Finite-difference least square method for solving Hamilton-Jacobi equations using neural networks

Carlos Esteve-Yagüe

Departmento de Matemáticas
Universidad de Alicante

c.esteve@ua.es

joint work with **Richard Tsai**, University of Texas at Austin
and **Alex Massucco**, University of Cambridge

Control, Problemas Inversos y Aprendizaje Automático, Almagro
2 of December, 2024

Universitat d'Alacant
Universidad de Alicante

Boundary value problem

$$\begin{cases} F(x, u, \nabla u, D^2 u) = 0 & \text{on } \Omega \subset \mathbb{R}^d \\ u(x) = g(x) & \text{on } \partial\Omega. \end{cases} \qquad \text{(BVP)}$$

Goal: Approximate the solution of (BVP) by means of a Neural Network.

For a hypothesis set $\mathcal{F} \subset C(\Omega)$, we consider the problem

$$\min_{u \in \mathcal{F}} \mathcal{J}(u), \qquad (1)$$

for some loss functional $\mathcal{J}(\cdot) : C(\Omega) \to \mathbb{R}$.

Questions:

- What hypothesis set $\mathcal{F}$? (NN architecture)
- What functional $\mathcal{J}(\cdot)$?
- What optimisation algorithm? (SGD or a variant)
- What type of solution are we looking for?

Boundary value problem

$$\begin{cases} F(x, u, \nabla u, D^2 u) = 0 & \text{on } \Omega \subset \mathbb{R}^d \\ u(x) = g(x) & \text{on } \partial\Omega. \end{cases} \tag{BVP}$$

Goal: Approximate the solution of (BVP) by means of a Neural Network.

For a hypothesis set $\mathcal{F} \subset C(\Omega)$, we consider the problem

$$\min_{u \in \mathcal{F}} \mathcal{J}(u), \tag{1}$$

for some loss functional $\mathcal{J}(\cdot) : C(\Omega) \to \mathbb{R}$.

Questions:

- What hypothesis set $\mathcal{F}$? (NN architecture)
- What functional $\mathcal{J}(\cdot)$?
- What optimisation algorithm? (SGD or a variant)
- What type of solution are we looking for?

Boundary value problem

$$\begin{cases} F(x, u, \nabla u, D^2 u) = 0 & \text{on } \Omega \subset \mathbb{R}^d \\ u(x) = g(x) & \text{on } \partial\Omega. \end{cases} \tag{BVP}$$

Goal: Approximate the solution of (BVP) by means of a Neural Network.

For a hypothesis set $\mathcal{F} \subset C(\Omega)$, we consider the problem

$$\min_{u \in \mathcal{F}} \mathcal{J}(u), \tag{1}$$

for some loss functional $\mathcal{J}(\cdot) : C(\Omega) \to \mathbb{R}$.

Questions:

- What hypothesis set $\mathcal{F}$? (NN architecture)
- What functional $\mathcal{J}(\cdot)$?
- What optimisation algorithm? (SGD or a variant)
- What type of solution are we looking for?

Boundary value problem

$$\begin{cases} F(x, u, \nabla u, D^2 u) = 0 & \text{on } \Omega \\ u(x) = g(x) & \text{on } \partial\Omega. \end{cases} \tag{BVP}$$

Hypothesis set: a parametric class of continuous functions

$$\mathcal{F} := \{\Phi(\cdot; \theta) : \Omega \to \mathbb{R} \quad : \quad \theta \in \mathbb{R}^p\}.$$

Universal approximation of NNs $\Longrightarrow \overline{\mathcal{F}} \approx C(\Omega)$.

Optimisation method: gradient descent

$$\theta_{t+1} = \theta_t - \gamma \nabla_\theta \mathcal{J}\left(\Phi(\cdot; \theta_t)\right), \tag{GD}$$

Loss functional: we want the following properties

1. Minimisers of $\mathcal{J}(\cdot)$ approximate the solution of (BVP).
2. Stationary points of (GD) are global minimisers.

Boundary value problem

$$\begin{cases} F(x, u, \nabla u, D^2 u) = 0 & \text{on } \Omega \\ u(x) = g(x) & \text{on } \partial\Omega. \end{cases} \tag{BVP}$$

Hypothesis set: a parametric class of continuous functions

$$\mathcal{F} := \{\Phi(\cdot;\theta) : \Omega \to \mathbb{R} \quad : \quad \theta \in \mathbb{R}^p\}.$$

Universal approximation of NNs $\Longrightarrow \overline{\mathcal{F}} \approx C(\Omega)$.

Optimisation method: gradient descent

$$\theta_{t+1} = \theta_t - \gamma \nabla_\theta \mathcal{J}\left(\Phi(\cdot;\theta_t)\right), \tag{GD}$$

Loss functional: we want the following properties

1. Minimisers of $\mathcal{J}(\cdot)$ approximate the solution of (BVP).
2. Stationary points of (GD) are global minimisers.

Let us look at the equation

$$\nabla_\theta \mathcal{J}\left(\Phi(\cdot;\theta)\right) = \mathcal{J}'(\Phi(\cdot;\theta)) \cdot \nabla_\theta \Phi(\cdot;\theta) = 0, \qquad \theta \in \mathbb{R}^p.$$

where $\mathcal{J}'(u) \in C(\Omega)^*$ is the Fréchet derivative of $\mathcal{J}(\cdot)$ at u.

