Optimisation Globale à Base de Particules & Formalisation

Gaëtan Serré

Journée Maths en Herbe

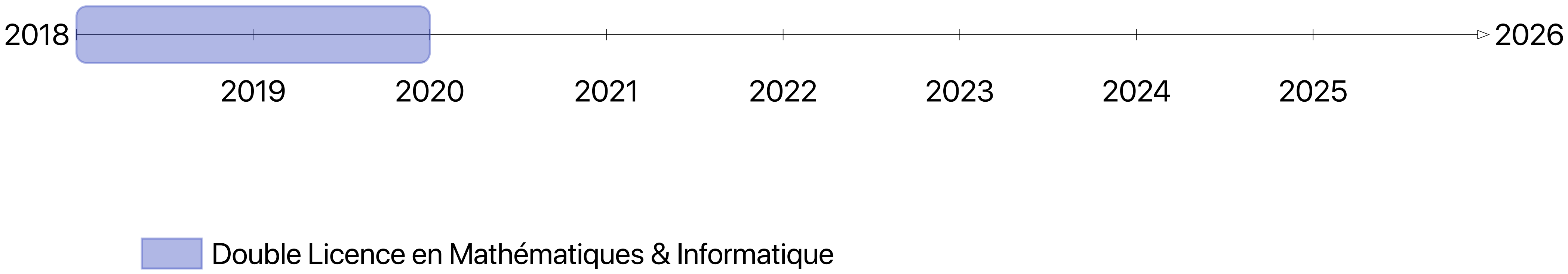
21 January 2026

Learning and Information Processing Systems

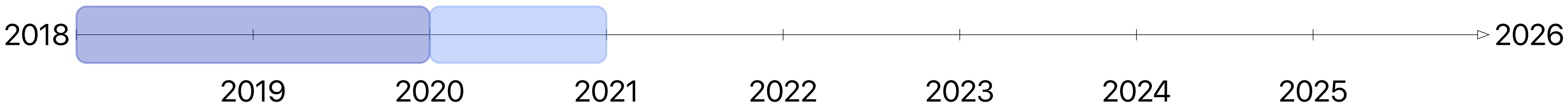
gaetan.serre@ens-paris-saclay.fr

école normale supérieure paris—saclay

Parcours

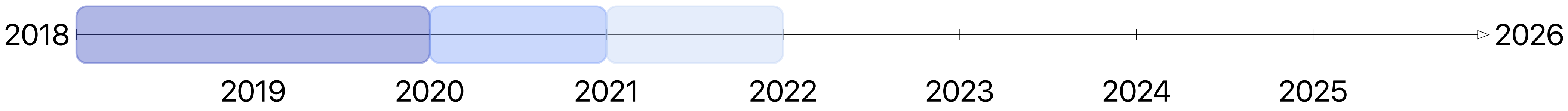

2018
2026
2019
2020
2021
2022
2023
2024
2025
Double Licence en Mathématiques & Informatique

Parcours



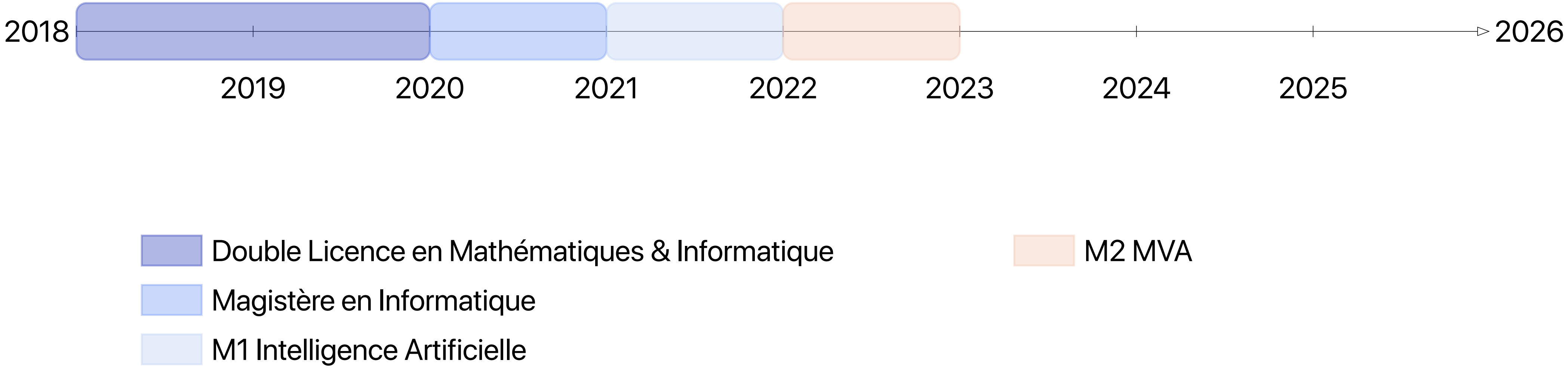
Double Licence en Mathématiques & Informatique
Magistère en Informatique

