Modeling Neural Random Field through Operator-based Diffusion

Guanyu Chen
College of Integrated Circuits
Zhejiang Uinversity
Hangzhou, Zhejiang
guanyu@zju.edu.cn

Abstract

1 Introduction

Definition 1. *Random field $\mathcal{M}(x,\omega)$, where $x \in D$ and $\omega \in \Omega$, is defined as:*

$$\begin{aligned} &\mathcal{M}(x,\cdot) \text{ is a random variable defined on the probability space } (\Omega,\mathcal{F},P), \\ &\mathcal{M}(\cdot,\omega) \text{ is a deterministic function of } x. \end{aligned} \tag{1}$$

Classical methods to simulate random field are based on polynomial chaos expansion B and Karhunen-Loeve expansion C, See [1]. Random field can be regarded as a Hilbert space($L^2(\Omega,H)$)-valued random variable.

Definition 2 ($L^p(\Omega,H)$ space)**.** *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space and H is a Hilbert space with norm $\|\cdot\|$. Then $\mathcal{L}^p(\Omega,H)$ with $1 \le p < \infty$ is the space of H-valued $\mathcal{F}$-measurable random vaiables $X : \Omega \to H$ with $\mathbf{E}[\|X\|^p] < \infty$ and a Banach space with norm:*

$$\|X\|_{\mathcal{L}^p(\Omega,H)} := \left(\int_\Omega \|X(\omega)\|^p dP(\omega)\right)^{\frac{1}{p}} = \mathbf{E}[\|X\|^p]^{\frac{1}{p}} \tag{2}$$

Then we can define the inner product:

$$\langle X,Y\rangle_{\mathcal{L}^2(\Omega,H)} := \int_\Omega \langle X(\omega),Y(\omega)\rangle_H dP(\omega) \tag{3}$$

Definition 3 (uncorrelated, covariance operator)**.** *Let H be a Hilbert space. A linear operator $\mathcal{C} : H \to H$ is the covariance of H-valued random variables X and Y if*

$$\langle \mathcal{C}_{XY}\phi,\psi\rangle_H = \mathrm{Cov}\left(\langle X,\phi\rangle_H, \langle Y,\psi\rangle_H\right), \forall \phi,\psi \in H \tag{4}$$

Example 1 ($H = \mathbb{R}^d$)**.** *In finite dimensional case $H = \mathbb{R}^d$, the covariance matrix conincides with the covariance operator.*

$$\begin{aligned} &\mathrm{Cov}\left(\langle X,\phi\rangle, \langle Y,\psi\rangle\right) = \mathrm{Cov}\left(\phi^T X, \psi^T Y\right) \\ =&\mathbf{E}\left[\phi^T(X-\mu_X)(Y-\mu_Y)^T\psi\right] = \phi^T\mathbf{E}\left[(X-\mu_X)(Y-\mu_Y)^T\right]\psi \\ =&\phi^T Cov(X,Y)\psi = \langle C_{XY}\phi,\psi\rangle \end{aligned} \tag{5}$$

Example 2 (X=Y)**.** *When $X = Y$ noted as $u(x, \omega) \in H$, the covariance operator: The covariance operator $\mathcal{C}$ can be defined as:*

$$\begin{aligned}
&Cov\left(\langle u, \phi\rangle_H, \langle u, \psi\rangle_H\right)\\
=&\mathbb{E}_m\left[\langle u, \phi\rangle_H \langle u, \psi\rangle_H\right]\\
=&\int_H \langle u, \phi\rangle_H \langle u, \psi\rangle_H m(du)\\
=&\int_D \left(\int_D Cov(u(x), u(y))\phi(x)dx\right)\psi(y)dy\\
=&\langle \mathcal{C}_u\phi, \psi\rangle_H
\end{aligned} \tag{6}$$

So that for $\forall x \in D$,

$$(\mathcal{C}_u\phi)(x) = \int_D Cov(u(x), u(y))\phi(y)dy \tag{7}$$

That is any $L^2(D)$-valued random variable $u(x)$ can defines a R.F. with $\mu(x)$ and $C(x, y)$ equal to the integral kernel of $\mathcal{C}$. The measure m is called probability measure.

Theorem 1. *For random field $u(x, \omega)$ with the covariance operator $\mathcal{C}$, suppose $\mathcal{C}$ is trace class with eigenpairs (λ_i, ϕ_i), then the second moment of $u(x, \omega)$ is given by:*

$$\mathbf{E}[\|u(x, \omega)\|_H^2] = \sum_{i=1}^{\infty} \lambda_i \leq \infty \tag{8}$$

Definition 4 (H-valued Gaussian random variable)**.** *Let H be a Hilbert space. An H-valued random variable $u(x, \omega)$ is Gaussian if $\langle u(x, \omega), \phi\rangle_H$ is a real-valued Gaussian random variable for all $\phi \in H$. Here the real-valued Gaussian Random Variable is defined as:*

$$\langle u, \phi\rangle_H \sim N(\langle \mu, \phi\rangle_H, \langle \mathcal{C}_u\phi, \phi\rangle_H) \tag{9}$$

This actually defines the Gaussian Measure m on H: $u \sim N(\mu, \mathcal{C}_u) := m$. The covariance operator of u is the symmetric, positive-definite operator $\mathcal{C}_u : H \to H$.