Stationary points

θ is a critical point if and only if one of the following holds:

- $\mathcal{J}'\left(\Phi(\cdot;\theta)\right) = 0$
- $\mathcal{J}'\left(\Phi(\cdot;\theta)\right) \neq 0$ and all the components of $\nabla_\theta \Phi(\cdot;\theta)$ are in the kernel of $\mathcal{J}'\left(\Phi(\cdot;\theta)\right)$.

Let us look at the equation

$$\nabla_\theta \mathcal{J}\left(\Phi(\cdot;\theta)\right) = \mathcal{J}'(\Phi(\cdot;\theta)) \cdot \nabla_\theta \Phi(\cdot;\theta) = 0, \qquad \theta \in \mathbb{R}^p.$$

where $\mathcal{J}'(u) \in C(\Omega)^*$ is the Fréchet derivative of $\mathcal{J}(\cdot)$ at u.

Stationary points

θ is a critical point if and only if one of the following holds:

- $\mathcal{J}'\left(\Phi(\cdot;\theta)\right) = 0$
- $\mathcal{J}'\left(\Phi(\cdot;\theta)\right) \neq 0$ and all the components of $\nabla_\theta \Phi(\cdot;\theta)$ are in the kernel of $\mathcal{J}'\left(\Phi(\cdot;\theta)\right)$.

- The specific form of $\nabla_\theta \Phi(\cdot;\theta)$ depends on the NN architecture.
- The form of $\mathcal{J}'\left(\Phi(\cdot;\theta)\right)$ depends on the choice of the functional.

Let us look at the equation

$$\nabla_\theta \mathcal{J}\left(\Phi(\cdot;\theta)\right) = \mathcal{J}'(\Phi(\cdot;\theta)) \cdot \nabla_\theta \Phi(\cdot;\theta) = 0, \qquad \theta \in \mathbb{R}^p.$$

where $\mathcal{J}'(u) \in C(\Omega)^*$ is the Fréchet derivative of $\mathcal{J}(\cdot)$ at u.

Stationary points

θ is a critical point if and only if one of the following holds:

- $\mathcal{J}'\left(\Phi(\cdot;\theta)\right) = 0$
- $\mathcal{J}'\left(\Phi(\cdot;\theta)\right) \neq 0$ and all the components of $\nabla_\theta \Phi(\cdot;\theta)$ are in the kernel of $\mathcal{J}'\left(\Phi(\cdot;\theta)\right)$.

Goal of the talk: construct $\mathcal{J} : C(\Omega) \to \mathbb{R}^+$ such that

1. any minimiser of $\mathcal{J}(u)$ approximates the solution of (BVP).
2. any solution of $\mathcal{J}'(u) = 0$ approximates the solution of (BVP).

We will only consider Hamilton-Jacobi PDEs.

Boundary value problem

$$\begin{cases} H(x, \nabla u) = 0 & \text{on } \Omega \subset \mathbb{R}^d \\ u(x) = g(x) & \text{on } \partial\Omega. \end{cases} \tag{HJ}$$

for some (non-linear) Hamiltonian $H(x, p) : \Omega \times \mathbb{R}^d \to \mathbb{R}$.

Physics Informed Neural Networks

$$\mathcal{J}(u) := \int_{\Omega} (H(x, \nabla u(x)))^2 \, dx + \int_{\partial\Omega} (u(x) - g(x))^2 \, dx$$

- Which continuous functions minimise $\mathcal{J}(u)$?
- What are the critical points? i.e. solutions to $\mathcal{J}'(u) = 0$?
- Is $\mathcal{J}(\cdot)$ even Fréchet differentiable in $C(\Omega)$?

Boundary value problem

$$\begin{cases} H(x, \nabla u) = 0 & \text{on } \Omega \subset \mathbb{R}^d \\ u(x) = g(x) & \text{on } \partial\Omega. \end{cases} \tag{HJ}$$

for some (non-linear) Hamiltonian $H(x, p) : \Omega \times \mathbb{R}^d \to \mathbb{R}$.

Physics Informed Neural Networks

$$\mathcal{J}(u) := \int_\Omega (H(x, \nabla u(x)))^2 \, dx + \int_{\partial\Omega} (u(x) - g(x))^2 \, dx$$

- Which continuous functions minimise $\mathcal{J}(u)$?
- What are the critical points? i.e. solutions to $\mathcal{J}'(u) = 0$?
- Is $\mathcal{J}(\cdot)$ even Fréchet differentiable in $C(\Omega)$?

Eikonal equation

$$\begin{cases} (\partial_x u)^2 - 1 = 0 & \text{in } (0,1) \\ u(0) = u(1) = 0 \end{cases} \tag{HJ}$$

$$\mathcal{J}(u) := \int_0^1 \left((\partial_x u)^2 - 1 \right)^2 dx + u(0)^2 + u(1)^2$$

Global minimisers:

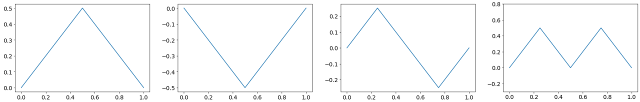
Hamilton-Jacobi equations

Eikonal equation

$$\begin{cases} (\partial_x u)^2 - 1 = 0 & \text{in } (0,1) \\ u(0) = u(1) = 0 \end{cases} \qquad \text{(HJ)}$$

$$\mathcal{J}(u) := \int_0^1 \left((\partial_x u)^2 - 1 \right)^2 dx + u(0)^2 + u(1)^2$$

Global minimisers:



...and many more

1. We are interested in the viscosity solution (the first plot above)
2. We need to add some regularity to the functional.
   something like $\mathcal{J}(u) + \gamma \mathcal{R}(u)$?
3. The regularisation has to single out the first plot from the second.

