussian random fields on Hilbert spaces. A natural next goal is to generalise their theory to our setting of Banach spaces.

Acknowledgements

The authors thank Onur Oktay (Usak University) for suggesting the idea of Proposition 4.9 and pointing us to the reference [31], and Nik Weaver (Washington University in St. Louis) for pointing out Theorem 8.49 in [32].

The authors would like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the programme “Mathematics of deep learning” where part of this work was undertaken. This work was supported by the EPSRC grant EP/R014604/1.

YK acknowledges support of the EPSRC (Fellowship EP/V003615/1) and the Cantab Capital Institute for the Mathematics of Information. CBS acknowledges support from the Philip Leverhulme Prize, the Royal Society Wolfson Fellowship, the EPSRC advanced career fellowship EP/V029428/1, EPSRC grants EP/S026045/1 and EP/T003553/1, EP/N014588/1, EP/T017961/1, the Wellcome Innovator Award RG98755, the European Union Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 777826 NoMADS, the Cantab Capital Institute for the Mathematics of Information and the Alan Turing Institute.

Rights retention and data access

There is no research data associated with this paper. For the purpose of open access, the authors have submitted a preprint to arxiv and will update it as appropriate.

References

- [1] R. J. Adler. *Random fields and geometry*. New York, NY: Springer New York, 2007.
- [2] R. J. Adler. “The geometry of random fields”. 2010.
- [3] P. Billingsley. *Probability and measure*. 3rd ed. John Wiley & Sons, 1995.
- [4] V. I. Bogachev. *Gaussian measures*. AMS, 1998.
- [5] R. Bonic, J. Frampton, and A. Tromba. “ Λ -manifolds”. *Journal of functional analysis* 3 (2 1969), pp. 310–320.
- [6] N. K. Chada, L. Roininen, and J. Suuronen. “Cauchy Markov Random Field Priors for Bayesian Inversion”. *Arxiv e-prints*, arXiv:2105.12488 (2021), arXiv:2105.12488.
- [7] N. Dunford and J. T. Schwartz. *Linear operators, part I: general theory*. Hoboken, NJ, 1958.

- [8] K. Floret. “Natural norms on symmetric tensor products of normed spaces”. *Note di matematica* 17 (1997), pp. 1X–188.
- [9] L. García-Lirola, C. Petitjean, and A. Rueda Zoca. “On the structure of spaces of vector-valued lipschitz functions”. arxiv:1606.05999. 2016.
- [10] B. C. Grecu and R. A. Ryan. “Schauder bases for symmetric tensor products”. *Publications of the research institute for mathematical sciences* 41.2 (2005), pp. 459–469.
- [11] M. Hairer. “Introduction to stochastic pdes”. *Lecture notes, university of warwick* (2009).
- [12] C. Heil. *A basis theory primer*. Birkhäuser, 2011.
- [13] O. Kallenberg. *Foundations of modern probability*. Cham: Springer International Publishing, 2021.
- [14] A. Klenke. “Brownian motion”. *Probability theory: a comprehensive course*. London: Springer London, 2014, pp. 457–508.
- [15] L. Kohaupt. “Introduction to a gram-schmidt-type biorthogonalization method”. *The rocky mountain journal of mathematics* 44.4 (2014), pp. 1265–1279.
- [16] H.-H. Kuo. *White Noise Distribution Theory*. CRC Press, 1996.
- [17] G. J. Lord, C. E. Powell, and T. Shardlow. *An introduction to computational stochastic pdes*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2014.
- [18] R. D. Mauldin. “A representation theorem for the second dual of $C[0, 1]$ ”. *Studia mathematica* 46 (1973), pp. 197–200.
- [19] J. Mercer. “XVI. Functions of positive and negative type, and their connection the theory of integral equations”. *Philosophical transactions of the royal society of london. series a, containing papers of a mathematical or physical character* 209.441-458 (1909), pp. 415–446.
- [20] J. Potthoff. “Sample properties of random fields I: continuity”. *Communications on stochastic analysis* 3.3 (2009), pp. 331–348.
- [21] J. Potthoff. “Sample properties of random fields I: separability and measurability”. *Communications on stochastic analysis* 3.1 (2009), pp. 143–153.
- [22] J. Potthoff. “Sample properties of random fields III: differentiability”. *Communications on stochastic analysis* 4.3 (2010), pp. 335–353.
- [23] R. A. Ryan. *Introduction to tensor products of banach spaces*. Springer, 2002.
- [24] J Ryll. “Schauder bases for the space of continuous functions on an n -dimensional cube”. *Commentationes mathematicae* 17.1 (1973).
- [25] J. Schauder. “Zur theorie stetiger abbildungen in funktionalräumen”. *Mathematische zeitschrift* 26.1 (1927), pp. 47–65.
- [26] I. J. Schoenberg. “Metric spaces and positive definite functions”. *Transactions of the american mathematical society* 44 (1938), pp. 522–536.
- [27] C Scovel and H Owhadi. “Conditioning gaussian measure on hilbert space”. *Journal of mathematical and statistical analysis* 1 (109 2018).
- [28] A. M. Stuart. “Inverse problems: A Bayesian perspective”. *Acta numerica*. Vol. 19. Cambridge University Press, 2010, pp. 451–559.
- [29] A. M. Stuart and A. L. Teckentrup. “Posterior consistency for gaussian process approximations of bayesian posterior distributions”. *Mathematics of computation* 87 (2018), pp. 721–753.
- [30] T. J. Sullivan. *Introduction to uncertainty quantification*. Cham: Springer International Publishing, 2015.
- [31] H. Triebel. *Theory of function spaces III*. Birkhäuser Basel, 2006.
- [32] N. Weaver. *Lipschitz algebras*. 2nd ed. World Scientific Publishing, 2018.