Proposition 1. *If $u(x, \omega)$ is a Gaussian random field, then $\mathcal{C}_u$ is trace class. Reversely, if $\mathcal{C}_u$ is a positive, symmetric, trace class operator in H, then there exists a Gasussian measure $m = N(0, \mathcal{C}_u)$ on H.*

Since we consider infinite dimensional case, unlike finite dimensions, not all translations preserve the measure. We need to cansider those directions in H along which translating a Gaussian measure does not change its essential nature (i.e., keeps it equivalent). First we have some definitions:

Definition 5. *Let m_1, m_2 be two measures on H.*

- *$m_1 \ll m_2$ means measure m_1 is absoutely continuous respect to m_2: if $m_1(A) = 0$ for all A s.t. $m_2(A) = 0$. This means that the support of m_1 is a subset of the support of m_2.*
- *If $m_1 \ll m_2, m_2 \ll m_1$, m_1, m_2 are said to be equivalent $m_1 \sim m_2$.*
- *If there exists a measurable set A s.t. $m_1(A) = 0$ and $m_2(A^c) = 1$, then m_1, m_2 are singular $m_1 \perp m_2$.*

Theorem 2 (Radon-Nikodym Theorem)**.** *Let $S = (H, \mathcal{B}(H))$ be a measurable space. m_1, m_2 are two σ-finite measures on S. if $m_1 \ll m_2$, then there exists a measurable function $f : H \to [0, \infty)$ s.t. $m_1(A) = \int_A f dm_2, \forall A \in \mathcal{B}(H)$. The function f is called the Radon-Nikodym derivative of m_1 with respect to m_2 and is denoted by $f = \frac{dm_1}{dm_2}$.*

Example 3. *In finite dimensional case, the Radon-Nikodym derivative is the probability density function. Like m_2 is the Lebesgue measure, $m_1(dx) = p(x)dx$, then the Radon-Nikodym derivative is the probability density function $p(x)$.*

Remark 1. *The KL-divergence is actually related to the Radon-Nikodym derivative:*

$$\mathrm{KL}(m_1||m_2) = \int_H \log\left(\frac{dm_1}{dm_2}(x)\right) dm_1(x) = \int_H \log\left(f(x)\right) f(x) dm_2(x) \tag{10}$$

So for $H = \mathbb{R}^d, dm_1 = p_1(x)dx, dm_2 = p_2(x)dx$, the KL-divergence can be written as:

$$\mathrm{KL}(m_1||m_2) = \int_{\mathbb{R}^d} \log\left(\frac{p_1(x)}{p_2(x)}\right) p_1(x)dx \tag{11}$$

The Radon-Nikodym derivative captures local density ratios, and the KL divergence is their global average.

Unlike the finite dimensional case, there is no natural Brownian Motion process in infinite dimensions. So for any Hilbert space H, we need to define the Brownian Motion process on H by using the Cameron-Martin space U. So if we want to define the Diffusion Process on infinite dimensional space, we need to consider the direction in H along which translating a measure preserves the absolutely continuity.

Definition 6 (Cameron-Martin Space)**.** *The Cameron-Martin space U is defined as:*

$$U := \{h \in H | m_h \ll m\}, m_h(A) = T_h^{\#} m(A) := m(A - h), \forall A \in \mathcal{B}(H) \tag{12}$$

where T_h is the translation operator: $T_h(u) = u + h$, $T_h^{\#} m$ is the push-forward measure of m by T_h,

Theorem 3 (Cameron-Martin Theorem)**.** *For $m = N(0, \mathcal{C})$, we have:*

$$m_h \sim m \textit{ if and only if } h \in \mathcal{C}^{\frac{1}{2}} H := U \tag{13}$$

Since $\mathcal{C}^{-1}$ is unbounded, U is a proper subspace of $U = \mathcal{C}^{1/2}H$. The inner product is defined as:

$$\langle \phi, \psi \rangle_U = \langle \mathcal{C}^{-1/2}\phi, \mathcal{C}^{-1/2}\psi \rangle_H \tag{14}$$

This can be generalized to general Gaussian measure cases:

Theorem 4 (Feldman-Hajek Theorem[2])**.** *The following statements hold:*
1. $m_1 = N(\mu_1, \mathcal{C}_1), m_2 = N(\mu_2, \mathcal{C}_2)$ are either singular or equivalent.
2. They are equivalent if and only if

- $\mathcal{C}_1^{1/2}(H) = \mathcal{C}_2^{1/2}(H) = H_0$, $\mu_1 - \mu_2 \in H_0$
- $(\mathcal{C}_1^{-1/2}\mathcal{C}_2^{1/2})(\mathcal{C}_1^{-1/2}\mathcal{C}_2^{1/2})^* - I$ *is a Hilbert-Schmidt operator.*

Specifically, if $\mathcal{C}_1 = \mathcal{C}_2 = \mathcal{C}$, we have the Radon Nikodym derivative:

$$\frac{dm_1}{dm_2}(u) = \exp\left(\langle \mu_1 - \mu_2, \mathcal{C}^{-1}(u - \mu_2)\rangle_H - \frac{1}{2}\|\mathcal{C}^{-1/2}(\mu_1 - \mu_2)\|_H^2\right) \tag{15}$$

Meanwhile, for Gaussian measure $m = N(\mu, \mathcal{C})$, we have the characteristic functional:

$$\hat{m}(\lambda) = \int_H e^{i\langle \lambda, u\rangle_H} m(du) = e^{i\langle \lambda, \mu\rangle_H - \frac{1}{2}\langle \lambda, \mathcal{C}\lambda\rangle_H}, \qquad \lambda \in H \tag{16}$$

First we assume define the Cylindrical Wiener Process $\hat{W}_t$ on H, but it does not take values in H cause I is not a trace class. Here we suppose $\hat{W}_t$ is H_1-valued:

$$\hat{W}_t = \sum_{i=1}^{\infty} \beta_i(t)\phi_i \tag{17}$$

where $\beta_i(t)$ are independent standart Brownian motions and ϕ_i are the eigenbasis of H. Normally, we can define the $\mathcal{C}$-Wiener Process on H as:

Definition 7 ($\mathcal{C}$-Wiener Process)**.** *A H-valued process W_t is called $\mathcal{C}$-Wiener process if: 1. $W_0 = 0$ 2. W_t has continuous trajectories: $\mathbb{R}^+ \to H$ 3. W_t has independent Gaussian increments: $m(W_t - W_s) = N(0, (t-s)\mathcal{C})$*

Therefore by applying the covariance operator $\mathcal{C}$ to cylindrical Wiener Process $\hat{W}_t$, we can define the H-valued $\mathcal{C}$-Wiener process W_t as:

$$W_t = \sum_{i=1}^{\infty} \sqrt{\lambda_i}\beta_i(t)\phi_i = \mathcal{C}^{1/2}\hat{W}_t \tag{18}$$

where λ_i are the eigenvalues of $\mathcal{C}$. So H is actually the Cameron-Matin space of H_1.

