Intrinsic Gaussian Vector Fields on Manifolds

Daniel Robert-Nicoud **Andreas Krause** **Viacheslav Borovitskiy**

ETH Zürich, Switzerland

Abstract

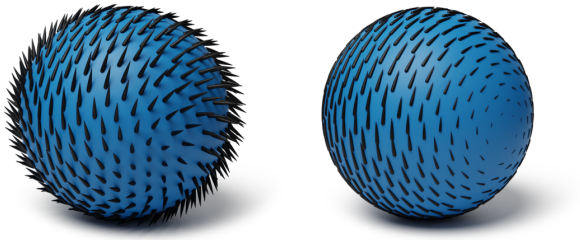
Various applications ranging from robotics to climate science require modeling signals on non-Euclidean domains, such as the sphere. Gaussian process models on manifolds have recently been proposed for such tasks, in particular when uncertainty quantification is needed. In the manifold setting, vector-valued signals can behave very differently from scalar-valued ones, with much of the progress so far focused on modeling the latter. The former, however, are crucial for many applications, such as modeling wind speeds or force fields of unknown dynamical systems. In this paper, we propose novel Gaussian process models for vector-valued signals on manifolds that are intrinsically defined and account for the geometry of the space in consideration. We provide computational primitives needed to deploy the resulting *Hodge–Matérn Gaussian vector fields* on the two-dimensional sphere and the hypertori. Further, we highlight two generalization directions: discrete two-dimensional meshes and "ideal" manifolds like hyperspheres, Lie groups, and homogeneous spaces. Finally, we show that our Gaussian vector fields constitute considerably more refined inductive biases than the extrinsic fields proposed before.

1 INTRODUCTION

Gaussian processes (Rasmussen and Williams, 2006) are a widely used class of Bayesian models in machine learning. They are known to perform well in small data scenarios and to provide well-calibrated uncertainty. Their notable applications include modeling spatial data (Chilès and Delfiner, 2012) and automated decision-making, e.g., optimization (Noskova and Borovitskiy, 2023; Shields et al., 2021; Snoek et al., 2012), or sensor placement (Krause et al., 2008).

Gaussian processes can be scalar- or vector-valued (Alvarez et al., 2012). The important special case of the latter is *Gaussian vector fields*. These can for example be used to model velocities or accelerations, either as a target in itself or as means for exploring an unknown dynamical system. When the input domain is Euclidean, vector fields are just vector-valued functions. However, when the domain is a submanifold of a Euclidean space, such as when modeling wind speeds or ocean currents on the surface of Earth, the situation can be quite different: geometry places additional constraints on vector fields that need to be accounted for. As illustrated in Figure 1, the values of a *vector field* ought to be tangential to the manifold, while those of a vector function can be arbitrary vectors.

In recent years, two different formalisms were proposed for defining Gaussian vector fields on manifolds. Lange-Hegermann (2018) approached the problem by considering linear constraints on vector-valued Gaussian processes, constraining them to lie in the tangent space of a submanifold of $\mathbb{R}^d$. Additionally, one can



(a) Vector Function (b) Vector Field

Figure 1: Vector fields on manifolds are not just vector functions, for them the vectors are always *tangential*.

Proceedings of the 27th International Conference on Artificial Intelligence and Statistics (AISTATS) 2024, Valencia, Spain. PMLR: Volume 238.

Code: https://github.com/DanielRobertNicoud/imv-gps. Also, see https://github.com/GPflow/GeometricKernels.

Correspondence to: daniel.robertnicoud@gmail.com and viacheslav.borovitskiy@gmail.com.

impose further linear constraints to the resulting fields, such as making them divergence-free. Hutchinson et al. (2021)—which is closer in spirit to this work—instead considered projecting vector-valued Gaussian processes to a submanifold of $\mathbb{R}^d$. While both these procedures can in principle produce any valid Gaussian vector field, they are fundamentally *extrinsic* and we will show that the fields one gets in practice introduce undesirable inductive biases.

To remedy this, we propose a new approach: fully intrinsic Gaussian vector fields based on the *Hodge Laplacian*[1] that we name *Hodge–Matérn Gaussian vector fields*. For some simple manifolds, namely for the two-dimensional sphere $\mathbb{S}_2$ and for the hypertori $\mathbb{T}^d$, we develop computational techniques that allow effortless use of these intrinsic fields in downstream applications.

The aforementioned computational techniques hinge on knowing the *eigenvalues* and *eigenfields* of the Hodge Laplacian. To this end, we describe how to:

a) derive these from the eigenpairs of the Laplace–Beltrami operator when the manifold is two-dimensional, using automatic differentiation only;

b) compute these on product manifolds in terms of the eigenvalues and eigenfields on the factors; and

c) get these explicitly in the cases of the circle $\mathbb{S}_1$, the hypertori $\mathbb{T}^d$, and the sphere $\mathbb{S}_2$.

We *conjecture* that (a) can also be used to define Gaussian vector fields on meshes, by changing the analytic notions into their appropriate discretizations. Furthermore, we *conjecture* that (c) can be done for more general *parallelizable* manifolds, e.g. on *Lie groups*, which may then facilitate the generalization to the very general class of *homogeneous spaces*. The latter includes many manifolds of interest which are poorly amenable to discretization because of their higher dimension.

By showing how our intrinsic Hodge–Matérn Gaussian vector fields improve over their naïve extrinsic counterparts on the two-dimensional sphere, we hope to motivate further research. First, into the development of practical intrinsic Gaussian vector fields on other domains. Second, into applying the proposed models in areas like climate/weather modeling and robotics.

[1]The concurrent work by Peach et al. (2024) studies Gaussian vector fields induced by the *connection Laplacian* rather than the Hodge Laplacian (see Appendix A.7 on the difference between these two notions of Laplacian). Importantly, they consider a very different setting: a priori unknown manifolds that are estimated from finite data, showcasing an impressive neuroscience application.

1.1 Gaussian Processes

Let X be a set. A random function f on X is called a *Gaussian process* (GP) with mean $\mu : X \to \mathbb{R}$ and covariance (or kernel) $k : X \times X \to \mathbb{R}$—denoted by $f \sim \mathrm{GP}(\mu, k)$—if for any finite set of points $\boldsymbol{x}$ in X we have $f(\boldsymbol{x}) \sim \mathrm{N}(\mu(\boldsymbol{x}), \mathrm{K}_{\boldsymbol{x}\boldsymbol{x}})$ where $\mathrm{K}_{\bullet\bullet'} = k(\bullet, \bullet')$. Without loss of generality, we usually assume $\mu(\cdot) = 0$.

Assuming a GP prior $f \sim \mathrm{GP}(0, k)$ and a Gaussian likelihood $\boldsymbol{y} \mid f = \mathcal{N}(\boldsymbol{y} \mid f(\boldsymbol{x}), \sigma_\varepsilon^2)$ with a fixed noise variance σ_ε^2, the posterior $f \mid \boldsymbol{y}$ is a GP (Rasmussen and Williams, 2006) with mean and covariance

$$\mu_{f|\boldsymbol{y}}(\cdot) = \mathrm{K}_{\cdot\boldsymbol{x}} \left(\mathrm{K}_{\boldsymbol{x}\boldsymbol{x}} + \sigma_\varepsilon^2 \mathbf{I}\right)^{-1} \boldsymbol{y}, \tag{1}$$

$$k_{f|\boldsymbol{y}}(\cdot, \cdot') = k(\cdot, \cdot') - \mathrm{K}_{\cdot\boldsymbol{x}} \left(\mathrm{K}_{\boldsymbol{x}\boldsymbol{x}} + \sigma_\varepsilon^2 \mathbf{I}\right)^{-1} \mathrm{K}_{\boldsymbol{x}\cdot'}. \tag{2}$$

Here, the function $\mu_{f|\boldsymbol{y}}$ can be used to draw *predictions* and the function $k_{f|\boldsymbol{y}}$ is used to quantify *uncertainty*.

When $X = \mathbb{R}^d$, *Matérn Gaussian processes* (Rasmussen and Williams, 2006; Stein, 1999) are most often used. Their respective kernels k_{ν,κ,σ^2} are the three-parameter family of *Matérn kernels*, whose limiting case $k_{\infty,\kappa,\sigma^2}$ for $\nu \to \infty$ is known as the heat (a.k.a. squared exponential, RBF, Gaussian, diffusion) kernel, which is arguably the most popular.

1.2 Gaussian Processes on Manifolds

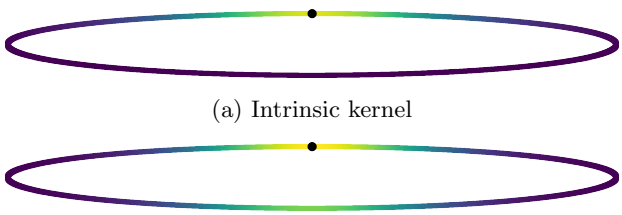
Now consider $X = \mathrm{M}$, where M is a compact Riemannian manifold. Throughout this paper, manifolds are always assumed to be connected. Using the SPDE-based characterization of Matérn Gaussian processes of Lindgren et al. (2011) and Whittle (1963), Borovitskiy et al. (2020) showed how to compute Matérn kernels on M in terms of the spectrum of the Laplace–Beltrami operator Δ:

$$k_{\nu,\kappa,\sigma^2}(x, x') = \frac{\sigma^2}{C_{\nu,\kappa}} \sum_{n=0}^{\infty} \Phi_{\nu,\kappa}(\lambda_n) f_n(x) f_n(x'), \tag{3}$$

where $\{f_n\}_{n=0}^{\infty}$ is an orthonormal basis of eigenfunctions of Δ such that $\Delta f_n = -\lambda_n f_n$,

$$\Phi_{\nu,\kappa}(\lambda) = \begin{cases} \left(\frac{2\nu}{\kappa} + \lambda\right)^{-\nu - d/2} & \nu < \infty, \\ e^{-\frac{\kappa^2}{2}\lambda} & \nu = \infty, \end{cases} \tag{4}$$

$d = \dim(\mathrm{M})$, and $C_{\nu,\kappa}$ is a normalization constant that ensures $\frac{1}{\operatorname{vol} \mathrm{M}} \int_{\mathrm{M}} k_{\nu,\kappa,\sigma^2}(x, x) \, \mathrm{d}x = \sigma^2$. There exist analytical (Azangulov et al., 2022; 2023) and numerical (Borovitskiy et al., 2020; Coveney et al., 2020) techniques for computing the eigenpairs λ_n, f_n, or bypassing the computation thereof. In the end, a truncated series from Equation (3) yields tractable Gaussian processes that respect the intrinsic geometry of the manifold (Rosa et al., 2023), as illustrated in Figure 2.



(a) Intrinsic kernel

(b) Extrinsic kernel

Figure 2: Comparing an intrinsic Matérn kernel ($\nu = \infty$) of Equation (3) to an *extrinsic* one, the restriction of a Euclidean Matérn kernel to the manifold. Note the latter induces high correlation between the points across the minor axis of the ellipse, despite them being far from each other in terms of the intrinsic distance.

Importantly, same as in the Euclidean case we have $k_{\infty,\kappa,\sigma^2}(x,x') \propto \mathcal{P}(\frac{\kappa^2}{2}, x, x')$ (Azangulov et al., 2022), where $\mathcal{P}(t,x,x')$ is the *heat kernel*: the solution of

$$\frac{\partial \mathcal{P}}{\partial t} = \Delta_x \mathcal{P}, \qquad \lim_{t\to 0} \mathcal{P}(t,x,x') = \delta(x-x'), \tag{5}$$

where Δ_x denotes the Laplace–Beltrami operator acting on the variable x and δ is the Dirac delta function; convergence takes place in the sense of distributions.

1.3 Gaussian Vector Fields on Manifolds

Following Hutchinson et al. (2021), we introduce the notions of Gaussian vector fields and their kernels.

Definition 1. *A* random vector field f *is a function mapping each* $x \in \mathrm{M}$ *to a random variable* $f(x)$ *with values in the* tangent space $T_x\mathrm{M}$ *of* M *at* x. *It is* Gaussian *if* $(f(x_1), \ldots, f(x_n)) \in T_{x_1}\mathrm{M} \oplus \ldots \oplus T_{x_n}\mathrm{M}$ *is jointly Gaussian for all* $x_1, \ldots, x_n \in \mathrm{M}$ *and* $n \in \mathbb{N}$.

When $\mathrm{M} = \mathbb{R}^d$, this reduces to vector-valued GPs with output dimension equal to the input dimension d, e.g. a concatenation of d independent scalar-valued GPs. If instead $\mathrm{M} \subseteq \mathbb{R}^d$ is an embedded manifold, concatenating d scalar-valued GPs does not lead to a vector field: in this case, $f(x) \in \mathbb{R}^d$ rather than $f(x) \in T_x\mathrm{M}$ as illustrated by Figure 1.

If $\boldsymbol{g}$ is a vector-valued GP on $\mathbb{R}^d$, then its values $\boldsymbol{g}(\boldsymbol{x})$ are Gaussian vectors. Thus, the kernel of $\boldsymbol{g}$ is a matrix-valued function $\boldsymbol{k}(\boldsymbol{x},\boldsymbol{x}') = \mathrm{Cov}(\boldsymbol{g}(\boldsymbol{x}), \boldsymbol{g}(\boldsymbol{x}'))$. Alternatively, it can be defined as a scalar-valued function $k((\boldsymbol{x},\boldsymbol{v}),(\boldsymbol{x}',\boldsymbol{u})) = \mathrm{Cov}(\boldsymbol{v}^\top \boldsymbol{g}(\boldsymbol{x}), \boldsymbol{u}^\top \boldsymbol{g}(\boldsymbol{x}'))$ where $\boldsymbol{x},\boldsymbol{x}',\boldsymbol{v},\boldsymbol{u} \in \mathbb{R}^d$. Reinterpreting the inner products $\boldsymbol{a}^\top \boldsymbol{b}$ as the linear functional $\bullet \to \boldsymbol{a}^\top \bullet$ applied to the vector $\boldsymbol{b}$ leads to a particularly elegant generalization. For a Gaussian vector field f on a manifold M we put

$$\boldsymbol{k}(x,x')(u,v) = k\big((x,v),(x',u)\big) \tag{6}$$
$$= \mathrm{Cov}\left(v(f(x)), u(g(x))\right), \tag{7}$$

where $x, x' \in M$, $u \in T_x^*\mathrm{M}, v \in T_{x'}^*\mathrm{M}$ and $T_x^*\mathrm{M}$ denotes the *cotangent* space of *covectors*, linear functionals on $T_x\mathrm{M}$. Together with a deterministic *mean vector field* μ, such a kernel k determines the distribution of f in a coordinate-free way, justifying the use of the standard notation $f \sim \mathrm{GP}(\mu, k)$.

In order to define Gaussian vector fields that can be used in practice, Hutchinson et al. (2021) introduced the notion of *projected Gaussian processes*. These are constructed by picking an isometric embedding $\phi : \mathrm{M} \to \mathbb{R}^D$ into some Euclidean space.[2] Then, a tangent space $T_x\mathrm{M}$ can be identified with a subspace of $T_{\phi(x)}\mathbb{R}^D \cong \mathbb{R}^D$. Thus, there exists a projection P_x from $\mathbb{R}^D$ to this subspace and $f(x) = \mathrm{P}_x\, \boldsymbol{g}(x)$ defines a valid Gaussian vector field for any vector-valued GP $\boldsymbol{g}$.

2 CHALLENGES

The class of projected GPs we reviewed in Section 1.3 is very large. In fact, any Gaussian vector field f can be represented as a projected GP, simply because $f = \mathrm{P}_x\, \boldsymbol{g}$ with $\boldsymbol{g} = f$, as proved in Hutchinson et al. (2021). However, in order to obtain f using this trick we need to know it in the first place.

To avoid this "chicken and egg" problem, we should construct $\boldsymbol{g}$ using the tools we already posses. Practically, this means stacking together scalar-valued GPs. The first challenge is to find scalar GPs that respect the geometry of the manifold. It can be solved by using the manifold Matérn GPs reviewed in Section 1.2. To obtain an expressive family, we can define $\boldsymbol{g} = \mathbf{A}\boldsymbol{h}$ where $h_j \sim \mathrm{GP}(0, k_{\nu,\kappa_j,\sigma^2})$ are independent and $\mathbf{A}$ is an arbitrary matrix. In fact, this is exactly what Hutchinson et al. (2021) propose to do in practice.

However, as can be seen in Figure 3, this construction can produce undesirable artifacts. Upon a closer examination, the reasons for this can be formalized.

Proposition 2. *With notation as above, if* $\mathrm{rk}\, \boldsymbol{A} > 1$, *there are* $x, x' \in \mathbb{S}_2$ *so that* $\angle(x,x') = 90°$ *but*

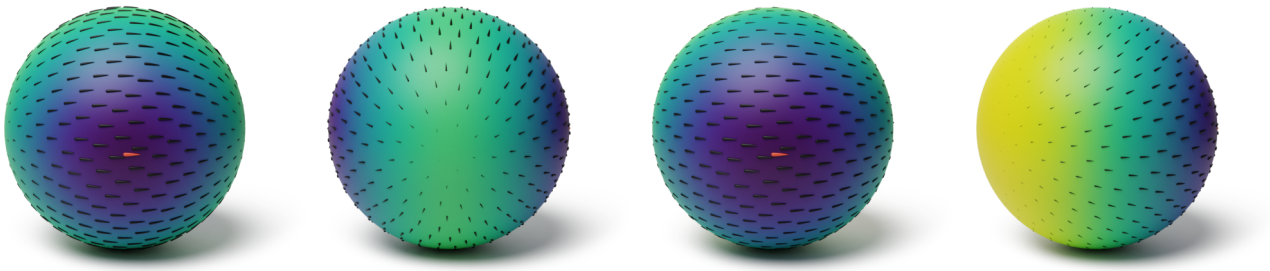
$$\|\mathrm{Cov}(f(x), f(x'))\| < \|\mathrm{Cov}(f(x), f(\tilde{x}))\|, \quad \kappa_j \to \infty, \tag{8}$$

where $\tilde{x} = -x$ *is antipodal to* x *(i.e.* $\angle(x,\tilde{x}) = 180°$*) and* $\|\bullet\|$ *denotes the Frobenius norm.*

For a proof, see Appendix B.1. Proposition 2 implies that, for large length scales κ, the covariance of a projected GP is non-monotonic in the intrinsic distance on the sphere—this is an undesirable trait, usually.

