Physics informed Gaussian processes and kernels : theory
with applications

Iain Henderson

INSA Toulouse/IMT

LIKE23, June 28^{th} 2023

Contact : iain.pl.henderson@gmail.com

Academic context

- Third year PhD student at the Mathematics Institute of Toulouse/INSA Toulouse, advised by Pascal Noble (Partial differential equations) and Olivier Roustant (Statistics and optimization).
- Funded by the SHOM (Service Hydrographique et Océanographique de la Marine), contact : Rémy Baraille

Outline

1. The problem
   - Regression under physical constraints
   - Gaussian process (regression)
2. Imposing physical constraints on Gaussian processes
   - The problem with PDEs...
   - Sobolev regularity of Gaussian processes
   - Linear PDE constraints on Gaussian processes
3. GPR for the wave equation
   - Wave equation tailored covariance functions
   - Solving inverse problems
   - Numerical application
4. Conclusion and perspectives

Outline of the presentation

1. The problem
   - Regression under physical constraints
   - Gaussian process (regression)
2. Imposing physical constraints on Gaussian processes
   - The problem with PDEs...
   - Sobolev regularity of Gaussian processes
   - Linear PDE constraints on Gaussian processes
3. GPR for the wave equation
   - Wave equation tailored covariance functions
   - Solving inverse problems
   - Numerical application
4. Conclusion and perspectives

Regression under physical constraints

- Aim : forecast of physical phenomena (oceanography)
  $\rightarrow$ unknown function u

- At our disposal : database w.r.t. u : $B = \{u(z_1), ..., u(z_n)\}$, probably limited.
- Physical model (Partial differential equation, PDE) :

$$\frac{\partial u}{\partial t} = D \frac{\partial^2 u}{\partial x^2} + f$$

- Objective : approximate $u(x, t)$ for all (x, t) (regression)

Regression under physical constraints

- Aim : forecast of physical phenomena (oceanography)
  $\rightarrow$ unknown function u

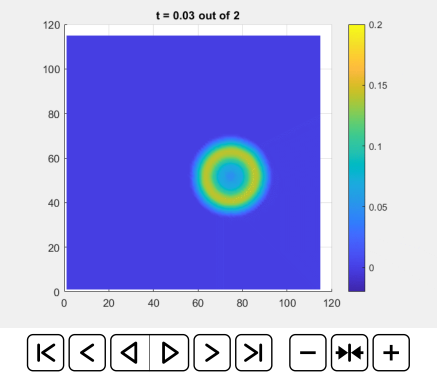
- At our disposal : database w.r.t. u : $B = \{u(z_1), ..., u(z_n)\}$, probably limited.
- Physical model (Partial differential equation, PDE) :

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} + f$$

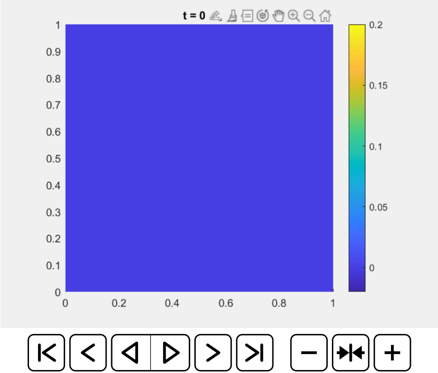
- Objective : approximate $u(x,t)$ for all (x,t) (regression)

- Idea : combine database and physics using probabilistic/Bayesian regression methods.

Direct numerical simulation



Reconstruction with GPR



Gaussian processes

- Gaussian process over $\mathcal{D} \subset \mathbb{R}^d$: $(U(z))_{z \in \mathcal{D}}$: family of Gaussian RVs such that the law of any vector of the form $(U(z_1), ..., U(z_n))$ is a Gaussian multivariate normal distribution.

Gaussian processes

- Gaussian process over $\mathcal{D} \subset \mathbb{R}^d$: $(U(z))_{z\in\mathcal{D}}$: family of Gaussian RVs such that the law of any vector of the form $(U(z_1), ..., U(z_n))$ is a Gaussian multivariate normal distribution.

  Mean $m(z) := \mathbb{E}[U(z)]$ and covariance $k(z, z') := \mathrm{Cov}(U(z), U(z'))$:

$$U(z) \sim GP(m, k)$$

  $\longrightarrow$ k is a positive semi-definite function : $k(z, z') = k(z', z)$ and $\forall (z_1, ..., z_n) \in \mathcal{D}^n, (k(x_i, x_j))_{1\leq i,j\leq n}$ is PSD.

- $\exists U \sim PG(0, k) \iff k$ is positive semi-definite.

Gaussian processes

- Gaussian process over $\mathcal{D} \subset \mathbb{R}^d$: $(U(z))_{z\in\mathcal{D}}$: family of Gaussian RVs such that the law of any vector of the form $(U(z_1), ..., U(z_n))$ is a Gaussian multivariate normal distribution.

  Mean $m(z) := \mathbb{E}[U(z)]$ and covariance $k(z, z') := \text{Cov}(U(z), U(z'))$:

$$U(z) \sim GP(m, k)$$

  $\longrightarrow$ k is a positive semi-definite function : $k(z, z') = k(z', z)$ and $\forall (z_1, ..., z_n) \in \mathcal{D}^n, (k(x_i, x_j))_{1\leq i,j\leq n}$ is PSD.

- $\exists U \sim PG(0, k) \iff k$ is positive semi-definite.
- Sample paths : given $\omega \in \Omega,\ U_\omega : z \mapsto U(z)(\omega)$.

Role of the covariance function

Let $\omega \in \Omega$, the associated sample path is $U(\omega) : z \mapsto U(z)(\omega)$.

Role of the covariance function

Let $\omega \in \Omega$, the associated sample path is $U(\omega) : z \mapsto U(z)(\omega)$.

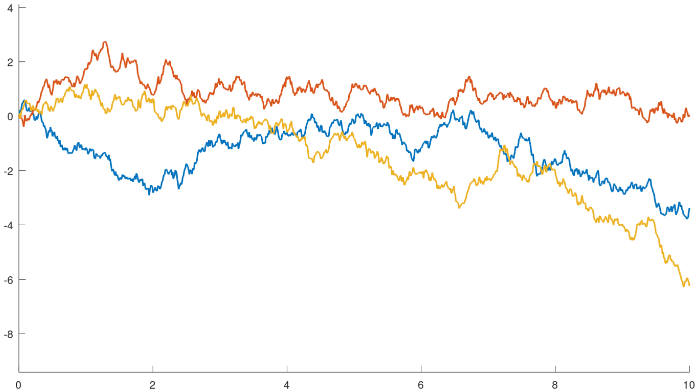


Figure 1 – Brownian motion : $k(x, y) = \min(x, y)$

Role of the covariance function

Let $\omega \in \Omega$ be a sample, the associated sample path is $U(\omega) : z \mapsto U(z)(\omega)$.

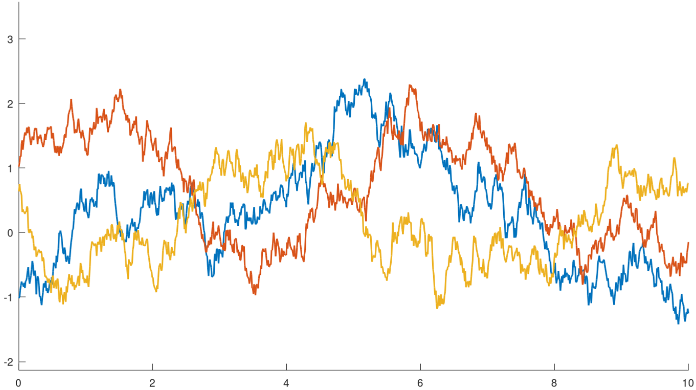


Figure 2 – Matérn 1/2 : $k(x, y) = \sigma^2 \exp(-|x-y|/\ell),\ \sigma = 1,\ \ell = 2$

Let $\omega \in \Omega$ be a sample, the associated sample path is $U(\omega) : z \mapsto U(z)(\omega)$.

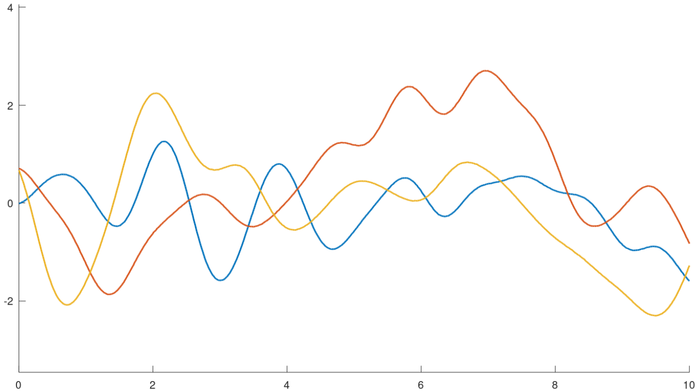


Figure 3 – Gaussian : $k(x, y) = \sigma^2 \exp(-|x - y|^2/2\ell^2), \sigma = 1, \ell = 0.5$

Role of the covariance function

Let $\omega \in \Omega$ be a sample, the associated sample path is $U(\omega) : z \mapsto U(z)(\omega)$.