Boundary value problem

$$\begin{cases} H(x, \nabla u) = 0 & \text{on } \Omega \subset \mathbb{R}^d \\ u(x) = g(x) & \text{on } \partial\Omega. \end{cases} \qquad \text{(HJ)}$$

for some (non-linear) Hamiltonian $H(x, p) : \Omega \times \mathbb{R}^d \to \mathbb{R}$.

Viscosity solution

$$\begin{cases} H(x, \nabla u_\varepsilon) - \varepsilon \Delta u_\varepsilon = 0 & \text{on } \Omega \subset \mathbb{R}^d \\ u_\varepsilon(x) = g(x) & \text{on } \partial\Omega. \end{cases} \qquad \text{(VHJ)}$$

has a unique classical solution $u_\varepsilon \in C^\infty(\Omega)$, for each $\varepsilon \neq 0$.

The viscosity solution of (HJ) is the point wise limit

$$u(x) = \lim_{\varepsilon \to 0^+} u_\varepsilon(x), \qquad \forall x \in \Omega.$$

Main idea

Numerical methods (such as F.D.) introduce what is known as **numerical diffusion**.

Why don't we replace $H(x, \nabla u(x))$ in

$$\mathcal{J}(u) := \int_\Omega (H(x, \nabla u(x)))^2 \, dx + \int_{\partial\Omega} (u(x) - g(x))^2 \, dx$$

by a numerical Hamiltonian?

$$\widehat{H}(x, D_\delta^+ u(x), D_\delta^- u(x))$$

where

$$D_\delta^+ u(x) = \frac{u(x + \delta I) - u(x)}{\delta} \quad \text{and} \quad D_\delta^- u(x) = \frac{u(x) - u(x - \delta I)}{\delta}.$$

Theorem (Crandall-Lions, 1984 and Barles-Souganidis, 1991)
For any $\delta > 0$, let $\Omega_\delta := \delta\mathbb{Z}^d \cap \Omega$.
If $\widehat{H}$ is **consistent and monotone**, then any solution u_δ to the discretised problem

$$\widehat{H}(x, D_\delta^+ u(x), D_\delta^- u(x)) = 0 \qquad x \in \Omega_\delta \tag{2}$$

converges, as $\delta \to 0$ to a viscosity solution of

$$H(x, \nabla u(x)) = 0 \qquad x \in \Omega.$$

Goal

Construct a **consistent and monotone** numerical Hamiltonian such that any critical point of

$$\widehat{\mathcal{R}}(u) := \sum_{x\in\Omega_\delta} \left[\widehat{H}(x, D_\delta^+ u(x), D_\delta^- u(x))\right]^2$$

is a solution of (2).

Loss functional

$$\mathcal{J}(u) := \widehat{\mathcal{R}}(u) + \int_{\partial\Omega} (u(x) - g(x))^2\, dx$$

Theorem (Crandall-Lions, 1984 and Barles-Souganidis, 1991)
For any $\delta > 0$, let $\Omega_\delta := \delta\mathbb{Z}^d \cap \Omega$.
If $\widehat{H}$ is **consistent and monotone**, then any solution u_δ to the discretised problem

$$\widehat{H}(x, D_\delta^+ u(x), D_\delta^- u(x)) = 0 \qquad x \in \Omega_\delta \tag{2}$$

converges, as $\delta \to 0$ to a viscosity solution of

$$H(x, \nabla u(x)) = 0 \qquad x \in \Omega.$$

Goal

Construct a **consistent and monotone** numerical Hamiltonian such that any critical point of

$$\widehat{\mathcal{R}}(u) := \sum_{x \in \Omega_\delta} \left[\widehat{H}(x, D_\delta^+ u(x), D_\delta^- u(x))\right]^2$$

is a solution of (2).

Loss functional

$$\mathcal{J}(u) := \widehat{\mathcal{R}}(u) + \int_{\partial\Omega} (u(x) - g(x))^2 \, dx$$

Theorem (Crandall-Lions, 1984 and Barles-Souganidis, 1991)
For any $\delta > 0$, let $\Omega_\delta := \delta\mathbb{Z}^d \cap \Omega$.
If $\widehat{H}$ is **consistent and monotone**, then any solution u_δ to the discretised problem

$$\widehat{H}(x, D_\delta^+ u(x), D_\delta^- u(x)) = 0 \qquad x \in \Omega_\delta \tag{2}$$

converges, as $\delta \to 0$ to a viscosity solution of

$$H(x, \nabla u(x)) = 0 \qquad x \in \Omega.$$

Goal

Construct a **consistent and monotone** numerical Hamiltonian such that any critical point of

$$\widehat{\mathcal{R}}(u) := \sum_{x \in \Omega_\delta} \left[\widehat{H}(x, D_\delta^+ u(x), D_\delta^- u(x))\right]^2$$

is a solution of (2).

Loss functional

$$\mathcal{J}(u) := \widehat{\mathcal{R}}(u) + \int_{\partial\Omega} (u(x) - g(x))^2 \, dx$$

Critical points

For $\delta > 0$, let $\Omega_\delta := \delta\mathbb{Z}^d \cap \Omega$.
For any $u \in C(\Omega)$, let $U := u|_{\Omega_\delta} = \{u(x) \ : \ \beta \in \mathcal{I}\}$.

