AI and PHYSICAL MODELS

Coordinators: F. Gamboa and S. Gratton

November 17, 2023

ANITI

Université de Toulouse

Chairs involved

Chairs

1. AI for physical models with geometrical tools (Fabrice Gamboa IMT-UT3)
2. Data Assimilation and Machine Learning (Serge Gratton IRIT-ENSEIHT)

Covering a total of

- Permanent researchers *(Inner circle only)*: **9**
- Post Doc student: **1**
- Ph. D students: **20**

Publications

- Journal articles 26
- Conference proceedings 8
- Books 1

Group outlines

How we work

1. Recurrent meeting of the whole group, each 2-3 weeks:
   - What's up ?
   - One hour work group around a member's talk
2. Semester one day workshop of the group. Example: First one, June 2021 organized in Foix by P. Boudier (NVIDIA).

Scientific contours

- At the crossroads of
  - Numerical Analysis (deterministic approaches)
  - Stochastic modelling and statistics
  - Optimisation
  - Computer sciences
- Strong interactions with the industrial world (majority of the PhDs are on topics proposed by industrial partners)

Partners

Industry

- RFence (startup specializing in intrusion detection)
- EDF ("Electricité De France") and CEA ("Commisariat de l'Energie Atomique")
- ATOS (1 CIFRE on DA and 1 on robustness)
- BRLi (1 CIFRE on the prediction of floods)
- Airbus (discussion of IA for simulations)
- RENAULT, SAFRAN, LIEBHERR, CONTINENTAL, NXP, ...

International collaborations

- Belgium, Columbia, Hong Kong, Ireland, Germany, Morocco, Uruguay, USA.

Theme perimeter: Improving physical models simulation via ML

Sensitivity analysis and Computer code experiments

- Sensitivity analysis backgrounds for dimension reduction, interpretability and fairness
- Geometric optimisation problems for reliability of solutions produced by solvers
- Bayesian assimilation of complex computer codes
- Robustness wrt adversarial attacks

Optimization for ML

- Robust methods for stochastic problem
- Methods beyond first order
- Complexity analysis

Solving physical equations with ML

- Architectures informed by Physics
- Training constrained by Physics
- Use of tools from statistics for variable selection and model reduction
- HPC techniques for large scale problems
- Prediction in chaotic systems

People (I)

Permanent

- T. Pellegrini, R. Chhaibi, P. Boudier, S. Gurol, A. Buttari, S. Zhang, C. Lapeyre
- Associated members: A. Lagnoux, C. Pellegrini, F. Costantino, N. Savy.

Postdoc

- A. Fillion

External collaborations

- Ph.L. Toint (Namur)
- A. Kopaničáková (Brown)

People (II)

PhD students

⚡: Lightning talk
*: CIFRE. ♦: Joint.

Sensitivity Analysis and Robustness

Clément Benesse
Louis Berry*
Virgile Foy*
Faouzi Hakimi*
Marouane Il Idrissi*
Baalu Ketema*
Eva Lawrence
Damien Remot*
Jérôme Stenger*

Optimization for ML

Anirban Bose ♦
Sadok Jerad
Ismail Khalfaoui
Alexey Lazarev ⚡
Mehdi Zouitine*

Surrogates in ML

Théo Beuzeville*
Valentin Mercier*
Vy Nguyen*,♦
Serigne Daouda Pene*
Mathis Peyron*
Justin Reverdi*,♦ ⚡

A selection of topics

I. Optimal optimization Methods in ML
   - Fast training algtorithms (*OMS 2021, SIMAX 2020, SIOPT 2020*)
   - Parallel solution of PDEs using physical domain decomposition techniques: improving scalability with global information exchange (*submitted*)

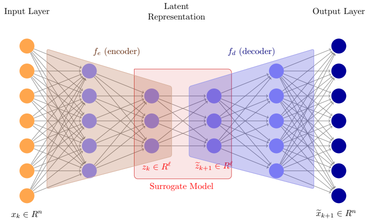
II. Prediction in time dependent physical systems
   - Outperforming Data Assimilation algorithms with ML (*QJRMS, 2021*)
   - Beating Data Assimilation algorithms with Data driven approach with VAE (*QJRMS, 2021*)

III. Fundamental and applied basis for sensitivity analysis (*IJUQ 2020, RESS 2021, Book SIAM 2021, Bernoulli 2022*)

IV. Reinforcement learning for satellite planning (NeurIPS 2023).

I. Improving DA with ML

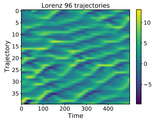
- DA refers to **prediction** of a time dependent variable combining models and data. DA: variational (projection), (small) ensemble



- DA crucially exploits **dimensionality reduction**.
- Use auto encoder to discover **a latent geometry in the physical space**
- Model reduction for DA using autoencoders and surrogate networks **jointly**.