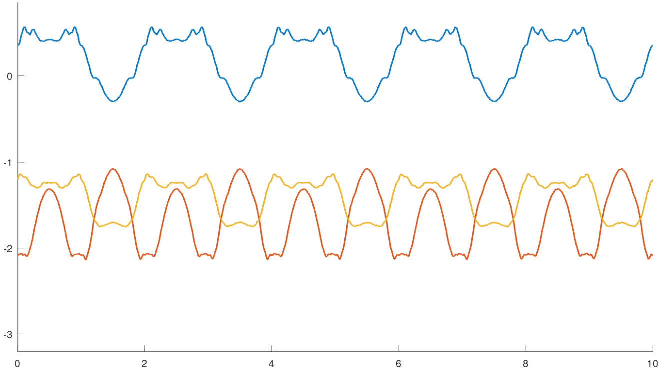


Figure 4 – Periodic Matérn 3/2 : $\sigma = 1,\ \ell = 1$ and
$k(x,y) = \sigma^2(1 + |\sin(\pi x) - \sin(\pi y)|/\ell)\exp(-|\sin(\pi x) - \sin(\pi y)|/\ell)$

Gaussian process regression (Kriging)

- Unknown function $z \in \mathcal{D} \longmapsto u(z)$, data $B = \{u(z_1), ..., u(z_n)\}$
- Model $z \longmapsto u(z)$ as a sample path of a GP (GP prior) $(U_z)_{z \in D} \sim PG(m(z), k(z, z'))$
- We condition U on the data : $V(z) = [U(z) | U(z_1) = u(z_1), ..., U(z_n) = u(z_n)]$ (GP posterior). We obtain

$$V(z) \sim PG(\tilde{m}(z), \tilde{k}(z, z'))$$

  $\tilde{m}$ et $\tilde{k}$ given by the Kriging formulas.
- Prediction/estimation : $\forall z \in \mathcal{D}$, we predict $u(z)$ with $\tilde{m}(z)$: $u(z) \simeq \tilde{m}(z)$, associated uncertainty $\tilde{k}(z, z) = \text{Var}(V(z))$

Kriging formulas

Denote $u_{obs} = (u(z_1), ..., u(z_n))$ the data, $K_{ij} := k(z_i, z_j)$ and $k(Z, z)_i := k(z_i, z)$. Then the a posteriori mean and covariance are given by

$$\begin{cases} \tilde{m}(z) & = k(Z,z)^T K^{-1} u_{obs} \in \text{Span}(k(z_1,\cdot), ..., k(z_n,\cdot)), \\ \tilde{k}(z,z') & = k(z,z') - k(Z,z)^T K^{-1} k(Z,z'). \end{cases}$$

Kriging formulas

Denote $u_{obs} = (u(z_1), ..., u(z_n))$ the data, $K_{ij} := k(z_i, z_j)$ and $k(Z, z)_i := k(z_i, z)$. Then the a posteriori mean and covariance are given by

$$\begin{cases} \tilde{m}(z) & = k(Z, z)^T K^{-1} u_{obs} \in \text{Span}(k(z_1, \cdot), ..., k(z_n, \cdot)), \\ \tilde{k}(z, z') & = k(z, z') - k(Z, z)^T K^{-1} k(Z, z'). \end{cases}$$

Assume that $Lu = 0$, L linear (prior knowledge). k is adapted to this constraint if $L\tilde{m} = 0$, i.e. $Lk(z, \cdot) = 0$ for all z.

Kriging formulas

Denote $u_{obs} = (u(z_1), ..., u(z_n))$ the data, $K_{ij} := k(z_i, z_j)$ and $k(Z, z)_i := k(z_i, z)$. Then the a posteriori mean and covariance are given by

$$\begin{cases} \tilde{m}(z) & = k(Z,z)^T K^{-1} u_{obs} \in \text{Span}(k(z_1, \cdot), ..., k(z_n, \cdot)), \\ \tilde{k}(z, z') & = k(z, z') - k(Z,z)^T K^{-1} k(Z, z'). \end{cases}$$

Assume that $Lu = 0$, L linear (prior knowledge). k is adapted to this constraint if $L\tilde{m} = 0$, i.e. $Lk(z, \cdot) = 0$ for all z.

Apply this when L is a PDE ! Conservation of mass, momentum, energy...

Outline of the presentation

Partial derivatives in dimension d : notations

Let $\alpha \in \mathbb{N}^d$, $\alpha = (\alpha_1, ..., \alpha_d)$. We denote

$$\partial^\alpha := (\partial_{x_1})^{\alpha_1} ... (\partial_{x_d})^{\alpha_d}. \tag{1}$$

Order of differentiation : $|\alpha| := \alpha_1 + ... + \alpha_d$.

Linear differential operator of order m :

$$L = \sum_{|\alpha| \leq m} a_\alpha \partial^\alpha \tag{2}$$

Solutions of partial differential equations

Consider a linear PDE $Lu = f$, or $\partial_t u = Lu$, $u(t = 0) = u_0$.
First difficulty : what *is* a solution to a PDE ?

Solutions of partial differential equations

Consider a linear PDE $Lu = f$, or $\partial_t u = Lu$, $u(t = 0) = u_0$.
First difficulty : what *is* a solution to a PDE ?

- Strong solution : PDE imposed pointwise, $(Lu)(x) = f(x)\ \forall x$.

Solutions of partial differential equations

Consider a linear PDE $Lu = f$, or $\partial_t u = Lu$, $u(t=0) = u_0$.
First difficulty : what *is* a solution to a PDE ?

- Strong solution : PDE imposed pointwise, $(Lu)(x) = f(x)\ \forall x$.
- Weak solution : PDE is true when "tested" against many finitely smooth functions.
  The PDE is turned into a problem of representation of continuous linear forms in some Sobolev space (Hilbert or Banach).
  Note : PDE dependent.

Solutions of partial differential equations

Consider a linear PDE $Lu = f$, or $\partial_t u = Lu$, $u(t=0) = u_0$.
First difficulty : what *is* a solution to a PDE ?

- Strong solution : PDE imposed pointwise, $(Lu)(x) = f(x)\ \forall x$.
- Weak solution : PDE is true when "tested" against many finitely smooth functions.
  The PDE is turned into a problem of representation of continuous linear forms in some Sobolev space (Hilbert or Banach).
  Note : PDE dependent.
- Distributional solution (L. Schwartz) : PDE is true when "tested" against compactly supported smooth functions. The PDE is turned into a problem of representation of continuous linear forms in some topological vector space (often non normable).
  Note : very general (minimal regularity assumptions).

Regularity theory for PDEs

Second difficulty : regularity theory.

Regularity theory for PDEs

Second difficulty : regularity theory.

- Elliptic (Poisson, Laplace) or parabolic (heat) PDEs : solution is (much) smoother than the source term.

Regularity theory for PDEs

Second difficulty : regularity theory.

- Elliptic (Poisson, Laplace) or parabolic (heat) PDEs : solution is (much) smoother than the source term.
- Hyperbolic PDEs : advection

$$\partial_t u + c\partial_x u = 0. \tag{3}$$

Solution given by $u(x,t) = u_0(x - ct)$.

Regularity theory for PDEs

Second difficulty : regularity theory.

- Elliptic (Poisson, Laplace) or parabolic (heat) PDEs : solution is (much) smoother than the source term.

- Hyperbolic PDEs : advection

$$\partial_t u + c\partial_x u = 0. \quad (3)$$

Solution given by $u(x,t) = u_0(x - ct)$.
$\rightarrow u(\cdot, t)$ as smooth as u_0.

Regularity theory for PDEs

Second difficulty : regularity theory.

- Elliptic (Poisson, Laplace) or parabolic (heat) PDEs : solution is (much) smoother than the source term.
- Hyperbolic PDEs : advection

$$\partial_t u + c\partial_x u = 0. \tag{3}$$

Solution given by $u(x,t) = u_0(x - ct)$.
$\rightarrow u(\cdot, t)$ as smooth as u_0.Also, what is the meaning of the PDE if u_0 is discontinuous ?

Regularity theory for PDEs

Second difficulty : regularity theory.

- Elliptic (Poisson, Laplace) or parabolic (heat) PDEs : solution is (much) smoother than the source term.
- Hyperbolic PDEs : advection

$$\partial_t u + c\partial_x u = 0. \tag{3}$$

Solution given by $u(x,t) = u_0(x - ct)$.
$\rightarrow u(\cdot, t)$ as smooth as u_0.Also, what is the meaning of the PDE if u_0 is discontinuous ?

Basic tool for regularity theory : Sobolev spaces.

Today's topic

Target function $u : \mathcal{D} \to \mathbb{R}$, solution of some linear PDE $Lu = 0$, regress u w.r.t. + dataset $\{u(z_1), ..., u(z_n)\}$ (or other linear forms).

Today's topic

Target function $u : \mathcal{D} \to \mathbb{R}$, solution of some linear PDE $Lu = 0$, regress u w.r.t. + dataset $\{u(z_1), ..., u(z_n)\}$ (or other linear forms).

Aim for today : identify the GP priors $U \sim GP(0, k)$ which take into account the (properties of the) PDE, at the level of the sample paths.

Today's topic

Target function $u : \mathcal{D} \to \mathbb{R}$, solution of some linear PDE $Lu = 0$, regress u w.r.t. + dataset $\{u(z_1), ..., u(z_n)\}$ (or other linear forms).

Aim for today : identify the GP priors $U \sim GP(0, k)$ which take into account the (properties of the) PDE, at the level of the sample paths. How to choose k such that

- such that the sample paths of U verify $LU_\omega = 0$, in the sense of distributions, almost surely.
- such that the sample paths of U have a given Sobolev regularity (e.g. that associated to the PDE), almost surely.

Today's topic

Target function $u : \mathcal{D} \to \mathbb{R}$, solution of some linear PDE $Lu = 0$, regress u w.r.t. + dataset $\{u(z_1), ..., u(z_n)\}$ (or other linear forms).

Aim for today : identify the GP priors $U \sim GP(0, k)$ which take into account the (properties of the) PDE, at the level of the sample paths. How to choose k such that

- such that the sample paths of U verify $LU_\omega = 0$, in the sense of distributions, almost surely.
- such that the sample paths of U have a given Sobolev regularity (e.g. that associated to the PDE), almost surely.