Let us define the function $F : \mathbb{R}^{|\mathcal{I}|} \to \mathbb{R}$ given by

$$\begin{aligned} F(U) &:= \sum_{\beta\in\mathcal{I}} \left[\widehat{H}(x_\beta, D_\delta^+ U_\beta, D_\delta^- U_\beta)\right]^2 \\ &= \sum_{x\in\Omega_\delta} \left[\widehat{H}(x, D_\delta^+ u(x), D_\delta^- u(x))\right]^2 = \widehat{\mathcal{R}}(u) \end{aligned}$$

Euler-Lagrange equation

Critical points

For $\delta > 0$, let $\Omega_\delta := \delta\mathbb{Z}^d \cap \Omega$.
For any $u \in C(\Omega)$, let $U := u|_{\Omega_\delta} = \{u(x) \ : \ \beta \in \mathcal{I}\}$.

Let us define the function $F : \mathbb{R}^{|\mathcal{I}|} \to \mathbb{R}$ given by

$$\begin{aligned} F(U) &:= \sum_{\beta\in\mathcal{I}} \left[\widehat{H}(x_\beta, D_\delta^+ U_\beta, D_\delta^- U_\beta)\right]^2 \\ &= \sum_{x\in\Omega_\delta} \left[\widehat{H}(x, D_\delta^+ u(x), D_\delta^- u(x))\right]^2 = \widehat{\mathcal{R}}(u) \end{aligned}$$

Euler-Lagrange equation

u is a critical point of $\widehat{\mathcal{R}}(\cdot)$ if and only if $\nabla F(U) = 0$.

Defining $w(x) = \widehat{H}(x, D_\delta^+ u(x), D_\delta^- u(x))$ and $W = w|_{\Omega_\delta}$, we have

$$\nabla F(U) = A_\delta(U)W,$$

where $A_\delta(U)$ is a linear operator in $\mathbb{R}^{|\mathcal{I}|}$.

Critical points

For $\delta > 0$, let $\Omega_\delta := \delta\mathbb{Z}^d \cap \Omega$.
For any $u \in C(\Omega)$, let $U := u|_{\Omega_\delta} = \{u(x) \ : \ \beta \in \mathcal{I}\}$.

Let us define the function $F : \mathbb{R}^{|\mathcal{I}|} \to \mathbb{R}$ given by

$$\begin{aligned} F(U) &:= \sum_{\beta\in\mathcal{I}} \left[\widehat{H}(x_\beta, D_\delta^+ U_\beta, D_\delta^- U_\beta)\right]^2 \\ &= \sum_{x\in\Omega_\delta} \left[\widehat{H}(x, D_\delta^+ u(x), D_\delta^- u(x))\right]^2 = \widehat{\mathcal{R}}(u) \end{aligned}$$

Euler-Lagrange equation

Goal: construct $\widehat{H}$ consistent and monotone and such that $A_\delta(U)$ is invertible for any grid function U on Ω_δ.

So $\nabla F(U) = A_\delta(U)W = 0$ implies $W = 0$, and then

$$\widehat{H}(x, D_\delta^+ u(x), D_\delta^- u(x)) = 0 \qquad \forall x \in \Omega_\delta.$$

For $\alpha > 0$, we consider a Lax-Friedrichs numerical Hamiltonian

$$\widehat{H}_\alpha(x, D_\delta^+ u(x), D_\delta^- u(x)) := H\left(x, \frac{D_\delta^+ u(x) + D_\delta^- u(x)}{2}\right) - \alpha \frac{D_\delta^+ u(x) - D_\delta^- u(x)}{2}$$

Known properties:

- $\widehat{H}_\alpha$ is consistent with H for all $\alpha \in \mathbb{R}$.
- For any $L > 0$, if

$$\alpha \geq C_H(L) := \max_{\substack{\|p\| \leq L \\ x \in \overline{\Omega}}} \|\nabla_p H(x,p)\|,$$

then $\widehat{H}_\alpha$ is monotone at each function u with Lipschitz constant L.

Euler-Lagrange equation:

The equation $\nabla F(U) = 0$ associated to $\widehat{H}_\alpha$ can be written as

$$-(A_\delta(U) + \alpha \Delta_\delta) W = 0,$$

- W is the grid function associated to $w(x) = \widehat{H}_\alpha(x, D_\delta^+ u(x), D_\delta^- u(x))$.
- $A_\delta(U)$ is a linear operator that can be computed in terms of $\nabla_p H(x,p)$.
- Δ_δ is the discretised Laplace operator associated to the grid Ω_δ.

For $\alpha > 0$, we consider a Lax-Friedrichs numerical Hamiltonian

$$\widehat{H}_\alpha(x, D_\delta^+ u(x), D_\delta^- u(x)) := H\left(x, \frac{D_\delta^+ u(x) + D_\delta^- u(x)}{2}\right) - \alpha \frac{D_\delta^+ u(x) - D_\delta^- u(x)}{2}$$

Known properties:

- $\widehat{H}_\alpha$ is consistent with H for all $\alpha \in \mathbb{R}$.
- For any $L > 0$, if

$$\alpha \geq C_H(L) := \max_{\substack{\|p\| \leq L \\ x \in \overline{\Omega}}} \|\nabla_p H(x, p)\|,$$

  then $\widehat{H}_\alpha$ is monotone at each function u with Lipschitz constant L.

Euler-Lagrange equation:

The equation $\nabla F(U) = 0$ associated to $\widehat{H}_\alpha$ can be written as

$$-(A_\delta(U) + \alpha \Delta_\delta) W = 0,$$

- W is the grid function associated to $w(x) = \widehat{H}_\alpha(x, D_\delta^+ u(x), D_\delta^- u(x))$.
- $A_\delta(U)$ is a linear operator that can be computed in terms of $\nabla_p H(x, p)$.
- Δ_δ is the discretised Laplace operator associated to the grid Ω_δ.

