RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

cea cnrs Inria

FRANCE 2030
PROGRAMME DE RECHERCHE
INTELLIGENCE ARTIFICIELLE

PDE-AI

Numerical analysis, optimal control and optimal transport for AI

Antonin Chambolle, CEREMADE, CNRS & Université Paris-Dauphine (PSL)
+ 9 partners in Paris area, Nice, Lyon, Toulouse, Bordeaux, Strasbourg

anr agence nationale de la recherche

Issues: develop mathematical analysis for Machine Learning and AI, leveraging theories of Partial Differential Equations (PDEs), optimal control and optimal transport.

Develop the knowledge and interest for AI of the wide and well established French mathematical analysis community.

The consortium gathers teams from:

- Université Paris-Dauphine, PSL
- Sorbonne Université
- Université Paris-Saclay
- ENSAE (Institut Polytechnique de Paris)
- Université Paris-Cité
- Université de Lyon / CNRS
- Université de Nice Côte d'Azur / INRIA
- Université de Toulouse
- Université de Bordeaux
- Université de Strasbourg (+Nancy)

Research topics & Objectives

Main research directions / objectives

Three main topics:

- Optimisation: dynamics of neural networks training
  - Study large neuron limits and the dynamics of (possibly stochastic) gradient descent schemes on such models (related to optimal control theory);
  - Study the dynamics / aggregation properties of modern architectures such as transformers;
  - Better understand ADAM-type (momentum) accelerations in a Euclidean and non-Euclidean framework;
  - Use tools from gradient flows in metric spaces (in particular, Wasserstein gradient flows) to better understand, analyse and develop learning of probability distributions and statistics.
  - Practical optimisation techniques: development of PnP methods with guarantees, particle methods, mean field methods
- New (very) deep architecture
  - Import tools from the numerical and theoretical analysis to study large recurrent NNs, or transformers-based architectures;
  - Develop a branch of control theory for the learning of large networks;
  - Develop and analyse specific networks with guarantees for difficult numerical analysis and scientific computing tasks or non-Euclidean data (e.g., graph neural networks, also point clouds, manifolds);
  - Better understand PINNs based methods, deep learning for quantum physics, etc

Main research directions / objectives

- Sampling and generative models
  - Develop new sampling strategies leveraging recent classes of optimal transport particle flows;
  - Integrate perturbed dynamics (noise injection, momentum, acceleration);
  - Develop optimal transport methods for the quantization of high dimensional distributions;
  - Investigate Sliced-Wasserstein metrics for the matching of high dimensional distributions;

- As an important outcome, we aim at shaping a community of analysts interested in machine learning problems and techniques, and developing and *promoting a scientific culture* viewing machine learning both as a tool and source of interesting scientific questions in the French mathematical analysis community.

Team roles, achievements, and competencies

Partner's roles, achievements, competencies

- Non convex optimisation: Bordeaux, Toulouse, Dauphine (CEREMADE + LAMSADE), zero order stochastic methods, accelerated first order methods, PnP type methods for machine learning regularized inverse problems, study of inertial methods such as ADAM in the discrete and continuous setting (Hermant-Aujol-Dossal-Rondepierre 2023, Hurault-Leclaire-Papdakis-Chambolle 2023)

- Optimal control: Nice, Strasbourg, Dauphine (Julia "control-toolbox" (https://control-toolbox.org/), Mazari-Nadin-Privat, 2022, Mazari-Ruiz Balet, 2022)

- Numerical analysis (Sorbonne U., Dauphine): ML for transport equation, Approximation properties of deep networks, error estimates (Ancelin, Despres, Jaouen: "Volume of fluid" interface advection schemes boosted by ML techniques, 2022), (Cohen, Dolbeault, Mula, Somacal: Nonlinear approximation spaces for inverse problems, 2023) (Chambolle-Pock 2021 for total variation)

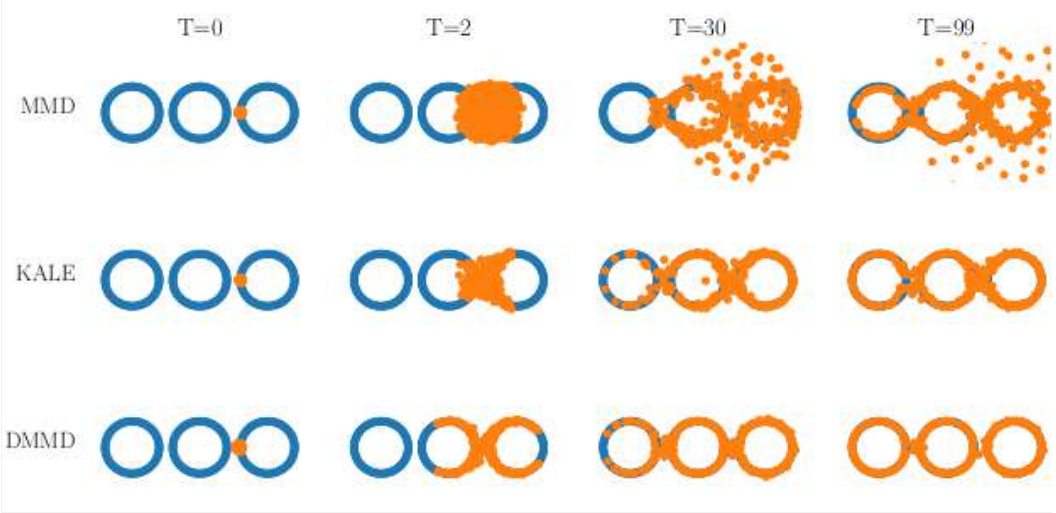
Partner's roles, achievements, competencies

