ANITI

ARTIFICIAL & NATURAL INTELLIGENCE TOULOUSE INSTITUTE

Large scale optimization for AI

Jérôme Bolte

Université Fédérale
Toulouse Midi-Pyrénées

J. Bolte (UT1, PI)

F. Malgouyres (UT3)

E. Pauwels (UT3)

Mathematical challenges

ANITI
ARTIFICIAL & NATURAL INTELLIGENCE
TOULOUSE INSTITUTE

- Design fast and simple algorithms for AI
- Provide theoretical guarantees and progress towards certification
- Mathematical models for neural networks and AI
- Towards new non convex theories

Deep Learning and Mathematics

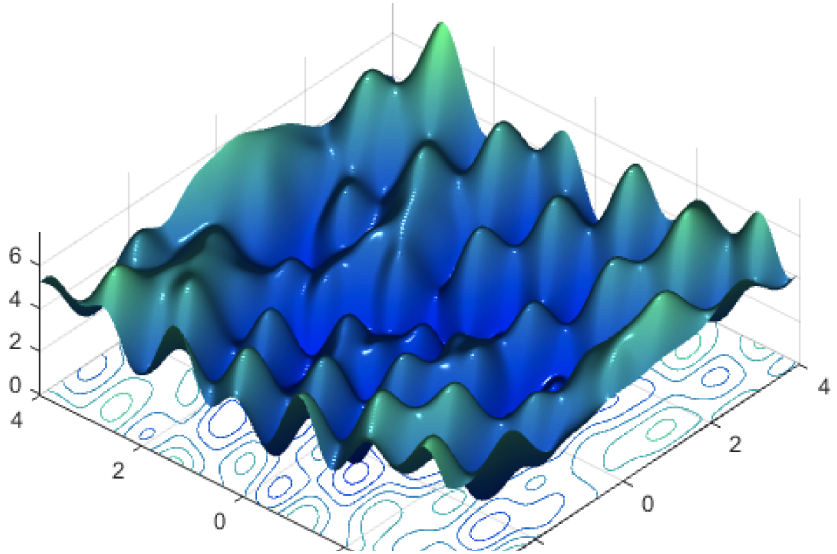


Figure: Find the lowest point in 2D can be delicate... In Deep Learning one can have millions of dimensions.

- Geometry and learning rates
- Geometry of large landscapes
- Non-convex theories for optimization
- Infinite dimensional models
- Symmetries and algorithms
- Stability of minimizers

From theory to coding

ANITI
ARTIFICIAL & NATURAL INTELLIGENCE
TOULOUSE INSTITUTE

- Mathematical models for training softwares
- Learning and activation functions: models for a slower but deeper learning?
- Construction of stable architectures
- Matching architectures to structured data

The extended team and collaborators

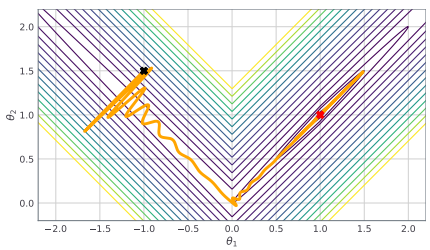
ANITI
ARTIFICIAL & NATURAL INTELLIGENCE
TOULOUSE INSTITUTE

- A. d'Aspremont, ENS
- A. Blanchet, IAST
- J. Bolte, UT1, PI
- Joaquim Bonna-Pelissier, PhD ANITI
- C. Castera, Phd, INP
- R. Dragomir, Phd UT1-ENS
- A. El Mehdi, PhD UT1/ANITI
- C. Févote, INP
- F. de Gournay, UT3
- L. Glaudin, Post-doc ANITI
- K. Kurdyka, U. de Savoie-Mont Blanc
- R. Rios-Zertuche, Post-Doc ANITI
- S. Sabach, The Technion
- L. Vanderberghe, UCLA
- M. Teboulle, Tel Aviv University
- P. Weiss, ITAV
- ...

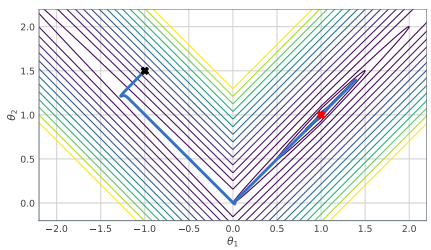
First algorithm: INDIAN "*An Inertial Newton Algorithm for Deep Learning*"

INDIAN algorithm with Castera, Bolte, Févotte and Pauwels

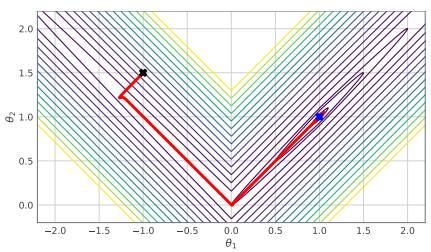
- Reach the best current performances,
- State of the art generalization properties,
- Avoid parasite oscillations.



(a) $\alpha = 0.5,\ \beta = 0.01$



(b) $\alpha = 0.5,\ \beta = 0.1$



(c) $\alpha = 1.3,\ \beta = 0.1$