For $\alpha > 0$, we consider a Lax-Friedrichs numerical Hamiltonian

$$\widehat{H}_\alpha(x, D_\delta^+ u(x), D_\delta^- u(x)) := H\left(x, \frac{D_\delta^+ u(x) + D_\delta^- u(x)}{2}\right) - \alpha \frac{D_\delta^+ u(x) - D_\delta^- u(x)}{2}$$

and the functional

$$\widehat{\mathcal{R}}(u) := \sum_{x \in \Omega_\delta} \left[\widehat{H}_\alpha(x, D_\delta^+ u(x), D_\delta^- u(x))\right]^2$$

Theorem

Let $u \in C(\Omega)$ be a critical point of $\widehat{\mathcal{R}}(\cdot)$ with Lipschitz constant $L > 0$. If

$$\alpha \frac{\lambda_1(\Omega_\delta)}{2d} > \max\left\{\|\nabla_p H(x,p)\| \ : \quad \|p\| \leq L, \ x \in \overline{\Omega}\right\}$$

then $\widehat{H}_\alpha(x, D_\delta^+ u(x), D_\delta^- u(x)) = 0$ for all $x \in \Omega_\delta$.

Remark: For regular domains $\Omega \subset \mathbb{R}^d$, we have $\lambda_1(\Omega_\delta) = O(d\delta)$.

$$\delta \mapsto \lambda_1(\Omega_\delta) \qquad \text{is increasing.}$$

Conclusion: Given Ω, $H(x,p)$ and $L > 0$, if we take the hyperparameters $\alpha > 0$ and $\delta > 0$ big enough

$$\alpha \frac{\lambda_1(\Omega_\delta)}{2d} > \max\left\{\|\nabla_p H(x,p)\| \ : \quad \|p\| \leq L, \ x \in \overline{\Omega}\right\}$$

then any critical point of $\widehat{\mathcal{R}}(\cdot)$ with Lipschitz constant $\leq L$ is a global minimiser and, therefore, approximates a viscosity solution.

Remarks:

- **Local result:** there might be other critical points with Lipschitz constant bigger than L.
- **Optimisation method:** SGD

$$\theta_{t+1} := \theta_t - \gamma \sum_{x \in \mathcal{X}_t} \nabla_\theta \left[\widehat{H}_\alpha \left(x, D_\delta^+ \Phi(x;\theta_t), D_\delta^- \Phi(x,\theta_t)\right)\right]^2,$$

where $\mathcal{X}_t \in \Omega_\delta^N$ is an i.i.d. sampling (mini-batch).

- **No fixed grid:** The NN is defined in the entire domain Ω, so we can vary the grid (e.g. reducing δ).

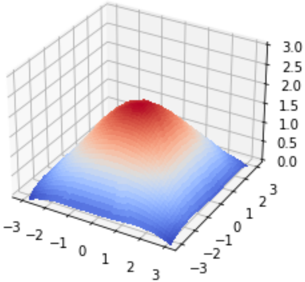
Numerical experiments

We consider the 2D Eikonal equation in $\Omega := (-3,3)^2$

$$\begin{cases} \|\nabla u\|^2 = 1 & \text{in } \Omega \\ u(x) = 0 & \text{on } \partial\Omega \end{cases}$$

The solution is the distance function to the boundary.

Training
We choose α and δ big enough as per Theorem 1



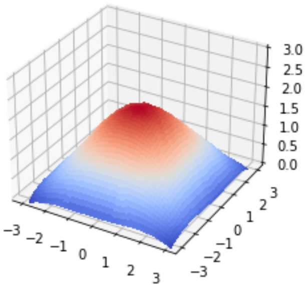
Good! ... but we regularized too much!

We consider the 2D Eikonal equation in $\Omega := (-3,3)^2$

$$\begin{cases} \|\nabla u\|^2 = 1 & \text{in } \Omega \\ u(x) = 0 & \text{on } \partial\Omega \end{cases}$$

The solution is the distance function to the boundary.

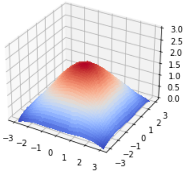
Training
We choose α and δ big enough as per Theorem 1



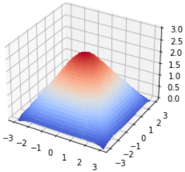
Good! ... but we regularized too much!

