Artificial And Natural Intelligence Toulouse Institute

2023-2024 highlights

see aniti.univ-toulouse.fr/recherche-ia/highlights-2023-2024/

Serge Gratton (Scientific Director)

November 24th, 2024

ANITI The Scientific committee and its bureau

•**The SC** meets every 3 months, **the board** every 2 weeks.
•**Responsibilities**: Manages the scientific program and strategy of the cluster, with validation by the Steering Committee.
•**Key decisions taken by the CS**:

- Launching integrative programs,
- Appointment of associate chairs,
- Inclusion of new members in existing chairs,
- Planning major cluster events e.g., ANITI Days, ANITI Seminars -- next is **December 6th**, 2p.m. at MFJA near B612
- Propose **highlights: please do send information the CS whenever something significant happens to you**

Current members:

- S. Gratton , N. Asher, J. Bolte, J.F. Bonnefon, J.M. Loubes, N. Mansard, Th. Schiex, Th. Serre, S. Thiébaux, R. VanRullen, L. Travé
- M. Allain-Moulet, G. Flandin, C. Merle, J. Sprauel
- R. Redon, C. Joffre

ANITI

Paper in machine learning avenues

- **Important core venues:** Neurips (9 papers), Math prog 2, ICLR, IJCAI, ECAI
- **Nature** collection: Nature Human Behaviour, Nature Communications, Nature Machine Intelligence
- Louis Goupil, Louise Travé-Massuyès, Elodie Chanthery, Thibault Kohler, Sébastien Delautier, Tree based Diagnosis Enhanced with Meta Knowledge Applied to Dynamic Systems. Safeprocess 2024 **Best Theory Paper** Paul M. Frank Award.
- IEEE Signal Processing **Society Sustained Impact Paper 2023 Award**, for Emmanuel Vincent, Rémi Gribonval, and Cédric Févotte for “Performance Measurement in Blind Audio Source Separation”, IEEE Transactions on Audio, Speech, and Language Processing, July 2006.

ANITI This year highlights

The **Lagrange Prize in Continuous Optimization** is awarded every three years by the Mathematical Optimization Society (MOS) and SIAM for an outstanding contribution in the area of continuous optimization.

Congratulation to Jérôme Bolte and Edouard Pauwels.

ANITI

This year highlights

IUF positions

Sylvain Cussat-Blanc

François Bachoc

Robin Bouclier

ANITI

Important projects / libraries

ERC – Advanced Grant GLOW (PI: VanRullen, 2023-2028): "GLoW: The Global Latent Workspace: towards AI models of flexible cognition" 2.5M€

Horizon 2025 WeatherGenerator (2025-2029) : Developing a foundation model for the Earth system, PI ECMWF, 15 partners incl. Météo-France and EXPLEARTH PI, 15M€

Former ANITI Chair : **ERC of Cesar Hidalgo**, Umberto Grandi, Ulle Endriss, Majia Setala, 9M€

One vision: Computational alignment of deep neural networks with humans Funding agency: **NSF**
Role: co-PI (**Serre**/Linsley) Award: $1,190,678 Duration: 2024–2028

Brain-inspired deep learning models of visual reasoning Funding agency: **ONR Grant**
Role: PI (**Serre**) Award: $2,478,465 Duration: 2023–2028

Libraries: xplique 76k download (648 l), scikit decide 155k (142 l.), Pinnochio 1.6M (1900 l.)

ANITI

Thomas Serre distinguished talk

ANITI

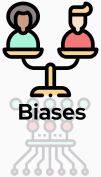
Franck Iutzeler distinguished talk

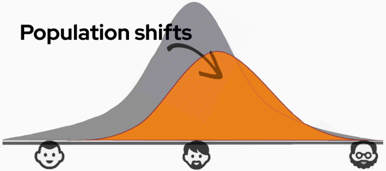
Reliable Machine Learning with Distributional Robustness

Franck IUTZELER
Univ. Toulouse III – Institut de Mathématiques & chair TRIAL

ANITI Days '24

- **Pressing issues** from **public** + **academics** + **industry**





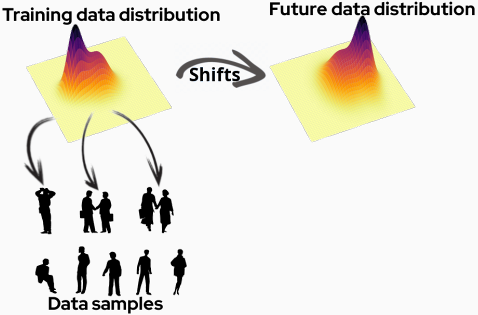
- **Mathematical modeling** of trustworthiness...
  - What **lifecycle changes** can one be **resilient** to?
  - How to **evaluate expected performance**?

