Control in moving interfaces and deep learning

Borjan Geshkovski (ESR6)
Advisor: Enrique Zuazua

4$^{\text{th}}$ ConFlex Workshop, Lacanau
August 04, 2021

Universidad Autónoma de Madrid

- Started fellowship on **July 10, 2018**.
- Defended PhD thesis on **May 14, 2021**.
- Current status: AI Researcher at Sherpa.ai
- Publications:

  1. *Null-controllability of perturbed porous medium gas flow*. ESAIM: COCV, 2020.
  2. *Controllability of one-dimensional viscous free boundary flows*. With E. Zuazua. SIAM J. Control Optim., 2021
  3. *Turnpike in Lipschitz-nonlinear optimal control*. With C. Esteve-Yagüe, D. Pighin and E. Zuazua. Submitted, 2020
  4. *Large-time asymptotics in deep learning*. With C. Esteve-Yagüe, D. Pighin and E. Zuazua. Submitted, 2021
  5. *Sparse approximation in learning via neural ODEs*. With C. Esteve-Yagüe Submitted, 2021.

Control of free boundary problems

Free boundary problems

Classical formulation of FBPs – a PDE for unknown v, and an ODE (or PDE) for the unknown velocity h of the interface.

Example. *the piston problem*[1]:

$$\begin{cases} \partial_t v - \partial_x^2 v + v\partial_x v = 0 & \text{in } (0,T)\times(-1,1) \setminus \{h(t)\} \\ \ddot{h}(t) = [\partial_x v](t, h(t)) & \text{in } (0,T) \\ v(t,h(t)) = \dot{h}(t) & \text{in } (0,T) \\ v(t,-1) = 0, \ v(t,1) = u(t) & \text{in } (0,T) \\ (v,h)_{|t=0} = (v_0, h_0) & \text{in } (-1,1)\setminus\{h_0\}. \end{cases}$$

FBPs model multi-physics phenomena such as

- Phase transitions
- Fluid-structure interaction
- Elasticity and contact problems
- Thin fluid films
- Water waves

[1] Vazquez & Zuazua, Comm. PDE '03; Liu, Takahashi, Tucsnak, COCV '13.

Controllability of one-dimensional viscous free boundary flows.
With E. Zuazua.
SIAM J. Control Optim., 2021
https://epubs.siam.org/doi/abs/10.1137/19M1285354

We consider

$$\begin{cases} \partial_t v - \partial_z^2 v + v\partial_z v = 0 & \text{in } (0,T)\times(0,h(t)) \\ \dot{h}(t) = v(t,h(t)) & \text{in } (0,T) \\ v(t,0) = u(t),\ \partial_z v(t,h(t)) = 0 & \text{in } (0,T) \\ (v,h)_{|t=0} = (v_0,h_0) & \text{in } (0,h_0). \end{cases}$$

Goal: Given $T>0$ and a free trajectory $(\overline{v},\overline{h})$, find a control u such that

$$h(T) = \overline{h}(T) \qquad \text{and} \qquad v(T,\cdot) = \overline{v}(T,\cdot) \quad \text{in } (0,\overline{h}(T)).$$

Theorem: Let $T>0$, $h_*>0$, $\overline{v}\in\mathbb{R}$ be such that $\overline{h}(t) = h_* + \overline{v}t > 0$ for all $t\in[0,T]$. There exits $r>0$ such that for all $h_0>0$ and $v_0\in H^1(0,h_0)$ satisfying

$$\|v_0-\overline{v}\|_{H^1} + |h_0-h_*| \leqslant r,$$

there exists a control $u\in C^0[0,T]$ such that unique solution $h\in C^1[0,T]$ and $v\in C^0(H^1)\cap L^2(H^2)$ satisfies

$$h(T) = \overline{h}(T) \quad \text{and} \quad v(T,\cdot) = \overline{v} \quad \text{in } (0,\overline{h}(T)).$$

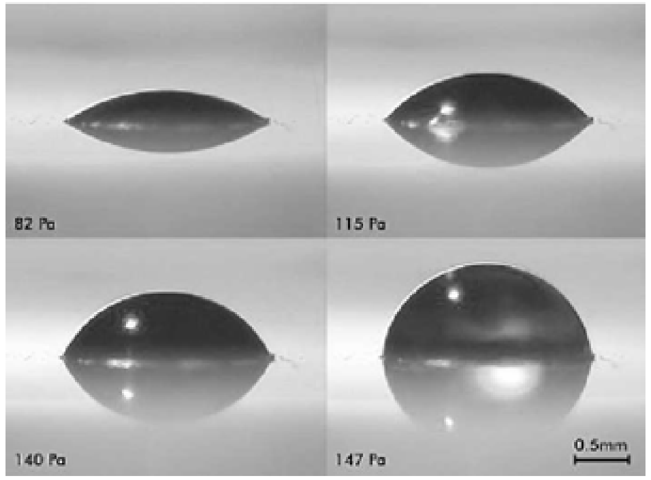
Null-controllability of perturbed porous medium gas flow.
ESAIM: COCV, 2020.
https://doi.org/10.1051/cocv/2020009

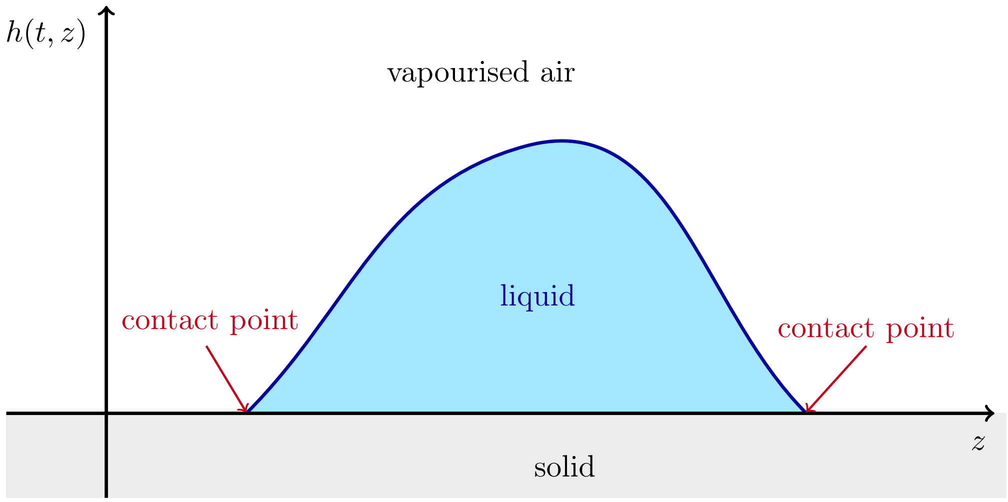
Model for porous medium gas flow and thin fluid film dynamics: *find* $h \geqslant 0$ *solving*

$$\partial_t h - \partial_z^2(h^m) = 0 \qquad \text{on} \;\; \mathbb{R} \times (0, +\infty)$$

for $m > 1$.

- Finite speed of propagation $\Rightarrow$ free boundary $\partial\{h(t) > 0\}$
- Add a distributed control, and control to self-similar Barenblatt attractor – stationary parabola in similarity variables. Need to linearize..
- Write equation in similarity variables and pressure $v = h^{m-1}$ – free boundary is now $\partial\{v(t) > 0\}$
- Transformation onto support of the parabola $\frac{1}{2}(1 - x^2)$, which is $[-1, 1]$.





Discussion leads us to consider[2] equation for perturbation around $\rho(x) = \frac{1}{2}(1 - x^2)$

$$\begin{cases} \partial_t y - \rho^{-\sigma}\, \partial_x(\rho^{\sigma+1}\, \partial_x y) = \mathcal{N}(y, \partial_x y) + u 1_\omega & \text{in } (0,T)\times(-1,1) \\ (\rho^{\sigma+1}\, \partial_x y)(t, \pm 1) = 0 & \text{in } (0,T) \\ y_{|t=0} = y_0 & \text{in } (-1,1), \end{cases}$$

where $\sigma = \frac{m-2}{m-1} > -1$.

Theorem: Let $T > 0$, $\omega \subsetneq (-1,1)$ open, nonempty, and $\sigma \in (-1,0)$. There exists $r > 0$ such that for every $\|y_0\|_{\mathcal{H}^1} \leqslant r$, there exists $u \in L^2((0,T)\times\omega)$ such that $y \in C^0(\mathcal{H}^1) \cap L^2(\mathcal{H}^2)$ satisfies $y(T,\cdot) = 0$ in $(-1,1)$.