Training

$\alpha = 2.5$
$\delta = 0.75$



$\alpha = 2$
$\delta = 0.5$

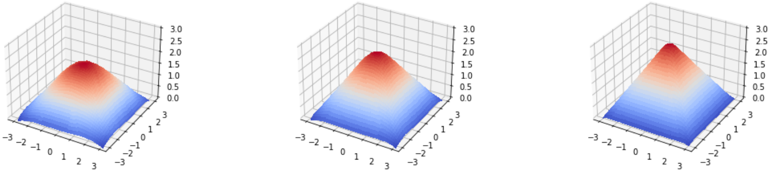


Training

$\alpha = 2.5$
$\delta = 0.75$

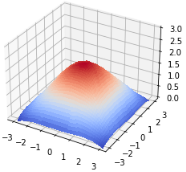
$\alpha = 2$
$\delta = 0.5$

$\alpha = 1.5$
$\delta = 0.3$

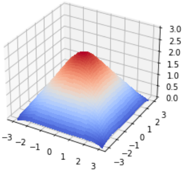


Training

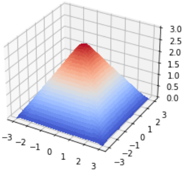
$\alpha = 2.5$
$\delta = 0.75$



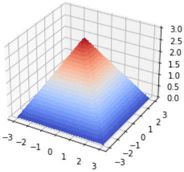
$\alpha = 2$
$\delta = 0.5$



$\alpha = 1.5$
$\delta = 0.3$



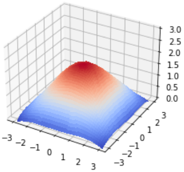
$\alpha = 1$
$\delta = 0.1$



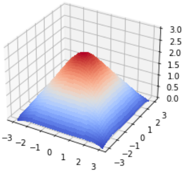
Example

Training

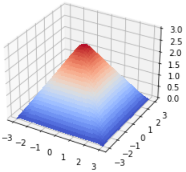
$\alpha = 2.5$
$\delta = 0.75$



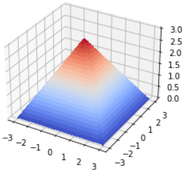
$\alpha = 2$
$\delta = 0.5$



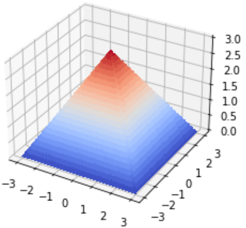
$\alpha = 1.5$
$\delta = 0.3$



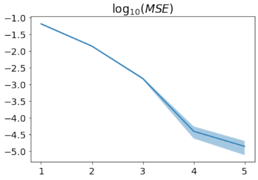
$\alpha = 1$
$\delta = 0.1$

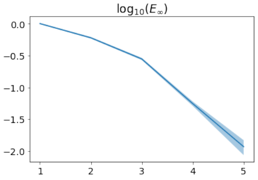


$\alpha = 0.5$
$\delta = 0.05$

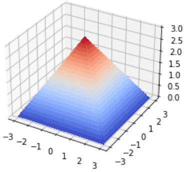


Error



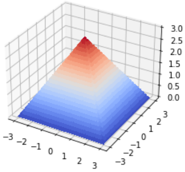


Let us start with $\Phi(x, \theta^*)$ from the previous example



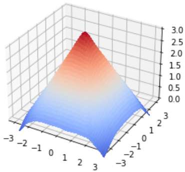
Question: can we recover the negative viscosity solution?

Let us start with $\Phi(x, \theta^*)$ from the previous example

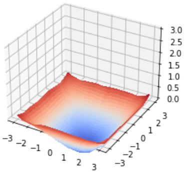


Question: can we recover the negative viscosity solution?

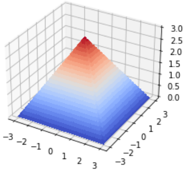
$\alpha = -2.5$
$\delta = 0.1$



$\alpha = -2.5$
$\delta = 0.75$

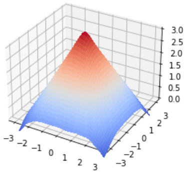


Let us start with $\Phi(x, \theta^*)$ from the previous example

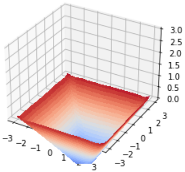


Question: can we recover the negative viscosity solution?

$\alpha = -2.5$
$\delta = 0.1$

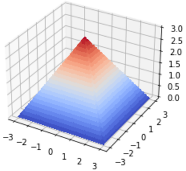


$\alpha = -1.5$
$\delta = 0.3$



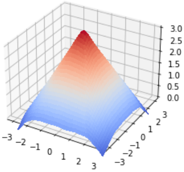
Example

Let us start with $\Phi(x, \theta^*)$ from the previous example

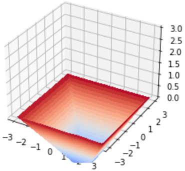


Question: can we recover the negative viscosity solution?

$\alpha = -2.5$
$\delta = 0.1$



$\alpha = -1$
$\delta = 0.1$

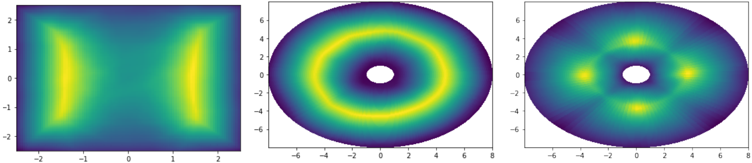


Non-trivial examples

We can consider other Eikonal equations in any domain $\Omega \subset \mathbb{R}^d$

$$\begin{cases} \|\nabla u\|^2 = f(x) & \text{in } \Omega \\ u(x) = g(x) & \text{on } \partial\Omega \end{cases}$$

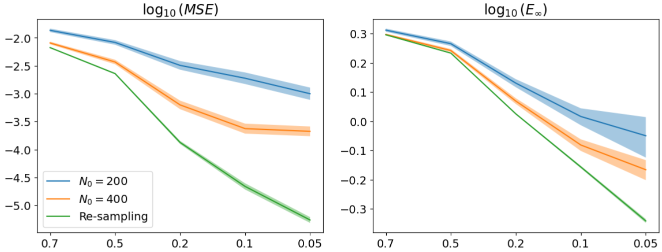
The solution is the distance function to the boundary in a non-homogeneous domain, determined by $f(x)$.



Question: how many collocation points are enough? (related to generalisation properties of the NN)

Two main observations:

1. Taking δ large is more data efficient.
2. Re-sampling the collocation points ate every iteration improves generalisation.