- Discrete sampling: ENSAE, Toulouse, Lyon, Paris-Saclay. Study of convergence and properties of Lloyd's algorithm for measure quantization. (Alternating descent.) (Mérigot, Santambrogio, Sarrazin, 2021: convergence to critical points), (Portales, Cazelles, Pauwels 2023: convergence when density subanalytic) Practical implementations in small dimension (does not scale, nice problem...)

- Wasserstein PCA / Probability distributions as a Riemanian (Otto) manifold (Toulouse [with Paris 1], Paris-Saclay) (Cazelles et al, 2018, Cazelles-Klein-Le Brigant, in preparation); Wasserstein barycenters for large data clustering and analysis (nonlinear / manifold K-means) (Bordeaux, Paris-Saclay, Lyon, Dauphine [Mokaplan])

Partner's roles, achievements, competencies

- Sampling through optimisation of Divergences / Wasserstein or KL gradient flows; Quantification / discrete to continuum limits (ENSAE, possible collaboration with Dauphine) (Xu, Korba, Slepcev, 2022, Carroccia, Chambolle, Slepcev 2020)



Research directions, current progress

Workpackages and current progress

Theme 1 – The dynamics of Neural Networks training:

- One PhD Student from Bordeaux (J.F. Aujol) collaborating with C. Royer and A. Chambolle (Julien Hermant, started 15th Oct, 2023): non-convex optimization

- One Postdoc in Paris-Dauphine (I. Waldspurger, C. Royer) on Nonconvex optimization on measure spaces (Annette Dumas, Oct 2024); one more post-doc should be next hired on Mean-field optimization (collaboration Dauphine / Paris-Saclay: Cardaliaguet, Ren, Kazeykina)

- In Strasbourg, Control in High Dimension and Learning: three PhD students started in 2023 but could not be funded by the project;

- Two Postdocs should be hired next fall in Nice on controllability and geometric properties of neural networks;
- Hirings are planned in Paris-Cité (development of the Keops library) but delayed for adminsitrative reasons.

Workpackages and current progress

Theme 2 – New very deep architectures

- Currently hiring an intern (starting April 2024) in Dauphine (Royer, Chambolle), hoping he will be a good PhD candidate (Eloi Martin), for working on neural PDE training and optimization methods, including blackbox hyperparameter optimization, collaboration with Strasbourg (Privat, Franck) is part of the project;

- Post-doc in Sorbonne Université: the team in SU had hired Borjan Gueshkovski (MIT) who studies the dynamics and behaviour very large networks (such as transformers), no new hiring until now;

- Post-doc in Lyon (E. Bretin and S. Masnou), collaboration with A. Chambolle, (Tokuhiro Eto, hopefully hired next year), to work on NNs based models for geometric flows.

Workpackages and current progress

Theme 3 - Sampling and generative models

- One PhD student ENSAE (A. Korba) (Clémentine Chazal, 2024), sampling as an optimization, design of new loss functions (scalable);
- One post-doc ENSAE (A. Korba) (Clément Bonet, 2023-24), Wasserstein Gradient flows; [2nd year]
- PhD student from Bordeaux (J. Bigot) with Toulouse (E. Cazelles) (Erell Gachon, 1st Nov. 2023): optimal transport for data analysis for medicine.
- PhD student (J. Digne) in Lyon (Camille Buonomo, Sept. 2024) on shape synthesis with topological constraints, and one post-doc (Anatole Gallouët) on Sliced Wasserstein distances for Generative AI.

Outcomes

Outcomes

- Optimisation schemes and techniques, with code;
- Novel architectures, better understanding of existing/new architecture and their training, guarantees in application oriented networks (PINNs, numerical simulation, quantum chemistry or other high dimensional scientific computing problems);
- PDEs based sampling methods, based on optimal transport theory and related tools;
- Development of an AI culture in the French mathematical analysis community, and a true interest in solving AI related problems (high dimensional analysis and PDEs, large scale optimization)

Dissemination

Mostly: scientific papers and code.

Typical example: Python KeOps package, for large scale data processing (such as scalable convolution operation on large point clouds), developed by the team at Université Paris-Cité (with INRIA / Montpellier), or the Julia control-toolbox from Toulouse and Nice's groups.

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

cea cnrs Inria

FRANCE 2030

PROGRAMME DE RECHERCHE

INTELLIGENCE ARTIFICIELLE

Retrouvez toutes nos actualités

anr agence nationale de la recherche