[2] Koch's Habilitation '99, Seis JDE '15.

Control of the Stefan problem with surface tension. With D. Maity, in preparation, 2021.

Unknowns ϑ (temperature) and h (melting interface) solve

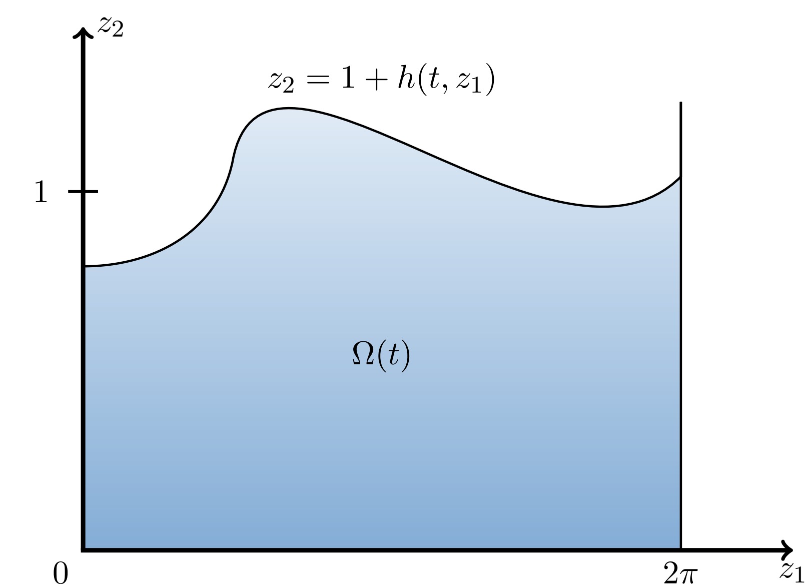
$$\begin{cases} \partial_t \vartheta - \Delta \vartheta = 0 & \text{in } (0,T)\times\Omega(t) \\ \partial_t h = -\sqrt{1+|\partial_{x_1} h|^2}\, \nabla\vartheta_{|_{\Gamma(t)}} \cdot \mathrm{n} & \text{on } (0,T)\times\mathbb{T} \\ \vartheta = u & \text{on } (0,T)\times\mathbb{T}\times\{0\} \\ \vartheta = -\sigma\kappa(h) & \text{on } (0,T)\times\Gamma(t) \\ (\vartheta,h)_{|t=0} = (\vartheta_0,h_0) & \text{in } \Omega(0)\times\mathbb{T}. \end{cases}$$

with $\sigma > 0$ surface tension; $\kappa(h)$ mean curvature, $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$.

Goal: Given $T > 0$ and $\sigma > 0$, **find a control** u actuating over $\mathbb{T}$ such that

$$h(T,\cdot) = 0 \ \text{ in } \ \mathbb{T} \qquad \text{and} \qquad \vartheta(T,\cdot) = 0 \text{ in } \ \mathbb{T}\times(0,1).$$

- Fluid:
  $\Omega(t) = \{(x_1,x_2)\in\mathbb{T}\times\mathbb{R} \ : 0 < x_2 < 1 + h(t,x_1)\}$
- Interface:
  $\Gamma(t) = \{(x_1,x_2)\in\mathbb{T}\times\mathbb{R} \ : x_2 = 1 + h(t,x_1)\}$

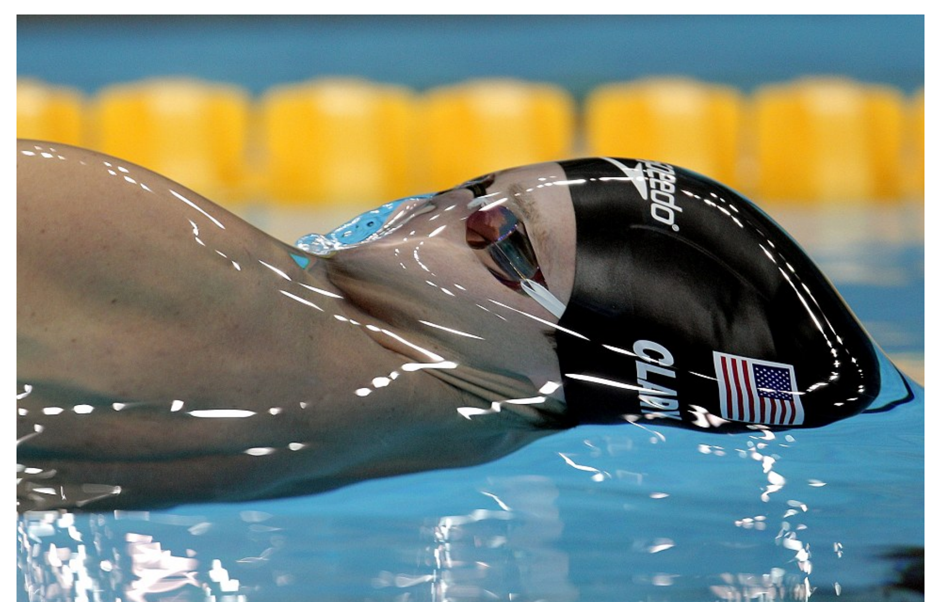


Transform, linearize, extend domain fictitiously to $\Omega := \mathbb{T} \times (-1,1)$:

$$\begin{cases} \partial_t y - \Delta y = u 1_\omega & \text{in } (0,T) \times \Omega \\ \partial_t h(t,x_1) = \partial_{x_2} y(t,x_1,1) & \text{on } (0,T) \times \mathbb{T} \\ y(t,x_1,-1) = 0 & \text{on } (0,T) \times \mathbb{T} \\ y(t,x_1,1) = \sigma \partial^2_{x_1} h(t,x_1) & \text{on } (0,T) \times \mathbb{T} \\ (y,h)_{|t=0} = (y_0,h_0) & \text{in } \Omega \times \mathbb{T}, \end{cases}$$

with $\omega \subset (-1,0)$. We look for $u \in L^2_{t,x}$ s.t. $y(T,\cdot) = 0$ in Ω and $h(T,\cdot) = 0$ in $\mathbb{T}$.

Surface tension: tendency of liquids to minimize their surface area. Adds an additional coupling between the liquid and solid.



Strategy:

- Fourier-decompose all unknowns w.r.t. $x_1 \in \mathbb{T}$:

$$\begin{cases} \partial_t \widehat{y}_n - \partial_{x_2} \widehat{y}_n + n^2 \widehat{y}_n = \widehat{u}_n 1_\omega & \text{in } (0,T) \times (-1,1) \\ \widehat{h}_n'(t) = \partial_{x_2} \widehat{y}_n(t,1) & \text{in } (0,T) \\ \widehat{y}_n(t,-1) = 0 & \text{in } (0,T) \\ \widehat{y}_n(t,1) = -\sigma n^2 \widehat{h}_n(t) & \text{in } (0,T) \\ (\widehat{y}_n, \widehat{h}_n)_{|t=0} = (\widehat{y}_{0,n}, \widehat{h}_{0,n}) & \text{in } (-1,1), \end{cases}$$

  for $n \in \mathbb{Z}$.
- For $n = 0$, much like Burgers system from Chapter 1.
- For $n \neq 0$: control $1d$ system uniformly w.r.t. n by
  - HUM for $\begin{bmatrix} \widehat{y}_n \\ \widehat{h}_n \end{bmatrix}$
  - coercivity inequality by computing spectrum $\{\lambda_{n,k}\}_{k=1}^{\infty}$, and after showing that $\inf_{n,k} |\lambda_{n,k+1} - \lambda_{n,k}| > 0$ and $\lambda_{n,k} \sim k^2 + n^2 + \mathcal{O}(k)$ and using

$$\int_0^T \left| \sum_{k=1}^{\infty} a_k e^{-\lambda_{n,k} t} \right|^2 dt \gtrsim_T \sum_{k=1}^{\infty} |a_k|^2 e^{-2\lambda_{n,k} T} \qquad \forall \{a_k\} \in \ell^2(\mathbb{N}) \qquad (1)$$

- Uniform control of all Fourier coefficients w.r.t. $n \in \mathbb{Z}$ (not σ!) yields linear result.