To sum up, in infinite dimensional diffusion model, suppose we have $u \sim m_0, \eta \sim m_1$, where m_0 is the initial measure and m_1 is the target measure, all defined on H. In the diffusion process, we add noise process η to u, getting $v = u + \eta$. Then $v \sim \nu = m_0 * m_1$. Here we need to do some assumptions on measure m_0 and m_1: m_0 needs to be supported on the Cameron-Martin space U of m_1, that is $m_0 \ll m_1$, m_0 should be more regular than m_1. Only in this way, we can define the Radon-Nikodym derivative of ν with respect to m_1, and the perturbed measure ν is equivalent to m_1.

2 Related Work

Some references: [3, 4, 5, 6, 7, 8]

[9] [10]

Kerrigan et al. [11] first studied the diffusion model in function space and KL approximation. In [12], Pidstrigach et all. formulate the diffusion model algorithm directly on infinitedimensional spaces, and proves that this formulation is well-posed and satisfies crucial theoretical guarantees. This work constructs the score function using conditional expectation and theoretically proves the existence of solutions to the proposed forward and reverse SDEs for arbitrary initial conditions. It also discusses the relationship between the choice of noise process W_t, loss norm$\|\cdot\|_K$, and data distribution μ_{data}, providing an explanation for why the diffusion model settings are appropriate for image data. In [13], Hagemann et al. deal with the infinite dimensional case by considering the limit of the finite-dimensional discrete stochastic differential equations and the reverse SDE theory. They define the isomorphism by KL expansion. In DDO[14], Lim et all. offered an alternative viewpoint with the score defined as a logarithmic derivative of a perturbed measure and sampling done by a Langevin process and its annealed version. They generate the notion of score and denoising score matching to infinite dimensions. Rissanen et al.[15] proposed an inverse heat dissipation model for inverse problems(IHDM). Images generated by the IHDM model are characterized by a coarse-to-fine structure, smooth interpolations, disentangled color and shape, and strong locality, all resulting from explicitly reversing the heat equation as a generative process. Lim, et al.[16] told the story from the perspective of stochastic evolution equation in Hilbert space, generalizing IHDM to infinite dimensions. In [17], Willcocks et al. defined a diffusion process in infinite dimensional image space based on neural oeprator. This work allows training and sampling at arbitrary resolution. In [18], Franzese consider the H-valued SDE, which is actually a SPDE. They discuss the limits of discretiztion and provide a theoretical framework for infinite dim diffusion between function space.

[19] [20] [21] [22]

3 For Stationary RF on $\mathbb{R}^d$

First we define stationary random field on $\mathbb{R}^d$ as:

Definition 8 (Stationary Random Field)**.** *A second-order random field $u(x) : x \in D$ is called stationary if the mean is constant and covariance function depends only on the difference $x - y$, i.e. $\mu(x) = \mu,\ C(x, y) = C(x - y)$.*

Then we can define the covariance operator $\mathcal{C}$ as:

$$\mathcal{C}\phi = \int_{\mathbb{R}^d} C(x - y)\phi(y)dy \tag{19}$$

We find that it is actually the convolution operator of $C(x)$ with $\phi(x)$.

Stationary random fields have some beautiful properties.

Theorem 5 (Wiener-Khinchin Theorem)**.** *There exists a stationary random field $u(x) : x \in D$ with mean μ and covariance function $c(x)$ that is mean square continuous if and only if the function $c(x) : \mathbb{R}^d \to \mathbb{R}$ is such that*

$$c(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{ik\cdot x} dF(k) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{ik\cdot x} S(k)dk = \left(\mathcal{F}^{-1}S\right)(x) \tag{20}$$

where $F(k)$ is some measure on $\mathbb{R}^d$ called spectral distribution and $\hat{S}(x)$ is the Fourier transform of $S(k)$, called spectral density. Reversely, $S(k) = (\mathcal{F}c)(k) = \hat{c}(k)$. If $S(k)$ is non-negative and integrable, then $c(x)$ is a valid covariance function.

Theorem 6 (Spectral Density of Random Field)**.** *Assume $u(x)$ has zero mean, then*

$$S_u(k) = \frac{1}{(2\pi)^d}\mathbb{E}[|\hat{u}(k)|^2] \tag{21}$$

By defining the pseudo-differential operators, the class of SPDEs is defined by:

$$\mathcal{L}_g u = W, \mathcal{L}_g = \mathcal{F}^{-1} g \mathcal{F} \tag{22}$$

where $g : \mathbb{R}^d \to \mathbb{C}$ must be a sufficiently regular and Hermitian-symmetric function, that is it must satisfy: $g(k) = \overline{g(-k)}$, $\bar{\cdot}$ denotes the complex conjugate. So if we have $\mathcal{L}_g u = W$, then:

$$u = \mathcal{L}_{\frac{1}{g}} W \tag{23}$$

Theorem 7. *The spectral density of $\mathcal{L}_g u$ and of u are related by:*

$$S_{\mathcal{L}_g u}(k) = |g(k)|^2 S_u(k) \tag{24}$$

Generally, if

$$\mathcal{L}_g u = w \tag{25}$$

where w is a GeRF source term, then $S_w(k) = |g(k)|^2 S_u(k)$. Therefore, when $w = W$, $S_u = \frac{1}{(2\pi)^d |g(k)|^2}\mathbb{E}[|\hat{W}(k)|^2] = \frac{1}{|g(k)|^2}$. Then,

$$u(x) = \mathcal{L}_{\frac{1}{g}} w(x) = \mathcal{L}_{\sqrt{\frac{S_u}{S_w}}} w(x) \tag{26}$$

Then consider the exitence of the function.

Theorem 8. *Let $w(x)$ be a real stationary GeRF over $\mathbb{R}^d$, and let $g : \mathbb{R}^d \to \mathbb{C}$ be a symbol function. Then for (25), there exists a unique stationary solution $u(x)$ if and only if: there exists $N \in \mathbb{N}$ s.t.*

$$\int_{\mathbb{R}^d} \frac{dS_w(k)}{|g(k)|^2 (1 + \|k\|^2)^N} < \infty \tag{27}$$

and

$$S_u(k) = |g(k)|^{-2} S_w(k) \tag{28}$$

Moreover, $S_u(k)$ is unique if and only if $|g| > 0$.