A A few facts about tensor products of Banach spaces

Our approach relies on tensor products of Banach spaces. We will briefly recall some important definitions and facts. In our exposition, we will follow [23] and [8].

Let E and F be Banach spaces. By $E \otimes F$ we denote the algebraic tensor product of E and F , i.e. the space of linear functionals on the space of bilinear forms on $E \times F$. For every $e \in E$, $f \in F$ we denote by $e \otimes f$ the following functional

$$(e \otimes f)(\mathcal{A}) = \langle \mathcal{A}, e \otimes f \rangle := \mathcal{A}(e, f),$$

where \mathcal{A} is an arbitrary bilinear form on $E \times F$. A typical tensor in $E \otimes F$ has the form

$$\omega = \sum_{i=1}^n \lambda_i e_i \otimes f_i, \quad (\text{A.1})$$

where $e_i \in E$ and $f_i \in F$ satisfy $\|e_i\| = \|f_i\| = 1$, $i = 1, \dots, n$, and λ_i are scalars.

There are many ways, in which the tensor product $E \otimes F$ can inherit the Banach space structure of E and F , giving rise to different topological tensor products. We will need the following one.

Definition A.1 (Projective tensor product). *Let E and F be Banach spaces and $E \otimes F$ their algebraic tensor product. For every tensor $\omega \in E \otimes F$ let*

$$\pi(\omega) := \inf \left\{ \sum_{i=1}^n |\lambda_i| : \omega = \sum_{i=1}^n \lambda_i e_i \otimes f_i \right\},$$

be the projective norm of ω , where the infimum is taken over all possible representations of ω in the form (A.1). The completion of $E \otimes F$ with respect to this norm is called the projective tensor product of E and F and denoted by

$$E \widehat{\otimes}_{\pi} F.$$

Definition A.2 (Injective tensor product). *Let E and F be Banach spaces and $E \otimes F$ their algebraic tensor product. For every tensor $\omega \in E \otimes F$ let*

$$\varepsilon(\omega) := \sup \left\{ \left| \sum_{i=1}^n \lambda_i \langle e_i, e^* \rangle \langle f_i, f^* \rangle \right| : e^* \in B_{E^*}, f^* \in B_{F^*} \right\},$$

satisfy $\|e_i\| = \|f_i\| = 1$ be the injective norm of ω , where $\omega = \sum_{i=1}^n \lambda_i e_i \otimes f_i$ is any representation of ω and B_{E^*}, B_{F^*} are the unit balls in E^* and F^* , respectively. The completion of $E \otimes F$ with respect to this norm is called the injective tensor product of E and F and denoted by

$$E \widehat{\otimes}_{\varepsilon} F.$$

The following result gives a useful representation of the elements of a projective tensor product.