Note :

- One may have $u \in H^m(\mathbb{R}^d)$ and $u \notin C^0(\mathbb{R}^d)$ ($u \in C^0$ if $m > d/2$). Hence, no continuity assumptions on the GP (!).

Today's topic

Target function $u : \mathcal{D} \to \mathbb{R}$, solution of some linear PDE $Lu = 0$, regress u w.r.t. + dataset $\{u(z_1), ..., u(z_n)\}$ (or other linear forms).

Aim for today : identify the GP priors $U \sim GP(0, k)$ which take into account the (properties of the) PDE, at the level of the sample paths.
How to choose k such that

- such that the sample paths of U verify $LU_\omega = 0$, in the sense of distributions, almost surely.
- such that the sample paths of U have a given Sobolev regularity (e.g. that associated to the PDE), almost surely.

Note :

- One may have $u \in H^m(\mathbb{R}^d)$ and $u \notin C^0(\mathbb{R}^d)$ ($u \in C^0$ if $m > d/2$). Hence, no continuity assumptions on the GP (!).
- No regularity assumptions on open set $\mathcal{D}$. Thus no fractional Sobolev spaces or Fourier methods.

Sobolev regularity of Gaussian processes

Derivatives with finite energy and Sobolev spaces

Some functions are "almost" differentiable : $h(x) = \max(0, 1 - |x|)$.

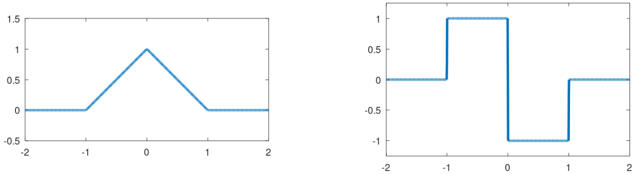
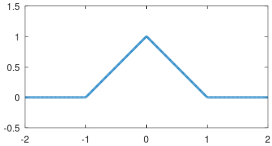


Figure 5 – Left : $h(x)$. Right : $h'(x)$ (hopefully).

Unfortunately, $h' \notin C^0$... but $h' \in L^2$ (finite energy) !

Some functions are "almost" differentiable : $h(x) = \max(0, 1 - |x|)$.



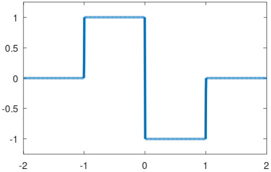


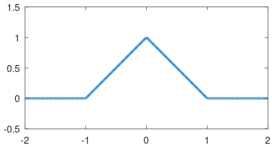
Figure 5 – Left : $h(x)$. Right : $h'(x)$ (hopefully).

Unfortunately, $h' \notin C^0$... but $h' \in L^2$ (finite energy) !
A function g is the weak derivative of h if for all $\varphi \in C_c^\infty(\mathbb{R})$,

$$\int_{\mathbb{R}} h(x)\varphi'(x)dx = -\int_{\mathbb{R}} g(x)\varphi(x)dx.$$

Derivatives with finite energy and Sobolev spaces

Some functions are "almost" differentiable : $h(x) = \max(0, 1 - |x|)$.



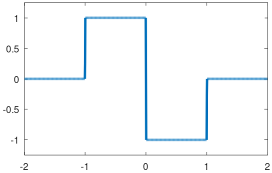


Figure 5 – Left : $h(x)$. Right : $h'(x)$ (hopefully).

Unfortunately, $h' \notin C^0$... but $h' \in L^2$ (finite energy) !
A function g is the weak derivative of h if for all $\varphi \in C_c^\infty(\mathbb{R})$,

$$\int_{\mathbb{R}} h(x)\varphi'(x)dx = -\int_{\mathbb{R}} g(x)\varphi(x)dx.$$

We then define

$$H^1(\mathbb{R}) := \{u \in L^2(\mathbb{R}) : u' \text{ exists in the weak sense and } u' \in L^2(\mathbb{R})\},$$

$$H^m(\mathcal{D}) := \{u \in L^2(\mathcal{D}) : \forall\ |\alpha| \leq m, \partial^\alpha u \text{ exists ITWS and } \partial^\alpha u \in L^2(\mathcal{D})\}.$$

Sobolev spaces and PDEs

- Energy functionals ("physical interpretation") :

- Energy functionals ("physical interpretation") :
  - Parabolic : $\partial_t - \Delta u = 0$, $(x, t) \in \mathbb{R}^d \times \mathbb{R}_+$.

$$\frac{1}{2}\partial_t \|u(\cdot, t)\|_{L^2}^2 = -\|\nabla u(\cdot, t)\|_{L^2}^2 < 0 \quad \text{(diffusion).} \tag{4}$$

Sobolev spaces and PDEs

- Energy functionals ("physical interpretation") :
  - Parabolic : $\partial_t - \Delta u = 0$, $(x, t) \in \mathbb{R}^d \times \mathbb{R}_+$.

$$\frac{1}{2}\partial_t \|u(\cdot, t)\|_{L^2}^2 = -\|\nabla u(\cdot, t)||_{L^2}^2 < 0 \quad \text{(diffusion)}. \tag{4}$$

  - Hyperbolic : wave, $\partial_{tt}^2 - \Delta u = 0$.

$$\partial_t \Big(\|\partial_t u(\cdot, t)\|_{L^2}^2 + \|\nabla u(\cdot, t)\|_{L^2}^2 \Big) = 0 \quad \text{(conservation)}. \tag{5}$$

  Likewise, advection : if $\partial_t u + \partial_x u = 0$, then $\partial_t \|u(\cdot, t)\|_{L^p} = 0$.

Sobolev spaces and PDEs

- Energy functionals ("physical interpretation") :
  - Parabolic : $\partial_t - \Delta u = 0$, $(x,t) \in \mathbb{R}^d \times \mathbb{R}_+$.

$$\frac{1}{2}\partial_t \|u(\cdot,t)\|_{L^2}^2 = -\|\nabla u(\cdot,t)\|_{L^2}^2 < 0 \quad \text{(diffusion)}. \tag{4}$$

  - Hyperbolic : wave, $\partial_{tt}^2 - \Delta u = 0$.

$$\partial_t \Big(\|\partial_t u(\cdot,t)\|_{L^2}^2 + \|\nabla u(\cdot,t)\|_{L^2}^2 \Big) = 0 \quad \text{(conservation)}. \tag{5}$$

  Likewise, advection : if $\partial_t u + \partial_x u = 0$, then $\partial_t \|u(\cdot,t)\|_{L^p} = 0$.

- Theory : Lax-Milgram for elliptic PDEs. More generally, Sobolev spaces are separable Banach spaces, reflexive when $1 < p < +\infty$.

- Numerical methods : finite element method, ...

L^2 regularity of Gaussian processes

Before understanding H^m regularity of GPs we need to understand their L^2 regularity. A GP is *measurable* if the map $(\omega, x) \mapsto U(x)(\omega)$ is measurable.

L^2 regularity of Gaussian processes

Before understanding H^m regularity of GPs we need to understand their L^2 regularity. A GP is *measurable* if the map $(\omega, x) \mapsto U(x)(\omega)$ is measurable.

- Integral criterion :

$$\mathbb{P}\big(U \in L^2(\mathcal{D})\big) = 1 \iff \int_{\mathcal{D}} k(x,x)dx < +\infty. \tag{6}$$

L^2 regularity of Gaussian processes

Before understanding H^m regularity of GPs we need to understand their L^2 regularity. A GP is *measurable* if the map $(\omega, x) \mapsto U(x)(\omega)$ is measurable.

- Integral criterion :

$$\mathbb{P}\big(U \in L^2(\mathcal{D})\big) = 1 \iff \int_{\mathcal{D}} k(x,x)dx < +\infty. \tag{6}$$

Comes from $\mathbb{E}[\int U(x)^2 dx] = \int \mathbb{E}[U(x)^2]dx = \int k(x,x)dx$.

L^2 regularity of Gaussian processes

Before understanding H^m regularity of GPs we need to understand their L^2 regularity. A GP is *measurable* if the map $(\omega, x) \mapsto U(x)(\omega)$ is measurable.

- Integral criterion :

$$\mathbb{P}\big(U \in L^2(\mathcal{D})\big) = 1 \iff \int_{\mathcal{D}} k(x,x)dx < +\infty. \tag{6}$$

  Comes from $\mathbb{E}[\int U(x)^2 dx] = \int \mathbb{E}[U(x)^2]dx = \int k(x,x)dx$.

- Spectral/Mercer criterion : denote $\mathcal{E}_k : L^2(\mathcal{D}) \to L^2(\mathcal{D})$ the operator

$$(\mathcal{E}_k f)(x) := \int k(x,y)f(y)dy. \tag{7}$$

L^2 regularity of Gaussian processes

Before understanding H^m regularity of GPs we need to understand their L^2 regularity. A GP is *measurable* if the map $(\omega, x) \mapsto U(x)(\omega)$ is measurable.

- Integral criterion :

$$\mathbb{P}\big(U \in L^2(\mathcal{D})\big) = 1 \iff \int_{\mathcal{D}} k(x,x)dx < +\infty. \tag{6}$$

  Comes from $\mathbb{E}[\int U(x)^2 dx] = \int \mathbb{E}[U(x)^2]dx = \int k(x,x)dx$.

