Physics-Informed Generative Neural Networks for Stochastic Differential Equations

Lucia CLAROTTO, Hugo GANGLOFF, Nicolas JOUVIN, Antoine REGARDIN

MIA-PS, INRAE

August 21, 2025

1 Methods and Mathematical definitions

Notations

- We will denote probability spaces by the usual triplet $(\Omega, \mathcal{A}, P)$, with Ω a set of outcomes, $\mathcal{A}$ a σ-algebra of measurable events and P a probability measure.
- $L^2(\Omega) = \{X : \Omega \longrightarrow \mathbb{R} \,|\, \text{X is measurable and } \mathbb{E}[X^2] < \infty\}$ is the Hilbert space of square-integrable (i.e. finit variance) random variables valued on Ω. We assume every random variable studied in this work is in this space.
- ξ will denote a normalized centered Gaussian variable, and $\xi(\omega)$ one realization ($\omega \in \Omega$). The bold notation $\boldsymbol{\xi}$ will correspond to a vector of such random variables.
- By the expression "Matérn (or Exponential) random field", we mean "Gaussian random field with Matérn (or Exponential) covariance function".
- the "x" in bold notation $\mathbf{x}$ will denote collocation points uniformally sampled over a specified spatial domain.
- The notation $\|.\|^2$ will be used to write the squared norm of the respective Hilbert space of the chosen elements, but when applied to finite vectors obtained from the discretizations of functions it will correspond to the standard 2-norm in $\mathbb{R}^d$:
  for $u = (u_1, \ldots, u_d) \in \mathbb{R}^d$, $\|u\|^2 := \|u\|_2^2 = \frac{1}{d}\Sigma_{i=1}^d |u_i|$ i.e. a squared mean of absolute values.
- "$\overset{d}{=}$" will denote the equality in distribution: for two real random variables X and Y, "$X \overset{d}{=} Y$" is equivalent to say that their probability measures coincide or that their cumulative distribution functions are equal in every point.

1.1 Random fields

Random function, random field Let D be a domain of $\mathbb{R}^d$, a random function Z is a function of two variables $x \in D, \omega \in \Omega$ valued in a set E. Thus, for a fixed ω_0, $Z(x; \omega_0)$ is a function defined in $D \longrightarrow E$ and for a fixed x_0, $Z(x_0; \omega)$ is a random variable. [2]
We will often mute the random event ω in the notation $Z(x)$.
Usually, the term *random field* is used for a random function defined on a multidimensional space ($x \in \mathbb{R}^d, d > 1$), but as it is the conventional term for geostatistics and SPDEs, we will use it on the unidimensional random functions we will study here.
A so-called **second-order** random function or random field is such that $\text{Var}[Z(x)] < \infty$.

Gaussian Random field [1] A Gaussian random field (GRF) is a collection of random variables $\{Z(x) : x \in D\}$, indexed by a domain $D \subseteq \mathbb{R}^d$, such that for any finite set of points $x_1, \ldots, x_n \in D$, the random vector $(X(x_1), \ldots, X(x_n))^\top$ follows a multivariate Gaussian distribution:

$$(X(x_1), \ldots, X(x_n))^\top \sim \mathcal{N}_n(\mu, \Sigma),$$

If $\mu(x) = 0$ for all $x \in D$, the field is called a *mean-zero Gaussian random field*.

Covariance The covariance is a fundamental property of a defined random function Z. It is the function $C(x, y) = \text{Cov}(Z(x) + Z(y))$. The covariance function C is positive-definite.
On a finite set of points $x_1, \ldots, x_n$, the covariance function corresponds to a matrix $\Sigma = [C(x_i, x_j)]_{i,j=1}^n$, with entries $C(x_i, x_j) = \text{Cov}(Z(x_i), Z(x_j))$.

On fixed spaces and domains, a random field can be defined by its covariance function, which will in this work be taken from well-known covariance function families like the Matérn kernel.