- Not only exploit the latent structure, but also **outperform the accuracy** of the results obtained by the best DA algorithms. Results on the Lorenz-96 system of equations (**Publised in QJRMS**), and on the 2-layer quasi geostrophic model of an operational code (OOPS).

Name	RMSE	Inflation	σ_Q/σ_{Q_ℓ}
ETKF-Q (Reference)	0.193	1.15	0.05
ETKF-Q with surrogate model	0.225	1.116	0.1
Latent ETKF-Q with AE (Ours)	**0.177**	**1.018**	$\mathbf{5.10^{-5}}$
Latent ETKF-Q with PCA	0.329	1.082	0.1



A possible invariant...

ODS autonomous $\Longrightarrow$ DA autonomous

Kalman filter analysis

The posterior mean and covariance depend on time

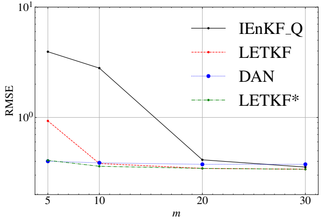
$$\Sigma_t^a = \left(H^\top R^{-1} H + \left(\Sigma_t^b\right)^{-1}\right)^{-1},$$
$$\mu_t^a = \mu_t^b + \Sigma_t^a H^\top R^{-1}\left(y_t - \mathscr{H}\left(\mu_t^b\right)\right),$$

but the analysis transformation does not !

$$a = \left(\mu_t^b, \Sigma_t^b\right), y_t \mapsto \left(\mu_t^a, \Sigma_t^a\right)$$

DA can be learned as a recurrent net.

$$a^\theta \in M \times Y \to M, \qquad \text{(analysis)}$$
$$b^\theta \in M \to M, \qquad \text{(propagation)}$$
$$g^\theta \in M \to \text{Gaussians}_X,$$



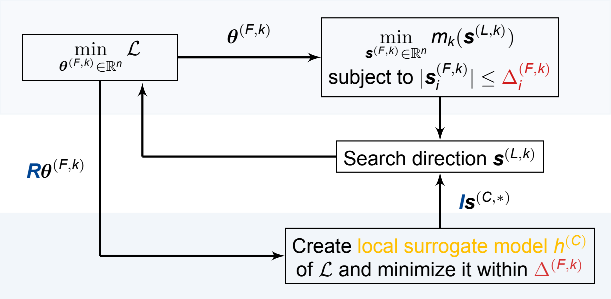
- Published in Journal of Advances in Modelling Earth Sciences (JAMES).

We consider the unconstrained nonlinear programming problem:
Given dataset $\mathcal{D} = \{(\boldsymbol{x}_j, \boldsymbol{c}_j)\}_{j=1}^{p}$, find parameters $\boldsymbol{\theta}$ of DDN as

$$\min_{\boldsymbol{\theta} \in \mathbb{R}^n} \mathcal{L}(\boldsymbol{\theta}) := \frac{1}{p} \sum_{j=1}^{p} \ell(\mathrm{DNN}(\boldsymbol{x}_j, \boldsymbol{\theta}), \boldsymbol{c}_j) + \mathcal{R}(\boldsymbol{\theta})$$

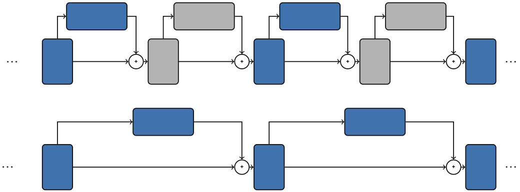
- Training is typically performed using 1st-order OFFO methods, e.g. AdaGrad, Adam, ...
- We propose a class of multilevel OFFO algorithms, which contains a momentum-less Adagrad
  - **OFFO:** reduced sensitivity to the sub-sampling noise
  - **Multi-level:** reduced computational cost

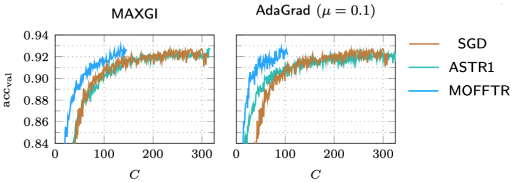
Multilevel Objective Function Free Trust-Region:



MOFFTR:

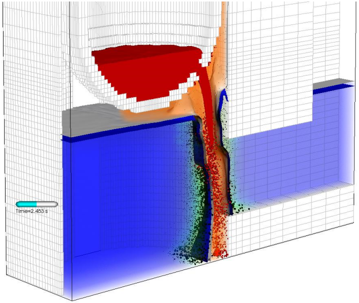
- Step size is controlled on all levels by means of $\Delta_i^{(F,k)} := |(\nabla \mathcal{L}(\boldsymbol{\theta}^{(F,k)}))_i| / \boldsymbol{w}_i^{(F,k)}$, where
$$\zeta_i (k+1)^{\nu} \leq \boldsymbol{w}_i^{(F,k)} \leq \kappa_w (k+1)^{\mu},$$
or
$$\boldsymbol{w}_i^{(k)} := \left(\zeta_i + \sum_{j=0}^{k} (\nabla \mathcal{L}(\boldsymbol{\theta}^{(F,k)}))^2 \right)^{\mu}$$
- Evaluation complexity (noiseless): At most $\mathcal{O}(\epsilon^{-2})$ iterations are needed to reduce $\|\nabla \mathcal{L}(\boldsymbol{\theta})\|$ below ϵ
- Numerically tested by training ResNets for supervised learning applications
- Published in SIOPT '23