Hence the key is the symbol function $g(k)$. The following theorem shows that solutions of SPDEs with White Noise source term is the starting point of more general solutions, when the source term can be any stationary GeRF.

Theorem 9. *Let $w(x)$ be a real stationary GeRF over $\mathbb{R}^d$ with covariance distribution $C_w(x)$. Let g be a symbol function over $\mathbb{R}^d$ such that $\frac{1}{g}$ is smooth with polynomially bounded derivatives of all orders. Then, there exists a unique stationary solution to (25) and its covariance distribution is given by*

$$C_u(x) = C_u^W * C_w(x) \tag{29}$$

where C_u^W is the covariance function of the solution to the SPDE with White Noise source term.

Proof. The proof is straightforward by using Wiener-Khinchin theorem. □

For any precision operator which is a polynomial in the Laplacian, $Q = p(-\Delta)$, such as the Matern operator with $\nu \in \mathbb{N}$, this results in a polynomial $F(Q) = p(\|k\|^2)$.

3.1 Matern Field

The important relationship that we will make use of is that a Gaussian field $u(x)$ with the Matern covariance is a solution to the linear fractional stochastic partial differential equation (SPDE):

$$\mathcal{L}^{\alpha/2}u(x) = (\kappa^2 - \Delta)^{\alpha/2}u(x) = W(x), \qquad x \in D \in \mathbb{R}^d, \alpha = \nu + d/2, \kappa > 0, \nu > 0, \quad (30)$$

where $\nu = \alpha - d/2, \rho = \frac{\sqrt{2\nu}}{\kappa}$ is the range parameter, Δ is the Laplacian operator, $W(x)$ is a spatial Gaussian white noise with unit variance. We will name any solution to Equ (30) a Matern field in the following.

Theorem 10 (Spectral Solution of Matern Field)**.** *The solution of u solved by Equ (30) is given by:*

$$u(x) = \mathcal{F}^{-1}\left[\frac{\hat{W}(k)}{(\kappa^2 + \|k\|^2)^{\alpha/2}}\right](x) \tag{31}$$

where $\mathcal{F}$ is defined in (A). And the covariance function of u is given by:

$$c(x) = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)}(\kappa\|x\|)^\nu K_\nu(\kappa\|x\|) \tag{32}$$

where $\nu = \alpha - d/2, \rho = \frac{\sqrt{2\nu}}{\kappa}, \sigma^2 = \frac{\Gamma(\nu)}{(4\pi)^{d/2}\kappa^{2\nu}\Gamma(\alpha)}$

Wiener-Khinchin theorem + Spectral Theorem.

3.2 Generalized Matern Field

Consider the following SPDE:

$$(\kappa^2 + (-\Delta)^\gamma)^{\alpha/2}u(x) = \mathcal{F}^{-1}\left((\kappa^2 + \|k\|^{2\gamma})^{\alpha/2}\mathcal{F}u\right)(x) = W(x), \quad x \in D \in \mathbb{R}^d \tag{33}$$

Hence the solution is:

$$u(x) = \mathcal{F}^{-1}\left[\frac{\hat{W}(k)}{(\kappa^2 + \|k\|^{2\gamma})^{\alpha/2}}\right](x) \tag{34}$$

Therefore the spectral density is:

$$S_u(k) = \frac{1}{(\kappa^2 + \|k\|^{2\gamma})^\alpha} \tag{35}$$

So when $\gamma = 1$, it becomes the Matern Field. Since the spectral density $S_u(k) \in L^2(\mathbb{R}^d)$ if and only if $\alpha\gamma > \frac{d}{2}$.

Generally, we can define the pseudo-differential operator through symbol function $g(k)$.

4 Spatial-Temporal General Random Field on $\mathbb{R}^d \times (0, T)$

4.1 Stein Model

Proposed in [23], we define the spatial-temporal white noise $\mathcal{W}(x, t)$ as Gaussian noise that is white in time but correlated in space.

$$\left(b(s^2 - \frac{d}{dt^2})^\beta + a(\kappa^2 - \Delta)^\alpha\right)^{\nu/2} u(x,t) = W, \qquad (x,t) \in D \times (0,T) \tag{36}$$

Consider when $D = \mathbb{R}^d$, where the space-time spectral density of the stationary solution is given by:

$$S_u(k_s, k_t) = \frac{1}{(b(s^2 + k_t^2)^\beta + a(\kappa^2 + k_s^2)^\alpha)^\nu} \tag{37}$$

We note the spatio-temporal symbol function as:

$$g(k_s, k_t) : (k_s, k_t) \to (b(s^2 + k_t^2)^\beta + a(\kappa^2 + k_s^2)^\alpha)^{\nu/2} \tag{38}$$

When α, β, ν are positive and $\frac{d}{\alpha\nu} + \frac{1}{\beta\nu} = 2$, [23] shows that the spectral density is finite and the corresponding random field is mean square continuous.

Theorem 11. *When $\kappa, s, a, b > 0$ and α, β, ν are not null, $g(k_s, k_t)$ satisfies Thm 9, then for any stationary GeRF X, the SPDE:*

$$\left(b(s^2 - \frac{d}{dt^2})^\beta + a(\kappa^2 - \Delta)^\alpha\right)^{\nu/2} U(x,t) = X(x,t) \tag{39}$$

has a unique stationary solution $U(x,t)$ with covariance function:

$$C_U(x,y,t,s) = C_U^W * C^X(x-y, t-s) \tag{40}$$

4.2 Evolution Equations Model

Here we consider the following model:

$$\frac{\partial^\beta u}{\partial t^\beta} + \mathcal{L}_g u = w(x,t) \tag{41}$$

where $\mathcal{L}_g$ is a pseudo-differential operator with symbol $g(k)$ and $w(x,t)$ is a stationary spatio-temporal GeRF.