Interplay of control and deep learning

Supervised learning: Find an approximation of an unknown function $f : \mathcal{X} \to \mathcal{Y}$ from a dataset

$$\{\vec{x}_i, \vec{y}_i\}_{i=1}^{N} \subset \mathcal{X} \times \mathcal{Y}$$

$\mathcal{X} \subset \mathbb{R}^d$; we distinguish:

- **Classification:** $\mathcal{Y}$ is a discrete set of m classes, e.g. $\mathcal{Y} = \{1, \dots, m\}$;
- **Regression:** $\mathcal{Y} \subset \mathbb{R}^m$

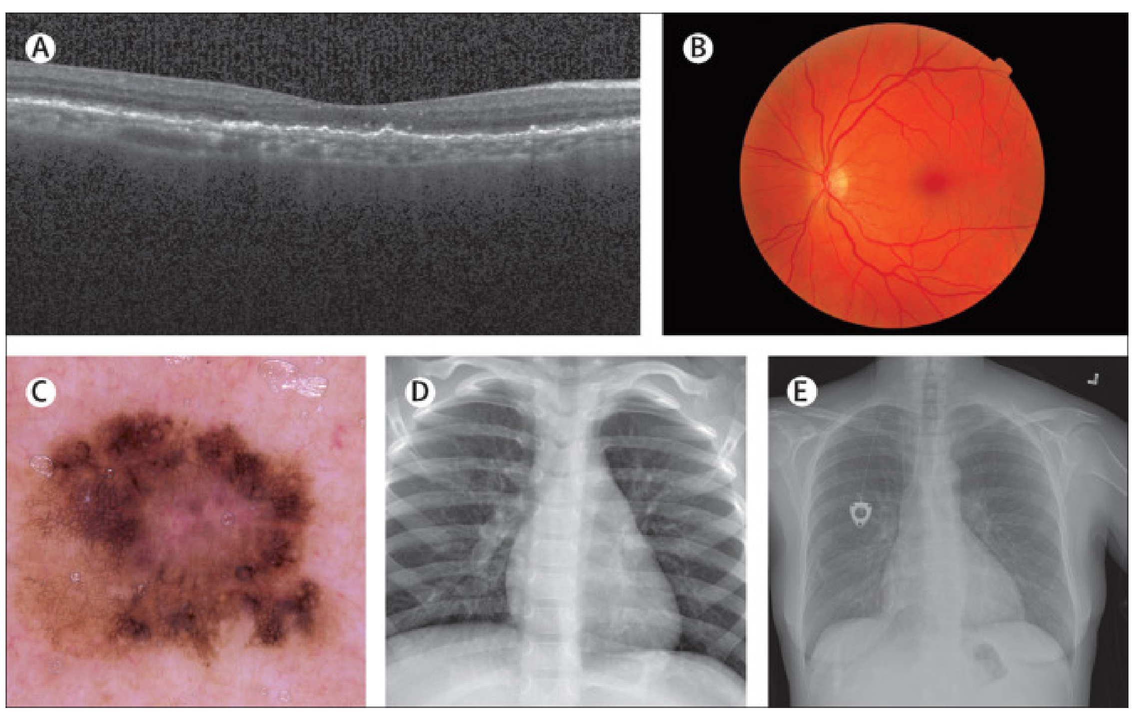


Figure: Classification ($f : \mathbb{R}^{9216} \to \{-1, +1\}$)

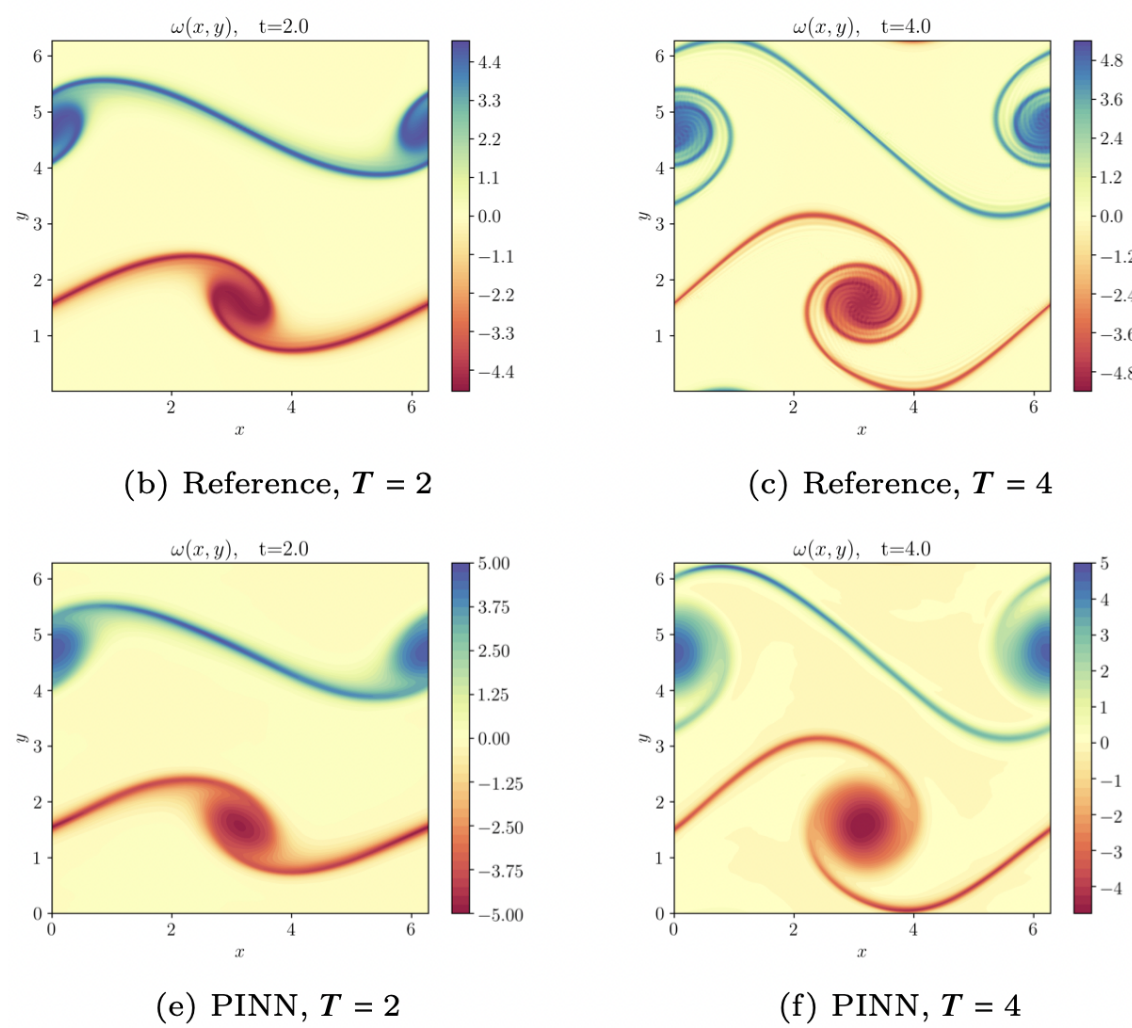


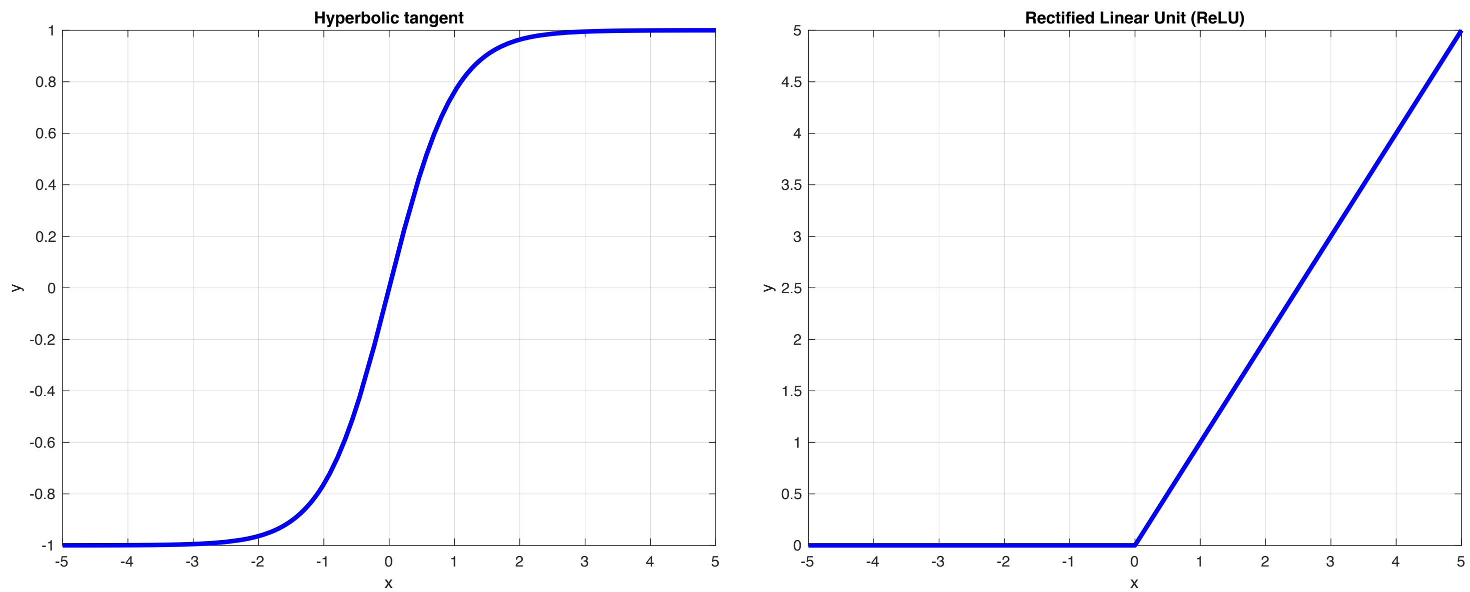
Figure: Regression ($f : [0, 4] \times [0, 6]^2 \to \mathbb{R}^2$)