III. GSA : (1) For understanding code failures

- Studied code: MC3D used by CEA. Simulates and computes the interaction between fuel and coolant for in severe nuclear accidents.
- Exploring the code behaviour: Runs with random samples in the input space.
- Problem: Large amount of runs (1/3) fail to converge



Work Objective: Analyze the sampled data to understand which of the inputs have the most influence on code failures.
Two methods:

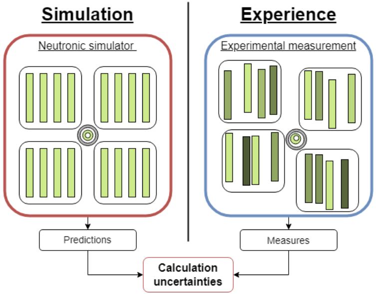
- Conditionally to code failure, goodness-of-fit tests between random input and input samples.
- Dependence measures between each input and occurrence of code failure.

III. GSA: (2.1) Calibration of (twin of) a nuclear reactor

How to model a **digital twin** in order to realize the **experimental validation** of an industrial simulator dedicated to the design of **nuclear core**?

Problematic :

- Large dimension of the calibration problem : 1000 dimensions.
- Time-consuming computer code : Days on multiple processors

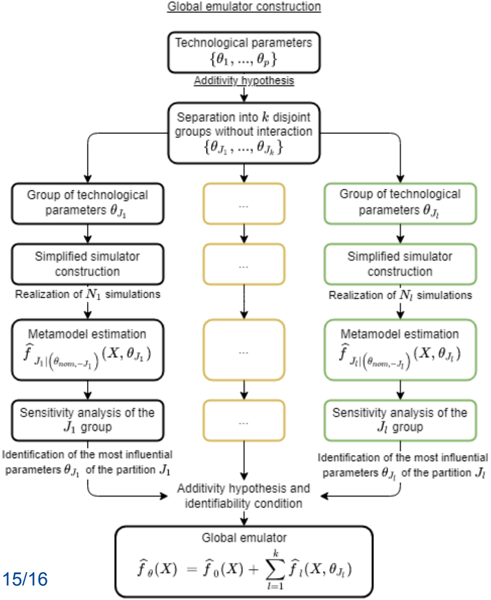


Proposition: Use of a Bayesian methodology to combine real measures and expert knowledge from simulators.

$$z_{x_i} = f_\theta(x_i) + e(x_i) + \delta(x_i)$$

where z are the measurements, f_θ the computer code, δ the simulator error, e the measurement error and θ the parameters to be calibrated.

III. GSA: (2.2) Bayesian calibration of a nuclear reactor



Methodology

- *Break multi-physics into smaller pieces:* Emulator as a sum of sub-emulators.
- *Sample recombination:* Use of **specific** methods from neutronics leading to a significative increase of the effective training datasets.

Results

Possibility to perform the **calibration** of the **digital twin**, based on a small number of Monte-Carlo simulations using **specific neutron methods** and **experimental measurements** on already fabricated reactors.

IV. Reinforcement Learning applied to satellite planning

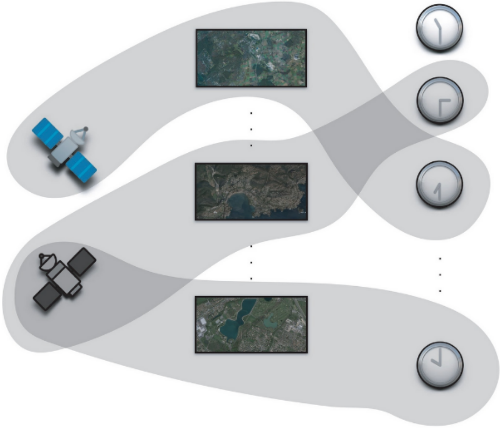


Figure: Satellite Constellation Scheduling: The objective is to match each observation with a suitable satellite and time window, while adhering to a multitude of complex constraints.

Objective

The objective is to explore deep reinforcement learning for developing heuristics for combinatorial optimization problems (COPS).

- Practical application: Satellite constellation scheduling, a complex COP.
- Alongside using Reinforcement Learning for combinatorial optimization problems, we also aim to develop tools to explain our agent's behavior (XAI).