$$g(k_s, k_t) = (ik_t)^\beta + g(k_s) \tag{42}$$

4.3 Advection-Diffusion SPDE

This is poeposed in [6]. The equation is given by:

$$\left[\frac{\partial}{\partial t} - \nabla \cdot (\Sigma \nabla) + \mu \nabla + C\right] u(x,t) = w(x,t) \tag{43}$$

where Σ is the diffusion matrix, μ is the advection velocity, C is the drift coefficient. Here we set the diffusion matrix as:

$$\Sigma = \frac{1}{\rho^2} \begin{pmatrix} \cos\theta & \sin\theta \\ -\gamma \sin\theta & \gamma\cos\theta \end{pmatrix}^T \begin{pmatrix} \cos\theta & \sin\theta \\ -\gamma \sin\theta & \gamma\cos\theta \end{pmatrix} \tag{44}$$

where ρ is the correlation length and γ is the anisotropy ratio, $\theta \in [0, \pi/2)$. With $\gamma = 1$, it becomes isotropic. Similarly the spectral density is given by:

$$\begin{aligned} S_u(k_s, k_t) &= \frac{S_w(k_s, k_t)}{|i(k_t + \mu k_s) + (C + k_s^T \Sigma k_s)|^2} \\ &= \frac{S_w(k_s, k_t)}{(k_t + \mu k_s)^2 + (C + k_s^T \Sigma k_s)^2} \\ &= \frac{S_w(k_s, k_t)}{|g_u|^2} \end{aligned} \tag{45}$$

By Wiener-Khinchin theorem, the covariance function is given by:

$$C_u(x,t) = \frac{1}{(2\pi)^d} \int S_w \frac{e^{-i\mu k_s t - (k_s^T \Sigma k_s + C)|t|}}{2(k_s^T \Sigma k_s + C)} e^{ik_s x} dk_s \tag{46}$$

Specifically, when $\mu = 0, \Sigma = 0$, the covariance function is given by:

$$C_u(x,t) = \frac{e^{-C|t|}}{2C} C_w(x,t) \tag{47}$$

However $\mu(x)$ may not be constant.

4.4 Generic class of non-stationary models

Similar to ADSPDE, we consider:

$$\frac{\partial u}{\partial t} + \left[-\nabla \cdot (\Sigma(x,t)\nabla) + \mu(x,t) \cdot \nabla + \kappa^2(x,t)\right]^{\alpha/2} u(x,t) = w(x,t) \tag{48}$$

where μ, Σ, κ are functions of x, t, and $w(x,t)$ is a GeRF driven by Equ (30).

5 Experiments

We must pay particular attention to the change in normalization factors when moving from the continuous to the discrete setting. Specifically:

- Regarding the simulation of white noise: We use $N(0, 1/h)$ to approximate the characteristic variance $\delta(x - y)$ of continuous white noise. However, when calculating the spectral density (67), we need to carefully incorporate appropriate normalization factors:

$$S_u(k) = \frac{1}{N\frac{1}{h}} \frac{\mathbb{E}\left[|DFT(W)(k)|^2\right]}{(\kappa^2 + \|k\|^2)^\alpha} \tag{49}$$

- Fourier transform definitions: Special attention must be given to the definition of the Fourier transform. Different normalization conventions will result in distinct normalization factors and, consequently, different results.

First we test the Matern Covariance function:

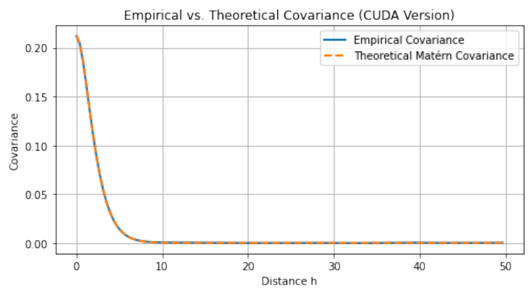


Figure 1: 1D stationary covariance function

5.1 1D random fields

Here are some examples of 1D random fields computed by classical methods in Fig 2.

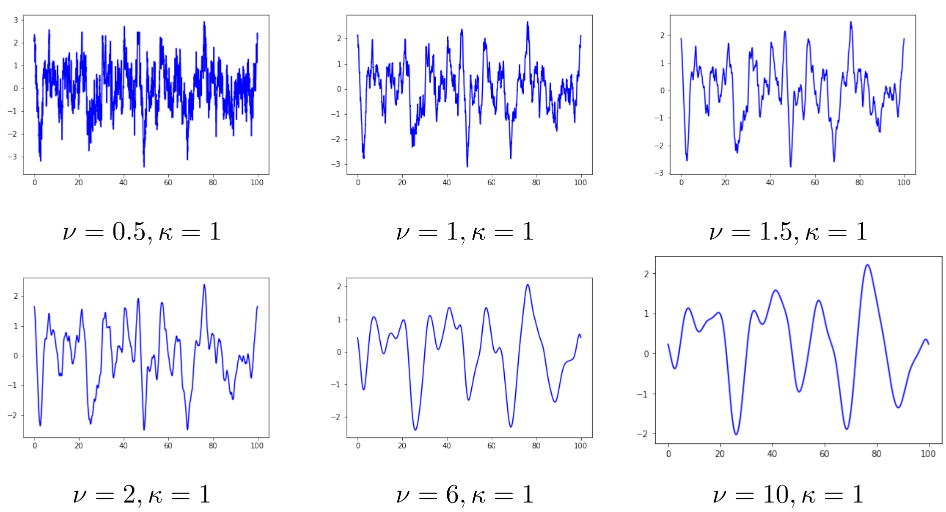
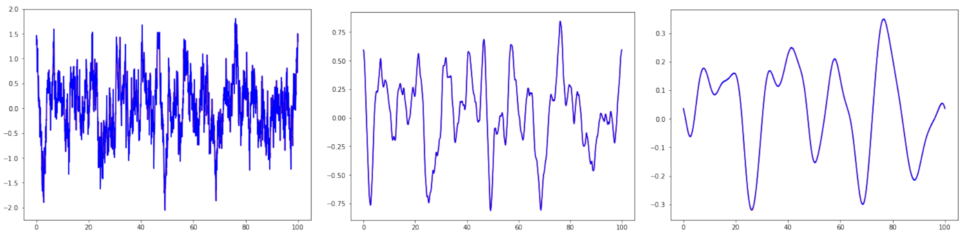


Figure 2: 1D Random Field (Normalized)

Use MLP to learn the spectral density of the random field, and then generate the random field by inverse Fourier transform.