Theorem A.3 ([23, Prop. 2.8]). *Let E and F be Banach spaces. Let $\omega \in E \widehat{\otimes}_{\pi} F$. Then there exist sequences $\{e_i\}_{i \in \mathbb{N}} \subset E$ and $\{f_i\}_{i \in \mathbb{N}} \subset F$ satisfying $\|e_i\| = \|f_i\| = 1$ such that $\omega = \sum_{i=1}^{\infty} \lambda_i e_i \otimes f_i$ and*

$$\pi(\omega) = \inf \left\{ \sum_{i=1}^n |\lambda_i| : \omega = \sum_{i=1}^{\infty} \lambda_i e_i \otimes f_i \right\},$$

where the infimum is taken over all possible representations of ω .

To every bilinear form \mathcal{A} on $E \times F$ corresponds a linear operator $A: E \rightarrow F^*$ defined as follows

$$\langle f, Ae \rangle = \mathcal{A}(e, f), \quad e \in E, f \in F.$$

Hence, one can also speak of tensor products in terms of linear operators. If E^* is a dual space, then under certain conditions the projective tensor product $E^* \widehat{\otimes}_{\pi} F$ can be identified with the space of nuclear operators $\mathcal{N}(E, F)$.

Definition A.4 (Nuclear operators on Banach spaces). *Let E and F be Banach spaces. An operator $N: E \rightarrow F$ is called nuclear if it can be written in the following form*

$$Ne = \sum_{i=1}^{\infty} \lambda_i \langle e, e_i^* \rangle f_i, \quad e \in E, \quad (\text{A.2})$$

where $\{e_i^*\}_{i \in \mathbb{N}} \subset E^*$ and $\{f_i\}_{i \in \mathbb{N}} \subset F$ satisfy $\|e_i^*\| = \|f_i\| = 1$ and $\{\lambda_i\}_{i \in \mathbb{N}} \subset \mathbb{R}$. The nuclear norm of N is given by

$$\|N\|_{\mathcal{N}} := \inf\{\|\lambda\|_{\ell^1} : Nx = \sum_{i=1}^{\infty} \lambda_i \langle x, e_i^* \rangle f_i \quad \forall x \in E\}, \quad (\text{A.3})$$

where the infimum is taken over all representations of the form (A.2).

Definition A.5 (Approximation property). *Let E be a Banach space. If for any compact set $K \subset E$ and every $\varepsilon > 0$ there exists a finite rank operator $S: E \rightarrow E$ such that for every $e \in K$ it holds that $\|e - Se\| \leq \varepsilon$, then E is said to have the approximation property. If, in addition, $\|S\| \leq 1$ then E is said to have the metric approximation property.*

Example A.6. *The spaces $C(K)$ (continuous functions on a compact K), $\mathcal{M}(K)$ (Radon measures on K), sequence spaces ℓ^p for $1 \leq p \leq \infty$, Lebesgue spaces $L^p(\mu)$ for $1 \leq p \leq \infty$ have the metric approximation property [23]. All Banach spaces with a basis can be equipped with an equivalent norm, under which they will have the metric approximation property.*

Theorem A.7 ([23, Cor. 4.8]). *Let E and F be Banach spaces. If either E^* or F has the approximation property, then*

$$\mathcal{N}(E, F) = E^* \widehat{\otimes}_{\pi} F.$$

If E and F possess bases, then the tensor products $E \widehat{\otimes}_{\pi} F$ and $E \widehat{\otimes}_{\varepsilon} F$ naturally inherit them. Let $\{e_i\}_{i \in \mathbb{N}}$ and $\{f_j\}_{j \in \mathbb{N}}$ be the bases of E and F , respectively. Let us order the tensor products $e_i \otimes f_j$ as shown in the following diagram (see [10, 23])