- Spectral/Mercer criterion : denote $\mathcal{E}_k : L^2(\mathcal{D}) \to L^2(\mathcal{D})$ the operator

$$(\mathcal{E}_k f)(x) := \int k(x,y)f(y)dy. \tag{7}$$

  If $\int k(x,x)dx < +\infty$, then (...) for some $(\phi_n) \subset L^2$, ONB (eigenvectors of $\mathcal{E}_k$) and $(\lambda_n) \subset \mathbb{R}_+$ (eigenvalues of $\mathcal{E}_k$)

$$k(x,y) = \sum_{n=0}^{+\infty} \lambda_n \phi_n(x)\phi_n(y) \quad \text{in } L^2(\mathcal{D}\times\mathcal{D}) \quad (\text{"Mercer"}). \tag{8}$$

Equivalently, set $\psi_n = \lambda_n^{1/2}\phi_n$, then $\|\psi_n\|_2^2 = \lambda_n$ and

$$k(x,y) = \sum_{n=0}^{+\infty} \psi_n(x)\psi_n(y) \quad \text{in } L^2(\mathcal{D}\times\mathcal{D}). \tag{9}$$

Thus (formally)

$$\int k(x,x)dx = \int \sum_{n=0}^{+\infty} \psi_n(x)^2 dx = \sum_{n=0}^{+\infty} \int \psi_n(x)^2 dx \tag{10}$$

$$= \sum_{n=0}^{+\infty} \|\psi_n\|_2^2 = \sum_{n=0}^{+\infty} \lambda_n = \mathrm{Tr}(\mathcal{E}_k) < +\infty \quad \text{(Trace class)}. \tag{11}$$

L^2 regularity of Gaussian processes

Equivalently, set $\psi_n = \lambda_n^{1/2}\phi_n$, then $\|\psi_n\|_2^2 = \lambda_n$ and

$$k(x, y) = \sum_{n=0}^{+\infty} \psi_n(x)\psi_n(y) \quad \text{in } L^2(\mathcal{D}\times\mathcal{D}). \tag{9}$$

Thus (formally)

$$\int k(x, x)dx = \int \sum_{n=0}^{+\infty} \psi_n(x)^2 dx = \sum_{n=0}^{+\infty} \int \psi_n(x)^2 dx \tag{10}$$

$$= \sum_{n=0}^{+\infty} \|\psi_n\|_2^2 = \sum_{n=0}^{+\infty} \lambda_n = \text{Tr}(\mathcal{E}_k) < +\infty \quad \text{(Trace class)}. \tag{11}$$

- RKHS imbedding criterion : observe that $H_k \subset L^2(\mathcal{D})$. Denote $\mathcal{I}$ the associated embedding, then $\mathcal{I}\mathcal{I}^* = \mathcal{E}_k$ is trace class ("Driscoll").

Sobolev regularity of Gaussian processes

Proposition 1 (H. [2022])

Let $(U(z))_{z\in\mathcal{D}} \sim GP(0,k)$ *be measurable, we have equivalence between*
(i) $\mathbb{P}(\{\omega \in \Omega : U(\omega) \in H^m(\mathcal{D})\}) = 1$

Sobolev regularity of Gaussian processes

Proposition 1 (H. [2022])

Let $(U(z))_{z\in\mathcal{D}} \sim GP(0,k)$ *be measurable, we have equivalence between*
(i) $\mathbb{P}(\{\omega \in \Omega : U(\omega) \in H^m(\mathcal{D})\}) = 1$
(ii) For all $|\alpha| \leq m$, $\partial^{\alpha,\alpha} k \in L^2(\mathcal{D}\times\mathcal{D})$ *and the integral operator* $\mathcal{E}_k^\alpha$

$$\mathcal{E}_k^\alpha : L^2(\mathcal{D}) \to L^2(\mathcal{D}), \quad \mathcal{E}_k^\alpha f(x) = \int_{\mathcal{D}} \partial^{\alpha,\alpha} k(x,y) f(y) dy$$

is trace class, with, $Tr(\mathcal{E}_k^\alpha) = \int_{\mathcal{D}} \partial^{\alpha,\alpha} k(x,x) dx < +\infty.$

Proposition 1 (H. [2022])

Let $(U(z))_{z\in\mathcal{D}} \sim GP(0,k)$ *be measurable, we have equivalence between*
(i) $\mathbb{P}(\{\omega\in\Omega : U(\omega)\in H^m(\mathcal{D})\}) = 1$
(ii) For all $|\alpha|\leq m$, $\partial^{\alpha,\alpha}k \in L^2(\mathcal{D}\times\mathcal{D})$ *and the integral operator* $\mathcal{E}_k^\alpha$

$$\mathcal{E}_k^\alpha : L^2(\mathcal{D}) \to L^2(\mathcal{D}), \quad \mathcal{E}_k^\alpha f(x) = \int_{\mathcal{D}} \partial^{\alpha,\alpha}k(x,y)f(y)dy$$

is trace class, with, $Tr(\mathcal{E}_k^\alpha) = \int_{\mathcal{D}} \partial^{\alpha,\alpha}k(x,x)dx < +\infty$.
(iii) There exists $(\phi_n)\subset L^2(\mathcal{D})$ *such that* $k(x,y) = \sum_n \phi_n(x)\phi_n(y)$ *in* $L^2(\mathcal{D}\times\mathcal{D})$. *Moreover, if* $|\alpha|\leq m$, *then* $\phi_n \in H^m(\mathcal{D})$ *and*

$$Tr(\mathcal{E}_k^\alpha) = \sum_{n=0}^{+\infty} \|\partial^\alpha\phi_n\|_2^2 < +\infty, \ \partial^{\alpha,\alpha}k = \sum_{n=0}^{+\infty} \partial^\alpha\phi_n \otimes \partial^\alpha\phi_n \text{ in } L^2(\mathcal{D}\times\mathcal{D})$$

Sobolev regularity of Gaussian processes

Proposition 1 (H. [2022])

Let $(U(z))_{z\in\mathcal{D}} \sim GP(0,k)$ *be measurable, we have equivalence between*
(i) $\mathbb{P}(\{\omega \in \Omega : U(\omega) \in H^m(\mathcal{D})\}) = 1$
(ii) For all $|\alpha| \le m$, $\partial^{\alpha,\alpha} k \in L^2(\mathcal{D}\times\mathcal{D})$ *and the integral operator* $\mathcal{E}_k^\alpha$

$$\mathcal{E}_k^\alpha : L^2(\mathcal{D}) \to L^2(\mathcal{D}), \quad \mathcal{E}_k^\alpha f(x) = \int_{\mathcal{D}} \partial^{\alpha,\alpha} k(x,y) f(y) dy$$

is trace class, with, $Tr(\mathcal{E}_k^\alpha) = \int_{\mathcal{D}} \partial^{\alpha,\alpha} k(x,x)dx < +\infty$.
(iii) There exists $(\phi_n) \subset L^2(\mathcal{D})$ *such that* $k(x,y) = \sum_n \phi_n(x)\phi_n(y)$ *in* $L^2(\mathcal{D}\times\mathcal{D})$. *Moreover, if* $|\alpha| \le m$, *then* $\phi_n \in H^m(\mathcal{D})$ *and*

$$Tr(\mathcal{E}_k^\alpha) = \sum_{n=0}^{+\infty} \|\partial^\alpha \phi_n\|_2^2 < +\infty, \ \partial^{\alpha,\alpha} k = \sum_{n=0}^{+\infty} \partial^\alpha \phi_n \otimes \partial^\alpha \phi_n \text{ in } L^2(\mathcal{D}\times\mathcal{D})$$

(iv) $H_k \subset H^m(\mathcal{D})$. *Note the imbedding* $\mathcal{I} : RKHS(k) \to H^m(\mathcal{D})$, *then* $Tr(\mathcal{I}\mathcal{I}^*) = \sum_{|\alpha|\le m} Tr(\mathcal{E}_k^\alpha) < +\infty$.

Sobolev spaces of non Hilbert type

Nonlinear Schrödinger equation, $p > 1$:

$$i\partial_t u + \Delta u = u|u|^{p-1}. \quad (12)$$

$$W^{1,p}(\mathbb{R}) := \{u \in L^p(\mathbb{R}) : u' \text{ exists in the weak sense and } u' \in L^p(\mathbb{R})\},$$
$$W^{m,p}(\mathcal{D}) := \{u \in L^p(\mathcal{D}) : \forall\ |\alpha| \leq m, \partial^\alpha u \text{ exists ITWS and } \partial^\alpha u \in L^p(\mathcal{D})\}.$$

Sobolev spaces of non Hilbert type

Nonlinear Schrödinger equation, $p > 1$:

$$i\partial_t u + \Delta u = u|u|^{p-1}. \tag{12}$$

$$W^{1,p}(\mathbb{R}) := \{u \in L^p(\mathbb{R}) : u' \text{ exists in the weak sense and } u' \in L^p(\mathbb{R})\},$$
$$W^{m,p}(\mathcal{D}) := \{u \in L^p(\mathcal{D}) : \forall\ |\alpha| \leq m, \partial^\alpha u \text{ exists ITWS and } \partial^\alpha u \in L^p(\mathcal{D})\}.$$

L^p regularity of GPs : C_p such that if $X \sim \mathcal{N}(0, \sigma^2)$, then $\mathbb{E}[|X|^p] = C_p \sigma^p$.

$$\mathbb{E}\left[\int_{\mathcal{D}} U(x)^p dx\right] = \int_{\mathcal{D}} \mathbb{E}[U(x)^p] dx = C_p \int_{\mathcal{D}} k(x,x)^{p/2} dx. \tag{13}$$

Sobolev spaces of non Hilbert type

Nonlinear Schrödinger equation, $p > 1$:

$$i\partial_t u + \Delta u = u|u|^{p-1}. \tag{12}$$

$$W^{1,p}(\mathbb{R}) := \{u \in L^p(\mathbb{R}) : u' \text{ exists in the weak sense and } u' \in L^p(\mathbb{R})\},$$
$$W^{m,p}(\mathcal{D}) := \{u \in L^p(\mathcal{D}) : \forall\ |\alpha| \leq m, \partial^\alpha u \text{ exists ITWS and } \partial^\alpha u \in L^p(\mathcal{D})\}.$$