$\nu = 0.5, \kappa = 1$ $\nu = 2, \kappa = 1$ $\nu = 10, \kappa = 1$

Figure 3: 1D Neural Random Field

Assume discrete samples of the stochastic simulator $\mathcal{M}(x, \omega)$ are given as:

$$\mathcal{T}_r := \{(x_i^r, \mathcal{M}(x_i^r, \omega^r)) : i = 1, 2, \cdots, N^r\}, \quad r = 1, 2, \cdots, R. \tag{50}$$

in the form of discrete evaluations of the stochastic simulator on R trajectories, where for every r, $\{x_i^r\}_{i=1}^{N^r}$ is an i.i.d. sample from the input distribution μ, and $\mathcal{M}(x_i^r, \omega^r)$ is the output of the stochastic simulator. Here we assume for notational simplicity that $N^r = N$ for all r.

That is we have R samples $u(x)$. By the definition of spectral density, we can have the estimation of $S_u(k)$:

$$S_u(k) = \frac{1}{(2\pi)^d} \frac{1}{R} \sum_{r=1}^{R} \left(\|(\mathcal{F}u_r)(k)\|^2 \right) \tag{51}$$

For the above example, the solution can be rewritten as:

$$u(x) = \mathcal{F}^{-1} \left[\hat{W}(k) (S_u(k))^{1/2} \right] (x) \tag{52}$$

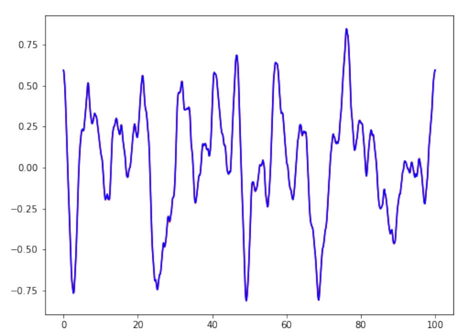
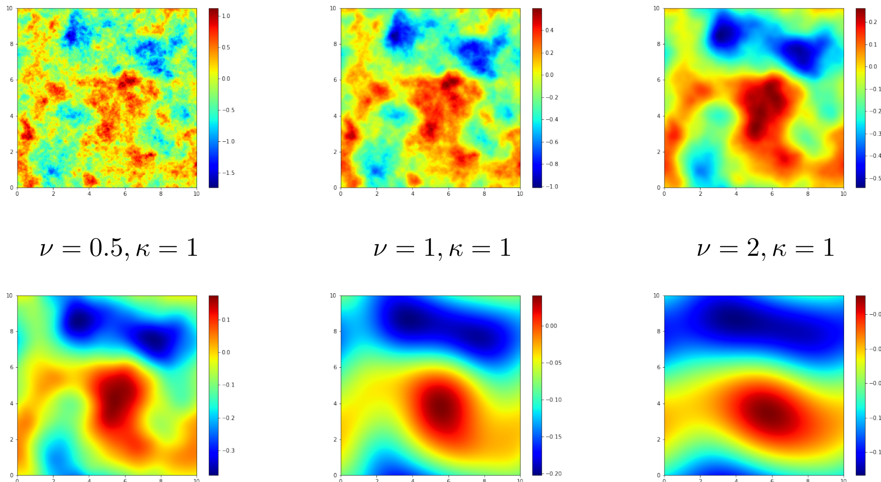


Figure 4: 1D Data-Driven Random Field

5.2 2D random fields

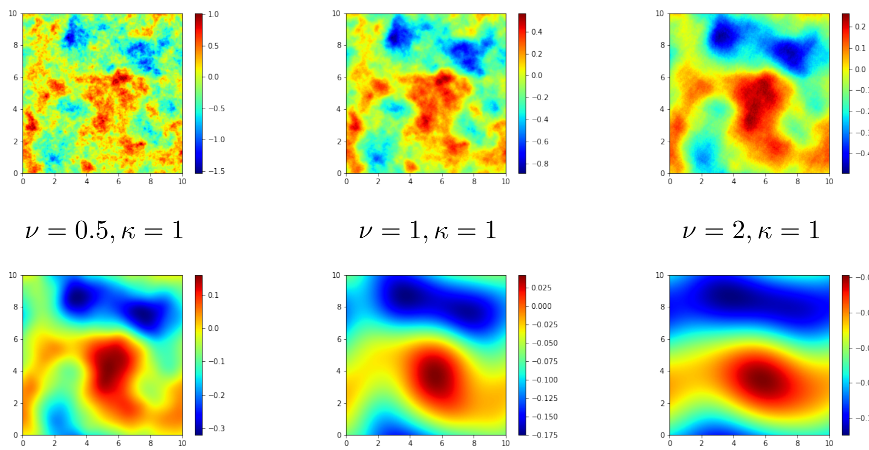
2D is similar. Neural 2D random field is vertified.



$\nu = 0.5, \kappa = 1$ $\quad$ $\nu = 1, \kappa = 1$ $\quad$ $\nu = 2, \kappa = 1$

$\nu = 3, \kappa = 1$ $\quad$ $\nu = 6, \kappa = 1$ $\quad$ $\nu = 10, \kappa = 1$

Figure 5: 2D Random Field



$\nu = 0.5, \kappa = 1$ $\quad$ $\nu = 1, \kappa = 1$ $\quad$ $\nu = 2, \kappa = 1$

$\nu = 3, \kappa = 1$ $\quad$ $\nu = 6, \kappa = 1$ $\quad$ $\nu = 10, \kappa = 1$

Figure 6: 2D Neural Random Field

6 Conclusion

Q1: For general SPDE, given the equation, how to simulate the random field? Or inversely, given the random field data, how to infer the equation?

Q2: What is the relationship between the equation and the covariance function? For $\forall \mathcal{M}(x, \omega)$, can we find a differential operator $\mathcal{L}$ and a noise $\mathcal{W}$ s.t. $\mathcal{L}\mathcal{M} = \mathcal{W}$?

Q3: Generally speaking, given the covariance function of $\mathcal{M}(x, \omega)$, how can we simulate the random field with DNN?