$$\begin{array}{ccccccc} e_1 \otimes f_1 & & e_1 \otimes f_2 & & e_1 \otimes f_3 & & \\ & & \uparrow & & \uparrow & & \\ e_2 \otimes f_1 & \rightarrow & e_2 \otimes f_2 & & e_2 \otimes f_3 & \dots & \\ & & & & \uparrow & \uparrow & \\ e_3 \otimes f_1 & \rightarrow & e_3 \otimes f_2 & \rightarrow & e_3 \otimes f_3 & & e_3 \otimes f_4 \\ & & & & & \uparrow & \\ e_4 \otimes f_1 & \rightarrow & e_4 \otimes f_2 & \rightarrow & e_4 \otimes f_3 & \rightarrow & e_4 \otimes f_4 \end{array} \quad (\text{A.4})$$

This ordering is called the *square ordering* and can be written as follows

$$e_1 \otimes f_1, e_2 \otimes f_1, e_2 \otimes f_2, e_1 \otimes f_2, e_3 \otimes f_1, e_3 \otimes f_2, e_3 \otimes f_3, e_2 \otimes f_3, \dots$$

Theorem A.8 ([23, Prop. 4.25]). *Let E and F be Banach spaces with bases $\{e_i\}_{i \in \mathbb{N}}$ and $\{f_j\}_{j \in \mathbb{N}}$, respectively. Then the sequence $e_i \otimes f_j$ with square ordering is a basis for both $E \widehat{\otimes}_{\pi} F$ and $E \widehat{\otimes}_{\varepsilon} F$, and is referred to as the tensor product basis.*

Let $\mathfrak{M}(\Omega, E)$ denote the space of vector measures on Ω with values in E . For every $\mu \in \mathfrak{M}(\Omega, E)$, let

$$|\mu|_1(D) := \sup \left\{ \sum_{i=1}^n \|\mu(D_i)\|_E : \{D_1, \dots, D_n\} \text{ is a partition of } D \right\}, \quad D \in \mathcal{B}\Omega,$$

denote the *variation* of μ . The norm on $\mathfrak{M}(\Omega, E)$, referred to as the *variation norm*, is given by

$$\|\mu\|_{\mathfrak{M}} := |\mu|_1(\Omega).$$

Definition A.9 (Radon-Nikodým property). *Let E be a Banach space. A measure $\mu \in \mathfrak{M}(\Omega, E)$ is said to have the Radon-Nikodým property if μ has bounded variation and there exists a Bochner-integrable function $g: \Omega \rightarrow E$ with respect to $|\mu|$, called the Radon-Nikodým derivative $\frac{d\mu}{d|\mu|}$, such that*

$$\mu(D) = \int_D g d|\mu| \quad \forall D \in \mathcal{B}\Omega.$$

Let $\mathfrak{M}_1(\Omega, E) \subset \mathfrak{M}(\Omega, E)$ be the subspace of all measures with the Radon-Nikodm property. By [23, Lem. 5.21], $\mathfrak{M}_1(\Omega, E)$ is complete under the variation norm.

Theorem A.10 ([23, Thm. 5.22]). *Let E be a Banach space. Then the projective tensor product $\mathcal{M}(\Omega) \widehat{\otimes}_\pi E$ is isometrically isomorphic to the Banach space $\mathfrak{M}_1(\Omega, E)$ of vector measures with the Radon-Nikodm property*

$$\mathcal{M}(\Omega) \widehat{\otimes}_\pi E = \mathfrak{M}_1(\Omega, E).$$

Now we turn to symmetric tensor products. We will follow [8]. Consider symmetric bilinear forms on $E \times E$, i.e. such that $\mathcal{A}(e, e') = \mathcal{A}(e', e)$ for all $e, e' \in E$. The algebraic dual of this space is called the symmetric tensor product of E with itself and will be denoted by

$$\otimes^{2,s} E.$$

Every element $\omega \in \otimes^{2,s} E$ has the following representation

$$\omega = \sum_{i=1}^n \lambda_i e_i \otimes e_i \tag{A.5}$$

where $e_i \in E$ satisfies $\|e_i\| = 1$, $i \in \mathbb{N}$, and λ_i are scalars.

The symmetric projective product $\widehat{\otimes}_{\pi_s}^s E$ and symmetric injective product $\widehat{\otimes}_{\varepsilon_s}^s E$ are defined analogously to Definitions A.1 and A.2.