Stationarity Following is the definition of stationarity we will use in this work:
A second-order random function Z is said to be **second-order stationary** if $\forall x, y, \tau \in \mathbb{R}$, $\mathbb{E}[Z(x)] = \mathbb{E}[Z(x + \tau)]$ and $\text{Cov}(Z(x), Z(y)) = C(x, y) = \text{Cov}(Z(x + \tau), Z(y + \tau))$. This brings:

$$\begin{aligned} \mathbb{E}[Z(x)] &= \mu \\ C(x, y) &= C(h) \end{aligned}$$

with $\mu \in \mathbb{R}^d$ a constant, and $h \in \mathbb{R}^d$ the variable defined as the vector $h = x - y$.[2]

Variogram, Correlogram The **theoretical variogram** of a random field Z is defined as

$$\gamma(h) \;=\; \tfrac{1}{2}\,\mathbb{E}\Big[\big(Z(x+h) - Z(x)\big)^2\Big], \quad h \in \mathbb{R}^d.$$

If $Z(x)$ is second-order stationary with covariance function, $C(h) = \text{Cov}(Z(x), Z(x+h))$, then $\gamma(h) = C(0) - C(h)$.

Suppose we observe $Z(x)$ at locations $x_1, x_2, \ldots, x_n \subset \mathbb{R}^d$. For a given lag h, the **empirical variogram estimator** is

$$\hat{\gamma}(h) \;=\; \frac{1}{2\,|N(h)|} \sum_{(i,j)\in N(h)} \big(Z(x_i) - Z(x_j)\big)^2,$$

where

- $N(h)$ is the set of pairs (i, j) such that $x_j - x_i \approx h$,
- $|N(h)|$ is the number of such pairs.

We will use this empirical estimator as a tool to study and compare the random fields generated by our models, as they may not be purely second-order stationary.
In practice, we will prefer to use the **correlogram** of the random fields. It is the plot of the autocorrelation $\rho(h)$ in function of a lag h, that is related to the variogram by the relation:

$$\rho(h) = 1 - \frac{\gamma(h)}{\sigma^2}$$

where $sigma^2$ is the total variance.

1.2 Random field expansions

Hermite polynomials The probabilistic Hermite polynomials $\{H_n(x)\}_{n=0}^{\infty}$ form a family of orthogonal polynomials with respect to the standard Gaussian weight. They are defined by the Rodrigues formula:

$$H_n(x) = (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n}\left(e^{-\frac{x^2}{2}}\right), \qquad n = 0, 1, 2, \dots$$

The first few Hermite polynomials are:

$$H_0(x) = 1, \quad H_1(x) = x, \quad H_2(x) = x^2 - 1, \quad H_3(x) = x^3 - 3x, \quad H_4(x) = x^4 - 6x^2 + 3.$$

We remark that the degree of a polynomial is equivalent to its index. They satisfy the recurrence relation

$$H_{n+1}(x) = xH_n(x) - nH_{n-1}(x), \qquad n \geq 1,$$

with initial conditions $H_0(x) = 1$, $H_1(x) = x$.

The Hermite polynomials are orthogonal with respect to the Gaussian measure $d\mu(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx$:

$$\int_{-\infty}^{\infty} H_n(x) H_m(x) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx = n!\, \delta_{nm}.$$

The infinite series of Hermite polynomials form a basis of the space of square-integrable functions with respect to the Gaussian measure. Thus, it is involved in the decomposition of random functions, as we will see.

Polynomial Chaos expansion [4] Polynomial Chaos Expansion (PCE) is a spectral method for representing a stochastic process or random variable in terms of orthogonal polynomials of standard random variables. Let $\boldsymbol{\xi} = (\xi_1, \dots, \xi_N)$ denote a vector of centered, normalized and mutually orthogonal (independant) Gaussian variables. Other types of random variables can be used

along with other types of basis functions, but we will only look at this case. Then any square-integrable random variable $X \in L^2(\Omega)$ can be expanded as:

$$X(\omega) = \sum_{\alpha \in \mathbb{N}_0^N} c_\alpha \, \Psi_\alpha(\boldsymbol{\xi}(\omega)), \quad \omega \in \Omega$$

where:

- $\alpha = (\alpha_1, \ldots, \alpha_N)$ is a multi-index,
- Ψ_α are multivariate orthogonal polynomials that form a functional basis. Here as we work with Gaussian variables, they are Hermite polynomials.
- $c_\alpha \in \mathbb{R}$ are deterministic, real-valued coefficients.