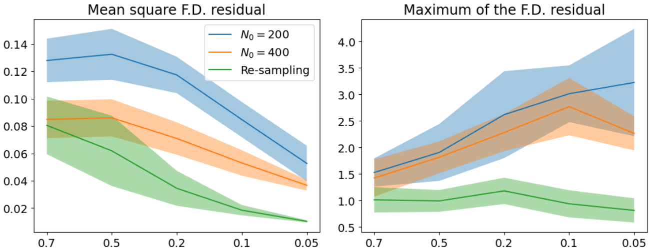


MSE and L^∞-error with respect to ground truth solution for eikonal equation in a 5D ball.

Question: how many collocation points are enough? (related to generalisation properties of the NN)

Two main observations:

1. Taking δ large is more data efficient.
2. Re-sampling the collocation points ate every iteration improves generalisation.

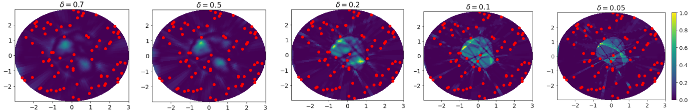


MSE and L^{∞}-error of the F.D. residual for eikonal equation in a 5D ball.

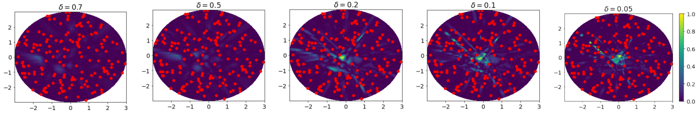
Data efficiency

F.D. residual for eikonal equation in a 2D ball

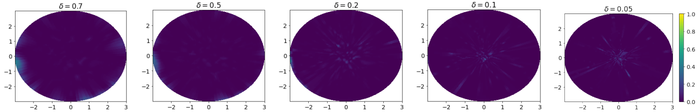
$N_0 = 80$



$N_0 = 160$



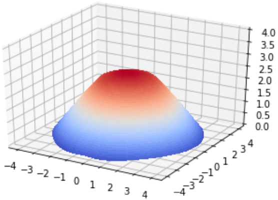
Re-sampling



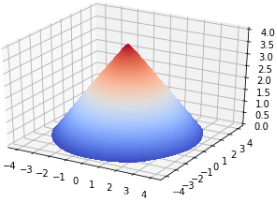
The sampling distribution

In high dimension, sampling the collocation points from a uniform distribution might not be the best idea

Eikonal equation in a 20-dimensional ball



Uniform sampling
Good accuracy in terms of MSE

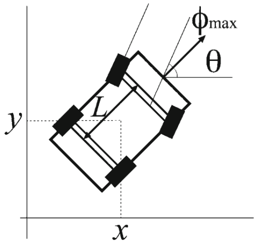


Radially uniform sampling
Good accuracy in terms of L^∞-error

We consider a mode for Reeds-Shepp's car

$$\begin{cases} \dot{x}(t) = \sigma a(t) \cos \omega(t) \\ \dot{y}(t) = \sigma a(t) \sin \omega(t) \\ \dot{\omega}(t) = \dfrac{b(t)}{\rho} \\ x(0) = x_0, \; y(0) = y_0, \; \omega(0) = \omega_0, \end{cases}$$

$(x, y, \omega) \in \mathbb{R}^2 \times [0, 2\pi)$ represent the car's position and orientation.



Problem: shortest path to the origin from the initial position.

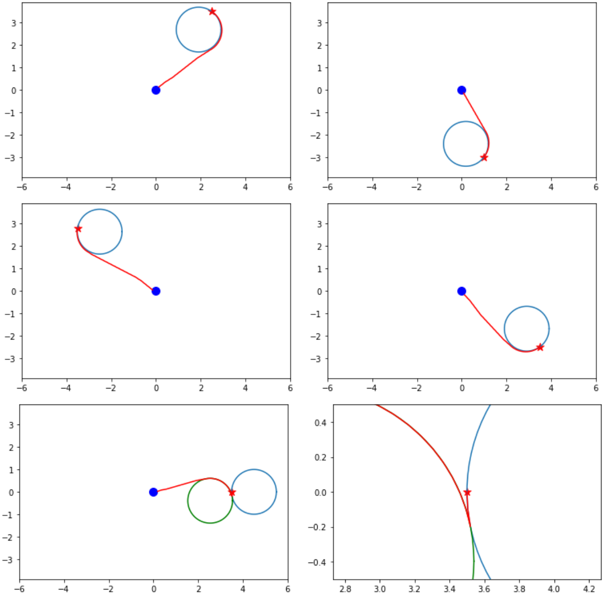
HJB equation:

$$H(x, y, \omega, \nabla u) = \sigma |\partial_x u \cos \omega + \partial_y u \sin \omega| + \frac{1}{\rho} |\partial_\omega u| - 1 = 0,$$

Domain: $\Omega := \mathbb{A}_{r,R} \times \mathbb{T}_{0,2\pi}$, where $\mathbb{A}_{r,R} := \{x \in \mathbb{R}^2 \; : \; r < \|x\|^2 < R\}$.
Boundary condition:

$$\begin{cases} u(x) = 0 & \|x\| = r \\ u(x) = R & \|x\| = R. \end{cases}$$

Optimal control problems with curvature constrained dynamics



We consider two Reeds-Shepp's cars (the Evader E and the Pursuer P)

$$\begin{cases} \dot{x}_e(t) = \sigma_e a_e(t)\cos\omega_e(t) \\ \dot{y}_e(t) = \sigma_e a_e(t)\sin\omega_e(t) \\ \dot{\omega}_e(t) = \dfrac{b_e(t)}{\rho_e} \\ x_e(0) = x_e,\ y_e(0) = y_e,\ \omega_e(0) = \omega_e, \end{cases}$$

$(x_e, y_e, \omega_e) \in \mathbb{R}^2 \times [0, 2\pi)$
represent the car's position of E.

$$\begin{cases} \dot{x}_p(t) = \sigma_p a_p(t)\cos\omega_p(t) \\ \dot{y}_p(t) = \sigma_p a_p(t)\sin\omega_p(t) \\ \dot{\omega}_p(t) = \dfrac{b_p(t)}{\rho_p} \\ x_p(0) = x_p,\ y_p(0) = y_p,\ \omega(0) = \omega_0, \end{cases}$$

$(x_p, y_p, \omega_p) \in \mathbb{R}^2 \times [0, 2\pi)$
represent the car's position of P.