Parcours

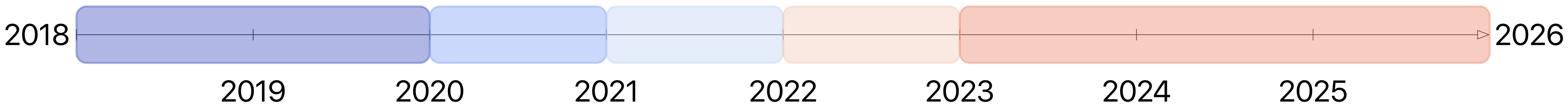


- Double Licence en Mathématiques & Informatique
- Magistère en Informatique
- M1 Intelligence Artificielle

Parcours


2018
2026
2019
2020
2021
2022
2023
2024
2025
Double Licence en Mathématiques & Informatique
Magistère en Informatique
M1 Intelligence Artificielle
M2 MVA

Parcours



Double Licence en Mathématiques & Informatique

Magistère en Informatique

M1 Intelligence Artificielle

M2 MVA

Doctorat en Mathématiques Appliquées

Optimisation Globale

Définition classique

Définition – *Optimisation globale*

Étant donné un ensemble compact $\Omega \subset \mathbb{R}^d$ et une fonction continue $f : \Omega \to \mathbb{R}$, le but de tout algorithme d'optimisation globale est de trouver un point $x^\star \in \Omega$ tel que

$$x^\star \in \underset{x \in \Omega}{\operatorname{argmin}} \ f(x).$$

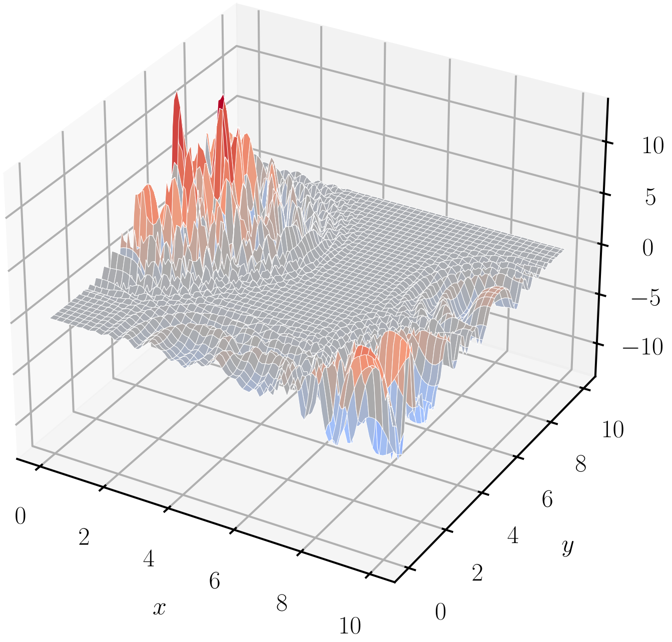
Optimisation Globale

Définition classique

Définition – *Optimisation globale*

Étant donné un ensemble compact $\Omega \subset \mathbb{R}^d$ et une fonction continue $f : \Omega \to \mathbb{R}$, le but de tout algorithme d'optimisation globale est de trouver un point $x^\star \in \Omega$ tel que

$$x^\star \in \operatorname*{argmin}_{x \in \Omega} f(x).$$



Exemple d'une fonction multimodale

Optimisation Globale

Lien avec l’échantillonnage

L’optimisation globale peut être vue comme un problème d’échantillonnage (Hwang, 1980; Kirkpatrick et al., 1983).

Optimisation Globale

Lien avec l'échantillonnage

L'optimisation globale peut être vue comme un problème d'échantillonnage (Hwang, 1980; Kirkpatrick et al., 1983).