The orthogonality property is

$$\mathbb{E}[\Psi_\alpha(\boldsymbol{\xi})\Psi_\beta(\boldsymbol{\xi})] = \langle \Psi_\alpha, \Psi_\beta \rangle = \delta_{\alpha\beta} \, \|\Psi_\alpha\|^2,$$

which enables efficient computation of the coefficients c_α. In practice, the infinite series is truncated to a finite order p:

$$X(\omega) \approx \widetilde{X}(\omega) = \sum_{|\alpha| \leq p} c_\alpha \, \Psi_\alpha(\boldsymbol{\xi}(\omega)).$$

This representation gives the source of randomness as polynomials valued at Gaussian variables, allowing us for example to simulate a random field we don't know an explicit expression for.
In the case of the expansion of a **univariate** distribution, the PC basis is unidimensional (in terms of random spaces), depending only of a single scalar Gaussian variable. Moreover, multi-index α is just one integer, and the sum is made on $(0, \ldots, p) = (0, \ldots, N)$.
We will see later that, in order to expand a random function depending on a space variable, we will implement a polynomial chaos expansion with functions $c_\alpha(x)$ as coefficients. This way, these coefficients carry an information about the deterministic part of the random function, separated from the information about the stochastic behaviour.

Multivariate Polynomial Chaos As written before, in the case of a univariate random variable (that can be a random function depending on a space variable), the PC expansion is a series indexed by positive integers. Each element of the series is a coefficient multiplied by one Hermite polynomial evaluated at a scalar Gaussian variable $c_n \Psi_n(\xi)$ ($\xi \sim \mathcal{N}(0,1)$), making the series an unidimensional representation.
This way, the polynomial of the n^{th} element of the sum is of degree n. For the case where we want to expand a multivariate random variable $U(\omega)$, the sum is made on a family of multi-indexes that assemble in each term several Hermite

polynomials evaluated at different independant Gaussian variables, making the expansion a multidimensional representation.
We have our centered, normalized and mutually independent Gaussian variables $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_N)$ and our Hermite polynomials $H_0, H_1, ...$ We will next use the notation from the book *Spectral Methods for Uncertainty Quantification* [4] to write our expansion:
We have a multi-index $\alpha = (\alpha_1, \ldots, \alpha_N)$, let $\lambda(p)$ be the sub-set of multi-indices such that they correspond to a p^{th} - order set (the sum of their indices is equal to p):

$$\lambda(p) = \left\{ \gamma : \sum_{i=1}^{N} \gamma_i = p \right\}$$

Following these definitions, we can construct the set of p^{th} - order polynomials to represent our multidimensional random space, by assembling Hermite polynomials evaluated at different ξ_i as such:

$$\left\{ \bigcup_{\gamma \in \lambda(p)} \prod_{i=1}^{N} \psi_{\gamma_i}(\xi_i) \right\}.$$

then, we need to assemble these sets in a series indexed by the multi-index, each term being weighted by an associated coefficient u_α. This gives a kind of weighted tensor product between the vectors of Hermite polynomials evaluated at the different random variables $(H_0(\xi_1), H_1(\xi_1), ...)^T$, $(H_0(\xi_2), H_1(\xi_2), ...)^T$, $\ldots$, $(H_0(\xi_N), H_1(\xi_N), ...)^T$.