Consider:

- Generative Model: GAN, VAE, Flow-based model
- Neural Operator: DeepONet, FNO, etc.

References

[1] Gabriel J. Lord, Catherine E. Powell, and Tony Shardlow. *An Introduction to Computational Stochastic PDEs*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2014.

[2] Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic Equations in Infinite Dimensions*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2 edition, 2014.

[3] Peter Whittle. On stationary processes in the plane. *Biometrika*, pages 434–449, 1954.

[4] Ricardo Carrizo Vergara, Denis Allard, and Nicolas Desassis. A general framework for spde-based stationary random fields. *Bernoulli*, 28(1):1–32, 2022.

[5] Finn Lindgren, Haakon Bakka, David Bolin, Elias Krainski, and Håvard Rue. A diffusion-based spatio-temporal extension of gaussian mat\'ern fields. *arXiv preprint arXiv:2006.04917*, 2020.

[6] Fabio Sigrist, Hans R Künsch, and Werner A Stahel. Stochastic partial differential equation based modelling of large space–time data sets. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 77(1):3–33, 2015.

[7] David Bolin and Kristin Kirchner. The rational spde approach for gaussian random fields with general smoothness. *Journal of Computational and Graphical Statistics*, 29(2):274–285, 2020.

[8] Emilio Porcu, Moreno Bevilacqua, Robert Schaback, and Chris Oates. The matérn model: A journey through statistics, numerical analysis and machine learning. 03 2023.

[9] Emilien Dupont, Yee Whye Teh, and Arnaud Doucet. Generative models as distributions of functions, 2022.

[10] Sarthak Mittal, Guillaume Lajoie, Stefan Bauer, and Arash Mehrjou. From points to functions: Infinite-dimensional representations in diffusion models, 2022.

[11] Gavin Kerrigan, Justin Ley, and Padhraic Smyth. Diffusion generative models in infinite dimensions, 2023.

[12] Jakiw Pidstrigach, Youssef Marzouk, Sebastian Reich, and Sven Wang. Infinite-dimensional diffusion models. *Journal of Machine Learning Research*, 25(414):1–52, 2024.

[13] Paul Hagemann, Sophie Mildenberger, Lars Ruthotto, Gabriele Steidl, and Nicole Tianjiao Yang. Multilevel diffusion: Infinite dimensional score-based diffusion models for image generation, 2024.

[14] Jae Hyun Lim, Nikola B. Kovachki, Ricardo Baptista, Christopher Beckham, Kamyar Azizzadenesheli, Jean Kossaifi, Vikram Voleti, Jiaming Song, Karsten Kreis, Jan Kautz, Christopher Pal, Arash Vahdat, and Anima Anandkumar. Score-based diffusion models in function space, 2025.

[15] Severi Rissanen, Markus Heinonen, and Arno Solin. Generative modelling with inverse heat dissipation, 2023.

[16] Sungbin Lim, EUN BI YOON, Taehyun Byun, Taewon Kang, Seungwoo Kim, Kyungjae Lee, and Sungjoon Choi. Score-based generative modeling through stochastic evolution equations in hilbert spaces. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, editors, *Advances in Neural Information Processing Systems*, volume 36, pages 37799–37812. Curran Associates, Inc., 2023.

[17] Sam Bond-Taylor and Chris G. Willcocks. ∞-diff: Infinite resolution diffusion with subsampled mollified states. In *International Conference on Learning Representations*, 2024.

[18] Giulio Franzese, Giulio Corallo, Simone Rossi, Markus Heinonen, Maurizio Filippone, and Pietro Michiardi. Continuous-time functional diffusion processes, 2023.

[19] Byoungwoo Park, Jungwon Choi, Sungbin Lim, and Juho Lee. Stochastic optimal control for diffusion bridges in function spaces, 2025.

[20] Elizabeth Louise Baker, Gefan Yang, Michael L. Severinsen, Christy Anna Hipsley, and Stefan Sommer. Conditioning non-linear and infinite-dimensional diffusion processes, 2024.

[21] Gefan Yang, Elizabeth Louise Baker, Michael L. Severinsen, Christy Anna Hipsley, and Stefan Sommer. Simulating infinite-dimensional nonlinear diffusion bridges, 2024.

[22] Kunwoo Na, Junghyun Lee, Se-Young Yun, and Sungbin Lim. Probability-flow ode in infinite-dimensional function spaces, 2025.

[23] Michael L Stein. Space–time covariance functions. *Journal of the American Statistical Association*, 100(469):310–321, 2005.

A Fourier Transform

Definition 9. *Define the Fourier transform of $u(x)$ as:*

$$\begin{cases} \mathcal{F}u(k) = \int_{\mathbb{R}^d} u(x)e^{-ikx}dx, \\ \mathcal{F}^{-1}u(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} u(k)e^{ikx}dk. \end{cases} \tag{53}$$

B Polynomial Chaos Expansion

A spectral expansion in $L_\mu(D)$ is called chasos expansion. By defining the inner product in $L_\mu(D)$ as:

$$\langle f, g\rangle_\mu = \int_D f(x)g(x)d\mu(x), \tag{54}$$

the space of $L_\mu(D)$ is a Hilbert space.

Theorem 12. *Let $\{\phi_i(x)\}_{i=1}^\infty$ be an orthonormal basis of $L_\mu(D)$, i.e.*

$$\begin{aligned} &1. \int_D \phi_i(x)\phi_j(x)d\mu(x) = \delta_{ij}, \\ &2. \phi_i(x) \textit{ is dense in } L_\mu(D). \end{aligned} \tag{55}$$

Then the chasos expansion of $\mathcal{M}(x,\omega) \in L_\mu(D)$ is given by:

$$\mathcal{M}(x,\omega) = \sum_{i=1}^\infty c_i\phi_i(x), \tag{56}$$

where $c_i(x) = \langle \mathcal{M}, \phi_i\rangle_\mu$ are the coefficients of the expansion.

C Karhunen-Loeve Expansion

Let $\mu(x)$ be the mean function and $C(x,y)$ be the covariance function of $\mathcal{M}(x,\omega)$. Assume that D is bounded, $C(x,y)$ is continuous and $\mathcal{M}(x,\cdot)$ has finite variables for all $x \in D$.