Another issue is that the projected GP construction does not provide a way to force divergence-free or curl-free inductive biases. These can be quite important for

[2] Such an embedding always exists by the *Nash embedding theorem*; it is not unique.



(a) Projected Matérn (b) Projected Matérn, rotated (c) Hodge–Matérn (d) Hodge–Matérn, rotated

Figure 3: GP regression from a single observation (the red vector) for a very large length scale κ. Black vectors represent the prediction $\mu_{f|\boldsymbol{y}}(\cdot)$, the background color shows the uncertainty $\left\|k_{f|\boldsymbol{y}}(\cdot,\cdot)\right\|$: yellow for high, blue for low. On (a) and (b) we use a projected Matérn GP of Hutchinson et al. (2021); on (c) and (d) we use our Hodge–Matérn Gaussian vector field, $\nu = \infty$ in both cases. Unnaturally, uncertainty in (a) and (b) is non-monotonous with respect to the distance on the sphere: it is considerably lower at the antipode than halfway to it.

certain types of vector field data (Berlinghieri et al., 2023), as we will clearly observe in Section 5.

To overcome these challenges, we introduce a fully intrinsic class of Gaussian vector fields on manifolds.

3 INTRINSIC GAUSSIAN VECTOR FIELDS

In this section, we present the main ideas behind the construction of the intrinsic Gaussian vector fields we propose. The mathematical formalism for this section is detailed in Appendix A.

3.1 The Hodge Heat Kernel

We start by generalizing the heat kernel, i.e. the Matérn kernel with $\nu = \infty$.

Let M be a compact, *oriented* Riemannian manifold. Then, *Hodge theory*—see e.g. Rosenberg (1997) for an approachable introduction—defines a generalization Δ of the Laplace–Beltrami operator that acts on vector fields on M instead of scalar functions, called the *Hodge Laplacian*. We consider the associated heat equation

$$\frac{\partial u}{\partial t}(t,x) = \Delta_x u(t,x), \tag{9}$$

where u is smooth in both variables and $u(t,x) \in T_x\mathrm{M}$. This equation admits a *fundamental solution* $\boldsymbol{\mathcal{P}}$: the *Hodge (heat) kernel* which for any choice of $t \in \mathbb{R}_{>0}$ and $x, x' \in M$ gives a function

$$\boldsymbol{\mathcal{P}}(t,x,x') : T_x^*\mathrm{M} \otimes T_{x'}^*\mathrm{M} \longrightarrow \mathbb{R}. \tag{10}$$

Considering t as a hyperparameter, we obtain a function $\boldsymbol{\mathcal{P}}_t(x,x')$ with the exact signature that a kernel of Gaussian vector field should posses, as by Section 1.3. We prove the following.

Theorem 3. *For any $t > 0$ there exists a Gaussian vector field whose kernel is $\boldsymbol{\mathcal{P}}_t$.*

Having found a suitable adaptation of the heat kernel to the vector field case, we turn to making it explicit. Similarly to the scalar case, a Hilbert space $\mathcal{L}^2(\mathrm{M}; T\mathrm{M})$ of square integrable vector fields can be defined. There is always an orthonormal basis $\{s_n\}_{n=0}^{\infty}$ of $\mathcal{L}^2(\mathrm{M}; T\mathrm{M})$ such that $\Delta s_n = -\lambda_n s_n$, i.e. s_n are eigenfields of the Hodge Laplacian. Moreover, we have $\lambda_n \geq 0$ for $n \geq 0$. The kernel $\boldsymbol{\mathcal{P}}_t$ can be computed in terms of the eigenfields just like its scalar counterpart (cf. (3)). Namely, as discussed in Appendix A.4, we have

$$\boldsymbol{\mathcal{P}}_t(x,x') = \sum_{n=0}^{\infty} e^{-t\lambda_n} s_n(x) \otimes s_n(x'), \tag{11}$$

Here, the notation means

$$(s_n(x) \otimes s_n(x'))(u,v) = u(s_n(x))\, v(s_n(x')) \tag{12}$$

for $u \in T_x^*\mathrm{M}$ and $v \in T_{x'}^*\mathrm{M}$.

3.2 Hodge–Matérn Kernels

Following Azangulov et al. (2022, 2023), we define Matérn kernels $\boldsymbol{k}_{\nu,\kappa,\sigma^2}$ as integrals of the heat kernel:

$$\boldsymbol{k}_{\nu,\kappa,\sigma^2}(x,x') = \frac{\sigma^2}{C_{\nu,\kappa}} \int_0^{\infty} t^{\nu-1+\frac{n}{2}} e^{-\frac{2\nu}{\kappa^2}t} \boldsymbol{\mathcal{P}}_t(x,x')\,\mathrm{d}t. \tag{13}$$

Fubini's theorem then readily implies the key formula

$$\boldsymbol{k}_{\nu,\kappa,\sigma^2}(x,x') = \frac{\sigma^2}{C_{\nu,\kappa}} \sum_{n=0}^{\infty} \Phi_{\nu,\kappa}(\lambda_n) s_n(x) \otimes s_n(x'), \tag{14}$$

where $\Phi_{\nu,\kappa}$ is as in (4) and $C_{\nu,\kappa}$ is a normalizing constant that ensures $\frac{1}{\mathrm{vol}\,\mathrm{M}} \int_{\mathrm{M}} \mathrm{tr}(\boldsymbol{k}_{\nu,\kappa,\sigma^2}(x,x))\,\mathrm{d}x = \sigma^2$. Same as the for Hodge heat kernel, these Hodge–Matérn kernels determine Gaussian vector fields.

Theorem 4. *For any $\nu, \kappa, \sigma^2 > 0$ there exists a Gaussian vector field whose kernel is $\boldsymbol{k}_{\nu,\kappa,\sigma^2}$.*

In practice, the series in Equation (11) should be truncated, with only a few terms corresponding to the smallest eigenvalues λ_n used to approximately compute the kernel, just like in the scalar case. Notice that the functions $\Phi_{\nu,\kappa}(\lambda)$ are all decreasing in λ, so that the most significant terms are the ones corresponding to the smallest eigenvalues.

3.3 Divergence-Free and Curl-Free Kernels

The celebrated *Helmholtz decomposition* (also known as the *fundamental theorem of vector calculus*) states that any vector field in $\mathbb{R}^d$ decomposes into the sum of its *divergence-free* and *curl-free* parts. The former, intuitively, has no sinks and sources; the latter has no vortexes. Many vector fields in physics are known to have only one of these parts. This suggests that divergence-free and curl-free Gaussian vector fields can be useful inductive biases (Berlinghieri et al., 2023).

For manifolds, the analog of Helmholtz decomposition is the Hodge decomposition, see e.g. Rosenberg (1997, Theorem 1.37). It states that any vector field u on M can be represented as a sum of three fields:

$$u = u_1 + u_2 + u_3, \tag{15}$$

where $u_1 = \nabla f_1$ for some function f_1, and thus is *pure divergence*—meaning that $\operatorname{div} u_1 \neq 0$ and $\operatorname{curl} u_1 = 0$—and in particular *curl-free*, $u_2 = \star\nabla f_2$, and thus is *pure curl* and *divergence-free*, and u_3 is a *harmonic form*, $\Delta u_3 = 0$, both curl- and divergence-free. The symbol $\star$ denotes the *Hodge star* operator, which we recall in Appendix A.1. For intuition on divergence and curl, see Appendix A.2.

Importantly, the orthonormal basis of eigenfields $\{s_n\}_{n=0}^{\infty}$ may be chosen in such a way that each s_n is in exactly one of the three classes above. Let $\mathcal{I}_{\mathrm{div}}$, $\mathcal{I}_{\mathrm{curl}}$, and $\mathcal{I}_{\mathrm{harm}}$ denote the index sets of the respective classes of eigenfields. Using a single class, we can define versions of Matérn Gaussian vector fields on the manifold M with the associated inductive bias.

Theorem 5. *There exists a Gaussian vector field $f^\bullet$ — where $\bullet \in \{\mathrm{div}, \mathrm{curl}, \mathrm{harm}\}$ —with kernel*

$$\boldsymbol{k}^{\bullet}_{\nu,\kappa,\sigma^2}(x,x') = \frac{\sigma^2}{C_{\nu,\kappa}} \sum_{n\in\mathcal{I}_\bullet} \Phi_{\nu,\kappa}(\lambda_n) s_n(x) \otimes s_n(x'). \tag{16}$$

What is more, $\operatorname{div} f^{\mathrm{curl}} = 0$, $\operatorname{curl} f^{\mathrm{div}} = 0$ *and* $\Delta f^{\mathrm{harm}} = 0$ *almost surely as long as $f^\bullet$ is smooth enough that* div *and* curl *are well-defined.*

3.4 Hodge-compositional Matérn Kernels

Combining the pure divergence, pure curl, and harmonic kernels, each with a separate set of hyperparameters, gives a more flexible family of kernels:

$$\sigma_1^2 \boldsymbol{k}^{\mathrm{div}}_{\nu,\kappa_1,1} + \sigma_2^2 \boldsymbol{k}^{\mathrm{curl}}_{\nu,\kappa_2,1} + \sigma_3^2 \boldsymbol{k}^{\mathrm{harm}}_{\nu,\kappa_3,1}. \tag{17}$$

By analogy with the concurrent paper by Yang et al. (2024), we call this family *Hodge-compositional Matérn kernels*. Unless prior knowledge suggests a more specialized choice, this is the family we recommend for use in practical applications, inferring all the hyperparameters κ_i, σ_i from data. Our experimental results in Section 5 support this recommendation.

3.5 Gaussian Process Regression

All kernels defined above fall under the umbrella framework of Gaussian vector fields described in Hutchinson et al. (2021). In practice, there are two ways to perform Gaussian process regression in this setting. The first one—if the manifold is embedded in $\mathbb{R}^D$—is to treat Gaussian vector fields as special cases of Gaussian vector functions in $\mathbb{R}^D$. The second is to introduce a *frame*, i.e. a coordinate choice (not necessarily smooth) in all of the tangent spaces, and describe all quantities in these coordinates. Depending on whether an embedding or a frame is available, either can be used. Both are merely modes of computation, not affecting the inductive biases of Gaussian vector fields and not introducing any error per se.

3.6 Kernel Evaluation and Sampling

As mentioned in the end of Section 3.2, given that eigenfields and eigenvalues are known, we can approximately evaluate the kernels by truncating the series in Equation (14). Such a truncation is a well-defined kernel, i.e. it corresponds to a Gaussian vector field.

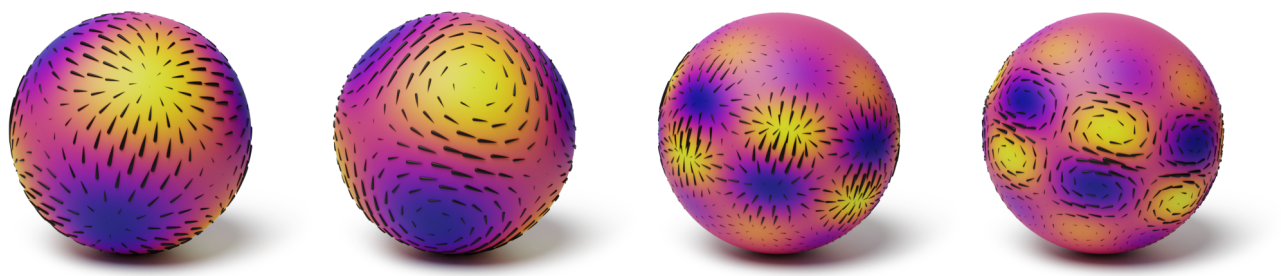
Proposition 6. *The Gaussian vector field*

$$f(x) = \frac{\sigma}{\sqrt{C_{\nu,\kappa}}} \sum_{n=0}^{L} w_n \sqrt{\Phi_{\nu,\kappa}(\lambda_n)} s_n(x), \tag{18}$$

where $w_n \overset{iid}{\sim} \mathcal{N}(0,1)$, corresponds to the kernel given by the truncation of Equation (14) with the sum $\sum_{n=0}^{\infty}$ therein substituted by the sum $\sum_{n=0}^{L}$.

Importantly, Equation (18) allows to approximately sample Hodge–Matérn Gaussian vector fields in an extremely computationally efficient way by simply drawing random $w_n \sim N(0,1)$. Efficiently sampling their respective posteriors can be performed using *pathwise conditioning* for Gaussian vector fields, as described in Hutchinson et al. (2021).

Remark 7. *Of course, direct analogs of Equation (18) also hold for the kernels of Theorem 5.*



(a) Eigenfield $\frac{1}{\sqrt{\lambda_3}}\nabla Y_{3,2}$ (b) Eigenfield $\frac{1}{\sqrt{\lambda_3}}\star\nabla Y_{3,2}$ (c) Eigenfield $\frac{1}{\sqrt{\lambda_7}}\nabla Y_{7,3}$ (d) Eigenfield $\frac{1}{\sqrt{\lambda_7}}\star\nabla Y_{7,3}$

Figure 4: Eigenfunctions on the sphere $\mathbb{S}_2$ (represented by color) and the respective eigenfields.

In summary, (approximate) sampling, kernel evaluation and differentiation reduce to knowing eigenvalues and eigenfields of the Hodge Laplacian on M. Thus, we proceed to discuss how to obtain those in practice.

4 EXPLICIT EIGEN-VALUES AND -FIELDS

The above allows defining intrinsic kernels on general compact oriented Riemannian manifolds. However, actually computing these kernel requires solving for eigenfields and eigenvalues of the Hodge Laplacian. Luckily, in some important cases this turns out to be tractable. We present them in this section.

4.1 Surfaces and the Sphere

The main case of interest here is the sphere $\mathbb{S}_2$. However, we start by considering the more general case of manifolds of dimension 2 (surfaces). We explain how to obtain the eigenfields and eigenvalues granted their scalar counterparts and a basis of harmonic vector fields are known.

Surfaces Suppose M is a compact, oriented Riemannian surface. We consider two intrinsic operators for this case. The first is the gradient of a scalar function, giving us a vector field. The second operator is the *Hodge star* operator $\star$ acting on vector fields, which in the case of surfaces is just a 90° rotation of a vector field in the positive direction, as shown in Appendix A.2.

Suppose we know all eigenfunctions $\{f_n\}_{n\geq 0}$ and their respective eigenvalues $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots$ of the Laplace–Beltrami operator on M. Further, assume that we have an orthonormal basis $\{g_j\}_{0\leq j\leq J}$ of the 0-eigenspace of the Hodge Laplacian.[3]

[3]By the Hodge decomposition theorem, these form a basis of the first de Rham cohomology group (a real vector space) of M, which is finite dimensional since M is compact by assumption.

Theorem 8. *For each $n \geq 1$, both ∇f_n and $\star\nabla f_n$ are eigenfields of the Hodge Laplacian, and the set*

$$\left\{ \frac{\nabla f_n}{\sqrt{\lambda_n}}, \frac{\star\nabla f_n}{\sqrt{\lambda_n}}, g_j \,\middle|\, n \geq 1 \text{ and } 0 \leq j \leq J \right\} \tag{19}$$

forms an orthonormal basis of $\mathcal{L}^2(\mathrm{M}; T\mathrm{M})$.

All of these operators can easily be computed numerically, e.g. via automatic differentiation, which makes pointwise evaluation and differentiation of the kernels an easy endeavor with modern computing systems.

The Sphere It is well known that the eigenfunctions of the Laplace–Beltrami operator on the sphere $\mathbb{S}_2$ are given by the *spherical harmonics* $Y_{\ell,m}$, for $\ell \geq 0, -\ell \leq m \leq \ell$, with eigenvalues $\lambda_\ell = \lambda_{\ell,m} = \ell(\ell+1)$. Additionally, the 0-eigenspace of the Hodge Laplacian on the sphere is trivial, see Appendix A.3. This, together with Theorem 8, allows us to compute the eigenfields. We visualize some of them in Figure 4.

The approximation of the full kernel can be made even more efficient via the *addition theorem*, e.g. De Vito et al. (2021, Section 7.3), that states

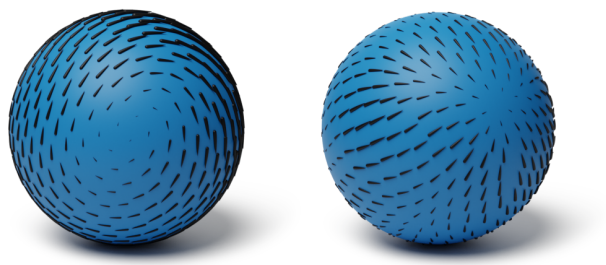
$$\sum_{-\ell\leq m\leq \ell} Y_{\ell,m}(x)Y_{\ell,m}(x') = \frac{2\ell+1}{4\pi}P_\ell(x\cdot x'), \tag{20}$$

where P_ℓ is the ℓ-th Legendre polynomial and the scalar product is taken in $\mathbb{R}^3$ after embedding the $\mathbb{S}_2$ as the standard unit sphere. This reduces the computations to a single simple function for each eigenvalue. As a result, we have the following.

Proposition 9. *Writing*

$$\widetilde{P}_{\ell,\nu,\kappa}(z) = \frac{2\ell+1}{4\pi\lambda_\ell}\Phi_{\nu,\kappa}(\lambda_\ell)P_\ell(z), \tag{21}$$

the pure divergence and pure curl Hodge–Matérn ker-



(a) Divergence-free field f^{curl} (b) Curl-free field f^{div}

Figure 5: Intrinsic Gaussian vector field samples on $\mathbb{S}_2$.

nels on $\mathbb{S}_2$ are given by

$$\boldsymbol{k}^{\text{div}}_{\nu,\kappa,\sigma^2}(x,x') = \frac{\sigma^2}{C^{\text{div}}_{\nu,\kappa}} \sum_{n=1}^{\infty} (\nabla_x \otimes \nabla_{x'}) \widetilde{P}_{\ell,\nu,\kappa}(x \cdot x'),$$

$$\boldsymbol{k}^{\text{curl}}_{\nu,\kappa,\sigma^2}(x,x') = \frac{\sigma^2}{C^{\text{curl}}_{\nu,\kappa}} \sum_{n=1}^{\infty} (\star\nabla_x \otimes \star\nabla_{x'}) \widetilde{P}_{\ell,\nu,\kappa}(x \cdot x').$$

The full Hodge–Matérn kernel is the mean of the two:

$$\boldsymbol{k}_{\nu,\kappa,\sigma^2} = \frac{1}{2}\left(\boldsymbol{k}^{\text{div}}_{\nu,\kappa,\sigma^2} + \boldsymbol{k}^{\text{curl}}_{\nu,\kappa,\sigma^2}\right). \tag{22}$$

Combining this with Proposition 6, we can sample Hodge–Matérn Gaussian vector fields. Thanks to Remark 7, we can also sample their pure divergence and pure curl counterparts, which we illustrate in Figure 5.