L^p regularity of GPs : C_p such that if $X \sim \mathcal{N}(0, \sigma^2)$, then $\mathbb{E}[|X|^p] = C_p \sigma^p$.

$$\mathbb{E}\left[\int_{\mathcal{D}} U(x)^p dx\right] = \int_{\mathcal{D}} \mathbb{E}[U(x)^p] dx = C_p \int_{\mathcal{D}} k(x,x)^{p/2} dx. \tag{13}$$

Likewise for Mercer decomposition : there exists $(\phi_n) \subset L^p(\mathcal{D})$ such that $\sum \|\psi_n\|_p^2 < +\infty$ and

$$k(x,y) = \sum_{n=0}^{+\infty} \psi_n(x)\psi_n(y) \quad \text{in} \quad L^p(\mathcal{D} \times \mathcal{D}) \quad (\text{“nuclear”}). \tag{14}$$

Sobolev regularity of Gaussian processes : Banach case $W^{m,p}, 1 < p < +\infty, m \in \mathbb{N}$

Proposition 2 (H. [2022])

Let $(U(z))_{z\in\mathcal{D}} \sim GP(0,k)$ *be measurable, we have equivalence between*
(i) $\mathbb{P}(\{\omega \in \Omega : U(\omega) \in W^{m,p}(\mathcal{D})\}) = 1$
(ii) For all $|\alpha| \leq m$, $\partial^{\alpha,\alpha}k \in L^p(\mathcal{D}\times\mathcal{D})$ *and the integral operator* $\mathcal{E}_k^\alpha$

$$\mathcal{E}_k^\alpha : L^q(\mathcal{D}) \to L^p(\mathcal{D}), \quad \mathcal{E}_k^\alpha f(x) = \int_{\mathcal{D}} \partial^{\alpha,\alpha}k(x,y)f(y)dy$$

is symmetric, nonnegative and nuclear : there exists $(\phi_n^\alpha) \subset L^p(\mathcal{D})$ *such that* $\partial^{\alpha,\alpha}k(x,y) = \sum_n \psi_n^\alpha(x)\psi_n^\alpha(y)$ *in* $L^p(\mathcal{D}\times\mathcal{D})$ *verifying*

$$\sum_{n=0}^{+\infty} \|\psi_n^\alpha\|_p^2 < +\infty \quad (+\textit{refinements if } 1 \leq p \leq 2)$$

(iii) For all $|\alpha| \leq m$, $\partial^{\alpha,\alpha}k \in L^p(\mathcal{D}\times\mathcal{D})$ *and* $\int_{\mathcal{D}}[\partial^{\alpha,\alpha}k(x,x)]^{p/2}dx < +\infty$.

Gaussian processes under linear distributional PDE constraints

Distributional formulations of PDEs

Let $\mathcal{D} \subset \mathbb{R}^d$ be an open set and consider the PDE

$$Lu = \sum_{|\alpha| \leq m} a_\alpha \partial^\alpha u = 0.$$

Distributional formulations of PDEs

Let $\mathcal{D} \subset \mathbb{R}^d$ be an open set and consider the PDE

$$Lu = \sum_{|\alpha| \leq m} a_\alpha \partial^\alpha u = 0.$$

Strong solution : $u \in C^m(\mathcal{D})$ et $(Lu)(x) = 0 \ \ \forall x \in \mathcal{D}$.

Distributional formulations of PDEs

Let $\mathcal{D} \subset \mathbb{R}^d$ be an open set and consider the PDE

$$Lu = \sum_{|\alpha| \leq m} a_\alpha \partial^\alpha u = 0.$$

Strong solution : $u \in C^m(\mathcal{D})$ et $(Lu)(x) = 0 \ \ \forall x \in \mathcal{D}$.

Multiply by a test function $\varphi \in C_c^\infty(\mathcal{D})$ and integrate over $\mathcal{D}$

$$\forall \varphi \in C_c^\infty(\mathcal{D}), \int_{\mathcal{D}} Lu(x)\varphi(x)dx = 0. \tag{15}$$

Distributional formulations of PDEs

Let $\mathcal{D} \subset \mathbb{R}^d$ be an open set and consider the PDE

$$Lu = \sum_{|\alpha| \leq m} a_\alpha \partial^\alpha u = 0.$$

Strong solution : $u \in C^m(\mathcal{D})$ et $(Lu)(x) = 0 \ \ \forall x \in \mathcal{D}$.

Multiply by a test function $\varphi \in C_c^\infty(\mathcal{D})$ and integrate over $\mathcal{D}$

$$\forall \varphi \in C_c^\infty(\mathcal{D}), \int_{\mathcal{D}} Lu(x)\varphi(x)dx = 0. \tag{15}$$

Formal adjoint : $L^* v = \sum_{|\alpha| \leq n} (-1)^{|\alpha|} \partial^\alpha (a_\alpha v)$. Successive integrations by parts yield distributional solutions :

$$\forall \varphi \in C_c^\infty(\mathcal{D}), \int_{\mathcal{D}} u(x) L^* \varphi(x) dx = 0. \tag{16}$$

Distributional formulations of PDEs

Let $\mathcal{D} \subset \mathbb{R}^d$ be an open set and consider the PDE

$$Lu = \sum_{|\alpha| \leq m} a_\alpha \partial^\alpha u = 0.$$

Strong solution : $u \in C^m(\mathcal{D})$ et $(Lu)(x) = 0 \quad \forall x \in \mathcal{D}$.

Multiply by a test function $\varphi \in C_c^\infty(\mathcal{D})$ and integrate over $\mathcal{D}$

$$\forall \varphi \in C_c^\infty(\mathcal{D}), \int_{\mathcal{D}} Lu(x)\varphi(x)dx = 0. \tag{15}$$

Formal adjoint : $L^* v = \sum_{|\alpha| \leq n} (-1)^{|\alpha|} \partial^\alpha (a_\alpha v)$. Successive integrations by parts yield distributional solutions :

$$\forall \varphi \in C_c^\infty(\mathcal{D}), \int_{\mathcal{D}} u(x) L^* \varphi(x) dx = 0. \tag{16}$$

Only requires that $u \in L^1_{loc}(\mathcal{D})$, i.e. $\int_K |u| < +\infty$ for all $K \subset \mathcal{D}$ compact.

PDE constrained (Gaussian) random fields

Proposition 3 (H. et al. [2023, to appear])

Let $\mathcal{D} \subset \mathbb{R}^d$ *be an open set and* $L := \sum_{|\alpha| \leq n} a_\alpha \partial^\alpha$ *with* $a_\alpha \in \mathcal{C}^{|\alpha|}(\mathcal{D})$*. Let* $U = \big(U(z)\big)_{z \in \mathcal{D}}$ *be a centered measurable second order random field with covariance function* $k(z, z')$*. Assume that* $\sigma : z \longmapsto k(z, z)^{1/2} \in L^1_{loc}(\mathcal{D})$*.*

PDE constrained (Gaussian) random fields

Proposition 3 (H. et al. [2023, to appear])

Let $\mathcal{D} \subset \mathbb{R}^d$ be an open set and $L := \sum_{|\alpha| \leq n} a_\alpha \partial^\alpha$ with $a_\alpha \in \mathcal{C}^{|\alpha|}(\mathcal{D})$. Let $U = \big(U(z)\big)_{z \in \mathcal{D}}$ be a centered measurable second order random field with covariance function $k(z, z')$. Assume that $\sigma : z \longmapsto k(z,z)^{1/2} \in L^1_{loc}(\mathcal{D})$.
1) Then $\mathbb{P}(\{\omega \in \Omega : U(\omega) \in L^1_{loc}(\mathcal{D})\}) = 1$ and $L^1_{loc}(\mathcal{D})$ et $k(z, \cdot) \in L^1_{loc}(\mathcal{D})$ for all $z \in \mathcal{D}$.
2) There is an equivalence between :

- $\mathbb{P}(\{\omega \in \Omega : L(U_\omega) = 0$ *in the sense of distributions*$\}) = 1$
- $\forall z \in \mathcal{D}, L(k(z, \cdot)) = 0$ *in the sense of distributions.*

PDE constrained (Gaussian) random fields

Proposition 3 (H. et al. [2023, to appear])

Let $\mathcal{D} \subset \mathbb{R}^d$ be an open set and $L := \sum_{|\alpha| \leq n} a_\alpha \partial^\alpha$ with $a_\alpha \in \mathcal{C}^{|\alpha|}(\mathcal{D})$. Let $U = \big(U(z)\big)_{z \in \mathcal{D}}$ be a centered measurable second order random field with covariance function $k(z, z')$. Assume that $\sigma : z \longmapsto k(z,z)^{1/2} \in L^1_{loc}(\mathcal{D})$.
1) Then $\mathbb{P}(\{\omega \in \Omega : U(\omega) \in L^1_{loc}(\mathcal{D})\}) = 1$ and $L^1_{loc}(\mathcal{D})$ et $k(z, \cdot) \in L^1_{loc}(\mathcal{D})$ for all $z \in \mathcal{D}$.
2) There is an equivalence between :

- $\mathbb{P}(\{\omega \in \Omega : L(U_\omega) = 0$ *in the sense of distributions*$\}) = 1$
- $\forall z \in \mathcal{D}, L(k(z, \cdot)) = 0$ *in the sense of distributions.*

Generalises a result from Ginsbourger et al. [2016] to PDE constraints.
Also inherited to conditioned GPs.