Neural networks

Neural network: for any $i \leqslant N$

$$\begin{cases} x_i^{k+1} = \sigma(w^k x_i^k + b^k) & \text{for } k \in \{0, \dots, N_{layers} - 1\} \\ x_i^0 = \vec{x}_i \in \mathbb{R}^d, & \end{cases} \qquad (\text{NN}_1)$$

- $w^k \in \mathbb{R}^{d_{k+1} \times d_k}$ and $b^k \in \mathbb{R}^{d_k}$ are *controls*;
- $N_{layers} \geqslant 1$ given **depth**; $d_k \geqslant 1$ called **widths** with $d_0 = d$ and $d_{N_{layers}} = m$.
- $\sigma \in \mathrm{Lip}(\mathbb{R})$ & $\sigma(0) = 0$ defined componentwise:



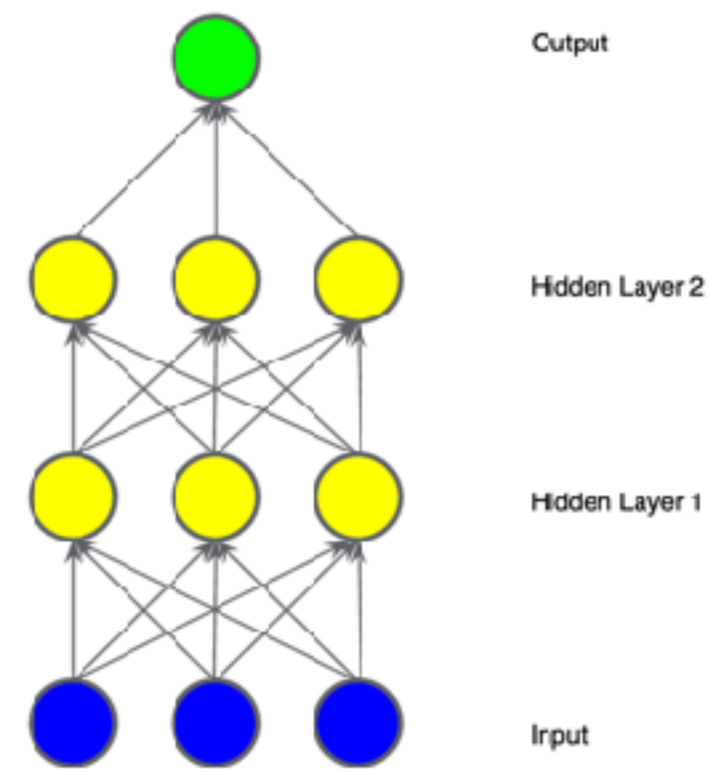


Figure: Sigmoid: $\tanh(x)$ and ReLU: $\max\{x, 0\}$

"Training" a NN

Training $\Longleftrightarrow$ **Optimization**: $\lambda > 0$ fixed,

$$\min_{\{w^k, b^k\}_{k=0}^{N_{layers}}} \underbrace{\frac{1}{N}\sum_{i=1}^{N} \text{loss}\left(Px_i^{N_{layers}}, \vec{y}_i\right)}_{\text{training error}} + \lambda \underbrace{\left\|\left\{w^k, b^k\right\}_k\right\|_{\ell^p}^p}_{\text{regularization}}$$

$P : \mathbb{R}^d \to \mathbb{R}^m$ is an affine map:

$$Px = w^{N_{layers}} x + b^{N_{layers}}.$$

1. **Regression:** $\vec{y}_i \in \mathcal{Y} \subset \mathbb{R}^m$ and

$$\text{loss}(x, y) = \|x - y\|^2$$

2. **Classification:** $\vec{y}_i \in \mathcal{Y} = \{-1, 1\}$ (so $m = 1$) and

$$\text{loss}(Px, y) = \max\left(0, 1 - yPx\right)$$

We shall assume that P is given, and arbitrary unless stated otherwise.

Residual neural networks

ResNets: fix $d_k \equiv d$; for any $i \leqslant N$

$$\begin{cases} \mathrm{x}_i^{k+1} = \mathrm{x}_i^k + h\sigma(w^k \mathrm{x}_i^k + b^k) & \text{for } k \in \{0, \ldots, N_{layers} - 1\} \\ \mathrm{x}_i^0 = \vec{x}_i \end{cases} \tag{ResNet}$$

where $h = 1$.

layer = timestep[3]; $h = \frac{T}{N_{layers}}$ for given $T > 0$:

$$\begin{cases} \dot{\mathrm{x}}_i(t) = \sigma(w(t)\mathrm{x}_i(t) + b(t)) & \text{for } t \in (0, T) \\ \mathrm{x}_i(0) = \vec{x}_i. \end{cases} \tag{nODE$_1$}$$

For (nODE$_1$), we shall henceforth assume $\sigma(\lambda x) = \lambda\sigma(x)$ for $\lambda > 0$ (positive homogeneity).


Deep residual learning for image recognition
K He, X Zhang, S Ren, J Sun - Proceedings of the IEEE ..., 2016 - openaccess.thecvf.com
Deeper neural networks are more difficult to train. We present a residual learning framework to ease the training of networks that are substantially deeper than those used previously. We explicitly reformulate the layers as learning residual functions with reference to the layer ...
Cited by 77708 Related articles All 55 versions

[3] Weinan E '17

Residual neural networks

In addition to (nODE_1), one can also consider variants:

- $$\begin{cases} \dot{\mathrm{x}}_i(t) = w(t)\sigma(\mathrm{x}_i(t)) + b(t) & \text{for } t \in (0, T) \\ \mathrm{x}_i(0) = \vec{x}_i. \end{cases} \tag{nODE$_2$}$$

- Also

$$\begin{cases} \dot{\mathrm{x}}_i(t) = w_1(t)\sigma(w_2\mathrm{x}_i(t) + b_2) + b_1(t) & \text{for } t \in (0, T) \\ \mathrm{x}_i(0) = \vec{x}_i \end{cases} \tag{nODE$_3$}$$

where $w_1 \in \mathbb{R}^{d \times d_1}$, $w_2 \in \mathbb{R}^{d_1 \times d}$.

Training is optimal control

Given $T, \lambda > 0$:

$$\inf_{[w,b]\in H^k(0,T;\mathbb{R}^{d_u})} \underbrace{\frac{1}{N}\sum_{i=1}^{N} \mathrm{loss}\Big(P\mathsf{x}_i(T), \vec{y}_i\Big)}_{=:\mathcal{E}(\mathsf{x}(T))} + \lambda \Big\| [w,b] \Big\|^2_{H^k(0,T;\mathbb{R}^{d_u})}$$

- $k = 0$ for (nODE$_2$), $k = 1$ for (nODE$_1$), (nODE$_3$) (L^2–regularization **might not be enough** for compactness)
- We henceforth suppose $P : \mathbb{R}^d \to \mathbb{R}^m$ is affine, and $\mathrm{loss} \in C^0(\mathbb{R}^m \times \mathcal{Y}; \mathbb{R}_+)$ is such that $\mathcal{E}$ attains its minimum 0.