4.2 Product Manifolds and Hypertori

Another tractable setting is product manifolds: if we know the eigenfields of the Hodge Laplacian for the factors, we can construct those of the product. We give an overview of this in general before deriving the spectrum of the circle from the scalar case and using it to resolve the case of the hypertori.

Product Manifolds Let M, N be two compact, oriented Riemannian manifolds with scalar manifold heat kernels $\mathcal{P}^{\mathrm{M}}, \mathcal{P}^{\mathrm{N}}$ and Hodge heat kernels $\boldsymbol{\mathcal{P}}^{\mathrm{M}}, \boldsymbol{\mathcal{P}}^{\mathrm{N}}$. The vector kernel on the product $\mathrm{M} \times \mathrm{N}$ is given by

$$\begin{aligned}\boldsymbol{\mathcal{P}}^{\mathrm{M}\times\mathrm{N}}_t(x,x') = \mathcal{P}^{\mathrm{M}}_t(x_1,x_1')\boldsymbol{\mathcal{P}}^{\mathrm{N}}_t(x_2,x_2') + \\ + \mathcal{P}^{\mathrm{N}}_t(x_2,x_2')\boldsymbol{\mathcal{P}}^{\mathrm{M}}_t(x_1,x_1')\end{aligned} \tag{23}$$

for $x = (x_1, x_2)$ and $x' = (x_1', x_2')$ and $x_1, x_1' \in \mathrm{M}$; $x_2, x_2' \in \mathrm{N}$. The details are explained in Appendix A.6.

Knowing the Hodge heat kernels, the other Hodge–Matérn kernels can be derived using Equation (13).

Circle The circle $\mathbb{S}_1$ is the only[4] compact Riemannian manifold of dimension 1. The Hodge star operator in this case gives an identification of scalar functions and vector fields on $\mathbb{S}_1$: there is a canonical global vector field v such that $\|v(x)\| \equiv 1$—in fact, there are exactly two of them, and a choice of orientation selects one—and a function $f(x)$ on $\mathbb{S}_1$ is identified with $f(x)v(x)$. The Hodge Laplacian reduces to the Laplace–Beltrami operator

$$\Delta(f(x)v(x)) = \Delta(f(x))v(x). \tag{24}$$

Thus, the spectrum of the Hodge Laplacian is the same as the spectrum of the Laplace–Beltrami operator under this identification. Similarly, the vector kernels coincide with their scalar counterparts, for which explicit formulas can be found in Borovitskiy et al. (2020).

Hypertori The d-dimensional (flat[5]) hypertorus is defined as the product of d circles $\mathbb{T}^d := \mathbb{S}_1 \times \cdots \times \mathbb{S}_1$. It has a global basis of tangent vector fields by the canonical identification $T\mathbb{T}^d \cong T\mathbb{S}_1 \otimes \cdots \otimes T\mathbb{S}_1$, so that we can write tangent fields on $\mathbb{T}^d$ as a vector of vector fields on $\mathbb{S}_1$. Putting together what was said above, we obtain the following, cf. Appendix A.6.

Proposition 10. *The eigenfields on $\mathbb{T}^d$ are*

$$\left(\prod_{i=1}^{d} f_{n_i}(x_i)\right)\boldsymbol{e_j} \in T\mathbb{S}_1 \otimes \cdots \otimes T\mathbb{S}_1 \cong T\mathbb{T}^d, \tag{25}$$

$1 \le j \le d$, with eigenvalue $\sum_{i=1}^{d} \lambda_{n_i}$, where $x_i \in \mathbb{S}_1$, and (f_n, λ_n) are eigenpairs of the Laplace–Beltrami operator on $\mathbb{S}_1$. In particular, with $k^{\mathbb{T}^d}$ the scalar Matérn kernel on $\mathbb{T}^d$, the vector Hodge–Matérn kernel is

$$\boldsymbol{k}^{\mathbb{T}^d}_{\nu,\kappa,\sigma^2}(x,x') = \frac{1}{d} k^{\mathbb{T}^d}_{\nu,\kappa,\sigma^2}(x_i,x_i')\boldsymbol{I}_d. \tag{26}$$

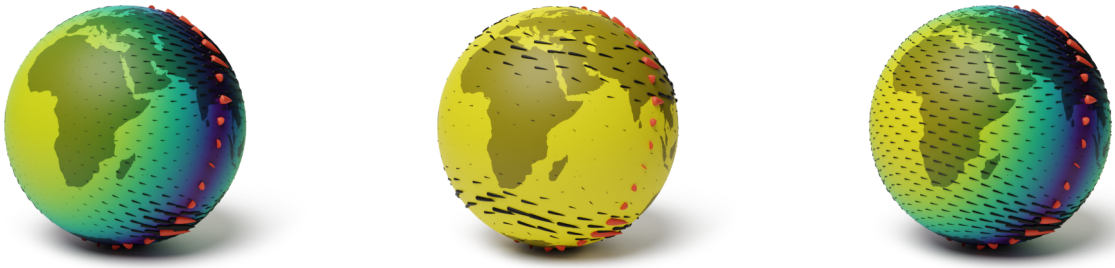
4.3 Possible Extensions

We propose two prospective directions into which our results could be extended.

Meshes Neither Hodge–Matérn kernels nor the eigenfields can be *analytically* computed on a general two-dimensional manifold. This is true even for their scalar counterparts (Borovitskiy et al., 2020). However, we expect that Hodge–Matérn kernels can be numerically approximated on surfaces discretized into meshes. To do this, one needs to apply suitable discrete counterparts of $\star$ and ∇ to numerically approximate the scalar eigenfunctions the Laplace–Beltrami operator and also take care of the harmonic forms.

[4] To be precise: any 1-dimensional compact Riemannian manifold is isometric to the circle of the same length.

[5] Not to be confused with the torus defined as a "donut" in $\mathbb{R}^3$, which has a different intrinsic geometry.



(a) Projected Matérn ($\nu = \frac{1}{2}$) (b) Ground truth (January 2010) (c) Div.-free Hodge-Matérn ($\nu = \frac{1}{2}$)

Figure 6: Interpolation of wind speed on the surface of Earth. The observations are the red vectors along a meridian. Figures (a) and (c) report predictive mean (black vectors) and uncertainty (color: yellow is high, blue is low). Note that in figure (a) sinks and sources are present, while the inductive bias of (c) prohibits that. We advise the reader to examine the global Figures 8 and 9, located in the appendix because of space limitations.

Lie Groups and Related Manifolds It is possible to obtain the scalar manifold Matérn kernels on homogeneous spaces via the representation theory of their symmetry groups (Azangulov et al., 2022). We conjecture that the the vector case can be treated similarly—in particular, in view of the work of Ikeda and Taniguchi (1978). This could result in explicit formulas for eigenfields and eigenvalues for homogeneous spaces given in terms of algebraic quantities only.

5 EXPERIMENTS

We complement the theoretical motivations for the use of intrinsic kernels on manifolds with a practical experiment on weather data from the ERA5 dataset (Hersbach et al., 2023). Further experiments on synthetically generated data are available in Appendix C.

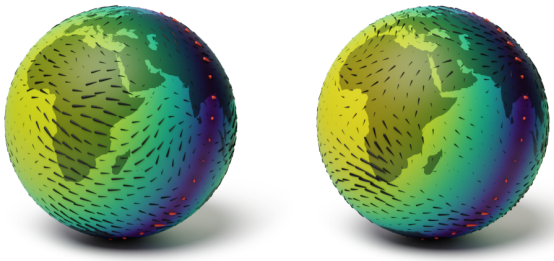
5.1 Setup

The dataset comes from the fifth generation ECMWF atmospheric reanalysis of the global climate (ERA5) (Hersbach et al., 2023). We took the monthly averaged reanalysis data for wind (u- and v-components) at the fixed 500hPa pressure level (corresponding to approximately 5.5km altitude) from January to December 2010. As the wind at the 500hPa pressure level at mid and high latitudes is approximately geostrophic and therefore, approximately divergence-free (Holton, 2004), we expect the family of divergence-free (i.e. pure curl) Hodge–Matérn kernels to provide good results.

The training and testing points are distributed along a great circle, as displayed in Figure 6b. An experiment was run for each month of data, totaling 12 experiments for each kernel.

We applied GP regression using pure noise, projected Matérn (P. M.), Hodge–Matérn (H.–M.), and divergence-free Hodge–Matérn (div-free H.–M.) kernels all with $\nu = \frac{1}{2}, \infty$. The hyperparameters are the noise variance for the first kernel, and the length scale κ, the variance σ^2, and the noise variance σ_ε^2 for the others. They were optimized by maximizing marginal log-likelihood. After visual inspection of the ground truth, we selected the best performing kernel—the divergence-free Hodge–Matérn kernel with $\nu = \frac{1}{2}$—and the projected Matérn-$\frac{1}{2}$ kernel, and ran four further experiments by fixing their length scale to high values $\kappa = 0.5$ and 1. Importantly, we also experimented with the Hodge-compositional Matérn (H.-C. M.) kernels of Section 3.4 to understand if these could detect the correct inductive bias.[6]

We report the mean and standard deviation of mean squared error (MSE) of the predicted mean and predictive negative log-likelihood (PNLL) of the ground truth against the predictive distribution, where each test point is considered independently of the others.



(a) Projected (b) Hodge

Figure 7: Posterior samples of the models featured in Figure 6 (note: different scaling of vectors).

Kernel	MSE		PNLL	
	Mean	Std	Mean	Std
Pure noise	2.07	0.18	2.89	0.10
P. M.–$\frac{1}{2}$	1.39	0.15	2.33	0.11
P. M.–∞	1.53	0.20	2.43	0.15
H.–M.–$\frac{1}{2}$	1.67	0.16	2.58	0.12
H.–M.–∞	1.76	0.17	2.63	0.14
div-free H.–M.–$\frac{1}{2}$	**1.10**	0.12	**2.16**	0.13
div-free H.–M.–∞	1.34	0.20	2.33	0.18
H.-C. M.–$\frac{1}{2}$	**1.09**	0.10	**2.16**	0.12
H.-C. M.–∞	1.34	0.20	2.33	0.18

Table 1: Mean squared error and predictive negative log-likelihood in the ERA5 wind data experiment.

Kernel	MSE		PNLL	
	Mean	Std	Mean	Std
div-free H.–M.–$\frac{1}{2}$	1.10	0.12	2.16	0.13
div-free H.–M.–∞	1.34	0.20	2.33	0.18
H.–M.–$\frac{1}{2}$ $\kappa = 0.5$	1.06	0.10	**2.15**	0.13
H.–M.–$\frac{1}{2}$ $\kappa = 1.0$	**1.05**	0.10	2.17	0.13
P. M.–$\frac{1}{2}$ $\kappa = 0.5$	1.36	0.13	2.32	0.10
P. M.–$\frac{1}{2}$ $\kappa = 1.0$	1.36	0.13	2.34	0.10

Table 2: Mean squared error and predictive negative log-likelihood in the ERA5 wind data experiment for divergence-free Hodge–Matérn and projected Matérn kernels with $\nu = \frac{1}{2}$ and fixed length scale κ.

5.2 Performance

Results are displayed in Tables 1 and 2. The data was scaled so that the training observations have unit mean norm, cf. Appendix C. Baselines are provided by fitting a pure noise kernel to the data and the projected kernels of Hutchinson et al. (2021). The best scores by far were obtained by fitting either of the divergence-free Hodge–Matérn or Hodge-compositional Matérn kernels with $\nu = \frac{1}{2}$, which outperformed the other kernels we considered on all metrics. Next to them were divergence-free Hodge–Matérn with $\nu = \infty$ and the projected Matérn kernel with $\nu = \frac{1}{2}$.

Two factors were of fundamental importance in obtaining good results: having a divergence-free inductive bias and allowing for lower degrees of smoothness by setting $\nu = \frac{1}{2}$. In particular, considering the divergence of the various kernels explains why the projected kernels tend to do better than the full Hodge-Matérn kernels in this situation: they have lower absolute divergence in expectation, cf. Appendix B.3, Figure 10.

The additional experiments where length scales were fixed to high values are reported in Table 2. We notice further test score improvements, but not significant ones. Figures 6 and 7 display posterior mean, standard deviations and samples from GP regressions fitted using projected Matérn and divergence-free Hodge–Matérn kernels, both with fixed $\kappa = 1$.

Analysis of the fitted hyperparameters of Hodge-compositional Matérn kernels shows that these were able to automatically detect the correct inductive bias. Specifically, the resulting Hodge-compositional kernel was virtually divergence-free, which enabled it to reach the same scores as the actual divergence-free Hodge–Matérn kernels. See Appendix C for more details.

[6]On the sphere $\mathbb{S}_2$, these are linear combinations of pure-divergence and pure-curl Hodge–Matérn kernels, without the harmonic part which vanishes on $\mathbb{S}_2$

6 CONCLUSION

In this work, we introduced a novel class of models—the *Hodge–Matérn* vector fields—for learning tangential vector fields on manifolds. These principled models solve the shortcomings present in the preexisting literature, providing improved inductive biases. We described computational techniques—kernel evaluation and differentiation, sampling—required for running Gaussian process regression and for downstream applications on important manifolds such as the two-dimensional sphere $\mathbb{S}_2$ and hypertori $\mathbb{T}^d$, and indicated further extension directions. We applied our methods both to synthetic and real-world data on $\mathbb{S}_2$, demonstrating that they can lead to improved performance. We hope our results will inspire new extensions and applications in areas like climate modeling and robotics.

ACKNOWLEDGMENTS

We are grateful to Gergana Gyuleva for her help with retrieving meaningful weather data for our experiments. We thank Maosheng Yang and Dr. So Takao for their insightful input on the drafts of the paper. We also thank Dr. Alexander Terenin for his guidance in using his Blender rendering scripts of his Ph.D. thesis repository.[7] Finally, we thank Dr. Alan Pinoy for providing the proof of Lemma 32. VB acknowledges support by an ETH Zürich Postdoctoral Fellowship.

[7]https://github.com/aterenin/phdthesis (Terenin, 2022).

REFERENCES

M. A. Alvarez, L. Rosasco, N. D. Lawrence, et al. Kernels for vector-valued functions: A review. *Foundations and Trends® in Machine Learning*, 4(3):195–266, 2012. Cited on page 1.

I. Azangulov, A. Smolensky, A. Terenin, and V. Borovitskiy. Stationary Kernels and Gaussian Processes on Lie Groups and their Homogeneous Spaces I: the compact case. *arXiv preprint arXiv:2208.14960*, 2022. Cited on pages 2–4, 8.

I. Azangulov, A. Smolensky, A. Terenin, and V. Borovitskiy. Stationary Kernels and Gaussian Processes on Lie Groups and their Homogeneous Spaces II: non-compact symmetric spaces. *arXiv preprint arXiv:2301.13088*, 2023. Cited on pages 2, 4.

N. Berline, E. Getzler, M. Vergne, and B. Orsted. Heat kernels and Dirac operators. *Jahresbericht der Deutschen Mathematiker Vereinigung*, 97(3):63–65, 1995. Cited on page 19.

R. Berlinghieri, B. L. Trippe, D. R. Burt, R. Giordano, K. Srinivasan, T. Özgökmen, J. Xia, and T. Broderick. Gaussian processes at the Helm (holtz): A more fluid model for ocean currents. *arXiv preprint arXiv:2302.10364*, 2023. Cited on pages 4, 5.

V. Borovitskiy, A. Terenin, P. Mostowsky, and M. P. Deisenroth. Matérn Gaussian processes on Riemannian manifolds. In *Advances in Neural Information Processing Systems*, 2020. Cited on pages 2, 7, 19.

R. Bryant. Vector-Valued Stone–Weierstrass Theorem? MathOverflow, 2020. https://mathoverflow.net/q/371618 (version: 2020-09-14). Cited on page 17.

Y. Canzani. Analysis on manifolds via the Laplacian. *Lecture Notes available at: http://www.math.harvard.edu/canzani/docs/Laplacian.pdf*, 2013. Cited on page 17.

J.-P. Chilès and P. Delfiner. *Geostatistics: Modeling Spatial Uncertainty*. Wiley Series in Probability and Statistics. Wiley & Sons, 2nd edition, 2012. Cited on page 1.

S. Coveney, C. Corrado, C. H. Roney, D. O'Hare, S. E. Williams, M. D. O'Neill, S. A. Niederer, R. H. Clayton, J. E. Oakley, and R. D. Wilkinson. Gaussian process manifold interpolation for probabilistic atrial activation maps and uncertain conduction velocity. *Philosophical Transactions of the Royal Society A*, 378(2173), 2020. Cited on page 2.

E. De Vito, N. Mücke, and L. Rosasco. Reproducing kernel Hilbert spaces on manifolds: Sobolev and diffusion spaces. *Analysis and Applications*, 19(03):363–396, 2021. Cited on page 6.

H. Hersbach, B. Bell, P. Berrisford, G. Biavati, A. Horányi, J. Muñoz Sabater, J. Nicolas, C. Peubey, R. Radu, I. Rozum, D. Schepers, A. Simmons, C. Soci, D. Dee, and J.-N. Thépaut. ERA5 monthly averaged data on pressure levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), 2023. DOI: 10.24381/cds.6860a573 (Accessed on 10-Oct-2023). Cited on page 8.

J. R. Holton. *An Introduction to Dynamic Meteorology*, volume 88 of *International Geophysics*. Elsevier Academic Press, 4th edition, 2004. Cited on page 8.

M. Hutchinson, A. Terenin, V. Borovitskiy, S. Takao, Y. Teh, and M. Deisenroth. Vector-valued Gaussian Processes on Riemannian Manifolds via Gauge Independent Projected Kernels. In *Advances in Neural Information Processing Systems*, volume 34, 2021. Cited on pages 2–5, 9, 16.