For example, in the 2D case ($\boldsymbol{\xi} = (\xi_1, \xi_2)$), the Hermite expansion can be expressed as:

$$\begin{aligned} U = {} & u_0\psi_0 + u_1\psi_1(\xi_1) + u_2\psi_1(\xi_2) \\ & + u_{11}\psi_2(\xi_1) + u_{21}\psi_1(\xi_2)\psi_1(\xi_1) + u_{22}\psi_2(\xi_2) \\ & + u_{111}\psi_3(\xi_1) + u_{211}\psi_1(\xi_2)\psi_2(\xi_1) + u_{221}\psi_2(\xi_2)\psi_1(\xi_1) \\ & + u_{222}\psi_3(\xi_2) + u_{1111}\psi_4(\xi_1) + \cdots \end{aligned}$$

To truncate the sum, instead of stopping the expansion at a set of multi-indices of order p we chose to set the size of our set of Hermite polynomials to a value M (we will later note N_{KL}), and then assemble the weighted tensor product with the fixed-size $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_N)$. We use this convention as it is more suited to the functions of the Python library *chaospy* we used.

Next, we will present another expansion, that is only defined for random fields (and not general random variables like the PC).

Karhunen-Loève expansion The Karhunen–Loève (KL) expansion is a spectral representation of a second-order random field in terms of deterministic

orthogonal functions and uncorrelated random variables. The idea is to diagonalize the covariance function operator of the random field in a functional basis in order to obtain an expansion. As it is computed directly from the structure of the random field, this expansion is more adapted to it, compared to the Polynomial Chaos which does not associate a unique functional basis to each random field, but uses a fixed orthogonal basis. Thus, generally the KL expansion requires more computation or more information to be computed but can give more information about the random field, as it is designed for this specific type of random variable. In fact, it provides a minimal representation of the random field, and few modes may be needed to capture most of its variance.

Let $X(s,\omega)$ be a centered stochastic process defined on a domain $D \subset \mathbb{R}^d$, with covariance function

$$C(x,y) = \mathbb{E}[X(x)X(y)], \qquad x,y \in D.$$

The KL expansion of X is given by

$$X(s,\omega) = \sum_{n=1}^{\infty} \sqrt{\lambda_n}\, \phi_n(s)\, \xi_n(\omega),$$

where:

- $\{\lambda_n, \phi_n\}_{n=1}^{\infty}$ are the eigenpairs of the covariance operator

  $$\int_D C(s,t)\, \phi_n(t)\, dt = \lambda_n \phi_n(s),$$

  with $\lambda_n \geq 0$ and $\{\phi_n\}$ forming an orthonormal basis in $L^2(D)$,

- $\{\xi_n(\omega)\}_{n=1}^{\infty}$ are uncorrelated random variables with zero mean and unit variance:

  $$\mathbb{E}[\xi_n] = 0, \qquad \mathbb{E}[\xi_n \xi_m] = \delta_{nm}.$$

  which will here be normalized centered orthogonal Gaussian variables.

Truncating the series after N terms yields the optimal (in the mean-square sense) finite-dimensional approximation, as the KL expansion is the minimal representation of X:

$$\widetilde{X}(s,\omega) = X_N(s,\omega) = \sum_{n=1}^{N} \sqrt{\lambda_n}\, \phi_n(s)\, \xi_n(\omega).$$

After presenting the Matérn and Exponential covariances, we will see that we know an explicit method to compute the KL expansion of an Exponential field over a segment in $\mathbb{R}$.

1.3 Stochastic Partial Differential Equations

SPDE [8] A stochastic partial differential equation (SPDE) is an equation of the form:

$$\mathfrak{D}u(t,x) \;=\; f(t,x,u(t,x)) \;+\; g(t,x,u(t,x))\,\dot{W}(t,x), \quad (t,x) \in [0,T] \times D,$$

Where

- $\mathfrak{D}$ is a (partial) differential operator in time and/or space,
- $u(t,x)$ is the unknown random field (solution), usually defined on a Sobolev space,
- f and g are given deterministic or random functions,
- $\dot{W}(t,x)$ denotes a *space-time white noise* or more generally a Gaussian random field. We will see later that we only had time to study this later case, with an exponential covariance function.
- like deterministic PDEs, SPDEs can present boundary conditions.