Interplay between **robust stochastic optimization** and **machine learning**

- ... fitting the **operational constraints** of AI
  - **Provably robust** but **not too pessimistic**
  - Efficient **open-source implementation**

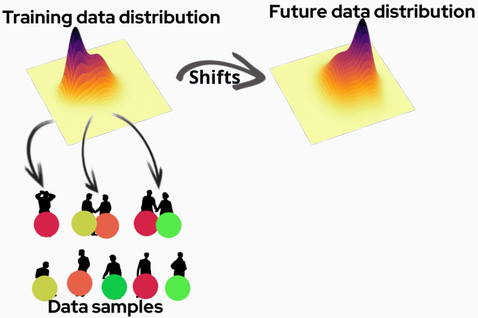
Bridge the **theoretical vision** of reliability with the public's **practical expectations**

- **Training** a model for **reliable performance**
  - Can only use collected **samples** from some unknown **training distribution**
  - Target a **low error** (eg. squared, logistic loss) on **average** for **future data**



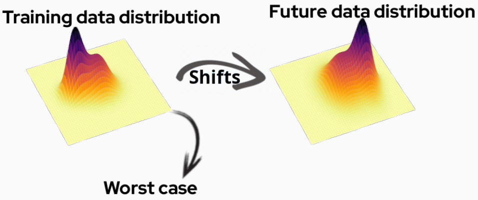
- **Training** a model for **reliable performance**
  - Can only use collected **samples** from some unknown **training distribution**
  - Target a **low error** (eg. squared, logistic loss) on **average** for **future data**

- **Classical** quantities optimized
  - **Empirical error over the data** → overly optimistic about future + replicates biases



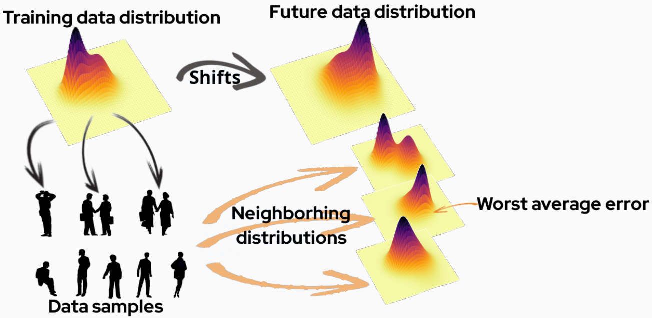
- **Training** a model for **reliable performance**
  - Can only use collected **samples** from some unknown **training distribution**
  - Target a **low error** (eg. squared, logistic loss) on **average** for **future data**

- **Classical** quantities optimized
  - **Worst error possible** → pessimistic + data agnostic



- **Training** a model for **reliable performance**
  - Can only use collected **samples** from some unknown **training distribution**
  - Target a **low error** (eg. squared, logistic loss) on **average** for **future data**

- A **sweet spot** in between: **Distributional robustness**
  - Infer a **distribution** with a **similar performance** as future → data-driven + trustworthy



Optimize model parameters

Average error of parametrized model on the robust distribution of data points

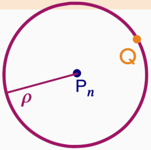
$$\min_{\theta} \sup_{Q} \mathbb{E}_{\xi \sim Q}\left[L_{\theta}(\xi) \right]$$

$$\text{s.t. } W(P_n, Q) \leq \rho$$

Eg. $L_{\theta}(\xi = (x, y)) = (\langle \theta, x \rangle - y)^2$

Worst distribution within some Wasserstein distance of the empirical distribution of the data

$$P_n := \frac{1}{n} \sum_{i=1}^{n} \delta_{\xi_i}$$



- The perfect tool for **statistical reliability**
  - **Generalization** of the performance to unseen samples
  - **Resilience to shifts** between training and future data
  - **Future performance** is controlled