PDE constrained (Gaussian) random fields

Proposition 3 (H. et al. [2023, to appear])

Let $\mathcal{D} \subset \mathbb{R}^d$ be an open set and $L := \sum_{|\alpha| \leq n} a_\alpha \partial^\alpha$ with $a_\alpha \in \mathcal{C}^{|\alpha|}(\mathcal{D})$. Let $U = \big(U(z)\big)_{z \in \mathcal{D}}$ be a centered measurable second order random field with covariance function $k(z, z')$. Assume that $\sigma : z \longmapsto k(z,z)^{1/2} \in L^1_{loc}(\mathcal{D})$.
1) Then $\mathbb{P}(\{\omega \in \Omega : U(\omega) \in L^1_{loc}(\mathcal{D})\}) = 1$ and $L^1_{loc}(\mathcal{D})$ et $k(z, \cdot) \in L^1_{loc}(\mathcal{D})$ for all $z \in \mathcal{D}$.
2) There is an equivalence between :

- $\mathbb{P}(\{\omega \in \Omega : L(U_\omega) = 0$ *in the sense of distributions*$\}) = 1$
- $\forall z \in \mathcal{D}, L(k(z, \cdot)) = 0$ *in the sense of distributions.*

Generalises a result from Ginsbourger et al. [2016] to PDE constraints. Also inherited to conditioned GPs. Example for $L = \partial_t + c\partial_x$:

$$k((x,t),(x',t')) := k_0(x - ct, x' - ct'). \tag{17}$$

Examples of kernels verifying $L(k(z,\cdot)) = 0 \;\; \forall z$

Given L, find k_L such that $L(k_L(\cdot, z)) = 0 \; \forall z$; $\Delta = \sum_{i=1}^{d} \partial^2_{x_i x_i}$.

Examples of kernels verifying $L(k(z,\cdot)) = 0 \ \forall z$

Given L, find k_L such that $L(k_L(\cdot, z)) = 0 \ \forall z$; $\Delta = \sum_{i=1}^d \partial^2_{x_i x_i}$.

- Laplace : $\Delta u = 0$ Mendes and da Costa Júnior [2012], Ginsbourger et al. [2016]
- Heat : $\partial_t - D\Delta u = 0$ Albert and Rath [2020]
- Div/Curl : $\nabla \cdot u = 0$, $\nabla \times u = 0$ Scheuerer and Schlather [2012],Owhadi [2023b]
- Continuum mechanics : Jidling et al. [2018]
- Helmholtz : $-\Delta u = \lambda u$ Albert and Rath [2020]
- (non)stationary Maxwell : Wahlstrom et al. [2013], Jidling et al. [2017],Lange-Hegermann [2018]
- 3D wave and transport equation : H. et al. [2023, to appear]
- See also "latent forces" : Álvarez et al. [2009], López-Lopera et al. [2021]

Examples of kernels verifying $L(k(z,\cdot)) = 0 \;\; \forall z$

Given L, find k_L such that $L(k_L(\cdot, z)) = 0 \; \forall z$; $\Delta = \sum_{i=1}^{d} \partial^2_{x_i x_i}$.

- Laplace : $\Delta u = 0$ Mendes and da Costa Júnior [2012], Ginsbourger et al. [2016]
- Heat : $\partial_t - D\Delta u = 0$ Albert and Rath [2020]
- Div/Curl : $\nabla \cdot u = 0$, $\nabla \times u = 0$ Scheuerer and Schlather [2012],Owhadi [2023b]
- Continuum mechanics : Jidling et al. [2018]
- Helmholtz : $-\Delta u = \lambda u$ Albert and Rath [2020]
- (non)stationary Maxwell : Wahlstrom et al. [2013], Jidling et al. [2017],Lange-Hegermann [2018]
- 3D wave and transport equation : H. et al. [2023, to appear]
- See also "latent forces" : Álvarez et al. [2009], López-Lopera et al. [2021]

Always based on representations of solutions of $Lu = 0$ of the form

$$u = Gf$$

Outline of the presentation

1. The problem
   - Regression under physical constraints
   - Gaussian process (regression)
2. Imposing physical constraints on Gaussian processes
   - The problem with PDEs...
   - Sobolev regularity of Gaussian processes
   - Linear PDE constraints on Gaussian processes
3. GPR for the wave equation
   - Wave equation tailored covariance functions
   - Solving inverse problems
   - Numerical application
4. Conclusion and perspectives

GPR and the wave equation H. et al. [2023]

Homogeneous 3D wave equation : $\Delta := \partial^2_{xx} + \partial^2_{yy} + \partial^2_{zz}$

$$\begin{cases} Lu & = \frac{1}{c^2}\partial^2_{tt}u - \Delta u = \Box u = 0, \quad (x,t) \in \mathbb{R}^3 \times \mathbb{R}^+, \\ u(x,0) & = u_0(x), \quad \partial_t u(x,0) = v_0(x). \end{cases} \tag{18}$$

Homogeneous 3D wave equation : $\Delta := \partial^2_{xx} + \partial^2_{yy} + \partial^2_{zz}$

$$\begin{cases} Lu & = \frac{1}{c^2}\partial^2_{tt}u - \Delta u = \Box u = 0, \quad (x,t) \in \mathbb{R}^3 \times \mathbb{R}^+, \\ u(x,0) & = u_0(x), \quad \partial_t u(x,0) = v_0(x). \end{cases} \tag{18}$$

The solution u is represented by (Krichhoff)

$$\begin{aligned} u(x,t) &= \int_S t v_0(x - c|t|\gamma) + u_0(x - c|t|\gamma) - c|t|\gamma \cdot \nabla u_0(x - c|t|\gamma) \frac{d\Omega}{4\pi} \\ &= (F_t * v_0)(x) + (\dot{F}_t * u_0)(x), \end{aligned} \tag{19}$$

where $F_t = \sigma_{ct}/4\pi c^2 t$ and $\dot{F}_t = \partial_t F_t$.

GPR and the wave equation H. et al. [2023]

Homogeneous 3D wave equation : $\Delta := \partial^2_{xx} + \partial^2_{yy} + \partial^2_{zz}$

$$\begin{cases} Lu & = \frac{1}{c^2}\partial^2_{tt}u - \Delta u = \Box u = 0, \quad (x,t) \in \mathbb{R}^3 \times \mathbb{R}^+, \\ u(x,0) & = u_0(x), \quad \partial_t u(x,0) = v_0(x). \end{cases} \tag{18}$$

The solution u is represented by (Krichhoff)

$$\begin{aligned} u(x,t) &= \int_S t v_0(x - c|t|\gamma) + u_0(x - c|t|\gamma) - c|t|\gamma \cdot \nabla u_0(x - c|t|\gamma) \frac{d\Omega}{4\pi} \\ &= (F_t * v_0)(x) + (\dot{F}_t * u_0)(x), \end{aligned} \tag{19}$$

where $F_t = \sigma_{ct}/4\pi c^2 t$ and $\dot{F}_t = \partial_t F_t$. Assume that u_0 and v_0 are unknown $\rightarrow u_0 \sim PG(0, k_u)$ and $v_0 \sim PG(0, k_v)$, independent. u given by (19) is a GP with covariance function

$$k((x,t),(x',t')) = [(F_t \otimes F_{t'}) * k_v](x,x') + [(\dot{F}_t \otimes \dot{F}_{t'}) * k_u](x,x'). \tag{20}$$

The kernel k verifies $\Box k((x,t),\cdot) = 0$ for all $(x,t) \in \mathbb{R}^3 \times \mathbb{R}_+$ (distribs !).

Estimation of physical parameters and initial conditions

- Reconstruction of initial conditions : the Kriging mean verifies $\Box\tilde{m} = 0$. Thus

$$\tilde{m}(\cdot, t = 0) \simeq u_0, \quad \partial_t \tilde{m}(\cdot, t = 0) \simeq v_0$$

Estimation of physical parameters and initial conditions

- Reconstruction of initial conditions : the Kriging mean verifies $\Box \tilde{m} = 0$. Thus

$$\tilde{m}(\cdot, t=0) \simeq u_0, \quad \partial_t \tilde{m}(\cdot, t=0) \simeq v_0$$

- The kernel k is parametrized by c, θ_u and θ_v ; θ_u and θ_v may contain physical information w.r.t. u_0 and v_0.
  Example : initial conditions with compact support yield

$$k_u(x, x') = k_u^0(x, x') \mathbb{1}_{B_R(x_0,R)}(x) \mathbb{1}_{B(x_0,R)}(x') \tag{21}$$

  Thus, $(x_0, R) \in \theta_u$. Likewise for v_0 (We can also encode symmetries).
  $\rightarrow$ can be estimated with the marginal likelihood.

Numerical application

Restrictive framework

Expensive convolutions (4D) $\rightarrow$ assume radial symmetry (explicit convolutions)

- Solve numerically the wave equation with $v_0 = 0$ and

$$u_0(x) = A\mathbb{1}_{[R_1,R_2]}(|x - x_0^*|)\left(1 + \cos\left(\frac{2\pi(|x - x_0^*| - \frac{R_1+R_2}{2})}{R_2 - R_1}\right)\right).$$

- Generate a database : finite difference scheme in $[0,1]^3$ with scattered sensors (LHS).
  $B = \{u(x_i, t_j) + \epsilon_{ij}, 1 \leq i \leq N_C, 1 \leq j \leq N_T\}, N_C = 30,\ N_T = 75.$

Numerical application

Restrictive framework

Expensive convolutions (4D) $\rightarrow$ assume radial symmetry (explicit convolutions)

- Solve numerically the wave equation with $v_0 = 0$ and

$$u_0(x) = A\mathbb{1}_{[R_1,R_2]}(|x - x_0^*|)\left(1 + \cos\left(\frac{2\pi(|x - x_0^*| - \frac{R_1+R_2}{2})}{R_2 - R_1}\right)\right).$$

- Generate a database : finite difference scheme in $[0,1]^3$ with scattered sensors (LHS).
  $B = \{u(x_i, t_j) + \epsilon_{ij}, 1 \leq i \leq N_C, 1 \leq j \leq N_T\}, N_C = 30,\ N_T = 75.$
- Kriging with

$$k_u(x, x') = k_{5/2}(x - x') \times \mathbb{1}_{B_R(x_0,R)}(x)\mathbb{1}_{B(x_0,R)}(x').$$

Data visualisation

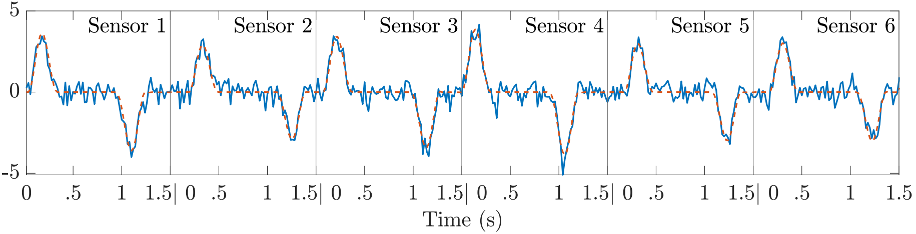


Figure 6 – Examples of captured signals. red : noiseless signals. Blue : noisy signals.