A. Ikeda and Y. Taniguchi. Spectra and eigenforms of the Laplacian on S^n and $P^n(\mathbb{C})$. *Osaka Journal of Mathematics*, 15(3):515–546, 1978. Cited on page 8.

A. Krause, A. Singh, and C. Guestrin. Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies. *Journal of Machine Learning Research*, 9(2), 2008. Cited on page 1.

M. Lange-Hegermann. Algorithmic linearly constrained Gaussian processes. *Advances in Neural Information Processing Systems*, 31, 2018. Cited on page 1.

J. M. Lee. *Manifolds and Differential Geometry*, volume 107 of *Graduate Studies in Mathematics*. American Mathematical Society, 2009. Cited on pages 14, 15, 19, 21.

F. Lindgren, H. Rue, and J. Lindström. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 73(4):423–498, 2011. Cited on page 2.

E. Noskova and V. Borovitskiy. Bayesian optimization for demographic inference. *G3: Genes, Genomes, Genetics*, 13(7), 2023. Cited on page 1.

V. K. Patodi. Curvature and the eigenforms of the Laplace operator. *Journal of Differential Geometry*, 5(1-2):233–249, 1971. Cited on page 16.

R. L. Peach, M. Vinao-Carl, N. Grossman, M. David, E. Mallas, D. Sharp, P. A. Malhotra, P. Vandergheynst, and A. Gosztolai. Implicit Gaussian process representation of vector fields over arbitrary latent manifolds. In *International Conference on Learning Representations*, 2024. Cited on page 2.

J. B. Prolla. On the Weierstrass–Stone Theorem. *Journal of Approximation Theory*, 78(3):299–313, 1994. Cited on page 17.

C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006. Cited on pages 1, 2.

P. Rosa, V. Borovitskiy, A. Terenin, and J. Rousseau. Posterior Contraction Rates for Matérn Gaussian Processes on Riemannian Manifolds. *arXiv preprint arXiv:2309.10918*, 2023. Cited on page 2.

S. Rosenberg. *The Laplacian on a Riemannian manifold: an introduction to analysis on manifolds*, number 31 in London Mathematical Society Student Texts. Cambridge University Press, 1997. Cited on pages 4, 5, 13–16, 18, 19.

N. Sharp, Y. Soliman, and K. Crane. The vector heat method. *ACM Transactions on Graphics*, 38(3):1–19, 2019. Cited on page 19.

B. J. Shields, J. Stevens, J. Li, M. Parasram, F. Damani, J. I. M. Alvarado, J. M. Janey, R. P. Adams, and A. G. Doyle. Bayesian reaction optimization as a tool for chemical synthesis. *Nature*, 590(7844):89–96, 2021. Cited on page 1.

J. Snoek, H. Larochelle, and R. P. Adams. Practical Bayesian Optimization of Machine Learning Algorithms. In *Advances in Neural Information Processing Systems*, volume 25, 2012. Cited on page 1.

M. L. Stein. *Interpolation of Spatial Data: Some Theory for Kriging*. Springer, 1999. Cited on page 2.

A. Terenin. Gaussian processes and statistical decision-making in non-Euclidean spaces. *arXiv preprint arXiv:2202.10613*, 2022. Cited on page 9.

P. Whittle. Stochastic processes in several dimensions. *Bulletin of the International Statistical Institute*, 40(2):974–994, 1963. Cited on page 2.

M. Yang, V. Borovitskiy, and E. Isufi. Hodge-Compositional Edge Gaussian Processes. In *International Conference on Artificial Intelligence and Statistics*, 2024. Cited on pages 5, 23.

CHECKLIST

1. For all models and algorithms presented, check if you include:
   (a) A clear description of the mathematical setting, assumptions, algorithm, and/or model.
   Yes
   (b) An analysis of the properties and complexity (time, space, sample size) of any algorithm.
   Yes
   (c) (Optional) Anonymized source code, with specification of all dependencies, including external libraries.
   Yes

2. For any theoretical claim, check if you include:
   (a) Statements of the full set of assumptions of all theoretical results.
   Yes
   (b) Complete proofs of all theoretical results.
   Yes
   (c) Clear explanations of any assumptions.
   Yes

3. For all figures and tables that present empirical results, check if you include:
   (a) The code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL).
   Yes
   (b) All the training details (e.g., data splits, hyperparameters, how they were chosen).
   Yes
   (c) A clear definition of the specific measure or statistics and error bars (e.g., with respect to the random seed after running experiments multiple times).
   Yes
   (d) A description of the computing infrastructure used. (e.g., type of GPUs, internal cluster, or cloud provider).
   Yes

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets, check if you include:
   (a) Citations of the creator If your work uses existing assets.
   Yes
   (b) The license information of the assets, if applicable.
   Yes

(c) New assets either in the supplemental material or as a URL, if applicable.
Not Applicable

(d) Information about consent from data providers/curators.
Not Applicable

(e) Discussion of sensible content if applicable, e.g., personally identifiable information or offensive content.
Not Applicable

5. If you used crowdsourcing or conducted research with human subjects, check if you include:

(a) The full text of instructions given to participants and screenshots.
Not Applicable

(b) Descriptions of potential participant risks, with links to Institutional Review Board (IRB) approvals if applicable.
Not Applicable

(c) The estimated hourly wage paid to participants and the total amount spent on participant compensation.
Not Applicable

A THEORY

In this section, we present the mathematical theory behind the Hodge–Matérn Gaussian vector fields introduced in the main body of the paper. Although we previously almost exclusively talked about vector fields on manifolds, this section is written in the more technically appropriate language of differential forms. On a Riemannian manifold, the metric provides an equivalence between 1-forms and vector fields, which immediately recovers the results of the main part of the article. The general theory provides more than just intrinsic kernels for 1-forms, though: we obtain kernels for differential forms of all degrees—hence intrinsic Gaussian differential forms—which can be refined to kernels for pure gradient, pure curl, and harmonic processes.

We will begin by introducing the Hodge star operator and link the abstract differential operators to the classical notions of gradients, curls, and divergences. The Hodge star operator will allow us to define the Hodge Laplacian and the heat equation on differential forms. The fundamental solution of the latter is the heat kernel (up to a normalization constant). We will then explore the cases of surfaces and products, and conclude by giving an overview of another possible generalization of the Laplacian to sections of a vector bundle that could be considered in future research.

We assume the reader has some basic familiarity with concepts of differential geometry, such as differential forms and the exterior derivative. All manifolds are assumed to be compact, connected, and oriented.

Most of the material found in this appendix can be found in the book by Rosenberg (1997). We only give proofs of original results.

A.1 The Hodge Star Operator

Let V be a finite-dimensional oriented vector space, say of dimension n, endowed with a non-degenerate inner product $\langle \cdot, \cdot \rangle$.

The k-th *exterior product* $\Lambda^k V$ of V is the quotient of the tensor product of k copies of V by the sign representation of the symmetric group:

$$v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(k)} \sim \operatorname{sgn}(\sigma) v_1 \otimes \cdots \otimes v_k \tag{27}$$

for $v_1, \ldots, v_k \in V$ and $\sigma \in \Sigma_k$ a permutation. We write $v_1 \wedge \cdots \wedge v_k$ for the equivalence class of $v_1 \otimes \cdots \otimes v_k$, and we call $\wedge$ the *wedge product*.

The inner product on V induces an inner product on the exterior product $\Lambda^k V$ for $0 \leq k \leq n$ by

$$\langle a, b \rangle = \det\left(\langle a_i, b_j \rangle_{i,j}\right) \tag{28}$$

for $a = a_1 \wedge \ldots \wedge a_k$ and $b = b_1 \wedge \ldots \wedge b_k$.

It is easily checked that $\Lambda^n V$ is 1-dimensional and has a canonical unit n-vector

$$\omega = e_1 \wedge \ldots \wedge e_n \in \Lambda^n V, \tag{29}$$

where $e_1, \ldots, e_n \in V$ is an oriented orthonormal basis.

The Hodge star operator is the linear operator

$$\star : \Lambda^k V \longrightarrow \Lambda^{n-k} V \tag{30}$$

mapping $b \in \Lambda^k V$ to the unique $\star b \in \Lambda^{n-k} V$ such that

$$a \wedge (\star b) = \langle a, b \rangle \omega \tag{31}$$

for any $a \in \Lambda^k V$.

From now on, let M be a compact oriented Riemannian manifold of dimension d_{M}. The cotangent bundle $T^*\mathrm{M}$ is the dual of the tangent bundle $T\mathrm{M}$. In other words, at each $x \in \mathrm{M}$ the fibre $T_x^*\mathrm{M}$ is the linear dual of the tangent space $T_x\mathrm{M}$ at x. One can then take the exterior products of the cotangent bundle, which are nothing more than the exterior product $\Lambda^k T_x^*\mathrm{M}$ at each $x \in \mathrm{M}$. Sections of this bundle—smooth functions that at each point $x \in \mathrm{M}$ associate an element of $\Lambda^k T_x^*\mathrm{M}$—are called *$k$-differential forms* on M. The space of k-differential forms is denoted by $\Omega^k(\mathrm{M})$.

The exterior derivative is an intrinsic differential operator

$$\mathrm{d} : \Omega^k(\mathrm{M}) \longrightarrow \Omega^{k+1}(\mathrm{M}) \tag{32}$$

defined in local coordinates by linear extension of

$$\mathrm{d}(f(x)\,\mathrm{d}x_I) = \sum_{i=1}^{d_{\mathrm{M}}} \frac{\partial f(x)}{\partial x_i}\,\mathrm{d}x_i \wedge \mathrm{d}x_I, \tag{33}$$

where $I = \{i_1, \ldots, i_k\} \subseteq \{1, \ldots, d_{\mathrm{M}}\}$ is a set of k indices and $\mathrm{d}x_I = \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k}$. In order to build intuition, one can notice that if X is a vector field on M and $f \in \Omega^0(\mathrm{M})$, then $\mathrm{d}f(X)$ is the directional derivative of f in direction X.

Applying the Hodge star operator to the cotangent bundle at each point of M (the inner product being given by the Riemannian metric), one obtains a Hodge star operator on differential forms

$$\star : \Omega^k(M) \longrightarrow \Omega^{n-k}(M). \tag{34}$$

The canonical unit n-vector in $\Lambda^n T^*\mathrm{M}$ is nothing else than the Riemannian volume form ω_g, and by definition we have $\star 1 = \omega_g$ and $\star \omega_g = 1$.

If $\alpha, \beta \in \Omega^k(\mathrm{M})$ are two k-forms, their (Hodge) inner product is

$$\langle \alpha, \beta \rangle_{\mathcal{L}^2(\mathrm{M})} = \int_{\mathrm{M}} \langle \alpha, \beta \rangle \omega_g = \int_{\mathrm{M}} \alpha \wedge \star \beta. \tag{35}$$

An easy computation shows that the dual of the exterior derivative $\mathrm{d} : \Omega^k(\mathrm{M}) \to \Omega^{k+1}(\mathrm{M})$ is given by

$$\mathrm{d}^\star = (-1)^{d_\mathrm{M} k+1} \star \mathrm{d} \star : \Omega^{k+1}(\mathrm{M}) \longrightarrow \Omega^k(\mathrm{M}). \tag{36}$$

A.2 Gradient, Divergence, and Curl on Surfaces

In this section, we will link the abstract operators $\star$, d, and $\mathrm{d}^\star$ to quantities which are in some sense more concrete. Before doing that, we start with a quick recap on vector fields, their curls and divergences in $\mathbb{R}^3$.

In $\mathbb{R}^3$, a basis for 1-forms is given by $\mathrm{d}x_1, \mathrm{d}x_2, \mathrm{d}x_3$, and a basis for 2-forms is $\mathrm{d}x_1 \wedge \mathrm{d}x_2, \mathrm{d}x_1 \wedge \mathrm{d}x_3, \mathrm{d}x_2 \wedge \mathrm{d}x_3$. The Hodge star operator sends the constant function 1 to the volume form $\mathrm{d}x_1 \wedge \mathrm{d}x_2 \wedge \mathrm{d}x_3$, and it maps

$$\star\, \mathrm{d}x_1 = \mathrm{d}x_2 \wedge \mathrm{d}x_3, \tag{37}$$
$$\star\, \mathrm{d}x_2 = -\,\mathrm{d}x_1 \wedge \mathrm{d}x_3, \tag{38}$$
$$\star\, \mathrm{d}x_3 = \mathrm{d}x_1 \wedge \mathrm{d}x_2, \tag{39}$$

and vice versa. Vector fields are identified with 1-forms by mapping $\mathrm{d}x_i \mapsto \boldsymbol{e}_i$, and also with 2-forms via the Hodge star operator.

With this in mind, we immediately see that the exterior derivative of a function

$$\mathrm{d}f = \partial_1 f\, \mathrm{d}x_1 + \partial_2 f\, \mathrm{d}x_2 + \partial_3 f\, \mathrm{d}x_3 \tag{40}$$

is identified with taking the gradient ∇f. A straightforward calculation also shows that taking the exterior derivative of a 1-form corresponds to taking the curl of the corresponding vector field, and the exterior derivative of a 2-form gives the divergence of the associated vector field. This and the fact that $\mathrm{d}^2 = 0$ recover the well known relations between ∇, curl, and div. Via the Hodge star operator, similar statements can be made about the $\mathrm{d}^\star$ operator.

On surfaces, the situation is a bit different since we only have two dimensions in which to move. The fact that our manifold is oriented gives us an orientation on each cotangent space $T_x^*\mathrm{M}$. Picking an oriented orthonormal local basis $\mathrm{d}x_1, \mathrm{d}x_2 \in T_x^*\mathrm{M}$, the volume form is given locally by $\mathrm{d}x_1 \wedge \mathrm{d}x_2$ and it follows that $\star\, \mathrm{d}x_1 = \mathrm{d}x_2$ and $\star\, \mathrm{d}x_2 = -\,\mathrm{d}x_1$. This corresponds to a rotation by 90° in the cotangent space.

In the case where our surface is embedded in $\mathbb{R}^3$, the choice of an orientation is equivalent to the choice of a global unit normal field for the manifold, i.e. a smooth choice of a unit normal vector for each point. Then, for tangent vectors the Hodge star is given extrinsically by a rotation by 90° around the unit normal at each point. This can also be written as the cross product of a tangent vector with the unit normal.

Building on the case of $\mathbb{R}^3$, we call

$$\nabla = \mathrm{d} : \Omega^0(\mathrm{M}) \longrightarrow \Omega^1(\mathrm{M}), \tag{41}$$
$$\mathrm{div} = \mathrm{d}^\star : \Omega^1(\mathrm{M}) \longrightarrow \Omega^0(\mathrm{M}), \tag{42}$$
$$\mathrm{curl} = \star\, \mathrm{d} : \Omega^1(\mathrm{M}) \longrightarrow \Omega^0(\mathrm{M}), \tag{43}$$

under the identification of 1-forms with vector fields and 2-forms with functions. This also corresponds—potentially up to a sign, depending on conventions—to the definition of divergence in Riemannian geometry using the Levi–Civita connection (Lee, 2009, Equation 13.11). Working in a local coordinate system on M, the explicit expression for these operators is

$$\nabla f(x) = \begin{pmatrix} \partial_1 f(x) \\ \partial_2 f(x) \end{pmatrix}, \tag{44}$$

$$\mathrm{div} \begin{pmatrix} v_1(x) \\ v_2(x) \end{pmatrix} = \partial_1 v_1(x) + \partial_2 v_2(x), \tag{45}$$

$$\mathrm{curl} \begin{pmatrix} v_1(x) \\ v_2(x) \end{pmatrix} = \partial_1 v_2(x) - \partial_2 v_1(x). \tag{46}$$

A.3 The Hodge Laplacian

Let M be a d-dimensional manifold. The *Hodge Laplacian* on differential forms is then defined as

$$\Delta := -\left(\mathrm{d}^\star\, \mathrm{d} + \mathrm{d}\, \mathrm{d}^\star\right). \tag{47}$$

Remark 11. *In some texts, including Rosenberg (1997), the Hodge Laplacian has the opposite sign.*

Remark 12. *For $k = 0$ (i.e. on functions) this recovers the Laplace–Beltrami operator*

$$\Delta = -\,\mathrm{d}\, \mathrm{d}^\star = -\nabla^* \nabla, \tag{48}$$

where ∇ is the Levi–Civita connection. Cf. also appendix A.7. In the dual context of vector fields, we obtain the classical divergence of the gradient

$$\Delta = \mathrm{div}\, \nabla. \tag{49}$$

The following deep results—found in Rosenberg (1997, Theorems 1.30, 1.37, and 1.45)—give us all we need to know about the spectrum of the Hodge Laplacian.

Theorem 13 (Hodge)**.** *All the eigenvalues of the Hodge Laplacian Δ on $\Omega^k(\mathrm{M})$ are non-negative, they have finite multiplicity, and they accumulate only at infinity. The eigenforms span a dense subset of $\Omega_{\mathcal{L}^2}(\mathrm{M})$. In particular, there exists an orthonormal basis of $\Omega_{\mathcal{L}^2}(\mathrm{M})$ consisting of smooth eigenforms of Δ.*

Remark 14. *The convention on eigenpairs is that ϕ is an eigenform of (minus) eigenvalue λ if $\Delta\phi = -\lambda\phi$.*

Theorem 15 (Hodge decomposition)**.** *The space of smooth k-forms decomposes as*

$$\Omega^k(\mathrm{M}) = \ker \Delta \oplus \mathrm{im}\, \mathrm{d} \oplus \mathrm{im}\, \mathrm{d}^\star. \tag{50}$$

The next result links the kernel of the Hodge Laplacian with a purely topological property of the manifold: de Rham cohomology. An accessible introduction is given in Rosenberg (1997, Section 1.4).

Theorem 16 (Hodge)**.** *The kernel of the Hodge Laplacian on k-forms is naturally isomorphic to the k-th de Rham cohomology group, which is a real vector space:*

$$\ker \Delta \cong H^k_{dR}(\mathrm{M}). \tag{51}$$

Remark 17. *The following facts about de Rham cohomology are often useful. Assume* M *is compact and connected.*

1. $H^0_{dR}(\mathrm{M}) \cong \mathbb{R}$*, spanned by the constant function* $f(x) = 1$*.*
2. $H^{d_\mathrm{M}}_{dR}(\mathrm{M}) \cong \mathbb{R}$*, spanned by the volume form.*
3. *All of the* $H^k_{dR}(\mathrm{M})$ *are finite dimensional, see e.g. Lee (2009, Theorem 10.17).*

An example that is exploited in the main body of this paper is the well known fact that

$$H^1_{dR}(\mathbb{S}_2) = 0. \tag{52}$$

This can be computed using the Mayer–Vietoris sequence, see e.g. Lee (2009, Section 10.1).