- **Limitations of classical WDRO**
  - **Numerically out-of-reach** in most situations **Eg.** Linear regression ✅ Neural nets ❌
  - **Modeling gaps** with reality **Eg.** High dimension ❌ Text, images ❌

- We proposed a **differentiable approximation** of Wasserstein distributional robustness
  - Built on a **entropic regularization** of the WDRO problem

    W. Azizian, F. Iutzeler, J. Malick : **Regularization for Wasserstein Distributionally Robust Optimization**, ESAIM: Control, Optimisation, and Calculus of Variations, 2023. ArXiv 2205.08826

  - Benefiting from **generalization** and **shift-resilience** guarantees

    W. Azizian, F. Iutzeler, J. Malick : **Exact Generalization Guarantees for (Regularized) Wasserstein Distributionally Robust Models**, NeurIPS 2023. ArXiv 2305.17076

  - Implemented in a **Python library** with both **sklearn** estimators and **torch** wrappers

    github.com/iutzeler/skwdro + pip / conda

    F. Vincent, W. Azizian, F. Iutzeler, J. Malick : **skwdro: a library for Wasserstein distributionally robust machine learning**, preprint, 2024. ArXiv 2410.21231

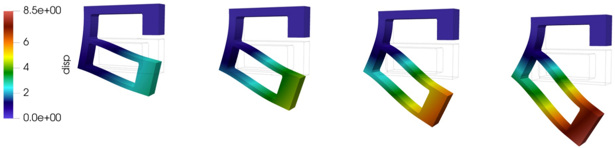
ANITI Alena Kopanicakova distinguished talk

ANITI

PRÉSENTATION
2024

Hybridization of large-scale numerical simulations with SciML

Parametric partial differential equations (PDEs)



- Example of application areas:
  - Uncertainty quantification
  - Design optimization
  - Inverse problems

Let $\Omega \in R^d, B \in R^p$, and $V = V(\Omega)$ be a Hilbert space.
The parametric problem is given in abstract strong form as:

For a given $\beta \in B$, find the solution $u(\beta) \in V$, s.t.

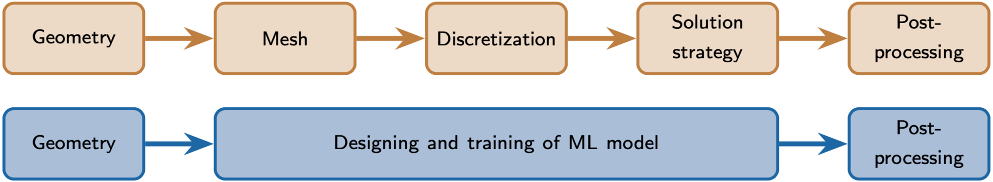
$$P(\beta)\, u(\beta) = f(\beta), \text{ in } V',$$

where $P(\beta): V \to V'$ is a differential operator and $f(\beta)$ is a linear continous form.

- Example of parametrizations:
  - Material parameters
  - Source terms
  - Boundary conditions
  - Geometry

Zhang, Kahana, Kopaničáková, Turkel, Ranade, Pathak, Karniadakis. Blending neural operators and relaxation methods in PDE numerical solvers. Nature Machine Intelligence. 2024
Kopaničáková, Karniadakis. DeepONet-based preconditioning strategies for solving parametric linear systems of equations. SIAM Journal on Scientific Computing. 2024

Numerical solution of parametric PDEs



Classical numerical methods:

- High-fidelity solution using advanced numerical methods, such as FEM, FDM, …
- High-computational cost, especially as problem size grows
- Capture physics on multiple scales
- Encounter challanges in data integration

Scientific machine-learning:

- Exceptionally cheap low-fidelity surrogates
- Requires expensive training phase
- Allows to discover unknown systems dynamics
- Seamless integration of data