Theorem 13. *The Karhunen-Loeve expansion of $\mathcal{M}(x,\omega)$ is given by:*

$$\mathcal{M}(x,\omega) = \mu(x) + \sum_{i=1}^\infty \lambda_i\phi_i(x)\xi_i(\omega), \tag{57}$$

where λ_i and $\phi_i(x)$ are the eigenvalues and eigenfunctions of the covariance operator $\mathcal{C}$:

$$(\mathcal{C}\phi_i)(x) = \int_D Cov(\mathcal{M}(x,\omega), \mathcal{M}(y,\omega))\phi_i(y)dy = \int_D C(x,y)\phi_i(y)dy = \lambda_i\phi_i(x). \tag{58}$$

where $C(x,y), x, y \in D$ is the covariance function of $\mathcal{M}(x,\omega)$. The KL-random variables $\xi_i(\omega)$ are the result of the projection of $\mathcal{M}(x,\omega)$ onto the eigenfunctions $\phi_i(x)$:

$$\xi_i(\omega) = \frac{1}{\sqrt{\lambda_i}}\int_D (\mathcal{M}(x,\omega) - \mu(x))\phi_i(x)dx. \tag{59}$$

Note that both $\{\phi_i(x)\}_{i=1}^\infty$ and $\{\xi_i(\omega)\}_{i=1}^\infty$ are orthonormal bases, one capturing the "spatial" variation of $\mathcal{M}(x,\omega)$ over D (in terms of x), the other capturing the stochastic variation of $\mathcal{M}(x,\omega)$ (in terms of ω).

Theorem 14 (Mercer's theorem)**.** *The covariance function $C(x,y)$ of $\mathcal{M}(x,\omega)$ can be expressed as:*

$$C(x,y) = \sum_{i=1}^\infty \lambda_i\phi_i(x)\phi_i(y). \tag{60}$$

It follows that the average variance of the random field over the domain D is equal to $\sum_{i=1}^\infty \lambda_i$.

In most cases, the integral eigenvalue problem in Equ (13) is difficult: analytically solved and high-dimensional.

D Some proofs

Proof of Thm 1.

$$
\begin{aligned}
\mathbb{E}\left[\|u(x,\omega)\|_H^2\right] &= \mathbb{E}\left[\sum_{i=1}^{\infty}\langle u,\phi_i\rangle_H^2\right] \\
&= \sum_{i=1}^{\infty}\mathbb{E}\left[\langle u,\phi_i\rangle_H^2\right] \\
&= \sum_{i=1}^{\infty}\langle \mathcal{C}_u\phi_i,\phi_i\rangle_H = \sum_{i=1}^{\infty}\lambda_i
\end{aligned}
\tag{61}
$$

□

Proof of Thm 6.

$$
\begin{aligned}
S_u(k) = (\mathcal{F}c)(k) &= \int_{\mathbb{R}^d} e^{-ikh}c(h)dh \\
&= \int_{\mathbb{R}^d} e^{-ik(x+h-x)}\mathbb{E}\left[u(x+h)u(x)\right]dh \\
&= \mathbb{E}\left[\int_{\mathbb{R}^d} e^{-ik(x+h)}u(x+h)e^{ikx}u(x)dh\right] \\
&= \frac{1}{(2\pi)^d}\mathbb{E}\left[\left|(\mathcal{F}u)(k)\right|^2\right]
\end{aligned}
\tag{62}
$$

□

Proof of Thm 10. Do Fourier transform on Equ (30):

$$
\left\{\mathcal{F}(\kappa^2-\Delta)^{\alpha/2}u\right\}(k) = (\kappa^2+\|k\|^2)^{\alpha/2}(\mathcal{F}u)(k) \tag{63}
$$

then we have

$$
(\mathcal{F}u)(k) = \hat{u}(k) = \frac{\hat{W}(k)}{(\kappa^2+\|k\|^2)^{\alpha/2}} \tag{64}
$$

Therefore, u can be written as:

$$
u(x) = \mathcal{F}^{-1}\left[\frac{\hat{W}(k)}{(\kappa^2+\|k\|^2)^{\alpha/2}}\right] \tag{65}
$$

Then the stationary covariance function of u is given by:

$$
c(x) = Cov(u(x),u(0)) \tag{66}
$$

By the definition of spectral density Equ (21) and Equ (64) we have:

$$
S_u(k) = \frac{1}{(2\pi)^d}\frac{\mathbb{E}\left[\left|\hat{W}(k)\right|^2\right]}{(\kappa^2+\|k\|^2)^{\alpha}} = \frac{1}{(\kappa^2+\|k\|^2)^{\alpha}} \tag{67}
$$

Then we have the variance of u:

$$
c(0) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} S_u(k)dk = \frac{\Gamma(\nu)}{(4\pi)^{d/2}\kappa^{2\nu}\Gamma(\alpha)} := \sigma^2 \tag{68}
$$

By Wiener-Khinchin theorem, we have:

$$
\begin{aligned}
c(x) = (\mathcal{F}^{-1}S_u)(x) &= \mathcal{F}^{-1}\left[\frac{1}{(\kappa^2+\|k\|^2)^{\alpha}}\right] \\
&= \frac{\|x\|^{\nu}K_\nu(\kappa\|x\|)}{(4\pi)^{d/2}2^{\nu-1}\kappa^{\nu}\Gamma(\alpha)} \\
&= \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)}(\kappa\|x\|)^{\nu}K_\nu(\kappa\|x\|)
\end{aligned}
\tag{69}
$$

□

Remark 2. *To make $c(0) = 1$, we can multiple a constant factor σ_1 to $S_u(k)$:*

$$\sigma_1^2 = \frac{\Gamma(\alpha)\kappa^{2\nu}(4\pi)^{d/2}}{\Gamma(\nu)} \tag{70}$$

Then the corresponding function of u is:

$$u(x) = \mathcal{F}^{-1}\left[\frac{\sigma_1 \hat{W}(k)}{(\kappa^2 + \|k\|^2)^{\alpha/2}}\right] \tag{71}$$