Large-time asymptotics in deep learning.
With C. Esteve-Yagüe, D. Pighin and E. Zuazua.
Submitted, 2021
https://arxiv.org/abs/2008.02491

- Set $\mathrm{x}^0 = [\vec{x}_1, \ldots, \vec{x}_N]$, $u = [w, b]$, and put both (nODE$_1$) and (nODE$_2$) in the form

$$\begin{cases} \dot{\mathrm{x}}(t) = \mathrm{f}(\mathrm{x}(t), u(t)) & \text{in } (0, T) \\ \mathrm{x}(0) = \mathrm{x}^0 \in \mathbb{R}^{d_x}. \end{cases} \tag{nODE}$$

- And so

$$\inf_{\substack{u \in H^k(0,T;\mathbb{R}^{d_u}) \\ \text{subject to (nODE)}}} \mathcal{E}(\mathrm{x}(T)) + \lambda \|u\|^2_{H^k(0,T;\mathbb{R}^{d_u})} \tag{SL$_1$}$$

Question: What happens to a minimizer u^T solving (SL$_1$), and corresponding state x^T to (nODE) when $T \to +\infty$?

Scaling

$$\inf_{\substack{u\in H^k(0,T;\mathbb{R}^{d_u})\\ \text{subject to (nODE)}}} \mathcal{E}(\mathrm{x}(T)) + \lambda \|u\|^2_{H^k(0,T;\mathbb{R}^{d_u})} \tag{SL$_1$}$$

Key idea: *Time-Scaling*.

- Assumptions on σ entail $\mathrm{f}(\mathrm{x}, u)$ positively homogeneous w.r.t. u, i.e. $\mathrm{f}(\mathrm{x}, \alpha u) = \alpha \mathrm{f}(\mathrm{x}, u)$ for $\alpha > 0$.
- Hence, given $u^T(t)$ and the solution $\mathrm{x}^T(t)$ to

$$\begin{cases} \dot{\mathrm{x}}^T(t) = \mathrm{f}\left(\mathrm{x}^T(t), u^T(t)\right) & \text{in } (0,T) \\ \mathrm{x}^T(0) = \mathrm{x}^0, \end{cases} \tag{2}$$

then $u^1(t) := Tu^T(tT)$ is such that $\mathrm{x}^1(t) := \mathrm{x}^T(tT)$ solves (2) for $t \in [0,1]$.

$$\Longrightarrow \qquad \boxed{\lambda \int_0^T \left\|u^T(t)\right\|^2 dt = \frac{\lambda}{T} \int_0^1 \left\|u^1(s)\right\|^2 ds}$$

Theorem: Fix $\lambda > 0$, let $P : \mathbb{R}^d \to \mathbb{R}^m$ be any surjective affine map. For any $T > 0$, let u^T be minimizer in (SL$_1$), x^T associated solution to (nODE). Assume that $\{\mathcal{E} = 0\}$ is reachable by (nODE). Then

1. $\exists C > 0$ independent of T such that
$$\mathcal{E}(\mathrm{x}^T(T)) \leqslant \frac{C}{T}.$$
2. Moreover, $\exists \{T_n\}_{n=1}^{+\infty}$ positive times and $\exists \mathrm{x}_\circ \in \mathbb{R}^{d_x}$, $\mathcal{E}(\mathrm{x}_\circ) = 0$, such that
$$\left\| \mathrm{x}^{T_n}(T_n) - \mathrm{x}_\circ \right\| \longrightarrow 0 \qquad \text{as } n \to +\infty.$$
3. Moreover
$$\left\| \frac{1}{T_n} u^{T_n}\left(\frac{\cdot}{T_n}\right) - u^* \right\|_{H^k(0,1;\mathbb{R}^{d_u})} \longrightarrow 0 \qquad \text{as } n \to +\infty$$
where u^* solves
$$\inf_{\substack{u \in H^k(0,1;\mathbb{R}^{d_u}) \\ \text{subject to (nODE) with } T=1 \\ \text{and} \\ \mathcal{E}(\mathrm{x}(1))=0}} \|u\|^2_{H^k(0,1;\mathbb{R}^{d_u})}.$$

Enhancing the decay

Question: Better quantitative estimates for the time T required to approach the zero training error regime $\mathcal{E}(\mathrm{x}(T)) = 0$?

- Consider $\mathrm{loss}(x, y) = \|x - y\|^2$ and so we recall

$$\mathcal{E}(\mathrm{x}(T)) := \frac{1}{N}\sum_{i=1}^{N} \|P\mathrm{x}_i(T) - \vec{y}_i\|^2$$

- We shall suppose $P : \mathbb{R}^d \to \mathbb{R}^m$ surjective, Lipschitz, but arbitrary
- and let $\overline{\mathrm{x}} \in \mathbb{R}^{d_x}$ s.t. $\overline{\mathrm{x}}_i \in P^{-1}(\{\vec{y}_i\})$ for $i \leqslant N$ be fixed.
- Augmented problem:

$$\inf_{\substack{u\in L^2(0,T;\mathbb{R}^{d_u}) \\ \text{subject to (nODE)}}} \mathcal{E}(\mathrm{x}(T)) + \int_0^T \|\mathrm{x}(t) - \overline{\mathrm{x}}\|^2 dt + \lambda \|u\|^2_{L^2(0,T;\mathbb{R}^{d_u})} \qquad \text{(SL*)}$$

Exponential decay

Theorem Fix $\lambda > 0$, and suppose that (nODE) is controllable with linear estimate of the cost. There exist $T^* > 0$ such that for any $T \geqslant T^*$, any solution (u^T, x^T) to (SL*)–(nODE) satisfies

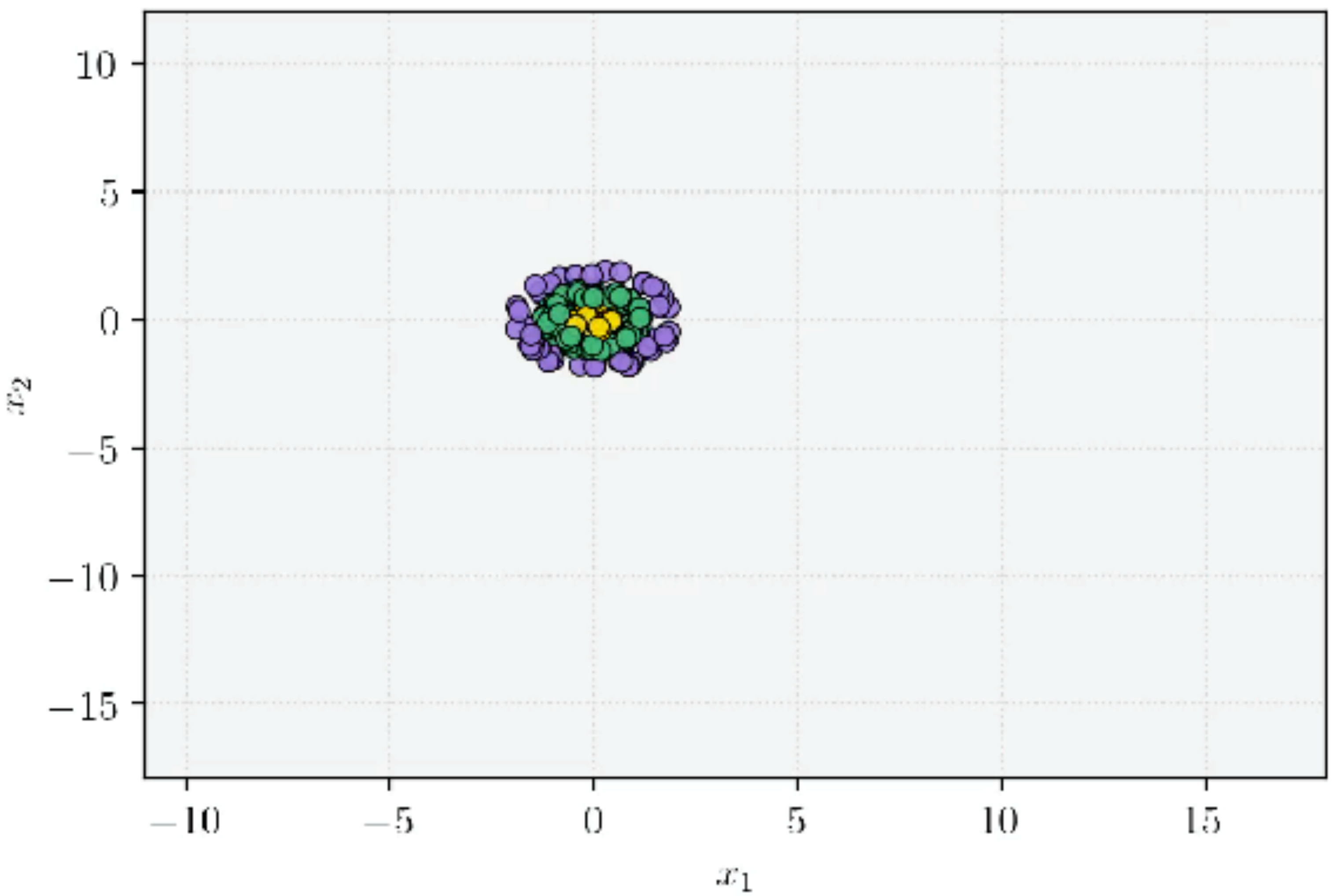
$$\mathcal{E}\left(\mathrm{x}^T(t)\right) + \left\|\mathrm{x}^T(t) - \overline{\mathrm{x}}\right\| \leqslant C_1 e^{-\mu t} \qquad \forall t \in [0, T]$$