Proposition 18. *The various spaces of eigenforms have the following relations.*

1. *The Hodge star* $\star : \Omega^k(\mathrm{M}) \to \Omega^{n-k}(\mathrm{M})$ *sends eigenforms of the Hodge Laplacian to eigenforms with the same eigenvalue, and it preserves their orthogonality and norm.*
2. *The exterior derivative* $\mathrm{d} : \Omega^k(\mathrm{M}) \to \Omega^{k+1}(\mathrm{M})$ *sends eigenforms in* $\operatorname{im} \mathrm{d}^\star$ *to eigenforms with the same eigenvalue (and is zero on the other eigenforms), it preserves orthogonality, and*
$$\| \mathrm{d}\phi\|_{\mathcal{L}^2(\mathrm{M})} = \sqrt{\lambda}\|\phi\|_{\mathcal{L}^2(\mathrm{M})} \tag{53}$$
for $\phi \in \operatorname{im} \mathrm{d}^\star \subseteq \Omega^k(\mathrm{M})$ *an eigenform of eigenvalue* $-\lambda$*.*
3. *The operator* $\mathrm{d}^\star : \Omega^k(\mathrm{M}) \to \Omega^{k-1}(\mathrm{M})$ *sends eigenforms in* $\operatorname{im} \mathrm{d}$ *to eigenforms with the same eigenvalue (and is zero on the other eigenforms), it preserves orthogonality, and*
$$\| \mathrm{d}^\star\phi\|_{\mathcal{L}^2(\mathrm{M})} = \sqrt{\lambda}\|\phi\|_{\mathcal{L}^2(\mathrm{M})} \tag{54}$$
for $\phi \in \operatorname{im} \mathrm{d} \subseteq \Omega^k(\mathrm{M})$ *an eigenform of eigenvalue* $-\lambda$*.*

Proof. For (1), it is immediate to see that $\Delta\star = \star\Delta$ so that $\star$ sends eigenforms to eigenforms with the same eigenvalue. If α, β are any two k-forms, then

$$\begin{aligned}
\langle\star\alpha, \star\beta\rangle_{\mathcal{L}^2(\mathrm{M})} &= \int_\mathrm{M} \star\alpha \wedge \star\star\beta && (55)\\
&= \int_\mathrm{M} \star\alpha \wedge \beta && (56)\\
&= \int_\mathrm{M} \beta \wedge \star\alpha = \langle\beta, \alpha\rangle_{\mathcal{L}^2(\mathrm{M})} && (57)\\
&= \langle\alpha, \beta\rangle_{\mathcal{L}^2(\mathrm{M})}, && (58)
\end{aligned}$$

showing that $\star$ is an isometry and concluding the point.

For (2), one easily checks that $\Delta\,\mathrm{d} = \mathrm{d}\Delta$ so that d sends eigenforms that are in $\operatorname{im} \mathrm{d}^\star$, i.e. the complement of $\ker \mathrm{d}$, to eigenforms with the same eigenvalue. Let $\phi, \psi \in \operatorname{im} \mathrm{d}^\star$ be eigenforms with eigenvalues $\lambda_\phi, \lambda_\psi$ respectively, then

$$\begin{aligned}
\langle \mathrm{d}\phi, \mathrm{d}\psi\rangle_{\mathcal{L}^2(\mathrm{M})} &= \langle\phi, \mathrm{d}^\star\,\mathrm{d}\psi\rangle_{\mathcal{L}^2(\mathrm{M})} && (59)\\
&= \langle\phi, -\Delta\psi\rangle_{\mathcal{L}^2(\mathrm{M})} && (60)\\
&= \lambda_\psi \langle\phi, \psi\rangle_{\mathcal{L}^2(\mathrm{M})}, && (61)
\end{aligned}$$

where in the middle equality we used the fact that $\psi \in \operatorname{im} \mathrm{d}^\star$ and that $\mathrm{d}^\star\,\mathrm{d}^\star = 0$. This concludes the proof of the point, and (3) is analogous. □

A.4 The Heat Equation and its Kernel

The heat equation for k-forms can now be defined as

$$\partial_t\alpha(t, x) = \Delta_x\alpha(t, x) \tag{62}$$

with a given initial condition $\alpha(0, x) = \beta(x) \in \Omega^k(\mathrm{M})$.

A *double (k-)form* over M is a smooth section of the bundle $\mathbb{R} \otimes \Lambda^k T^*\mathrm{M} \otimes \Lambda^k T^*\mathrm{M}$ over $\mathrm{M} \times \mathrm{M}$, where the fibre above $(x, x') \in \mathrm{M} \times \mathrm{M}$ is $\mathbb{R} \otimes \Lambda^k T^*_x\mathrm{M} \otimes \Lambda^k T^*_{x'}\mathrm{M}$.

A *heat kernel* for k-forms is a double form $\mathcal{P}_t(x, y)$ such that

1. $(\partial_t - \Delta_x)\mathcal{P}_t(x, x') = 0$ and
2. $\lim_{t\to\infty} \int_M \langle\mathcal{P}_t(x, x'), \alpha(x')\rangle_{x'}\,\mathrm{d}x' = \alpha(x)$ for any $\alpha \in \Omega^k(\mathrm{M})$, where the pointwise inner product and the integration are taken with respect to x', and integration is against the volume form of M.

Theorem 19. *Let* $\phi_i \in \Omega^k(\mathrm{M})$ *be an orthonormal basis of k-eigenforms of the Hodge Laplacian with (minus) eigenvalues* $0 \le \lambda_1 \le \lambda_2 \le \cdots$*. The heat kernel on k-forms exists, it is unique, and it can be expressed by the following sum over eigenforms:*

$$\mathcal{P}_t(x, x') = \sum_{n=0}^{\infty} e^{-\lambda_n t}\phi_n(x) \otimes \phi_n(x'). \tag{63}$$

Proof. See Rosenberg (1997, Proposition 3.1) for functions, the discussion at the end of Rosenberg (1997, Section 3.2) and Patodi (1971) for the general statement on forms. □

Given the heat kernel, the solution for the heat equation

$$(\partial_t + \Delta_x)\alpha(t,x) = 0 \tag{64}$$

with initial condition $\alpha(0,x) = \beta(x)$ is given by

$$\begin{aligned}\alpha(t,x) &= \left(e^{-t\Delta}\beta\right)(x) && (65)\\ &:= \int_{\mathrm{M}} \langle \mathcal{P}_t(x,x'), \beta(x')\rangle_{x'}\, \mathrm{d}x' && (66)\\ &= \langle \mathcal{P}_t(x,\cdot), \beta(\cdot)\rangle_{\mathcal{L}^2(\mathrm{M})} && (67)\end{aligned}$$

for $t > 0$. We now show that the heat kernel is a valid kernel for Gaussian processes.

Proposition 20. *The heat kernel satisfies the following properties.*

1. *The kernel $\mathcal{P}_t(x,x')$ is symmetric: for $x, x' \in \mathrm{M}$ we have $\mathcal{P}_t(x,x') = \mathcal{P}_t(x',x)$ under the canonical identification*

$$\Lambda^k T_x^*\mathrm{M} \otimes \Lambda^k T_{x'}^*\mathrm{M} \cong \Lambda^k T_{x'}^*\mathrm{M} \otimes \Lambda^k T_x^*\mathrm{M}. \tag{68}$$

2. *The propagator $e^{-t\Delta}$ satisfies the semigroup property $e^{-(t+s)\Delta} = e^{-t\Delta}e^{-s\Delta}$.*

3. *The kernel $\mathcal{P}_t(x,x')$ is positive semi-definite: for all $\alpha \in \Omega^k_{\mathcal{L}^2}(\mathrm{M})$ we have*

$$\langle \alpha(x), \langle \mathcal{P}_t(x,x'), \alpha(x')\rangle_{x'}\rangle_x \geq 0. \tag{69}$$

Proof. Statement (1) is immediate from Theorem 19. Alternatively, it can be proved from the fact that the Hodge Laplacian is self-adjoint. Statement (2) follows in the same way as in Rosenberg (1997, pp. 28–29), as does statement (3). □

As an immediate corollary in the vector setting, we obtain the following.

Theorem 3. *For any $t > 0$ there exists a Gaussian vector field whose kernel is $\boldsymbol{\mathcal{P}}_t$.*

Proof. The statement follows from Hutchinson et al. (2021, Theorem 4), provided we prove that our kernel satisfies Definition 3 in the cited work. The link between the two definitions is given by

$$k(\alpha_x, \beta_{x'}) = \left\langle \alpha_x, \langle \mathcal{P}_t(x,x'), \beta_{x'}\rangle^{\mathrm{pt}}_{x'}\right\rangle^{\mathrm{pt}}_x, \tag{70}$$

where the scalar products on the right-hand side are taken pointwise, not integrating over the manifold.

Fibrewise bilinearity is obvious, we are left to prove that our kernel is positive semi-definite in the sense of Hutchinson et al. (2021, Definition 3). In order to do so, for $x_i \in \mathrm{M}$ and $\alpha_{x_i} \in T^*_{x_i}\mathrm{M}$, $1 \leq i \leq n$, consider the sequence

$$\alpha_i^m(x) = \left\langle \mathcal{P}_{\frac{1}{m}}(x, x_i), \alpha_{x_i}\right\rangle^{\mathrm{pt}}_{x_i}. \tag{71}$$

Then we have

$$\begin{aligned}&\langle \mathcal{P}_t(x,x'), \alpha_i^m(x')\rangle_{x'} = && (72)\\ &= \left\langle \mathcal{P}_t(x,x'), \left\langle \mathcal{P}_{\frac{1}{m}}(x',x_i), \alpha_{x_i}\right\rangle^{\mathrm{pt}}_{x_i}\right\rangle_{x'} && (73)\\ &= \left\langle \left\langle \mathcal{P}_t(x,x'), \mathcal{P}_{\frac{1}{m}}(x',x_i)\right\rangle_{x'}, \alpha_{x_i}\right\rangle^{\mathrm{pt}}_{x_i} && (74)\\ &\xrightarrow{m\to\infty} \langle \mathcal{P}_t(x,x_i), \alpha_{x_i}\rangle^{\mathrm{pt}}_{x_i} && (75)\end{aligned}$$

and similarly

$$\begin{aligned}&k(\alpha_{x_i}, \alpha_{x_j}) = \\ &\quad = \lim_{m\to\infty} \left\langle \alpha_i^m(x), \langle \mathcal{P}_t(x,x'), \alpha_j^m(x')\rangle_{x'}\right\rangle_x.\end{aligned} \tag{76}$$

Therefore, letting $\alpha^m(x) = \sum_{i=1}^n \alpha_i^m(x)$ we obtain

$$\begin{aligned}&\sum_{i=1}^n\sum_{j=1}^n k(\alpha_{x_i}, \alpha_{x_j}) = && (77)\\ &\quad = \lim_{m\to\infty} \langle \alpha^m(x), \langle \mathcal{P}_t(x,x'), \alpha^m(x')\rangle_{x'}\rangle_x && (78)\\ &\quad \geq 0, && (79)\end{aligned}$$

since each term in the sequence is non-negative. □

Proposition 6. *The Gaussian vector field*

$$f(x) = \frac{\sigma}{\sqrt{C_{\nu,\kappa}}} \sum_{n=0}^{L} w_n \sqrt{\Phi_{\nu,\kappa}(\lambda_n)}\, s_n(x), \tag{18}$$

where $w_n \overset{iid}{\sim} \mathcal{N}(0,1)$, corresponds to the kernel given by the truncation of Equation (14) with the sum $\sum_{n=0}^{\infty}$ therein substituted by the sum $\sum_{n=0}^{L}$.

Proof. The mean of the random variable $f(x)$ is obviously 0. The covariance is easily computed to be

$$\begin{aligned}&\mathrm{Cov}(f(x), f(x')) = \\ &\quad = \frac{\sigma^2}{C_{\nu,\kappa}} \sum_{n=0}^{L} \Phi_{\nu,\kappa}(\lambda_n) s_n(x) \otimes s_n(x').\end{aligned} \tag{80}$$

This concludes the proof. □

As a direct consequence of Theorem 19, Proposition 20, and of the Hodge decomposition theorem, we also have the following.

Theorem 5. *There exists a Gaussian vector field $f^\bullet$ — where $\bullet \in \{\mathrm{div}, \mathrm{curl}, \mathrm{harm}\}$ —with kernel*

$$\boldsymbol{k}^\bullet_{\nu,\kappa,\sigma^2}(x,x') = \frac{\sigma^2}{C_{\nu,\kappa}} \sum_{n \in \mathcal{I}_\bullet} \Phi_{\nu,\kappa}(\lambda_n) s_n(x) \otimes s_n(x'). \quad (16)$$

What is more, $\mathrm{div}\, f^{\mathrm{curl}} = 0$, $\mathrm{curl}\, f^{\mathrm{div}} = 0$ *and* $\Delta f^{\mathrm{harm}} = 0$ *almost surely as long as $f^\bullet$ is smooth enough that* div *and* curl *are well-defined.*

There are some situations where it is only necessary to have partial information on the eigenforms of the Hodge Laplacian in order to gain full knowledge of the spectrum. We will now give the examples of surfaces—where it is enough to know the spectrum in the scalar case and the harmonic 1-forms—and product manifolds—where we only need to know the spectrum on the factors separately.

A.5 Surfaces

Suppose now that $\dim \mathrm{M} = 2$ and assume that we have an orthonormal basis $\{f_n\}_n$ of eigenfunctions of the Laplace–Beltrami operator on M with eigenvalues $\{-\lambda_n\}_n$ where $0 = \lambda_1 < \lambda_2 \leq \lambda_3 \leq \cdots$ (where we notice that the only eigenfunction with eigenvalue 0 is the constant), as well as an orthonormal basis of 1-eigenforms $\{\alpha_j\}_{0 \leq j \leq J}$ of the first de Rham cohomology $H^1_{dR}(\mathrm{M})$. Then, thanks to Hodge decomposition (Theorem 15) and Proposition 18, we have that an orthonormal basis of 1-eigenforms of $\Omega^1(\mathrm{M})$ is given by

$$\left\{ \frac{\mathrm{d}f_n}{\sqrt{\lambda_n}}, \frac{\star \mathrm{d}f_n}{\sqrt{\lambda_n}}, \alpha_j \,\middle|\, n \geq 1 \text{ and } 0 \leq j \leq J \right\}. \quad (81)$$

An orthonormal basis of 2-forms is simply given by $\star f_i = f_i \omega$, where $\omega \in \Omega^2(\mathrm{M})$ is the canonical volume form. Similarly, we can also recover all the 0- and 2-eigenforms from knowledge of the 1-eigenforms.

In the dual case we have the following result from the main body of the paper.

Theorem 8. *For each $n \geq 1$, both ∇f_n and $\star \nabla f_n$ are eigenfields of the Hodge Laplacian, and the set*

$$\left\{ \frac{\nabla f_n}{\sqrt{\lambda_n}}, \frac{\star \nabla f_n}{\sqrt{\lambda_n}}, g_j \,\middle|\, n \geq 1 \text{ and } 0 \leq j \leq J \right\} \quad (19)$$

forms an orthonormal basis of $\mathcal{L}^2(\mathrm{M}; T\mathrm{M})$.

A.6 Products

In this section, let M, N be two oriented, connected, compact Riemannian manifolds. We recall the Stone–Weierstrass theorem, see e.g. Prolla (1994).

Theorem 21 (Stone–Weierstrass)**.** *Let $A \subseteq C^\infty(\mathrm{M})$ be a subalgebra of the algebra of smooth functions on* M *which contains all constant functions and which* separates points*, i.e. such that for each $x, y \in \mathrm{M}$ there is a function $f \in A$ with $f(x) \neq f(y)$. Then A is dense in $C^\infty(\mathrm{M})$.*

With this, we get the following, e.g. Canzani (2013, Section 4.6).

Proposition 22. *For $X \in \{\mathrm{M}, \mathrm{N}\}$, let $\{f_i^X\}_i \subseteq C^\infty(X)$ be an orthonormal basis of eigenfunctions of the Laplace–Beltrami operator on X with eigenvalues $-\lambda_i^X$. Then*

$$\left\{ h_{i,j}(x_\mathrm{M}, x_\mathrm{N}) \coloneqq f_i^\mathrm{M}(x_\mathrm{M}) g_j^\mathrm{N}(x_\mathrm{N}) \,\middle|\, i, j \right\}, \quad (82)$$

where $(x_\mathrm{M}, x_\mathrm{N}) \in \mathrm{M} \times \mathrm{N}$, is an orthonormal basis of eigenfunctions of the Laplace–Beltrami operator on $\mathrm{M} \times \mathrm{N}$ with eigenvalues $-(\lambda_i^\mathrm{M} + \lambda_j^\mathrm{N})$.

Proof. It is straightforward to check that the functions $h_{i,j}$ are eigenfunctions with the prescribed eigenvalues, and that they are orthonormal. We need to show that these functions are dense in $\mathcal{L}^2(\mathrm{M} \times \mathrm{N})$. The sets $\{f_i^X\}_i \subseteq C^\infty(X)$ are both dense in the spaces of smooth functions on the respective manifolds, so that the $h_{i,j}$ are dense in the subalgebra spanned by the products $C^\infty(\mathrm{M}) \cdot C^\infty(\mathrm{N})$. By the Stone–Weierstrass theorem, this subalgebra is dense in $C^\infty(\mathrm{M} \times \mathrm{N})$, and thus, also in $\mathcal{L}^2(\mathrm{M} \times \mathrm{N})$, which concludes the proof. □

Corollary 23. *For the heat kernel on $\mathrm{M} \times \mathrm{N}$ we have*

$$\mathcal{P}_t^{\mathrm{M}\times\mathrm{N}}(x, x') = \mathcal{P}_t^\mathrm{M}(x_\mathrm{M}, x'_\mathrm{M}) \mathcal{P}_t^\mathrm{N}(x_\mathrm{N}, x'_\mathrm{N}), \quad (83)$$

for $x = (x_\mathrm{M}, x_\mathrm{N}) \in \mathrm{M} \times \mathrm{N}$ and similarly for x'.