Problem: P minimises the time to catch E, and E maximises the time until it gets caught by P.

HJI equation: we define $(X, Y) \in \mathbb{R}^2$ as $X = x_E - x_P$ and $Y = y_e - y_p$

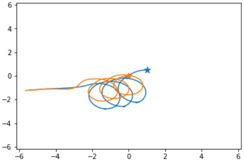
$$\begin{aligned} H(X, Y, \omega_e, \omega_p, \nabla u) \quad := \quad & \sigma_p \left|\partial_X u \cos\omega_p + \partial_Y u \sin\omega_p\right| + \frac{1}{\rho_p}\left|\partial_{\omega_p} u\right| \\ & -\sigma_e \left|\partial_X u \cos\omega_e + \partial_Y u \sin\omega_e\right| - \frac{1}{\rho_e}\left|\partial_{\omega_e} u\right|, \end{aligned}$$

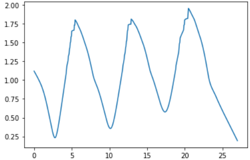
Domain: $\Omega := \mathbb{A}_{r,R} \times \mathbb{T}^2_{0,2\pi}$, where $\mathbb{A}_{r,R} := \{x \in \mathbb{R}^2 \ : \ r < \|x\|^2 < R\}$.
Boundary condition:

$$\begin{cases} u(x) = 0 & \|x\| = r \\ u(x) = R & \|x\| = R. \end{cases}$$

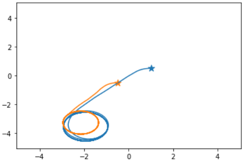
Velocities | Game trajectories | Time versus the distance

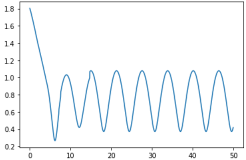
$$[\sigma_e, \rho_e] = [0.8, 1]$$
$$[\sigma_p, \rho_p] = [1, 1.2]$$



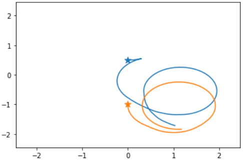


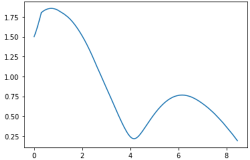
$$[\sigma_e, \rho_e] = [0.8, 1]$$
$$[\sigma_p, \rho_p] = [1, 1]$$





$$[\sigma_e, \rho_e] = [0.8, 1]$$
$$[\sigma_p, \rho_p] = [1, 0.8]$$





Conclusions:

1. We address a BVP for a HJ equation through a minimisation problem

$$\min_u \mathcal{J}(u) = \int_\Omega \left[\widehat{H}_\alpha(x, D_\delta^+ u(x) D_\delta^- u(x))\right]^2 dx + \int_{\partial\Omega} (u(x) - g(x))^2 dx$$

2. By choosing a suitable numerical Hamiltonian $\widehat{H}_\alpha(x, D_\delta^+ u(x) D_\delta^- u(x))$, we can ensure that any critical point approximates the viscosity solution.
3. The minimiser can be approximated by a NN trained through SGD.
4. We can start with α and δ large and then reduce them to refine the numerical solution.

Open questions:

1. What is the best **sampling distribution** for the collocation points?
   - In high dimension, uniform sampling is not effective.
   - Do we need more collocation points near the singular set?
   - Can we use the causality of the PDE to design a suitable sampling distribution?
2. **Sample complexity:** how many collocation points we need to achieve a good approximation?
   - for smaller values of δ we need more collocation points.
   - since the viscosity solution has typically a rather simple structure, we need less collocation points than grid points.
   - for more complex NN architectures we need more collocation points.
3. What about the **NN architecture**?
   - Is there a specific architecture that uses the structure of the solution to approximate it with less parameters?
4. Other non-linear PDEs?
   - We can consider any PDE.
   - A suitable numerical scheme, e.g. FD, FEM, etc.
   - Address the discretized problem by means of DL.
   - Analyse the optimality condition for the associated loss functional.

Thanks for the attention **Preprint:** arXiv:2406.10758

Open questions:

1. What is the best **sampling distribution** for the collocation points?
   - In high dimension, uniform sampling is not effective.
   - Do we need more collocation points near the singular set?
   - Can we use the causality of the PDE to design a suitable sampling distribution?
2. **Sample complexity:** how many collocation points we need to achieve a good approximation?
   - for smaller values of δ we need more collocation points.
   - since the viscosity solution has typically a rather simple structure, we need less collocation points than grid points.
   - for more complex NN architectures we need more collocation points.
3. What about the **NN architecture**?
   - Is there a specific architecture that uses the structure of the solution to approximate it with less parameters?
4. Other non-linear PDEs?
   - We can consider any PDE.
   - A suitable numerical scheme, e.g. FD, FEM, etc.
   - Address the discretized problem by means of DL.
   - Analyse the optimality condition for the associated loss functional.

Thanks for the attention **Preprint:** arXiv:2406.10758