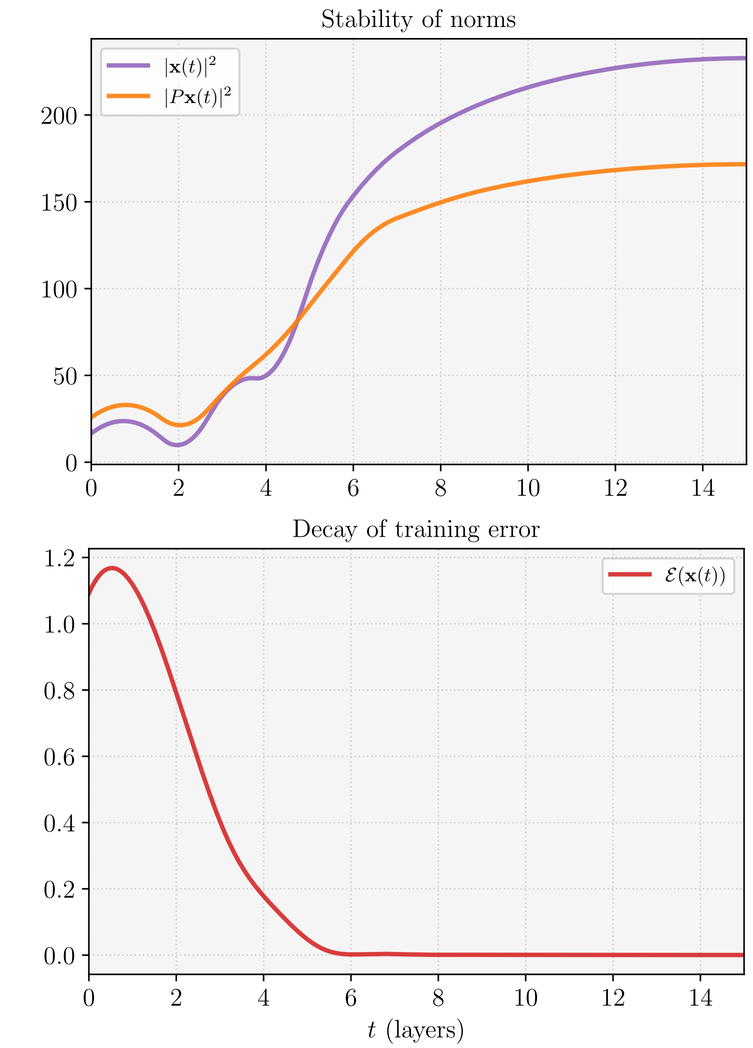
and

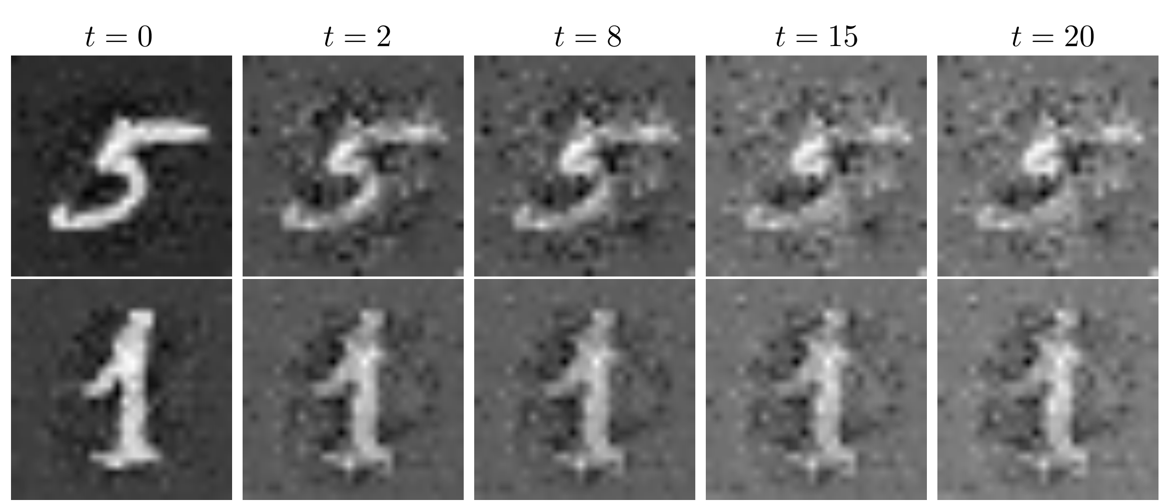
$$\left\|u^T(t)\right\| \leqslant C_2 e^{-\mu t} \qquad \text{for a.e. } t \in [0, T]$$

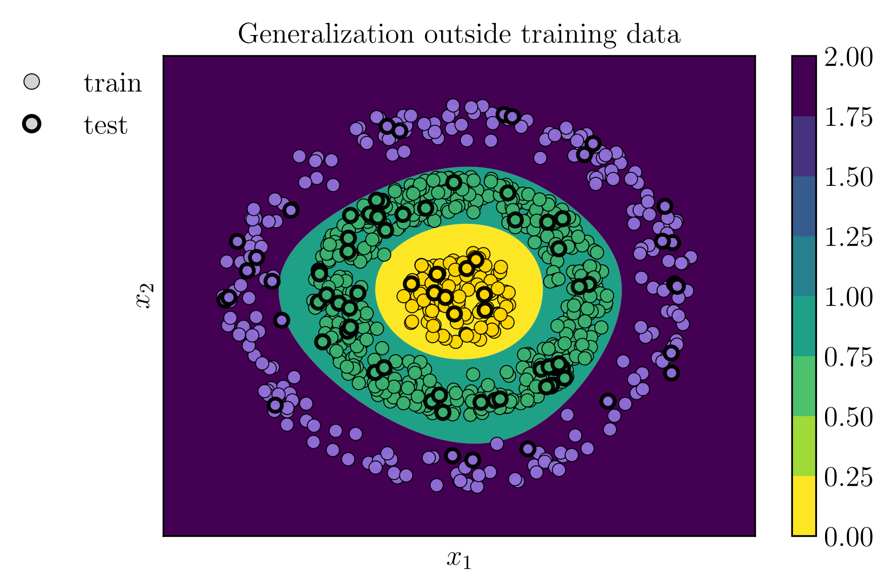
for some $C_1, C_2, \mu > 0$, all independent of T.

- Akin to *universal approximation*: given tolerance $\varepsilon > 0$, there exists $T_\varepsilon > 0$ (number of layers) and control parameters u^ε such that the neural network output is ε–close to the desired target.
- One difference with universal approximation is that our parameters may be computed explicitly via a training procedure.


10
5
0
−5
−10
−15
−10
−5
0
5
10
15
x_1
x_2


Stability of norms
$|\mathbf{x}(t)|^2$
$|P\mathbf{x}(t)|^2$
Decay of training error
$\mathcal{E}(\mathbf{x}(t))$
t (layers)


$t = 0$
$t = 2$
$t = 8$
$t = 15$
$t = 20$


Generalization outside training data
train
test
x_1
x_2
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00

Turnpike in Lipschitz-nonlinear optimal control.
With C. Esteve-Yagüe, D. Pighin and E. Zuazua.
Submitted, 2020
https://arxiv.org/abs/2011.11091

- Theorem is a special manifestation of the **turnpike property** in optimal control and economics.
- *For suitable optimal control problems in a sufficiently large T, any optimal solution (u^T, x^T) remains, during most of the time, $\mathcal{O}(e^{-t} + e^{-(T-t)})$–close to the optimal solution of a corresponding "static" optimal control problem.*
  Optimal static solution is referred to as the turnpike – the name stems from the idea that a turnpike is the fastest route between two points which are far apart, even if it is not the most direct route.
- Since $f(x, 0) = 0$ for all x, $\overline{x}_i$ may be seen as the turnpike for Px_i. Since this is a steady state, we do not see an exit from the turnpike and we stabilize.