The more general case of differential forms is similar. We need the following result, which will play the role of the Stone–Weierstrass theorem (Bryant, 2020). Recall that if A is an algebra and M is an A-module, an A-submodule of M is a linear subspace M' of M such that for each $a \in A$ and $m' \in M$ we have $a \cdot m' \in M'$.

Proposition 24. *Let $A \subseteq C^\infty(\mathrm{M})$ be as in Theorem 21 and let $E \subseteq \Omega^k(\mathrm{M})$ be an A-submodule which is such that for each point $x \in \mathrm{M}$ we have*

$$\mathrm{span}\{e(x) \mid e \in E\} = \Lambda^k T_x^* \mathrm{M}. \quad (84)$$

Then E is dense in $\Omega^k(\mathrm{M})$ with respect to the supremum norm.

Proof. First notice that we can choose a finite subset $e_1, \ldots, e_n \in E$ that spans $\Lambda^k T^* \mathrm{M}$ everywhere. Indeed, consider the collection of open sets given by

$$\{x \in \mathrm{M} \mid e'_1(x), \ldots, e'_{n'}(x) \text{ spans } \Lambda^k T_x^* \mathrm{M}\} \quad (85)$$

for all possible choices of $e'_1, \dots, e'_{n'} \in E$. By our assumptions, this covers M. Thus, by compactness we can choose a finite sub-collection that also covers M. Taking all the elements of E appearing in this finite sub-collection gives us the desired set. Every form $\alpha \in \Omega^k(\mathrm{M})$ can then be written as

$$\alpha(x) = \sum_{i=1}^{n} f_i(x) e_i(x) \tag{86}$$

for some smooth f_i (e.g. using bump functions). Each of these functions f_i can be approximated by elements of A, and since all sums of sections of the form $a \cdot e$ for $a \in A$ and $e \in E$ are in E (as it is an A-module), it follows that we can approximate α as well. □

Proposition 25. *For $X \in \{\mathrm{M}, \mathrm{N}\}$, let $\{\phi_i^{X,k}\}_i \subseteq \Omega^k(X)$ be an orthonormal basis of the eigenfields of the Hodge Laplacian with eigenvalues $\lambda_i^{X,k}$. The set*

$$\left\{ \phi_i^{\mathrm{M},k_\mathrm{M}} \wedge \phi_j^{\mathrm{N},k_\mathrm{N}} \mid k_\mathrm{M} + k_\mathrm{N} = k, i, j \right\} \tag{87}$$

is an orthonormal basis of $\Omega^k(\mathrm{M} \times \mathrm{N})$ with eigenvalues $\lambda_i^{\mathrm{M},k_\mathrm{M}} + \lambda_j^{\mathrm{N},k_\mathrm{N}}$.

Proof. Similar to Proposition 22. Checking that these forms are eigenforms of the Hodge Laplacian with the prescribed eigenvalues and that they are orthonormal is straightforward. To show that their span is dense in $\Omega^k(\mathrm{M} \times \mathrm{N})$, we notice that for $x = (x_\mathrm{M}, x_\mathrm{N}) \in \mathrm{M} \times \mathrm{N}$ we have

$$\Lambda^k T_x^*(\mathrm{M} \times \mathrm{N}) \cong \bigoplus_{k_\mathrm{M} + k_\mathrm{N} = k} T_{x_\mathrm{M}}^* \mathrm{M} \wedge T_{x_\mathrm{N}}^* \mathrm{N}. \tag{88}$$

It follows that the span our set of forms can play the role of E in Proposition 24, with the role of A being played by the products of eigenfunctions

$$A = \left\{ \phi_i^{\mathrm{M},0} \wedge \phi_j^{\mathrm{N},0} \mid i, j \right\}. \tag{89}$$

The statement follows. □

Corollary 26. *The heat kernel on $\Omega^k(\mathrm{M} \times \mathrm{N})$ is given by*

$$\begin{aligned} &\mathcal{P}_t^{\Omega^k(\mathrm{M}\times\mathrm{N})}(x, x') = \\ &\quad = \sum_{i+j=k} \mathcal{P}_t^{\Omega^i(\mathrm{M})}(x_\mathrm{M}, x'_\mathrm{M}) \mathcal{P}_t^{\Omega^j(\mathrm{N})}(x_\mathrm{N}, x'_\mathrm{N}) \end{aligned} \tag{90}$$

for $x = (x_\mathrm{M}, x_\mathrm{N}) \in \mathrm{M} \times \mathrm{N}$ and similarly for x'.

A direct application of this is Hodge–Matérn kernels on the torus, which we give in the context of vector fields.

Proposition 10. *The eigenfields on $\mathbb{T}^d$ are*

$$\left(\prod_{i=1}^{d} f_{n_i}(x_i) \right) \boldsymbol{e_j} \in T\mathbb{S}_1 \otimes \cdots \otimes T\mathbb{S}_1 \cong T\mathbb{T}^d, \tag{25}$$

$1 \le j \le d$, with eigenvalue $\sum_{i=1}^d \lambda_{n_i}$, where $x_i \in \mathbb{S}_1$, and (f_n, λ_n) are eigenpairs of the Laplace–Beltrami operator on $\mathbb{S}_1$. In particular, with $k^{\mathbb{T}^d}$ the scalar Matérn kernel on $\mathbb{T}^d$, the vector Hodge–Matérn kernel is

$$\boldsymbol{k}_{\nu,\kappa,\sigma^2}^{\mathbb{T}^d}(x, x') = \frac{1}{d} k_{\nu,\kappa,\sigma^2}^{\mathbb{T}^d}(x_i, x'_i) \boldsymbol{I}_d. \tag{26}$$

Proof. We have a canonical identification

$$T\mathbb{S}_1 \cong \mathbb{S}_1 \times \mathbb{R} \tag{91}$$

of the tangent bundle of the circle with a trivial bundle. It follows that we also have

$$T\mathbb{T}^d \cong \mathbb{T}^d \times \mathbb{R}^d \tag{92}$$

so that we can work in global coordinates as if we were working in $\mathbb{R}^d$.

The Hodge star operator gives us an identification of vector fields on the circle with functions on the circle. By Proposition 25 we obtain the desired form for the eigenfields on $\mathbb{T}^d$.

The unnormalized form of the Hodge–Matérn kernel is given by

$$\begin{aligned} &\sum_{j=1}^{d} \sum_{\boldsymbol{n} \in \mathbb{N}^d} e^{-t\lambda_{\boldsymbol{n}}} \left(\prod_{i=1}^{d} f_{n_i}(x_i) f_{n_i}(x'_i) \right) \boldsymbol{e}_j \otimes \boldsymbol{e}_j = \\ &\quad = \sum_{\boldsymbol{n} \in \mathbb{N}^d} e^{-t\lambda_{\boldsymbol{n}}} \left(\prod_{i=1}^{d} f_{n_i}(x_i) f_{n_i}(x'_i) \right) \boldsymbol{I}_d, \end{aligned} \tag{93}$$

where $\boldsymbol{n} = (n_1, \dots, n_d)$ and $\lambda_{\boldsymbol{n}} = \sum_{i=1}^d \lambda_{n_i}$. In the first part of the expression, we recognize the unnormalized form of the scalar manifold Matérn kernel on $\mathbb{T}^d$. We then notice that the normalization factors of the scalar kernel and the vector one will differ by a factor d by taking the trace of the identity matrix, which concludes the proof. □

A.7 An Alternative: Connection Laplacian

There is another possible natural extension of the Laplace–Beltrami operator to differential forms, called the connection (or Bochner) Laplacian. In fact, it can be extended to much more than just differential forms: it exists as soon as we have a vector bundle on a manifold with a nicely behaved connection and inner product. An accessible introduction to this notion for the case of differential forms can be found in Rosenberg

(1997, Section 2.2), while an extensive treatise on it, including the existence of a heat kernel, in Berline et al. (1995).

There is a formal relationship between the Hodge Laplacian and the connection Laplacian, called the Weitzenböck formula, see Rosenberg (1997, Section 2.2.2) or Lee (2009, Section 13.12): their difference is given by a term depending only on the curvature of the manifold. The Hodge theorems make it easier to work with the Hodge Laplacian by exploiting the relationships between the exterior derivative and the Hodge star operator, which is the reason why we focused on that operator in the present work. However, one gets valid kernels for GPs using the connection Laplacian as well, and this latter operator has some very nice properties that could be fruitful to use. For example, the heat kernel of the connection Laplacian asymptotically acts on vectors by parallel transport, see Berline et al. (1995, Theorem 2.30). This fact was exploited in Sharp et al. (2019) to provide efficient computational methods for parallel transport on manifolds.

B FURTHER RESULTS

This section contains the proofs of some additional results on Gaussian vector fields that were stated in the main body of the paper or that we found of general interest but could not include because of space limitations.

The first result formally shows that projected GPs built from stacking copies of known intrinsic scalar kernels will generally lead to undesirable uncertainty patterns. The second subsection studies normalization constants for the kernels of Hodge–Matérn Gaussian vector fields. The third and last subsection aims to quantify the divergence of samples from different Gaussian vector fields.

B.1 Limitations of Projected GPs

We work on the standard unit sphere $\mathbb{S}_2 \subset \mathbb{R}^3$. Let $\boldsymbol{g} = \mathbf{A}\boldsymbol{h}$ where $h_j \sim \mathrm{GP}(0, k_{\nu,\kappa_j,\sigma^2})$ for $i \in \{1,2,3\}$ are independent and $\mathbf{A} \in \mathbb{R}^{3\times 3}$ is an arbitrary matrix. Let f be the associated projected vector GP.

Proposition 2. *With notation as above, if* $\operatorname{rk} \boldsymbol{A} > 1$, *there are* $x, x' \in \mathbb{S}_2$ *so that* $\angle(x,x') = 90°$ *but*

$$\|\mathrm{Cov}(f(x), f(x'))\| < \|\mathrm{Cov}(f(x), f(\tilde{x}))\|, \quad \kappa_j \to \infty, \tag{8}$$

where $\tilde{x} = -x$ *is antipodal to* x *(i.e.* $\angle(x,\tilde{x}) = 180°$*) and* $\|\cdot\|$ *denotes the Frobenius norm.*

Proof. For simplicity of notation, we assume the κ_j are all equal. Otherwise, we can formally regard $\kappa_j \to \infty$ as $\min_j \kappa_j \to \infty$.

Take arbitrary $x, x' \in \mathbb{S}_2$. Then

$$\begin{aligned}\mathrm{Cov}(f(x), f(x')) &= \mathrm{Cov}(\mathrm{P}_x\, \boldsymbol{g}(x), \mathrm{P}_{x'}\, \boldsymbol{g}(x')) && (94)\\ &= \mathrm{P}_x\, \mathrm{Cov}(\boldsymbol{g}(x), \boldsymbol{g}(x'))\, \mathrm{P}_{x'}^\top && (95)\\ &= \mathrm{P}_x\, \mathbf{A}\, \mathrm{Cov}(\boldsymbol{h}(x), \boldsymbol{h}(x'))\mathbf{A}^\top\, \mathrm{P}_{x'}^\top\,. && (96)\end{aligned}$$

Since $\mathrm{Cov}(\boldsymbol{h}(x), \boldsymbol{h}(x')) = k_{\nu,\kappa,\sigma^2}(x,x')\mathbf{I}$ and for arbitrary x, x' we have $k_{\nu,\kappa,\sigma^2}(x,x') \to \sigma^2$ when $\kappa \to \infty$ (Borovitskiy et al., 2020), we obtain

$$\lim_{\kappa\to\infty} \mathrm{Cov}(f(x), f(x')) = \sigma^2\, \mathrm{P}_x\, \mathbf{A}\mathbf{A}^\top\, \mathrm{P}_{x'}^\top =: \mathbf{C}_{xx'}. \tag{97}$$

Without loss of generality, we can assume $\sigma^2 = 1$. To analyze $\mathbf{C}_{xx'}$ we construct the singular value decomposition (SVD) $\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top$ of the matrix $\mathbf{A}$. Here $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices and $\boldsymbol{\Sigma}$ is a diagonal matrix with non-negative entries. Then

$$\mathbf{A}\mathbf{A}^\top = \mathbf{U}\boldsymbol{\Sigma}\underbrace{\mathbf{V}^\top\mathbf{V}}_{\mathbf{I}}\boldsymbol{\Sigma}\mathbf{U}^\top = \mathbf{U}\underbrace{\boldsymbol{\Sigma}^2}_{=:\boldsymbol{\Lambda}}\mathbf{U}^\top = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^\top \tag{98}$$

where the right-hand side is the eigendecomposition of the matrix $\mathbf{A}\mathbf{A}^\top$. We denote $\lambda_i = \boldsymbol{\Lambda}_{ii}$ and, without any loss of generality, assume that $\lambda_1 \le \lambda_2 \le \lambda_3$. By assumption, $\lambda_2, \lambda_3 > 0$.

The columns $\mathbf{U}_{\cdot j}$ of $\mathbf{U}$, $j = 1,..,3$, form an orthonormal basis of $\mathbb{R}^3$. We choose $x = \mathbf{U}_{\cdot 1}$ and $x' = \mathbf{U}_{\cdot 2}$. Then

$$\mathrm{P}_x\, \mathbf{U}_{\cdot j} = \mathrm{P}_{\tilde{x}}\, \mathbf{U}_{\cdot j} = \begin{cases} 0 & \text{for } j = 1,\\ 1 & \text{for } j = 2,3.\end{cases} \tag{99}$$

$$\mathrm{P}_{x'}\, \mathbf{U}_{\cdot j} = \begin{cases} 0 & \text{for } j = 2,\\ 1 & \text{for } j = 1,3.\end{cases} \tag{100}$$

It follows that

$$\begin{aligned}\mathbf{C}_{xx'} &= (\mathbf{0}, \mathbf{U}_{\cdot 2}, \mathbf{U}_{\cdot 3})\, \boldsymbol{\Lambda}\, (\mathbf{U}_{\cdot 1}, \mathbf{0}, \mathbf{U}_{\cdot 3})^\top && (101)\\ &= \lambda_1 \mathbf{0}\, \mathbf{U}_{\cdot 1}^\top + \lambda_2 \mathbf{U}_{\cdot 2}\, \mathbf{0}^\top + \lambda_3 \mathbf{U}_{\cdot 3}\, \mathbf{U}_{\cdot 3}^\top && (102)\\ &= \lambda_3 \mathbf{U}_{\cdot 3}\, \mathbf{U}_{\cdot 3}^\top. && (103)\end{aligned}$$

Analogously, $C_{x\tilde{x}} = \lambda_2 \mathbf{U}_{\cdot 2}\, \mathbf{U}_{\cdot 2}^\top + \lambda_3 \mathbf{U}_{\cdot 3}\, \mathbf{U}_{\cdot 3}^\top$. Thus, we have

$$\|\mathbf{C}_{xx'}\|_F = \sqrt{\operatorname{tr} \mathbf{C}_{xx'}\mathbf{C}_{xx'}^\top} = \lambda_3 \tag{104}$$

and similarly $\|\mathbf{C}_{x\tilde{x}}\|_F = \sqrt{\lambda_2^2 + \lambda_3^2}$, proving the claim. □

B.2 Kernel Normalization Constants

In Section 3.2, we defined the normalization constant for the kernels of Hodge–Matérn Gaussian vector fields implicitly by requiring that

$$\frac{1}{\operatorname{vol} \mathrm{M}} \int_{\mathrm{M}} \operatorname{tr}\big(\boldsymbol{k}(x,x)\big)\, \mathrm{d}x = \sigma^2. \tag{105}$$

We will now explain what this normalization means in practice and make these constants explicit for the kernels of Hodge–Matérn Gaussian vector fields and projected Matérn Gaussian vector fields.

Proposition 27. *Assume (105) holds, then*

$$\frac{1}{\operatorname{vol} \mathrm{M}} \mathbb{E}_{f \sim \mathrm{GP}(0, \boldsymbol{k})} \left[\|f\|^2_{\mathcal{L}^2(\mathrm{M})} \right] = \sigma^2. \tag{106}$$

Proof. We have

$$\begin{aligned}
&\frac{1}{\operatorname{vol} \mathrm{M}} \mathbb{E}_{f \sim \mathrm{GP}(0, \boldsymbol{k})} \left[\|f\|^2_{\mathcal{L}^2(\mathrm{M})} \right] = && (107)\\
&= \frac{1}{\operatorname{vol} \mathrm{M}} \mathbb{E}_{f \sim \mathrm{GP}(0, \boldsymbol{k})} \left[\int_{\mathrm{M}} \|f(x)\|_2^2 \, \mathrm{d}x \right] && (108)\\
&= \frac{1}{\operatorname{vol} \mathrm{M}} \int_{\mathrm{M}} \mathbb{E}_{f \sim \mathrm{GP}(0, \boldsymbol{k})} \left[\|f(x)\|_2^2 \right] \mathrm{d}x && (109)\\
&= \frac{1}{\operatorname{vol} \mathrm{M}} \int_{\mathrm{M}} \mathbb{E}_{f(x) \sim \mathcal{N}(0, \boldsymbol{k}(x,x))} \left[\|f(x)\|_2^2 \right] \mathrm{d}x && (110)\\
&= \frac{1}{\operatorname{vol} \mathrm{M}} \int_{M} \operatorname{tr} \left(\boldsymbol{k}(x,x) \right) \mathrm{d}x && (111)\\
&= \sigma^2, && (112)
\end{aligned}$$

where the last equality holds by (105). □

Proposition 28. *The constant $C_{\nu,\kappa}$ for the Hodge–Matérn kernel $\boldsymbol{k}_{\nu,\kappa,\sigma^2}$ is given by*

$$C_{\nu,\kappa} = \frac{1}{\operatorname{vol} \mathrm{M}} \sum_{n=0}^{\infty} \Phi_{\nu,\kappa}(\lambda_n), \tag{113}$$

where the sum runs over all of the eigenfields of the Hodge Laplacian. Similar formulas are valid for the pure divergence, pure curl, and harmonic kernels by restricting the sum to the appearing eigenfields.