The following representation holds.

Theorem A.11 ([8, Prop. 2.2]). *Let E be a Banach space and $\omega \in \widehat{\otimes}_{\pi_s}^s E$. Then there exists a sequence $\{e_i\}_{i \in \mathbb{N}} \subset E$ satisfying $\|e_i\| = 1$ such that $\omega = \sum_{i=1}^{\infty} \lambda_i e_i \otimes e_i$ and*

$$\pi_s(\omega) = \inf \left\{ \sum_{i=1}^{\infty} |\lambda_i| : \omega = \sum_{i=1}^{\infty} \lambda_i e_i \otimes e_i \right\},$$

where the infimum is taken over all possible representations of ω .

Remark A.12. *By [8, Prop 2.3], the symmetric projective product $\widehat{\otimes}_{\pi_s}^s E$ is a complemented subspace of the “full” projective product $E \widehat{\otimes}_\pi E$.*

Similarly as the full projective product can be identified with the space of nuclear operators, the symmetric projective product can be identified with a subspace of this space.

Theorem A.13 (similar to [8, Prop. 4.3]). *Let E be a Banach space and suppose that its dual E^* has the approximation property. Then the symmetric projective tensor product $\widehat{\otimes}_{\pi_s}^s E^*$ can be identified with the following subspace of the space of nuclear operators $\mathcal{N}(E, E^*)$*

$$\mathcal{N}^s(E, E^*) := \{N \in \mathcal{N}(E, E^*) : Ne = \sum_{i=1}^{\infty} \lambda_i \langle e, e_i^* \rangle e_i^*, \quad e \in E\}, \tag{A.6}$$

where $\{e_i^*\}_{i \in \mathbb{N}} \subset E^*$ is some sequence that satisfies $\|e_i^*\| = 1$ and $\sum_{i=1}^n |\lambda_i| < \infty$.

A nuclear operator $N \in \mathcal{N}^s(E, E^*)$ is called positive semidefinite if

$$\langle e, Ne \rangle_{E, E^*} \geq 0, \quad e \in E.$$

It is clear that $N \in \mathcal{N}^s(E, E^*)$ is positive semidefinite if and only if the expansion coefficients λ_i in (A.6) satisfy $\lambda_i \geq 0$ for all n .

Theorem A.14 ([8, Thm. 4.6]). *Let E be a Banach space such that its dual E^* has the metric approximation property. Then the following embedding*

$$\widehat{\otimes}_{\pi_s}^s E^* \hookrightarrow (\widehat{\otimes}_{\varepsilon_s}^s E)^*$$

is a metric injection.

Let $e \otimes_s e'$ denote the symmetric tensor product of $e, e' \in E$

$$e \otimes_s e' := \frac{1}{2}(e \otimes e' + e' \otimes e).$$

If $\{e_i\}_{i \in \mathbb{N}}$ is a basis in E then a basis in $\widehat{\otimes}_{\pi_s}^s E$ and $\widehat{\otimes}_{\varepsilon_s}^s E$ is obtained with the following ordering of the symmetric tensor products $e_i \otimes_s e_j$, compare with (A.4)

$$\begin{aligned} & e_1 \otimes e_1 \\ & e_2 \otimes_s e_1 \quad \rightarrow \quad e_2 \otimes e_2 \\ & e_3 \otimes_s e_1 \quad \rightarrow \quad e_3 \otimes_s e_2 \quad \rightarrow \quad e_3 \otimes e_3 \\ & e_4 \otimes_s e_1 \quad \rightarrow \quad e_4 \otimes_s e_2 \quad \rightarrow \quad e_4 \otimes_s e_3 \quad \rightarrow \quad \dots \end{aligned} \tag{A.7}$$

Theorem A.15 ([10]). *Let $\{e_i\}_{i \in \mathbb{N}}$ be a basis of a Banach space E . Then the sequence $e_i \otimes_s e_j$ with the above ordering is a basis for both $\widehat{\otimes}_{\pi_s}^s E$ and $\widehat{\otimes}_{\varepsilon_s}^s E$.*