Sparse approximation in learning via neural ODEs.
With C. Esteve-Yagüe
Submitted, 2021.
https://arxiv.org/abs/2102.13566

What about L^1–regularization?

Theorem: Fix $M > 0$. Consider

$$\inf_{\substack{u \in L^1(0,T;\mathbb{R}^{d_u}) \\ \|u\|_{L^\infty(0,T)} \leqslant M \\ \text{subject to (nODE}_2)}} \int_0^T \mathcal{E}(x(t))dt + \lambda \|u\|_{L^1(0,T;\mathbb{R}^{d_u})}.$$

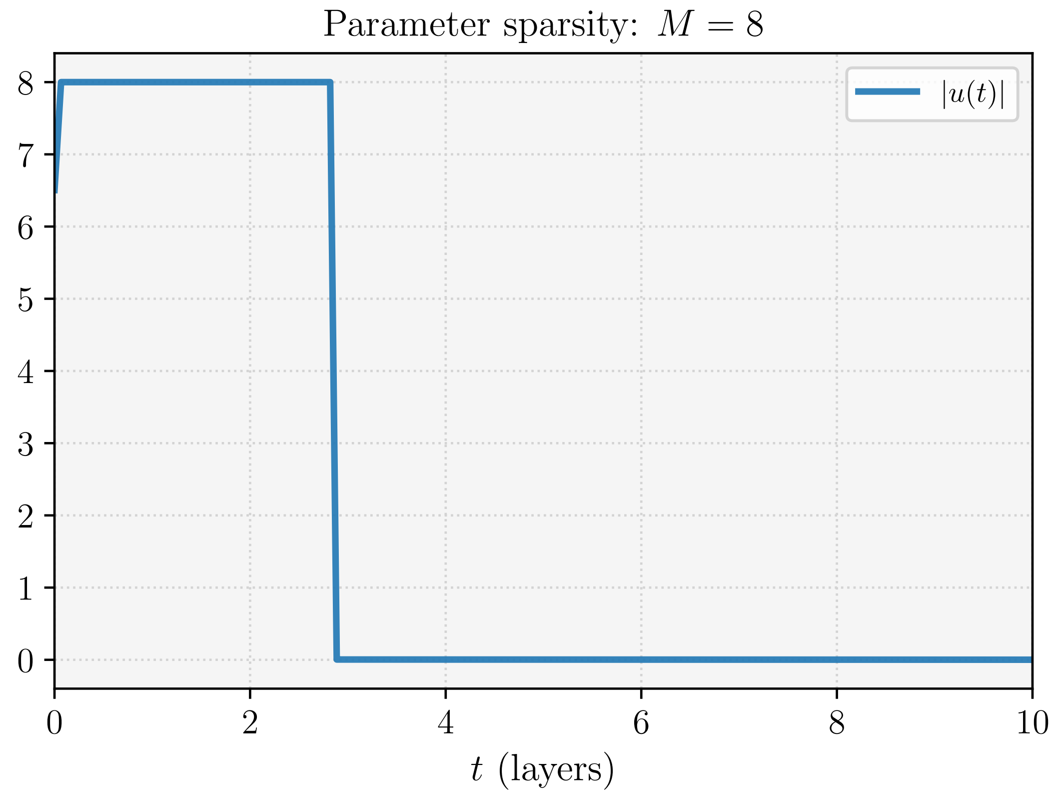
Then there exists $T^* > 0$ such that for any $T > T^*$, any optimal u^T satisfies

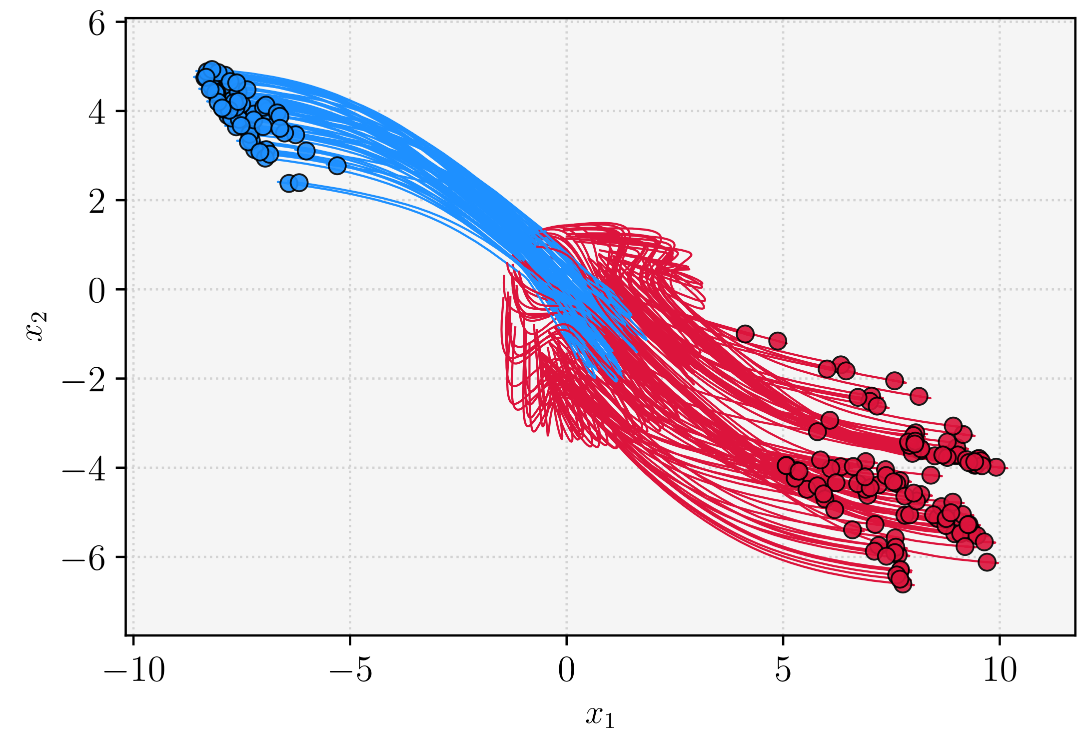
$$\begin{aligned} \left\| u^T(t) \right\| &= M, \qquad \text{for a.e. } t \in (0, T^*) \\ \left\| u^T(t) \right\| &= 0, \qquad \text{for a.e. } t \in (T^*, T). \end{aligned}$$

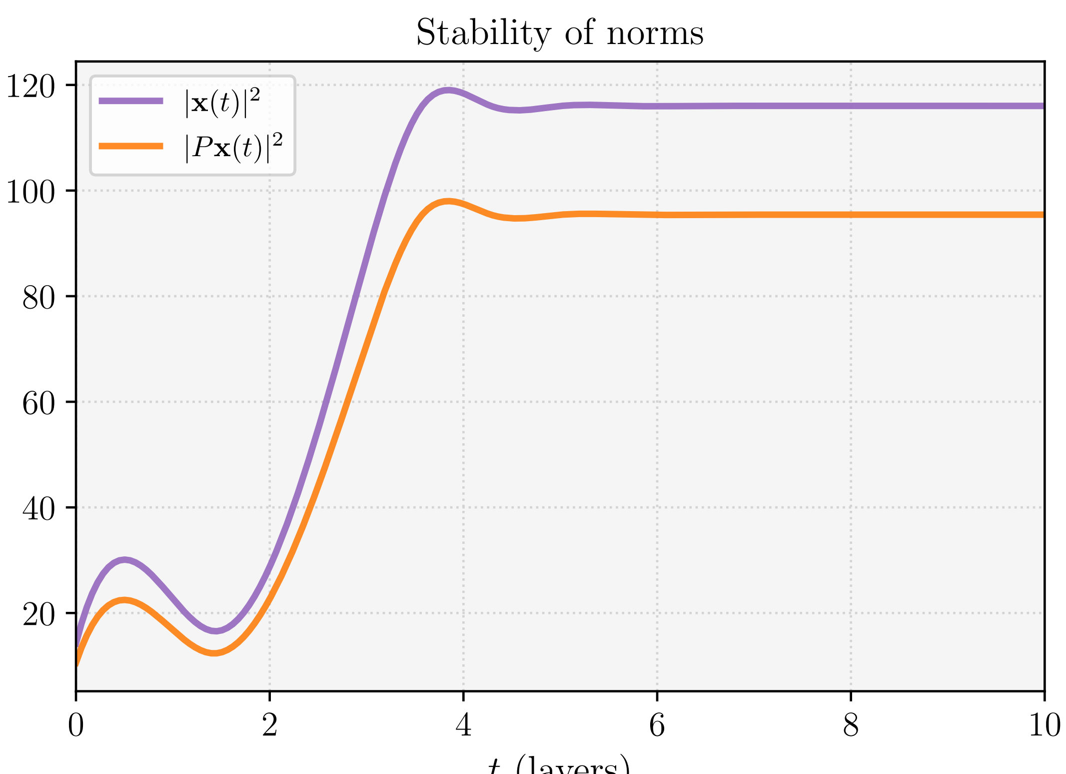
If moreover (nODE$_2$) is controllable, then there exist $C(M) > 0$ and $T(M) > 0$ such that