Proof. We have

$$\begin{aligned}
&\int_{\mathrm{M}} \operatorname{tr}(\boldsymbol{k}_{\nu,\kappa,\sigma^2}(x,x)) \, \mathrm{d}x = && (114)\\
&= \int_{\mathrm{M}} \operatorname{tr} \left(\frac{\sigma^2}{C_{\nu,\kappa}} \sum_{n=0}^{\infty} \Phi_{\nu,\kappa}(\lambda_n) s_n(x) \otimes s_n(x) \right) \mathrm{d}x && (115)\\
&= \frac{\sigma^2}{C_{\nu,\kappa}} \sum_{n=0}^{\infty} \Phi_{\nu,\kappa}(\lambda_n) \int_{\mathrm{M}} \operatorname{tr} \left(s_n(x) \otimes s_n(x) \right) \mathrm{d}x && (116)\\
&= \frac{\sigma^2}{C_{\nu,\kappa}} \sum_{n=0}^{\infty} \Phi_{\nu,\kappa}(\lambda_n) \int_{\mathrm{M}} \|s_n(x)\|_2^2 \, \mathrm{d}x && (117)\\
&= \frac{\sigma^2}{C_{\nu,\kappa}} \sum_{n=0}^{\infty} \Phi_{\nu,\kappa}(\lambda_n), && (118)
\end{aligned}$$

since $\|s_n\|_{\mathcal{L}^2(\mathrm{M})} = 1$. Notice that here the trace is taken with respect to the metric on M. The result follows immediately by requiring the left-hand side of the equation to equal $\sigma^2 \operatorname{vol} \mathrm{M}$. □

Proposition 29. *The appropriately normalized projected Matérn kernel (with trivial coregionalization matrix) on a manifold* M *of dimension d embedded in $\mathbb{R}^N$ is given by*

$$\boldsymbol{k}^{\pi}_{\nu,\kappa,\sigma^2}(x,x') = \frac{1}{d} k_{\nu,\kappa,\sigma^2}(x,x') P_x^T \boldsymbol{I}_N P_{x'}, \tag{119}$$

where k_{ν,κ,σ^2} is the scalar manifold Matérn kernel.

Proof. Let $x = x'$ and pick a coordinate system such that $T_x\mathrm{M} = \operatorname{span}\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_d\}$. Then we have

$$P_x^T \boldsymbol{I}_N P_{x'} = \operatorname{diag}(\underbrace{1, \dots, 1}_{d}, 0, \dots, 0). \tag{120}$$

It follows that

$$\begin{aligned}
&\frac{1}{\operatorname{vol} \mathrm{M}} \int_{\mathrm{M}} \operatorname{tr} \left(\boldsymbol{k}^{P}_{\nu,\kappa,\sigma^2}(x,x) \right) \mathrm{d}x = && (121)\\
&= \frac{1}{\operatorname{vol} \mathrm{M}} \int_{\mathrm{M}} k_{\nu,\kappa,\sigma^2}(x,x) \, \mathrm{d}x && (122)\\
&= \sigma^2, && (123)
\end{aligned}$$

as desired. □

B.3 Divergence of Gaussian Vector Fields

We will now study the distribution of the (pointwise) divergence of the Hodge–Matérn Gaussian vector fields and projected Matérn GPs. Similar techniques can be used to compute the full distribution for $\mathrm{d}\alpha$ and $\mathrm{d}^\star\alpha$ for $\alpha \sim \mathrm{GP}(0, \boldsymbol{k}_{\nu,\kappa,\sigma^2})$, which turns out to be another Gaussian process.

We fix a compact, oriented Riemannian manifold M of dimension $d_{\mathrm{M}} \geq 1$ and we look at degree $k = 1$ differential forms, although the computations straightforwardly generalize to all other k. We write $f_n \in C^\infty(\mathrm{M}) = \Omega^0(\mathrm{M})$ for a basis of eigenfunctions with eigenvalues $-\lambda_n$, for $n \geq 0$, where $0 = \lambda_0 < \lambda_1 \leq \cdots$. We also set $\phi_n, n \in \mathbb{N}$ a basis of 1-eigenforms with eigenvalues including the eigenforms

$$\frac{1}{\sqrt{\lambda_n}} \, \mathrm{d}f_n \tag{124}$$

for $n \geq 1$. We use $C^k_{\nu,\kappa}$ to denote the normalization constant for the Hodge–Matérn kernel on k-forms (with $k = 0$ being the case of functions). For $\alpha \sim \mathrm{GP}(0, \boldsymbol{k})$ a Gaussian differential form, we write $\operatorname{div} \alpha(x)$ for the random variable $\mathrm{d}^\star\alpha(x)$, where $x \in \mathrm{M}$. We assume $\operatorname{div} \alpha(x)$ is well-defined, i.e. the Gaussian process is smooth enough, which places restrictions on the parameter ν. We leave the precise nature of these out of the scope of this paper.

Proposition 30. *For the Hodge–Matérn Gaussian form $\alpha_{\nu,\kappa,\sigma} \sim \mathrm{GP}(0, \boldsymbol{k}_{\nu,\kappa,\sigma^2})$ on 1-forms we have*

$$\mathrm{Var}(\operatorname{div} \alpha_{\nu,\kappa,\sigma}(x)) = \frac{\sigma^2}{C^1_{\nu,\kappa}} \sum_{n=1}^{\infty} \lambda_n \Phi_{\nu,\kappa}(\lambda_n) f_n(x)^2 \quad (125)$$

whenever $\alpha_{\nu,\kappa,\sigma}$ is smooth enough for the divergence to be well-defined.

Proof. By Proposition 6 we have

$$\begin{aligned}
\operatorname{div}\alpha_{\nu,\kappa,\sigma}(x) &= \mathrm{d}^\star \alpha_{\nu,\kappa,\sigma}(x) && (126)\\
&= \mathrm{d}^\star \left(\frac{\sigma}{\sqrt{C^1_{\nu,\kappa}}} \sum_{n=0}^{\infty} \sqrt{\Phi_{\nu,\kappa}(\lambda_n)} w_n \phi_n(x) \right) && (127)\\
&= \frac{\sigma}{\sqrt{C^1_{\nu,\kappa}}} \sum_{n=0}^{\infty} \sqrt{\Phi_{\nu,\kappa}(\lambda_n)} w_n \, \mathrm{d}^\star \phi_n(x) && (128)\\
&= \frac{\sigma}{\sqrt{C^1_{\nu,\kappa}}} \sum_{n=1}^{\infty} \sqrt{\frac{\Phi_{\nu,\kappa}(\lambda_n)}{\lambda_n}} w_n \, \mathrm{d}^\star \, \mathrm{d} f_n(x) && (129)\\
&= \frac{\sigma}{\sqrt{C^1_{\nu,\kappa}}} \sum_{n=1}^{\infty} \sqrt{\lambda_n \Phi_{\nu,\kappa}(\lambda_n)} w_n f_n(x), && (130)
\end{aligned}$$

where in the third line we used that $\mathrm{d}^\star \phi_n(x) = 0$ unless $\phi_n(x)$ is of the form (124) and in the last line we used the fact that $\mathrm{d}^\star \, \mathrm{d} f_n(x) = -\Delta f_n(x) = \lambda_n f_n(x)$. This gives the full distribution for $\operatorname{div} \alpha_{\nu,\kappa,\sigma}(x)$, and in particular the desired formula for the variance. Note that Equation (130) can fail to converge when ν is not large enough (i.e. $\alpha_{\nu,\kappa,\sigma}$ is not smooth enough). □

Corollary 31. *On the sphere, in terms of vector fields, we have*

$$\mathrm{Var}(\operatorname{div} f_{\nu,\kappa,\sigma}(x)) = \frac{\sigma^2}{2} \frac{\sum_{n=1}^{\infty} \lambda_n \Phi_{\nu,\kappa}(\lambda_n)}{4\pi \big(C^0_{\nu,\kappa} - \Phi_{\nu,\kappa}(0) \big)} \quad (131)$$

for any $x \in \mathbb{S}_2$.

Proof. By symmetry, $\mathrm{Var}(\operatorname{div} f_{\nu,\kappa,\sigma}(x))$ does not depend on x, i.e. it is a constant. Therefore, we have

$$\begin{aligned}
4\pi \, \mathrm{Var}(\operatorname{div} f_{\nu,\kappa,\sigma}(x)) &= && (132)\\
&= \int_{\mathrm{M}} \mathrm{Var}(\operatorname{div} f_{\nu,\kappa,\sigma}(x')) \, \mathrm{d}x' && (133)\\
&= \int_{\mathrm{M}} \frac{\sigma^2}{C^1_{\nu,\kappa}} \sum_{n=1}^{\infty} \lambda_n \Phi_{\nu,\kappa}(\lambda_n) f_n(x')^2 \, \mathrm{d}x' && (134)\\
&= \frac{\sigma^2}{C^1_{\nu,\kappa}} \sum_{n=1}^{\infty} \lambda_n \Phi_{\nu,\kappa}(\lambda_n) && (135)
\end{aligned}$$

since $\|f_n\|_{\mathcal{L}^2(\mathrm{M})} = 1$. The statement follows by noticing that each eigenvalue in the spectrum on functions appears twice in the spectrum on vector fields—except for 0, which does not appear. □

We now look at the projected kernel. Suppose $\phi : \mathrm{M} \to \mathbb{R}^N$ is an isometric embedding and write $\boldsymbol{k}^{\pi}_{\nu,\kappa,\sigma^2}$ for the projected kernel obtained by projecting the vector kernel in $\mathbb{R}^N$ where each component is a scalar manifold Matérn kernel with the same hyperparameters ν, κ, σ^2. In this case, we will talk about vectors and gradients instead of differential forms (although a formulation in terms of 1-forms is also possible).

We write $P : \phi^*(T\mathbb{R}^N) \to T\mathrm{M}$ for the orthogonal projection to the tangent bundle of M. The proof of the following helpful lemma was provided by Alan Pinoy in a private communication.

Lemma 32. *Let $f : \mathrm{M} \to \mathbb{R}$ be a smooth function and let $\boldsymbol{w} \in \mathbb{R}^N$ be a fixed vector. Then*

$$\operatorname{div}(f(x) P_x \boldsymbol{w}) = \big(\nabla f(x) + f(x) H(x) \big) \cdot \boldsymbol{w}, \quad (136)$$

where H is the mean curvature vector of the embedding defined in Equation (138).

Proof. We denote by ∇ the Levi–Civita connection of M and by $\overline{\nabla}$ that of $\mathbb{R}^N$. Recall that the (vector) *second fundamental form* of the embedding is given by

$$\mathrm{I\!I}(u, v) = \overline{\nabla}_u v - \nabla_u v \in \phi^*(T\mathbb{R}^N), \quad (137)$$

cf. Lee (2009, Section 4.2) (for the scalar version). Intuitively, it measures the infinitesimal curvature of M inside of $\mathbb{R}^N$. Notice that it is always orthogonal to the tangent space of M. The *mean curvature vector* is the trace of the second fundamental form: if $e_1, \ldots, e_{d_\mathrm{M}}$ is a local orthonormal frame of M, then

$$H(x) = \sum_{i=1}^{d_\mathrm{M}} \mathrm{I\!I}(e_i, e_i). \quad (138)$$

With these definitions at hand, we have

$$\operatorname{div}(f(x) P_x \boldsymbol{w}) = \nabla f(x) \cdot \boldsymbol{w} + f(x) \operatorname{div}(P_x \boldsymbol{w}). \quad (139)$$

We use the standard differential geometric notation $v(f)$ for the derivative of f in direction v, where f is a function and v is a vector field. For this last term, with g denoting the metric on M, we compute

$$\begin{aligned}
\operatorname{div}(P_x \boldsymbol{w}) &= \operatorname{tr}\big(\nabla (P_x \boldsymbol{w}) \big) && (140)\\
&= \sum_{i=1}^{d_\mathrm{M}} g(\nabla_{e_i} (P_x \boldsymbol{w}), e_i) && (141)\\
&= \sum_{i=1}^{d_\mathrm{M}} e_i \big(g(P_x \boldsymbol{w}, e_i) \big) - g(P_x \boldsymbol{w}, \nabla_{e_i} e_i) && (142)\\
&= \sum_{i=1}^{d_\mathrm{M}} e_i (P_x \boldsymbol{w} \cdot e_i) - P_x \boldsymbol{w} \cdot \nabla_{e_i} e_i && (143)\\
&= \sum_{i=1}^{d_\mathrm{M}} e_i (\boldsymbol{w} \cdot e_i) - \boldsymbol{w} \cdot \nabla_{e_i} e_i && (144)
\end{aligned}$$

$$= \sum_{i=1}^{d_{\mathrm{M}}} \overline{\nabla}_{e_i} \boldsymbol{w} \cdot e_i + \boldsymbol{w} \cdot \overline{\nabla}_{e_i} e_i - \boldsymbol{w} \cdot \nabla_{e_i} e_i \quad (145)$$

$$= \sum_{i=1}^{d_{\mathrm{M}}} \boldsymbol{w} \cdot \left(\overline{\nabla}_{e_i} e_i - \nabla_{e_i} e_i\right) \quad (146)$$

$$= \boldsymbol{w} \cdot \left(\sum_{i=1}^{d_{\mathrm{M}}} \mathrm{I\!I}(e_i, e_i)\right) \quad (147)$$

$$= \boldsymbol{w} \cdot H(x), \quad (148)$$

where (142) comes from the fact that ∇ is the Levi–Civita connection, (143) from the fact that ϕ is an isometric embedding, and (145) comes from the fact that $\overline{\nabla}$ is the Levi–Civita connection of $\mathbb{R}^N$ (the metric being given by the dot product). Inserting this in (139) concludes the proof. □

Proposition 33. *For the projected Matérn GP* $f^{\pi}_{\nu,\kappa,\sigma^2} \sim \mathrm{GP}(0, \boldsymbol{k}^{\pi}_{\nu,\kappa,\sigma^2})$ *we have*

$$\mathrm{Var}(\operatorname{div} f^{\pi}_{\nu,\kappa,\sigma^2}) = \quad (149)$$

$$= \frac{\sigma^2}{d_{\mathrm{M}} C^0_{\nu,\kappa}} \left(\sum_{n=1}^{\infty} \lambda_n \Phi_{\nu,\kappa}(\lambda_n) \left\| \frac{\nabla f_n(x)}{\sqrt{\lambda_n}} \right\|_2^2 + \right. \quad (150)$$

$$\left. + \sum_{n=0}^{\infty} \Phi_{\nu,\kappa}(\lambda_n) f_n(x)^2 \left\| H(x) \right\|_2^2 \right) \quad (151)$$

whenever $f^{\pi}_{\nu,\kappa,\sigma^2}$ *is smooth enough for the divergence to be well-defined.*

Proof. Once again, by Proposition 6 we have

$$f^{\pi}_{\nu,\kappa,\sigma^2}(x) = \\ = \frac{\sigma}{\sqrt{d_{\mathrm{M}} C^0_{\nu,\kappa}}} \sum_{n=0}^{\infty} \sqrt{\Phi_{\nu,\kappa}(\lambda_n)} f_n(x) P_x \boldsymbol{w}_n \quad (152)$$

where $\boldsymbol{w}_n \sim \mathcal{N}(0, \boldsymbol{I}_N)$ is a sequence of i.i.d. multivariate normal vectors. Taking the divergence by applying Lemma 32 to each summand, we obtain

$$\operatorname{div} f^{\pi}_{\nu,\kappa,\sigma^2}(x) = \\ = \frac{\sigma}{\sqrt{d_{\mathrm{M}} C^0_{\nu,\kappa}}} \sum_{n=1}^{\infty} \sqrt{\Phi_{\nu,\kappa}(\lambda_n)} \nabla f_n(x) \cdot \boldsymbol{w}_n + \\ + \frac{\sigma}{\sqrt{d_{\mathrm{M}} C^0_{\nu,\kappa}}} \sum_{n=0}^{\infty} \sqrt{\Phi_{\nu,\kappa}(\lambda_n)} f_n(x) H(x) \cdot \boldsymbol{w}_n. \quad (153)$$

Since the $\boldsymbol{w}_n$ are independent and

$$\|\nabla f_n(x) + f_n(x) H(x)\|^2 = \\ = \|\nabla f_n(x)\|^2 + f_n(x)^2 \|f_n(x) H(x)\|^2 \quad (154)$$

as $\nabla f_n(x)$ and $H(x)$ are orthogonal, the statement follows. □

Corollary 34. *On the sphere we have*

$$\mathrm{Var}(\operatorname{div} f^{\pi}_{\nu,\kappa,\sigma^2}) = \\ = \frac{\sigma^2}{2} \left(\frac{\sum_{n=1}^{\infty} \lambda_n \Phi_{\nu,\kappa}(\lambda_n)}{4\pi C^0_{\nu,\kappa}} + 1 \right) \quad (155)$$

for any $x \in \mathbb{S}_2$.

Proof. By the same symmetry argument as in Corollary 31 and using the fact that for the standard unit sphere embedding we have $H(x) = \boldsymbol{n}(x)$ the unit normal vector, we obtain

$$4\pi \, \mathrm{Var}(\operatorname{div} f^{\pi}_{\nu,\kappa,\sigma^2}) = \\ = \frac{\sigma^2}{2C^0_{\nu,\kappa}} \left(\sum_{n=1}^{\infty} \lambda_n \Phi_{\nu,\kappa}(\lambda_n) + 4\pi C^0_{\nu,\kappa} \right), \quad (156)$$

which is the stated result. □

C ADDITIONAL EXPERIMENTAL DETAILS

C.1 Weather Modeling

In the weather modeling experiments, the signal was normalized before fitting by scaling it by a constant scalar value so that the mean norm of the training observations is 1.

The training points were selected by taking the observations at longitudes 90°E and 90°W, and then picking one every 180 of them, spaced regularly. This gives a final training set of 34 points. The 1220 testing points were generated randomly.[8]

Using the results of Appendix B.3, it was possible to quantify the variance of the divergence of the Gaussian vector fields arising from the (prior) kernels that were fitted in the experiments. The results are displayed in Figure 10, confirming that the absolute divergence was higher for the Hodge heat and Hodge–Matérn kernels, which explains the worse performance of these intrinsic kernels against the projected kernels for this experiment.

An examination of the fitted hyperparameters revealed that the typical length scale for the divergence-free Hodge–Matérn kernels is between $\kappa = 0.1$ and $\kappa = 0.3$. In some rare exceptions, hyperparameter fitting converged to a local optimum with large length scale, but performance was not overly affected. After visual inspection of the ground truth, we fitted divergence-free Hodge–Matérn kernels with fixed length scales of $\kappa = 0.5$ and $\kappa = 1$. The results are reported in Table 2. We notice potential small improvements in performance, but not statistically significant ones.