Zhang, Kahana, Kopaničáková, Turkel, Ranade, Pathak, Karniadakis. Blending neural operators and relaxation methods in PDE numerical solvers. Nature Machine Intelligence. 2024
Kopaničáková, Karniadakis. DeepONet-based preconditioning strategies for solving parametric linear systems of equations. SIAM Journal on Scientific Computing. 2024

ANITI

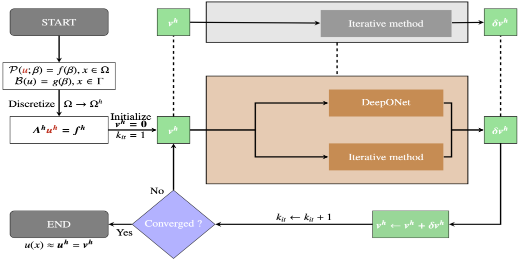
Hybridization of iterative methods with SciML

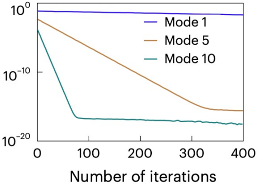
Goal:

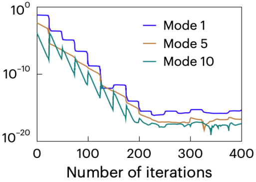
- High accuracy
- Low computational cost
- Algorithmic scalability

Observations:

- « Classical » iterative methods eliminate the high-frequency components of the error
- Operator learning approaches suffer from the spectral bias, i.e., eliminate the low-frequency components of the error quickly, while not capable to remove high-frequency components of the error







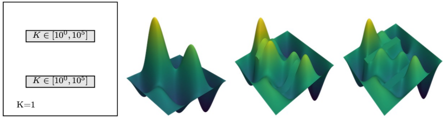
Zhang, Kahana, Kopaničáková, Turkel, Ranade, Pathak, Karniadakis. Blending neural operators and relaxation methods in PDE numerical solvers. Nature Machine Intelligence. 2024

Kopaničáková, Karniadakis. DeepONet-based preconditioning strategies for solving parametric linear systems of equations. SIAM Journal on Scientific Computing. 2024

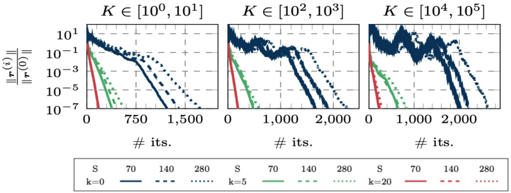
Numerical results (DeepONet augmented ASM)

Diffusion with jumping coefficients:

$$-\nabla \cdot (K(\boldsymbol{x},\boldsymbol{\theta})\nabla u(\boldsymbol{x})) = f(\boldsymbol{x},\boldsymbol{\theta}), \forall \boldsymbol{x} \in \Omega,$$
$$u(\boldsymbol{x}) = 0, \text{ on } \partial\Omega,$$

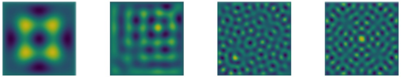


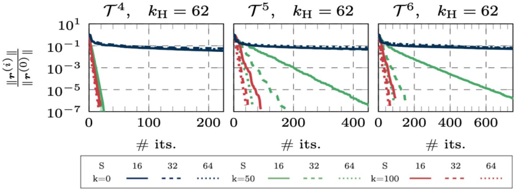
Convergence of GMRES (50) preconditioned with two-level ASM (S subdomains). Symbol k denotes a number of TB functions.



Helmholtz equation:

$$-\Delta u(\boldsymbol{x}) - k_H^2 u(\boldsymbol{x}) = f(\boldsymbol{x},\boldsymbol{\theta}), \forall \boldsymbol{x} \in \Omega,$$
$$u(\boldsymbol{x}) = 0, \text{ on } \partial\Omega,$$





Zhang, Kahana, Kopaničáková, Turkel, Ranade, Pathak, Karniadakis. Blending neural operators and relaxation methods in PDE numerical solvers. Nature Machine Intelligence. 2024
Kopaničáková, Karniadakis. DeepONet-based preconditioning strategies for solving parametric linear systems of equations. SIAM Journal on Scientific Computing. 2024