B Proof of Proposition 4.10

Proof. (i) Fix $y \in X$ and consider the function $x \mapsto c_\alpha(x, y)$. The following estimate holds for any $x, x' \in X$ such that $d(x, y) \leq d(x', y)$

$$\begin{aligned} \left| e^{-d^{2\alpha}(x, y)} - e^{-d^{2\alpha}(x', y)} \right| &= e^{-d^{2\alpha}(x, y)} \left| 1 - e^{-(d^{2\alpha}(x', y) - d^{2\alpha}(x, y))} \right| \\ &= e^{-d^{2\alpha}(x, y)} \left(1 - e^{-(d^{2\alpha}(x', y) - d^{2\alpha}(x, y))} \right) \\ &\leq e^{-d^{2\alpha}(x, y)} (d^{2\alpha}(x', y) - d^{2\alpha}(x, y)) \\ &\leq e^{-d^{2\alpha}(x, y)} (d^\alpha(x', y) + d^\alpha(x, y)) (d^\alpha(x', y) - d^\alpha(x, y)) \\ &\leq e^{-d^{2\alpha}(x, y)} (d^\alpha(x', y) + d^\alpha(x, y)) d^\alpha(x, x'), \end{aligned}$$

where in the second line we used the inequality $e^t - 1 \geq t$ valid for all $t \in \mathbb{R}$, and in the last line we use the reverse triangle inequality with respect to the metric d^α . Proceeding similarly in the case $d(x', y) \leq d(x, y)$, we get

$$\left| e^{-d^{2\alpha}(x, y)} - e^{-d^{2\alpha}(x', y)} \right| \leq e^{-(\min\{d(x, y), d(x', y)\})^{2\alpha}} (d^\alpha(x', y) + d^\alpha(x, y)) d^\alpha(x, x'). \tag{B.1}$$

Dividing this estimate by $d^\alpha(x, x')$ we get

$$\sup_{p, p' \in X} \frac{\left| e^{-d^{2\alpha}(p, y)} - e^{-d^{2\alpha}(p', y)} \right|}{d^\alpha(p, p')} \leq \sup_{x, x' \in X} e^{-(\min\{d(x, y), d(x', y)\})^{2\alpha}} (d^\alpha(x', y) + d^\alpha(x, y)) \leq c_\alpha,$$

where the constant c_α depends of the diameter of X . Hence, $c_\alpha(\cdot, y) \in \text{Lip}(X^\alpha)$. Similarly, dividing (B.1) by $d^\gamma(x, x')$ with $0 < \gamma < \alpha$, we get

$$\frac{\left| e^{-d^{2\alpha}(x, y)} - e^{-d^{2\alpha}(x', y)} \right|}{d^\gamma(x, x')} \leq e^{-(\min\{d(x, y), d(x', y)\})^{2\alpha}} (d^\alpha(x', y) + d^\alpha(x, y)) d^{\alpha-\gamma}(x, x') \rightarrow 0$$

as $d(x, x') \rightarrow 0$. Therefore, $c_\alpha(\cdot, y) \in \text{lip}(X^\gamma)$ for all $0 < \gamma < \alpha$.

Using the canonical identification of $\text{Lip}(X^\alpha \times X^\alpha)$ with the space of vector-valued Lipschitz functions $\text{Lip}(X^\alpha, \text{Lip}(X^\alpha))$ and repeating the computations with appropriate minor modifications, we obtain the claim.

- (ii) Fix $y \in X$ and let $x, x' \in X$ be such that $d(x', y) \leq d(x, y)$ and $d^\alpha(x, y) = d^\alpha(x, x') + d^\alpha(x', y)$ (that is, x' lies on a d^α -geodesic connecting x and y). Then

$$\begin{aligned}
\left| e^{-d^{2\alpha}(x', y)} - e^{-d^{2\alpha}(x, y)} \right| &= e^{-d^{2\alpha}(x, y)} \left| e^{-(d^{2\alpha}(x', y) - d^{2\alpha}(x, y))} - 1 \right| \\
&= e^{-d^{2\alpha}(x, y)} \left(e^{-(d^{2\alpha}(x', y) - d^{2\alpha}(x, y))} - 1 \right) \\
&\geq e^{-d^{2\alpha}(x, y)} (d^{2\alpha}(x, y) - d^{2\alpha}(x', y)) \\
&= e^{-d^{2\alpha}(x, y)} (d^\alpha(x, y) + d^\alpha(x', y)) (d^\alpha(x, y) - d^\alpha(x', y)).
\end{aligned}$$