[8] Specifically, we used the Poisson disk sampling routine `create_poisson_disk_samples` on the sphere from https://github.com/aterenin/phdthesis.

In the last experiment on this data, we fitted the Hodge-compositional Matérn kernels, i.e. linear combinations of pure-divergence and pure-curl Hodge–Matérn kernels, as reported also in Table 1. We see that these kernels recover almost exactly the results of fitting a divergence-free Hodge–Matérn kernel. A detailed analysis of the fitted hyperparameters shows that the resulting length scales and variances put full weight on the divergence-free part of the kernel with an almost exact match of length scales. The only exceptions are when the length scale converges to a local optimum, which fully explains the minimal advantage in performance of the linear combination kernel over the divergence-free Hodge–Matérn kernel with $\nu = \frac{1}{2}$. This supports the fact that Hodge-compositional Matérn kernels are able to automatically recover appropriate inductive biases in such situations, mirroring their discrete counterparts in Yang et al. (2024).

C.2 Synthetic Experiments

In addition to the weather modeling experiment detailed in the main body of the paper, we ran various experiments on synthetically generated data.

Namely, we considered samples drawn from each of the following GPs on the sphere: Hodge–Matérn ($\nu = \infty, \kappa = 0.5$), projected Matérn ($\nu = \frac{1}{2}, \kappa = 0.5$), Hodge–Matérn ($\nu = \frac{1}{2}, \kappa = 0.5$), and curl-free Hodge–Matérn ($\nu = \frac{1}{2}, \kappa = 0.5$). In addition to this, we also used the "rotation" vector field of Figure 1b, which is given by

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} y \\ -x \\ 0 \end{pmatrix}. \tag{157}$$

This vector field is pure curl, as it can obtained as the curl of the function $f(x, y, z) = z$ on the sphere. For each experiment, 30 training points were selected uniformly at random from the northern hemisphere, and 100 testing points were selected—also uniformly at random—from the southern hemisphere. Each experiment was repeated 10 times: for the rotation vector field the training and testing points were resampled at each experiment, while for the others a new sample was drawn for each experiment.

In each experiment, we fitted the following GPs: pure noise, projected Matérn (P. M.), Hodge–Matérn (H.–M.), divergence-free Hodge–Matérn (div-free H.–M.), and curl-free Hodge–Matérn (curl-free H.–M.), all with $\nu = \frac{1}{2}, \infty$. All kernels also fitted variance and an additive noise.

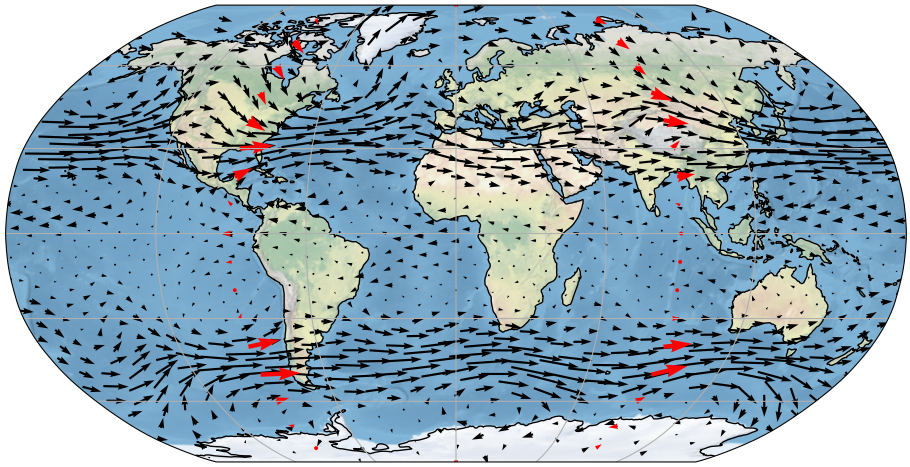
The results are reported in Tables 3 and 4. We see that for each experiment based on samples, the respective kernel performed best. On the rotation vector field, Hodge–Matérn and divergence-free Hodge–Matérn vastly outperformed all other kernels.

Kernel	H.–M.–$\frac{1}{2}$ sample		H.–M.–∞ sample		P. M.–$\frac{1}{2}$ sample		Rotation field		curl-free H.–M.–$\frac{1}{2}$ sample	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Pure noise	0.17	0.04	1.18	0.32	0.22	0.06	0.68	0.02	0.08	0.02
P. M.–$\frac{1}{2}$	0.14	0.03	0.87	0.42	**0.16**	0.05	0.06	0.02	0.07	0.02
P. M.–∞	0.19	0.04	0.71	0.26	0.20	0.07	0.34	0.36	0.07	0.03
H.–M.–$\frac{1}{2}$	**0.14**	0.03	0.84	0.38	0.16	0.05	0.02	0.01	0.07	0.02
H.–M.–∞	0.17	0.04	**0.65**	0.25	0.20	0.06	0.00	0.00	0.08	0.02
curl-free H.–M.–$\frac{1}{2}$	0.20	0.06	1.15	0.47	0.22	0.07	0.68	0.02	**0.05**	0.01
curl-free H.–M.–∞	0.16	0.04	1.11	0.37	0.23	0.06	0.69	0.03	0.08	0.02
div-free H.–M.–$\frac{1}{2}$	0.15	0.05	1.00	0.60	0.19	0.05	0.01	0.00	0.08	0.02
div-free H.–M.–∞	0.16	0.04	0.81	0.46	0.18	0.04	**0.00**	0.00	0.08	0.02

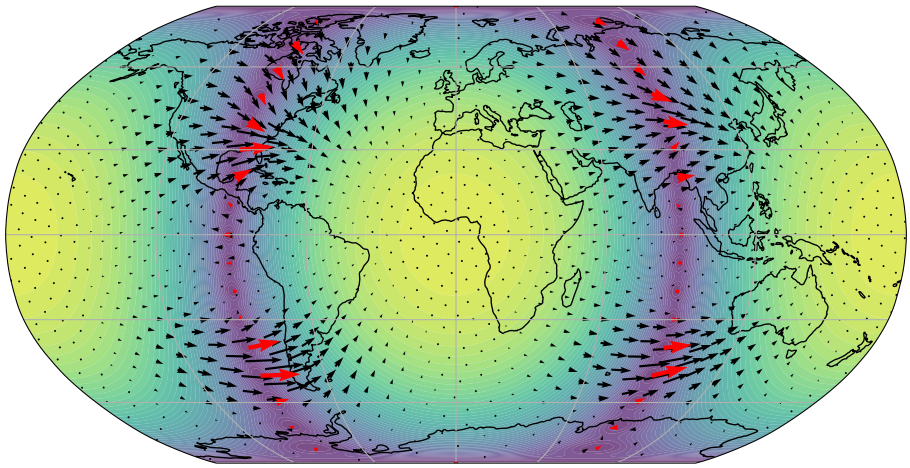
Table 3: MSE for synthetic experiments. The columns are datasets, the rows are models.

Kernel	H.–M.–$\frac{1}{2}$ sample		H.–M.–∞ sample		P. M.–$\frac{1}{2}$ sample		Rotation field		curl-free H.–M.–$\frac{1}{2}$ sample	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Pure noise	0.41	0.25	2.39	0.34	0.67	0.29	1.76	0.04	-0.31	0.25
P. M.–$\frac{1}{2}$	0.15	0.26	2.18	0.78	**0.31**	0.27	-0.65	0.20	-0.51	0.34
P. M.–∞	0.61	0.39	1.47	0.63	0.58	0.40	-3.43	5.47	-0.38	0.73
H.–M.–$\frac{1}{2}$	**0.13**	0.24	2.12	0.71	0.33	0.28	-1.41	0.15	-0.58	0.28
H.–M.–∞	0.41	0.25	**1.27**	0.48	0.53	0.31	-9.42	0.04	-0.31	0.25
curl-free H.–M.–$\frac{1}{2}$	0.66	0.59	2.66	0.71	0.67	0.37	1.77	0.03	**-0.79**	0.21
curl-free H.–M.–∞	0.38	0.30	2.55	0.58	0.73	0.32	1.77	0.05	-0.31	0.25
div-free H.–M.–$\frac{1}{2}$	0.25	0.31	2.37	1.16	0.48	0.31	-2.20	0.17	-0.37	0.31
div-free H.–M.–∞	0.33	0.25	1.52	0.66	0.46	0.18	**-9.63**	0.00	-0.31	0.25

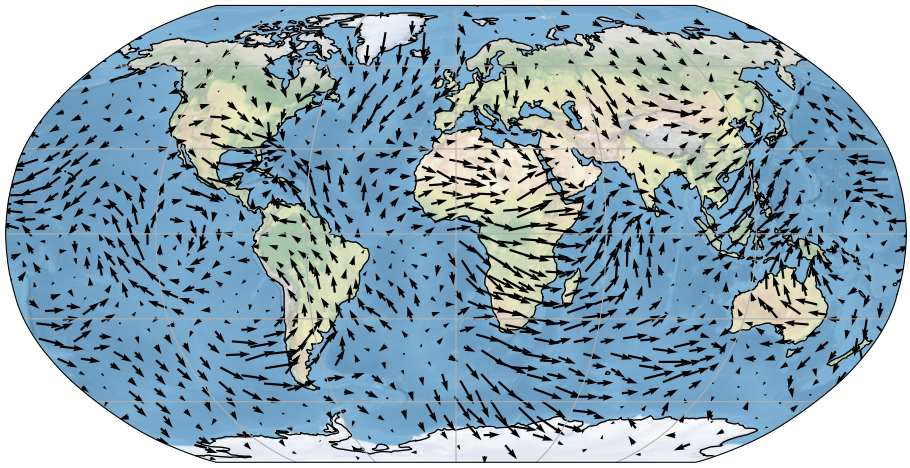
Table 4: Predictive NLL for synthetic experiments. The columns are datasets, the rows are models.



(a) Ground truth (January 2010) and observations.

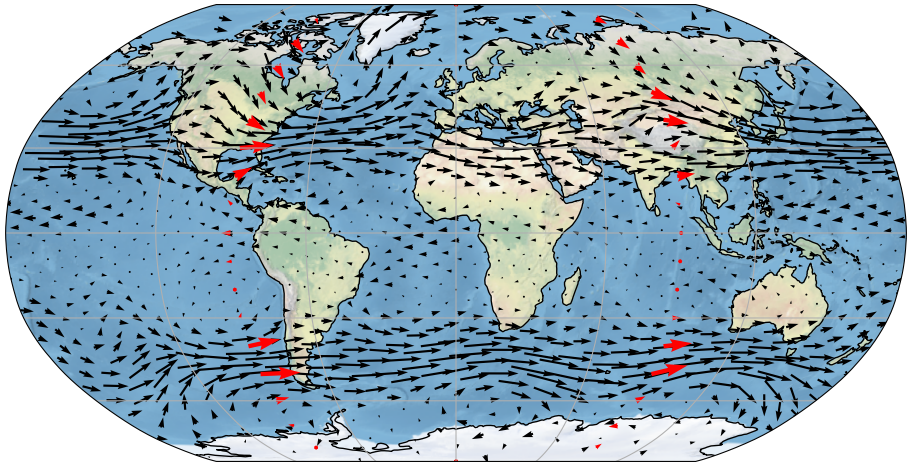


(b) Predictive mean and uncertainty (blue is low and yellow is high).

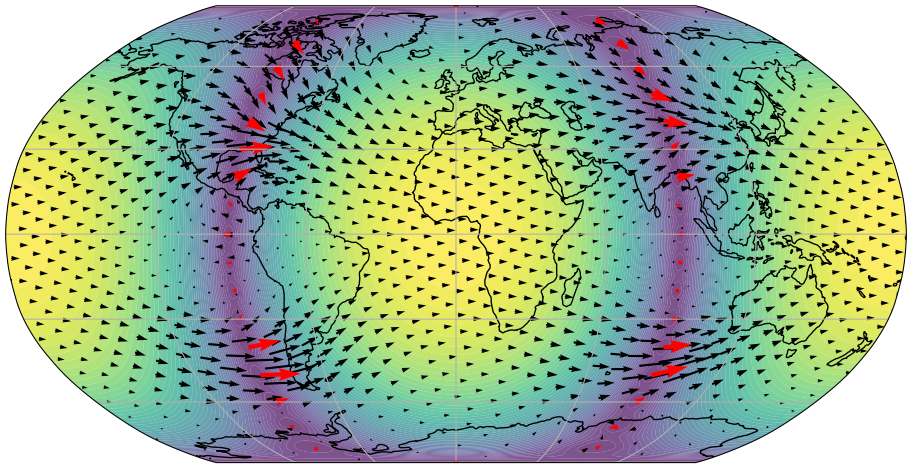


(c) Posterior sample.

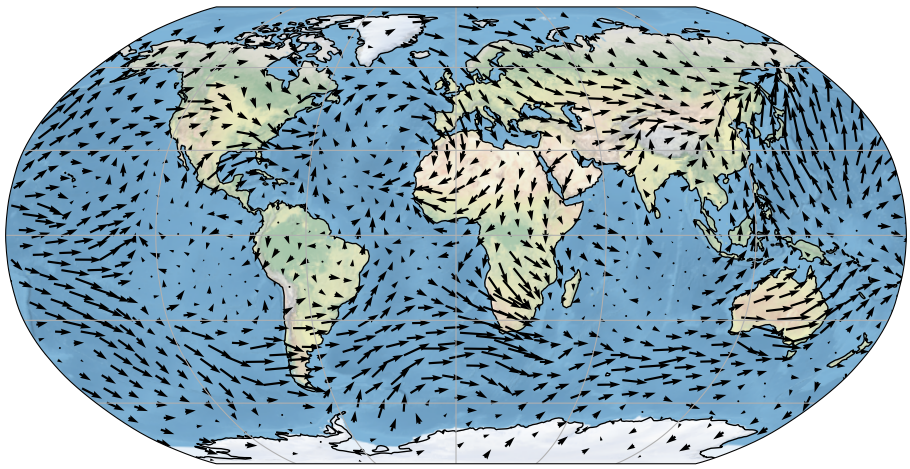
Figure 8: Robinson projection of Figures 6a, 6b and 7a displaying the ground truth, observations, predictive mean, uncertainty, and a posterior sample of the GP with projected Matérn kernel with $\nu = \frac{1}{2}$ and length scale $\kappa = 0.5$. The vectors in the sample are scaled independently from the ground truth and predictive mean. Visually, the sample does not have structures reminiscent of that of the ground truth.



(a) Ground truth (January 2010) and observations.



(b) Predictive mean and uncertainty (blue is low and yellow is high).



(c) Posterior sample.

Figure 9: Robinson projection of Figures 6b, 6c and 7b displaying the ground truth, observations, predictive mean, uncertainty, and a posterior sample of the GP with divergence-free Hodge–Matérn kernel with $\nu = \frac{1}{2}$ and length scale $\kappa = 0.5$. The vectors in the sample are scaled independently from the ground truth and predictive mean. Visually, and to the contrary of Figure 8, the sample appears to have structures reminiscent of that of the ground truth, such as a strong west to east current in the southern hemisphere.

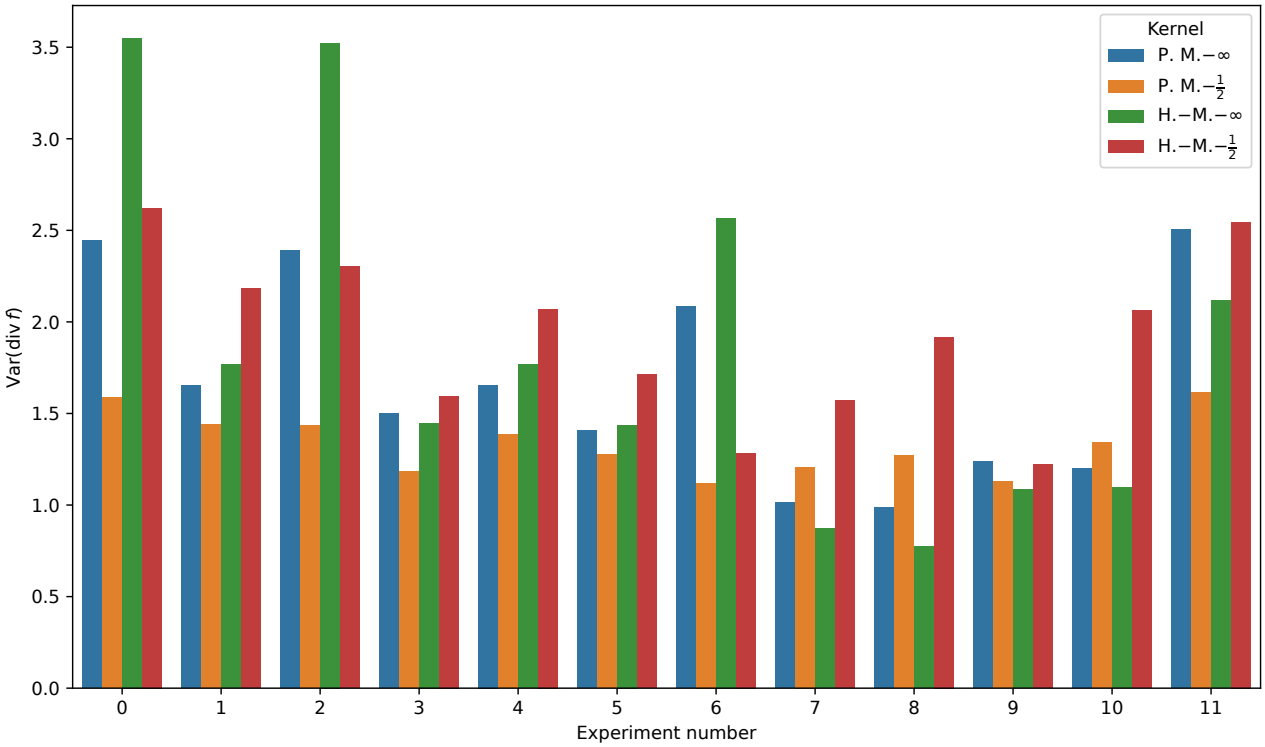


Figure 10: Variance of divergence for the prior kernels with fitted hyperparameters in the weather modeling experiments. Note that the variance of the divergence does not depend on the input location in this case. The divergence-free kernels and the kernels with fixed length scale are not represented here.