Formally, the term $\dot{W}(t,x)$ is not a classical function but a distribution, so SPDEs are often interpreted in a *weak* sense, like we will see in the next definition.
Note that we will not study time-dependant SPDEs, so we will write such equations depending only on a space variable.

Pointwise-defined & Distribution-based solutions of random fields [5] Let Z be a random field over $\mathbb{R}^d$.

A *pointwise-defined solution* means that $Z(x)$ is a well-defined random variable for every point $x \in \mathbb{R}^d$. The random field must have sufficient regularity (e.g., continuous or Hölder-continuous sample paths) for pointwise evaluation. For example: Matérn fields with smoothness parameter $\nu > \frac{1}{2}$, or Brownian motion.

A *distribution-based solution* is defined only in the sense of distributions, allowing the studied random function Z to not be defined at pointwise values. For a test function ϕ in a suitable space, the field acts as

$$Z(\phi) := \int_{\mathbb{R}^d} Z(x)\,\phi(x)\,dx.$$

For example, the Gaussian white noise has infinite variance at single points, implying that SPDEs that involve it may not present classical pointwise solutions. This brings the concept of *Generalized random fields* which are a generalization of random fields, defined in the sense of distributions. We will not study these, we just need to acknowledge that generalized random fields such as the Gaussian white noise bring more mathematical complexity to stochastic problems.

Collocation methods [3] If we want to numerically solve a SDPE like previously defined (but simplified to match the problems we will address):

$$\mathfrak{D}u(x) \;=\; f(x, u(t,x)) \;+\; W(x), \quad x \in D,$$

We can use two approaches: a collocation method or a stochastic Finite-Elements Method. Collocation methods are mesh-free, they are applied to the strong form of the equation and by enforcing the SPDE at discrete points, whereas sFEM are applied to a weak form. We will only look at resolutions of strong form equations in this report, assuming the noise W is a known random field that can be evaluated pointwise, thus we only need to define collocation methods:
The key idea of a **collocation method** is to approximate $u(x)$ by a finite expansion

$$u_N(x) = \sum_{j=1}^{N} c_j \, \varphi_j(x),$$

where $\{\varphi_j\}_{j=1}^N$ are chosen basis functions (such as polynomials, like in the PC or KL methods presented above). A set of collocation points $\{x_1, \ldots, x_N\} \subset D$ is then chosen.

Instead of enforcing the SPDE in an integrated weak sense (as in the finite element method), the equation is imposed pointwise at the collocation points:

$$\mathfrak{D}u_N(x_i) = f(x_i) + W(x_i), \quad i = 1, \ldots, N.$$

This yields a linear system for the coefficients $\{c_j\}$:

$$A\mathbf{c} = \mathbf{f} + \boldsymbol{W},$$

where A is constructed from evaluations of $\mathfrak{D}\varphi_j$ at the collocation points, $\mathbf{f} = (f(x_1), \ldots, f(x_N))^\top$, and $\boldsymbol{W}$ is the discretized noise vector.

1.4 Matérn random fields

Matérn Covariance A **Matérn Gaussian random field** is a mean-zero Gaussian random field $\{X(s) : s \in \mathbb{R}^d\}$ whose covariance function is of the form:[6]

$$C(h) = \sigma^2 \, \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\|h\|}{\rho} \right)^{\nu} K_\nu \left(\frac{\|h\|}{\rho} \right), \quad h \in \mathbb{R}^d, \tag{1}$$

where

- $\sigma^2 > 0$ is the variance,
- $\rho > 0$ is the range parameter,
- $\nu > 0$ is the smoothness parameter,
- K_ν is the modified Bessel function of the second kind of order ν,
- $\Gamma(\cdot)$ denotes the gamma function.

We will keep the parameters as constants, making such Matérn fields *stationary*. The parameter ν controls the mean-square differentiability of the field: Z is k-times mean-square differentiable if and only if $\nu > k + \frac{1}{2}$. When $\nu = \frac{1}{2}$, the Matérn covariance is the same as the exponential covariance structure, and as $\nu \to \infty$, the Matérn covariance converges to the squared exponential covariance.