Estimation of physical parameters

N_{sensors}	3	5	10	15	20	25	30	Target
$\lvert\hat{x}_0 - x_0^*\rvert$	0.204	0.003	0.004	0.008	0.003	0.004	0.015	0
$\hat{R}$	0.386	0.432	0.462	0.431	0.414	0.471	0.452	0.25
$\lvert\hat{c} - c^*\rvert$	0.084	0.004	0.005	0.005	0.006	0.001	0.004	0
$\hat{\sigma}^2_{\mathrm{noise}}$	0.917	0.879	0.93	0.99	0.361	0.988	0.377	0.2025
$\hat{\ell}$	0.02	0.02	0.025	0.02	0.035	0.024	0.032	~ 0.05
$\hat{\sigma}^2$	2.367	3.513	4.903	3.168	4.446	4.619	4.79	Unknown
$e^{\mathrm{u}}_{1,\mathrm{rel}}$	1.275	0.157	0.128	0.168	0.11	0.103	0.248	0
$e^{\mathrm{u}}_{2,\mathrm{rel}}$	1.056	0.095	0.082	0.124	0.088	0.064	0.213	0
$e^{\mathrm{u}}_{\infty,\mathrm{rel}}$	1.037	0.132	0.128	0.198	0.136	0.101	0.321	0

Table 1 – Estimation of hyperparameters and relative errors

Reconstruction of initial condition

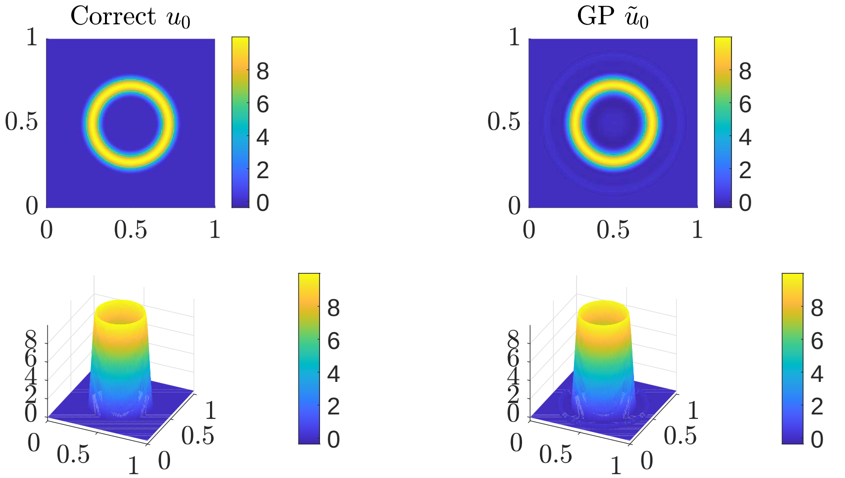


Figure 7 – True u_0 (left column) vs GPR u_0 (right column). 15 sensors were used. The images correspond to 3D slices at $z = 0.5$.

Outline of the presentation

Extension to non linear PDEs

- Non linear constraints on $k(z, \cdot)$: not realistic (+ GP interpretation not valid).
- Alternative : in Chen et al. [2021], the nonlinear PDE constraint in applied pointwise on $\tilde{m}$: modification of the RKHS optimization problem as

$$\inf_{v \in \mathcal{H}_k} \|v\|_{\mathcal{H}_k} \quad \text{s.c.} \quad \mathcal{N}(v(z_i), \nabla v(z_i), ...) = \ell_i \quad \forall i \in \{1, ..., n\}$$

  Generalizes an approach desribed in Wendland [2004].
- Coupling of this approach with strict linear constraints : Owhadi [2023b] (div/curl/periodicity).

Conclusion and perspectives

Overall conclusions :

- GPR : at the intersection of machine learning, statistical and Bayesian approaches and functional analysis.

Overall conclusions :

- GPR : at the intersection of machine learning, statistical and Bayesian approaches and functional analysis.

Some research perspectives :

- Insert the Sobolev regularity results in the analysis of GPR for PDEs, e.g. of Chen et al. [2021].
- Current research : draw links between numerical methods for PDEs (finite differences) and some GPR regimes.

Thank you for your attention !

Contact : henderso@insa-toulouse.fr,iain.pl.henderson@gmail.com

C. G. Albert and K. Rath. Gaussian process regression for data fulfilling linear differential equations with localized sources. *Entropy*, 22(2), 2020. ISSN 1099-4300. URL `https://www.mdpi.com/1099-4300/22/2/152`.

Yifan Chen, Bamdad Hosseini, Houman Owhadi, and Andrew M. Stuart. Solving and learning nonlinear PDEs with Gaussian processes. *Journal of Computational Physics*, 447 :110668, 2021. ISSN 0021-9991. doi : https://doi.org/10.1016/j.jcp.2021.110668. URL `https://www.sciencedirect.com/science/article/pii/S0021999121005635`.

D. Ginsbourger, O. Roustant, and N. Durrande. On degeneracy and invariances of random fields paths with applications in Gaussian process modelling. *Journal of Statistical Planning and Inference*, page 170 :117 – 128, 2016.

Iain H. Sobolev regularity of Gaussian random fields. working paper or preprint, October 2022. URL `https://hal.science/hal-03769576`.

References II

Iain H., Pascal Noble, and Olivier Roustant. Wave equation-tailored Gaussian process regression with applications to related inverse problems. working paper or preprint, January 2023. URL `https://hal.science/hal-03941939`.

Iain H., Pascal Noble, and Olivier Roustant. Characterization of the second order random fields subject to linear distributional PDE constraints. *Bernoulli*, 2023, to appear. URL `https://hal.science/hal-03770715`.

C. Jidling, N. Wahlström, A. Wills, and T. B. Schön. Linearly constrained Gaussian processes. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL `https://proceedings.neurips.cc/paper/2017/file/71ad16ad2c4d81f348082ff6c4b20768-Paper.pdf`.

C. Jidling, J. Hendriks, N. Wahlstrom, A. Gregg, T. Schon, C. Wensrich, and A. Wills. Probabilistic modelling and reconstruction of strain. *Nuclear Instruments & Methods in Physics Research Section B-beam Interactions With Materials and Atoms*, 436 :141–155, 2018.

M. Lange-Hegermann. Algorithmic linearly constrained Gaussian processes. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018. URL `https://proceedings.neurips.cc/paper/2018/file/68b1fbe7f16e4ae3024973f12f3cb313-Paper.pdf`.

A. F. López-Lopera, N. Durrande, and M. Álvarez. Physically-inspired Gaussian process models for post-transcriptional regulation in drosophila. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 18 :656–666, 2021.

References IV

M. Álvarez, D. Luengo, and N. D. Lawrence. Latent force models. In D. van Dyk and M. Welling, editors, *Proceedings of the Twelth International Conference on Artificial Intelligence and Statistics*, volume 5 of *Proceedings of Machine Learning Research*, pages 9–16, Hilton Clearwater Beach Resort, Clearwater Beach, Florida USA, 16–18 Apr 2009. PMLR. URL https://proceedings.mlr.press/v5/alvarez09a.html.

Fábio Macêdo Mendes and Edson Alves da Costa Júnior. Bayesian inference in the numerical solution of Laplace's equation. *AIP Conference Proceedings*, 1443(1) :72–79, 2012. doi : $10.1063/1.3703622$. URL https://aip.scitation.org/doi/abs/10.1063/1.3703622.

References V

Houman Owhadi. Do ideas have shape ? idea registration as the continuous limit of artificial neural networks. *Physica D : Nonlinear Phenomena*, 444 :133592, 2023a. ISSN 0167-2789. doi : https://doi.org/10.1016/j.physd.2022.133592. URL https://www.sciencedirect.com/science/article/pii/S0167278922002962.

Houman Owhadi. Gaussian process hydrodynamics, 2023b.

C. E. Rasmussen and C.K.I. Williams. *Gaussian Processes for Machine Learning*. the MIT Press, 2006. ISBN 026218253X. URL www.{G}aussianProcess.org/gpml.

M. Scheuerer and M. Schlather. Covariance models for divergence-free and curl-free random vector fields. *Stochastic Models*, 28 :433 – 451, 2012.

References VI

Michael Scheuerer. Regularity of the sample paths of a general second order random field. *Stochastic Processes and their Applications*, 120 (10) :1879–1897, 2010. ISSN 0304-4149. doi : https://doi.org/10.1016/j.spa.2010.05.009. URL https://www.sciencedirect.com/science/article/pii/S0304414910001328.

Ingo Steinwart. Convergence types and rates in generic Karhunen-Loeve expansions with applications to sample path properties. *Potential Anal.*, 51(3) :361–395, 2019.

N. Wahlstrom, M. Kok, T. B. Schön, and F. Gustafsson. Modeling magnetic fields using Gaussian processes. *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 3522–3526, 2013.