Dividing both sides by $d^\beta(x, x')$, $\alpha \leq \beta \leq 1$, we get

$$\begin{aligned}
\frac{\left| e^{-d^{2\alpha}(x', y)} - e^{-d^{2\alpha}(x, y)} \right|}{d^\beta(x, x')} &\geq e^{-d^{2\alpha}(x, y)} (d^\alpha(x, y) + d^\alpha(x', y)) \frac{d^\alpha(x, y) - d^\alpha(x', y)}{d^\beta(x, x')} \\
&= e^{-d^{2\alpha}(x, y)} (d^\alpha(x, y) + d^\alpha(x', y)) \frac{d^\alpha(x, x')}{d^\beta(x, x')}.
\end{aligned}$$

Consider two sequences $\{x_n, x'_n\}_{n \in \mathbb{N}} \subset X$ such that $d(x_n, y) \rightarrow 0$ while $d(x'_n, y) \leq d(x_n, y)$ and $d^\alpha(x_n, y) = d^\alpha(x_n, x'_n) + d^\alpha(x'_n, y)$. That is, $x_n, x'_n \rightarrow y$ along the d^α -geodesic connecting x and y . Then we have

$$\begin{aligned}
\sup_{p, p' \in X} \frac{\left| e^{-d^{2\alpha}(p, y)} - e^{-d^{2\alpha}(p', y)} \right|}{d^\beta(x, x')} &\geq \limsup_{n \in \mathbb{N}} \left(e^{-d^{2\alpha}(x_n, y)} (d^\alpha(x_n, y) + d^\alpha(x'_n, y)) \frac{d^\alpha(x_n, x'_n)}{d^\beta(x_n, x'_n)} \right) \\
&= \infty
\end{aligned}$$

for $\alpha < \beta \leq 1$. Hence, $c_\alpha(\cdot, y) \notin \text{Lip}(X^\beta)$. Similarly, taking $\beta = \alpha$, we get

$$\begin{aligned}
\sup_{\substack{p, p' \in X \\ 0 < d(p, p') < \delta}} \frac{\left| e^{-d^{2\alpha}(p, y)} - e^{-d^{2\alpha}(p', y)} \right|}{d^\beta(x, x')} &\geq \limsup_{n \in \mathbb{N}} \left(e^{-d^{2\alpha}(x_n, y)} (d^\alpha(x_n, y) + d^\alpha(x'_n, y)) \frac{d^\alpha(x_n, x'_n)}{d^\alpha(x_n, x'_n)} \right) \\
&= \tilde{c}_\alpha > 0 \quad \text{for all } \delta > 0,
\end{aligned}$$

where the constant \tilde{c}_α depends of the diameter of X . Therefore, $c_\alpha(\cdot, y) \notin \text{lip}(X^\alpha)$.

Using the canonical identification of $\text{Lip}(X^\alpha \times X^\alpha)$ with $\text{Lip}(X^\alpha, \text{Lip}(X^\alpha))$, we extend this to the function $c_\alpha(\cdot, \cdot)$, which completes the proof. \square

Remark B.1. One can see from the proof that the result actually holds for any metric space X (which has to be path-connected for the second statement).

Corollary B.2. As a corollary, we get the following estimate valid for all $x, x', y \in X$ and any $0 < \alpha \leq 1$

$$\begin{aligned}
(d^\alpha(x', y) + d^\alpha(x, y)) e^{-(\max\{d(x, y), d(x', y)\})^{2\alpha}} &\leq \frac{\left| e^{-d^{2\alpha}(x, y)} - e^{-d^{2\alpha}(x', y)} \right|}{d^\alpha(x, x')} \\
&\leq (d^\alpha(x', y) + d^\alpha(x, y)) e^{-(\min\{d(x, y), d(x', y)\})^{2\alpha}}.
\end{aligned}$$