Matérn equation A Matérn random field $X(s)$, $s \in \mathbb{R}^d$ can be defined as the stationary solution to the SPDE:

$$\left(\kappa^2 - \Delta\right)^{\alpha/2} X(s) \stackrel{d}{=} W(s), \quad s \in \mathbb{R}^d, \tag{2}$$

where

- $\kappa > 0$ is a spatial scale parameter,
- $\alpha > 0$ controls the smoothness of the field,
- $W(s)$ denotes Gaussian white noise on $\mathbb{R}^d$.

Looking at the definition of the Matérn covariance function 1, we have:

- $\kappa = \frac{1}{\rho}$
- σ^2 is the marginal variance, related to (κ, ν, d) via

$$\sigma^2 = \frac{\Gamma(\nu)}{\Gamma(\alpha)(4\pi)^{d/2}\kappa^{2\nu}}.$$

We will now use the notation $\mathfrak{D} := \kappa^2 - \Delta$ for the differential operator associated with the Matérn equation, and we will only study the 1-dimensional case:

$$\mathfrak{D}^{\alpha/2} X(s) \stackrel{d}{=} W(s), \quad s \in \mathbb{R} \tag{3}$$

1.4.1 Simplified Matérn equation

If our goal is to solve the Matérn equation

$$(\kappa - \Delta)^{\frac{\alpha}{2}} X \stackrel{d}{=} W$$

With W a Gaussian white noise – i.e. a generalized random field, which are not defined at pointwise values, and thus cannot be handled by the neural network models we will present later.
We thus propose to approach first a simpler problem, of which we know the true solution and that does not include a generalized random field such as W so that we can test our method ideas.

From the theory [5], we know that the solution $X_{\alpha/2}$ of our goal problem $\mathfrak{D}^{\alpha/2}X_{\alpha/2} \stackrel{d}{=} W$ is a Gaussian random field with Matèrn covariance of regularity $\nu = \alpha - \frac{\dim}{2}$ (where *dim* is the dimension of the functions — here we always have dim = 1). With $\alpha \in \{1, 2\}$, we know how to solve this, by least-squares and Galerkin methods (we did not work on this part of the project during the internship). For above values, we know that we can solve it iteratively:

$$\mathfrak{D}X_\alpha \stackrel{d}{=} X_{\alpha-2} \Leftrightarrow X_\alpha \stackrel{d}{=} \mathfrak{D}^{-1}X_{\alpha-2}$$

So, to obtain an easier problem to solve, we propose to first focus on solving an iteration step like this, as it only includes well-known Matèrn gaussian random fields:
we will first solve $\mathfrak{D}X \stackrel{d}{=} Y$, with Y a GRF of Matèrn covariance of regularity $\nu = \frac{1}{2}$, which is equivalent to an Exponential covariance $C(d) = \sigma^2 exp(-\kappa d)$.
In this case, we already know that the solution X is a GRF of Matèrn covariance with $\nu = \frac{5}{2}$.

1.4.2 KL expansion of an Exponential random field

We saw that the Matérn covariance function 1 with $\nu = \frac{1}{2}$ is the Exponential covariance function. As detailed in the previous paragraph, we want to solve a SPDE implying an Exponential random field Y as a source term with a model. This implies that we need to simulate it with a convenient source of randomness, in order to compare it with the model evaluated with this same source.
To do so we propose to use the explicit method for its KL expansion presented by the article *An efficient, high-order perturbation approach for flow in random porous media via Karhunen–Loeve and polynomial expansions* by Dongxiao Zhang, Zhiming Lu [9]:

We want the KL expansion at rank N of a field Y with $C_Y(x_1, y_1) = \sigma_Y^2 \exp\left(-\frac{|x_1 - y_1|}{\eta}\right)$ on a segment $[-\frac{L}{2}, \frac{L}{2}]$ of size $L \in \mathbb{R}$:

$$\widetilde{Y}(\omega) = \sum_{n=1}^{N} \sqrt{\lambda_n}\, \phi_n(s)\, \xi_n(\omega)$$