Holger Wendland. *Scattered data approximation*, volume 17. Cambridge university press, 2004.

Comparison with previous results

- Steinwart [2019] If $H_k \simeq H^t$ then $\mathbb{P}(U \in H^s) = 1$ if $t - s > d/2$. When $s \in \mathbb{N}$, reduces to our result as the imbedding $H^t \to H^s$ is Hilbert-Schmidt if $t - s > d/2$ (Maurey's theorem).
- Scheuerer [2010] : if for all $|\alpha| \leq m$ and $x \in \mathcal{D}$, $\partial^{\alpha,\alpha}k(x,)$ exists and $x \mapsto \partial^{\alpha,\alpha}k(x,x)$ is continuous and

$$\int_{\mathcal{D}} \partial^{\alpha,\alpha}k(x,x)dx < +\infty, \tag{22}$$

then $\mathbb{P}(U \in H^m) = 1$. We removed the continuity assumptions and added the Gaussianity assumption. We obtain a NSC.

Reproducing kernel Hilbert spaces

Let $k : \mathcal{D} \times \mathcal{D} \to \mathbb{R}$ a psotive semi-definite function. We define H_k as

$$H_k := \left\{ \sum_{i=1}^{+\infty} a_i k(z_i, \cdot) \text{ where } (a_i) \subset \mathbb{R}, (z_i) \subset \mathcal{D} \text{ and } \sum_{i,j=1}^{+\infty} a_i a_j k(z_i, z_j) < +\infty \right\}$$

endowed with the inner product

$$\left\langle \sum_{i=1}^{+\infty} a_i k(x_i, \cdot), \sum_{j=1}^{+\infty} b_j k(y_j, \cdot) \right\rangle := \sum_{i,j=1}^{+\infty} a_i b_j k(x_i, y_j).$$

Reproducing kernel Hilbert spaces

Let $k : \mathcal{D} \times \mathcal{D} \to \mathbb{R}$ a psotive semi-definite function. We define H_k as

$$H_k := \left\{ \sum_{i=1}^{+\infty} a_i k(z_i, \cdot) \text{ where } (a_i) \subset \mathbb{R}, (z_i) \subset \mathcal{D} \text{ and } \sum_{i,j=1}^{+\infty} a_i a_j k(z_i, z_j) < +\infty \right\}$$

endowed with the inner product

$$\left\langle \sum_{i=1}^{+\infty} a_i k(x_i, \cdot), \sum_{j=1}^{+\infty} b_j k(y_j, \cdot) \right\rangle := \sum_{i,j=1}^{+\infty} a_i b_j k(x_i, y_j).$$

The function k verifies the reproducing properties

$$\langle k(z, \cdot), k(z', \cdot) \rangle = k(z, z') \text{ and } \langle k(z, \cdot), f \rangle = f(z) \ \forall f \in H_k$$

Reproducing kernel Hilbert spaces

Let $k : \mathcal{D} \times \mathcal{D} \to \mathbb{R}$ a psotive semi-definite function. We define H_k as

$$H_k := \left\{ \sum_{i=1}^{+\infty} a_i k(z_i, \cdot) \text{ where } (a_i) \subset \mathbb{R}, (z_i) \subset \mathcal{D} \text{ and } \sum_{i,j=1}^{+\infty} a_i a_j k(z_i, z_j) < +\infty \right\}$$

endowed with the inner product

$$\left\langle \sum_{i=1}^{+\infty} a_i k(x_i, \cdot), \sum_{j=1}^{+\infty} b_j k(y_j, \cdot) \right\rangle := \sum_{i,j=1}^{+\infty} a_i b_j k(x_i, y_j).$$

The function k verifies the reproducing properties

$$\langle k(z, \cdot), k(z', \cdot) \rangle = k(z, z') \text{ and } \langle k(z, \cdot), f \rangle = f(z) \ \forall f \in H_k$$

Details on F_t and $\dot{F}_t$

$\longrightarrow F_t = \sigma_{ct}/4\pi c^2 t$ means that

Details on F_t and $\dot{F}_t$

$\longrightarrow F_t = \sigma_{ct}/4\pi c^2 t$ means that

$$\int_{\mathbb{R}^3} f(x)F_t(dx) = \frac{t}{4\pi}\int_0^{2\pi}\int_0^{\pi} f(ct\gamma)\sin\theta d\theta d\varphi = \frac{t}{4\pi}\int_{S(0,1)} f(ct\gamma)d\Omega$$

where γ is the unit length vector $\gamma = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)^T$.

Details on F_t and $\dot{F}_t$

$\longrightarrow$ $F_t = \sigma_{ct}/4\pi c^2 t$ means that

$$\int_{\mathbb{R}^3} f(x)F_t(dx) = \frac{t}{4\pi}\int_0^{2\pi}\int_0^{\pi} f(ct\gamma)\sin\theta d\theta d\varphi = \frac{t}{4\pi}\int_{S(0,1)} f(ct\gamma)d\Omega$$

where γ is the unit length vector $\gamma = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)^T$.

$\longrightarrow$ Convolution between functions and measures :

$$(f * g)(x) = \int_{\mathbb{R}^3} g(x-y)f(y)dy \qquad (\mu * g)(x) = \int_{\mathbb{R}^3} g(x-y)\mu(dy)$$

Details on F_t and $\dot{F}_t$

$\longrightarrow$ $F_t = \sigma_{ct}/4\pi c^2 t$ means that

$$\int_{\mathbb{R}^3} f(x) F_t(dx) = \frac{t}{4\pi} \int_0^{2\pi} \int_0^{\pi} f(ct\gamma) \sin\theta d\theta d\varphi = \frac{t}{4\pi} \int_{S(0,1)} f(ct\gamma) d\Omega$$

where γ is the unit length vector $\gamma = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)^T$.

$\longrightarrow$ Convolution between functions and measures :

$$(f * g)(x) = \int_{\mathbb{R}^3} g(x-y) f(y) dy \qquad (\mu * g)(x) = \int_{\mathbb{R}^3} g(x-y) \mu(dy)$$

$\longrightarrow$ $\dot{F}_t = \partial_t F_t$ means that

$$\begin{aligned} \langle \dot{F}_t, f \rangle &= \partial_t \int f(x) dF_t(x) \\ &= \frac{1}{4\pi} \int_{S(0,1)} f(ct\gamma) d\Omega + \frac{c}{4\pi} \int_{S(0,1)} \nabla f(ct\gamma) \cdot \gamma d\Omega \end{aligned}$$

Radial symmetry formulas

$$[(F_t \otimes F_{t'}) * k_v](x, x') = \frac{\mathrm{sgn}(tt')}{16c^2 rr'} \sum_{\varepsilon,\varepsilon' \in \{-1,1\}} \varepsilon\varepsilon' K_{\mathrm{v}}\big((r + \varepsilon ct)^2, (r' + \varepsilon' c|t'|)^2\big)$$

$$[(\dot{F}_t \otimes \dot{F}_{t'}) * k_u](x, x') = \frac{1}{4rr'} \sum_{\varepsilon,\varepsilon' \in \{-1,1\}} (r + \varepsilon ct)(r' + \varepsilon' c|t'|) k_{\mathrm{u}}\big((r + \varepsilon ct)^2, (r' + \varepsilon' c|t'|)^2\big)$$

A word on GPR and neural networks

- Certain Gaussian processes as limits of one layer neural netwokrs with infinitely many neurons (Rasmussen and Williams [2006], Section 4.2.3).

- Regression with neural networks as GPR with a kernel learnt from data (Owhadi [2023a] ; Mallat, collège de France).

- GPR : "only" current contender (physics informed) neural networks, cf Chen et al. [2021] for a discussion.

Point source localization

Assume that $u_0 \equiv 0$ and that v_0 is almost a point source at x_0^* : we use the kernels

$$k_v^R(x,x') = k_v(x,x') \frac{\mathbb{1}_{B(x_0,R)}(x)}{4\pi R^3/3} \frac{\mathbb{1}_{B(x_0,R)}(x')}{4\pi R^3/3} \quad (23)$$

$$k((x,t),(x',t')) = [(F_t \otimes F_{t'}) * k_v^R](x,x') \quad (24)$$

with $R \ll 1$. Hyperparameters of k : (θ_v, x_0, R, c) We fix θ_v, R and c to the "right values" : $\mathcal{L}(\theta) = \mathcal{L}(x_0)$.

Question : behaviour of $x_0 \mapsto \mathcal{L}(x_0)$?

Minimize negative marginal likelihood ≡ GPS localization

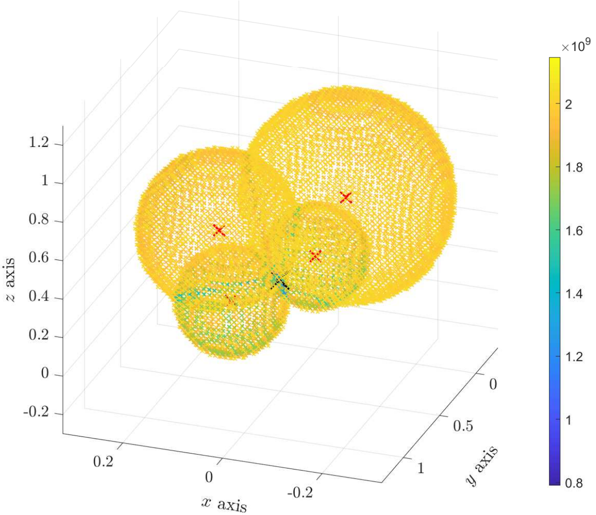


Figure : negative log marginal likelihood.

Displayed values : less than 2.035×10^9.

$\times$: sensor locations.

$\times$: source location.

See H. et al. [2023] for study/proofs.