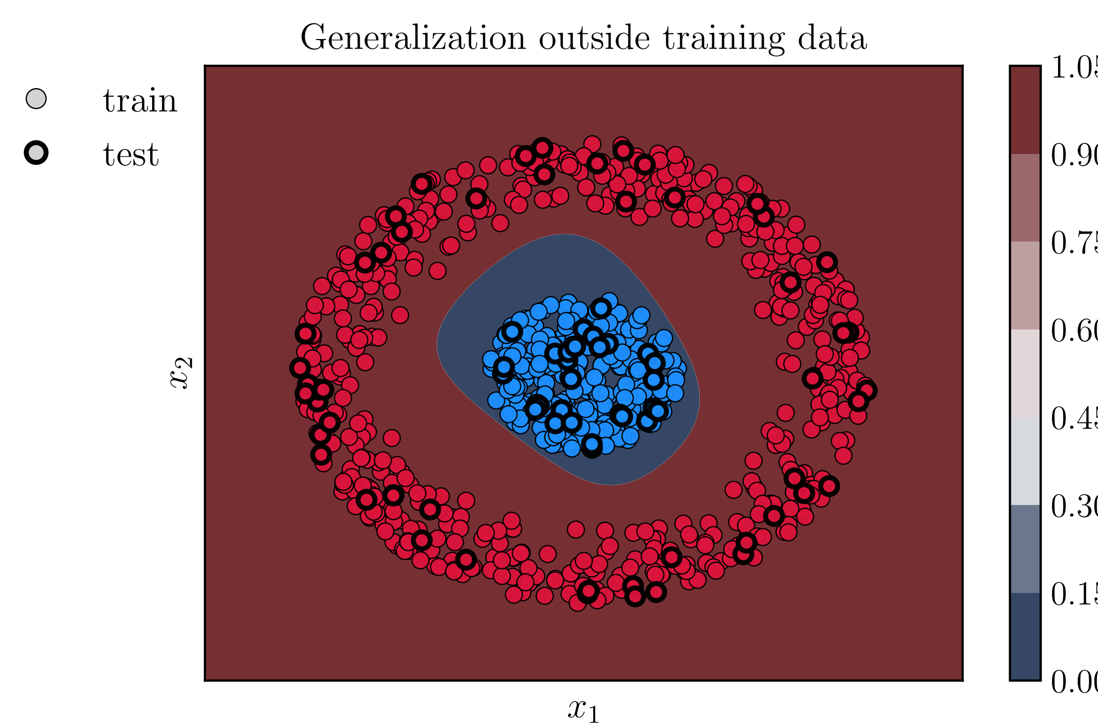
$$T^* \leqslant T(M) \qquad \text{and} \qquad \mathcal{E}(x(T^*)) \leqslant \frac{C(M)}{T} \tag{3}$$

L^1–regularization

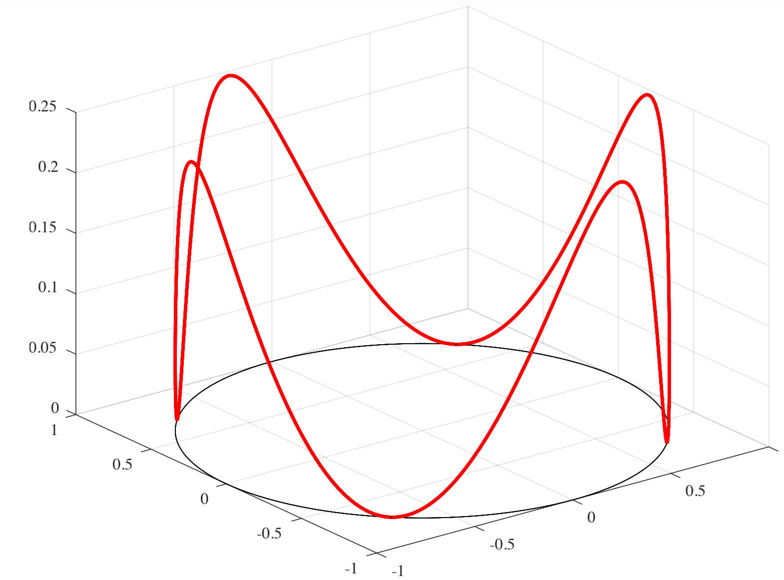








Coming soon...



> *Optimal controller design via Brunovsky's normal form.*
> With E. Zuazua.
> 2021

$$\begin{cases} y'(t) - Ay(t) = bu(t) & \text{in } (0,T), \\ y(0) = y_0, \end{cases}$$

$A \in \mathcal{M}_{n\times n}(\mathbb{R})$, $b \in \mathbb{R}^n$ Kalman rank.

Consider

$$\mathfrak{C}^*(b,T) := \inf_{\|y_0\|=1} \|\Gamma_b(y_0)\|_{L^2(0,T)},$$

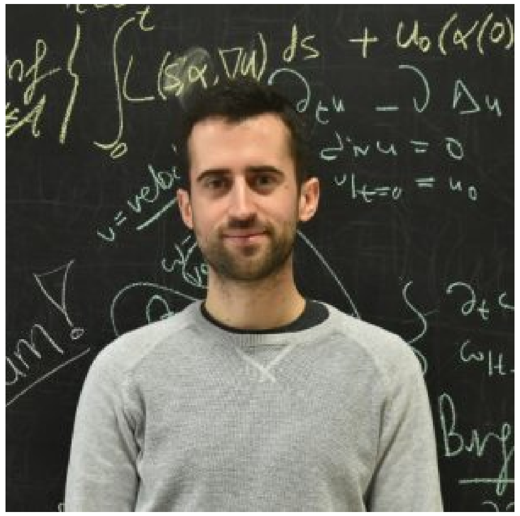
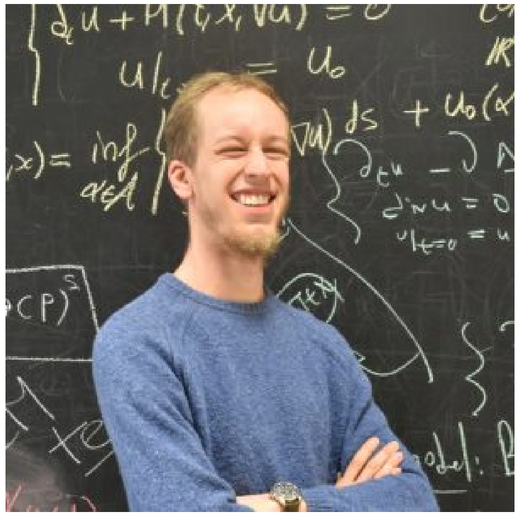
where $\mathbb{R}^n \ni y_0 \mapsto \Gamma_b(y_0) = u \in L^2(0,T)$ is the "datum to minimal L^2–norm exact control". We show

$$\boxed{\min_{b\in\mathbb{S}^{n-1}} \mathfrak{C}^*(b,T)} \Longleftrightarrow \boxed{\max_{b\in\mathbb{S}^{n-1}} \lambda_1\left(P(b)P(b)^\top\right)} \tag{4}$$

where $P(b)P(b)^\top = \sum_{j=1}^n p_j(A)bb^\top p_j(A)^\top$ and $p_j(A) \sim A^{n-j} + l.o.t$

Thank you for your attention!

Colaborators:


- E. Zuazua (FAU/ Deusto/UAM), C. Esteve-Yagüe (UAM/Deusto), D. Pighin (PhD @ UAM, 2020), D. Maity (TIFR Bangalore).

ConFlex
www.conflex.org

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 765579.