λ_n and $\phi_n(x)$ can be found analytically, and the article gives their expressions:

$$\lambda_n = \frac{2\eta\sigma_Y^2}{\eta^2 w_n^2 + 1},$$

and

$$\phi_n(x) = \frac{1}{\sqrt{(\eta^2 w_n^2 + 1)L/2 + \eta}} \big[\eta w_n \cos(w_n x) + \sin(w_n x)\big],$$

Where w_n are positive roots of the characteristic equation:

$$(\eta^2 w^2 - 1)\sin(wL) = 2\eta w \cos(wL).$$

Once we have computed these values we have an expression to generate the field, and we can simulate it by drawing a realization of a vector of gaussians $\boldsymbol{\xi}$ that the expansion will call. If we want to compare a model with the field, we only need to feed the same realizations of the vector of gaussians to both.

1.5 Physics-Informed Neural Networks

Physics-Informed Neural Networks (PINNs) are a class of deep learning models that integrate data and physical laws, expressed as partial differential equations (PDEs), into the training process [7]. Instead of relying on labeled data (that we will not have here), PINNs embed the governing equations into the loss function, enforcing that the neural network's predictions satisfy the underlying physical dynamics.

Formally, given a PDE (only space-dependant, like the ones we will study) of the form:

$$\mathfrak{D}u(x) = f(x), \quad x \in D\,(\text{a spatial domain}),\ t \in [0,T],\ f \in L^2(D)$$

a PINN approximates the solution $u(x)$ by a neural network u_θ (θ being the set of parameters of the neural network), and the loss function typically combines data mismatch with the residual of the PDE:

$$\mathcal{L}(\theta) = \underbrace{\frac{1}{N_d}\sum_{i=1}^{N_d} \left\| u_\theta(\mathbf{x}_i) - u_i \right\|^2}_{\text{data loss}} + \underbrace{\frac{1}{N_f}\sum_{j=1}^{N_f} \left\| \mathfrak{D}[u_\theta(\mathbf{x}_j)] - f(\mathbf{x}_j) \right\|^2}_{\text{physics loss}}.$$

By minimizing $\mathcal{L}(\theta)$, PINNs exploit both data and prior knowledge of physical laws, making them effective for solving forward and inverse problems in scientific computing. In our case, we will only rely on the physic-informed part of the loss and not use data:

$$\mathcal{L}(\theta) = \frac{1}{N_f}\sum_{j=1}^{N_f} \left\| \mathfrak{D}[u_\theta(\mathbf{x}_j)] - f(\mathbf{x}_j) \right\|^2. \tag{4}$$

with N_f the number of collocation points, and $\|.\|^2$ the squared mean of absolute values.

References

[1] R. J. Adler. *The Geometry of Random Fields*. John Wiley & Sons, 1981.

[2] Denis Allard. *Statistiques spatiales: introduction à la géostatistique*. 2012.

[3] Ivo Babuška, Fabio Nobile, and Raul Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 45(3):1005–1034, 2007.

[4] O.P. Le Maître and O.M. Knio. *Spectral Methods for Uncertainty Quantification*. 2010.

[5] Finn Lindgren, Håvard Rue, and Johan 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.

[6] Bertil Matérn. *Spatial Variation*, volume 49. Meddelanden fran Statens Skogsforskningsinstitut, 1960.

[7] Maziar Raissi, Paris Perdikaris, and George E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*.

[8] John B. Walsh. *An Introduction to Stochastic Partial Differential Equations*, volume 1180 of *Lecture Notes in Mathematics*. Springer, Berlin, 1986.

[9] Dongxiao Zhang and Zhiming Lu. An efficient, high-order perturbation approach for flow in random porous media via karhunen–loeve and polynomial expansions. *Journal of Computational Physics*, 2003.