$\mathrm{Boltzmann}_f(\kappa)$ a la densité suivante :

$$\rho_\kappa(x) = \frac{\exp(-\kappa f(x))}{\int_\Omega \exp(-\kappa f(y)) \, \mathrm{d}y}.$$

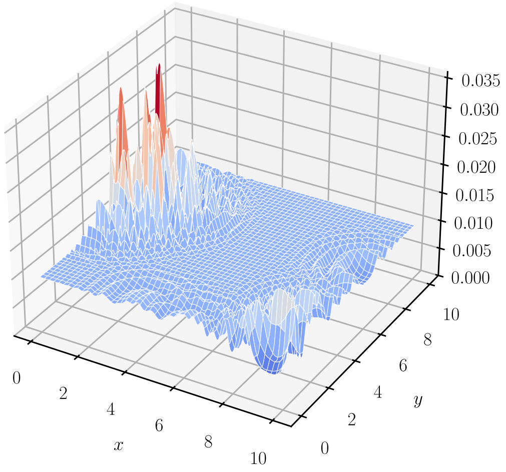
Optimisation Globale

Lien avec l'échantillonnage

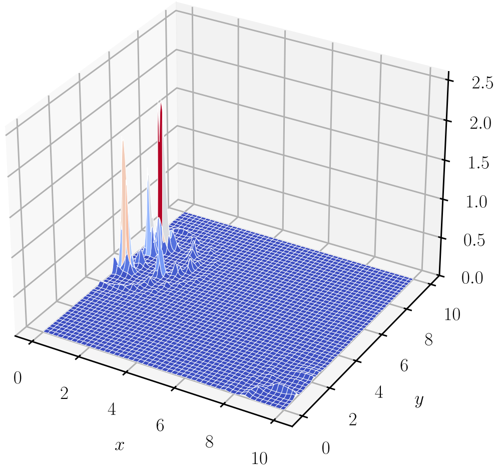
L'optimisation globale peut être vue comme un problème d'échantillonnage (Hwang, 1980; Kirkpatrick et al., 1983).

$\mathrm{Boltzmann}_f(\kappa)$ a la densité suivante :

$$\rho_\kappa(x) = \frac{\exp(-\kappa f(x))}{\int_\Omega \exp(-\kappa f(y))\,\mathrm{d}y}.$$



$\rho_{0.1}$



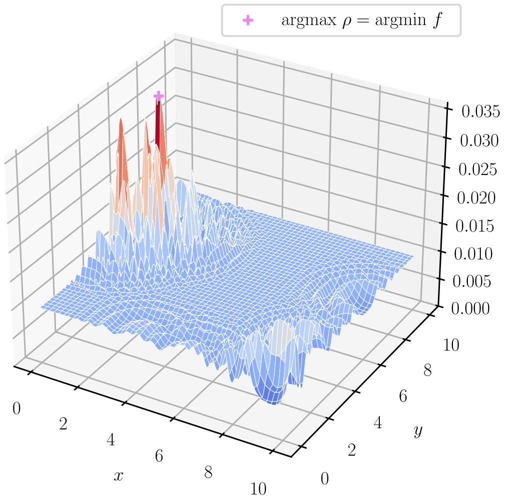
$\rho_{0.5}$

Optimisation Globale

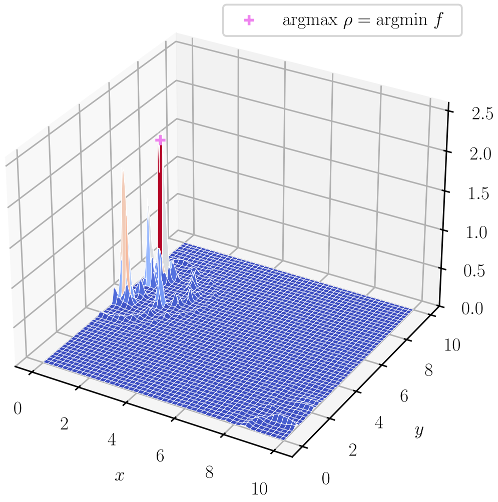
Lien avec l'échantillonnage

L'optimisation globale peut être vue comme un problème d'échantillonnage (Hwang, 1980; Kirkpatrick et al., 1983).

La limite de $\mathrm{Boltzmann}_f(\kappa)$ quand $\kappa \to \infty$ est supportée par l'ensemble des minimiseurs de f.



$\rho_{0.1}$



$\rho_{0.5}$

Systèmes de particules

Une approche moderne de l'optimisation globale

Concevoir des systèmes de particules en interaction qui se comportent asymptotiquement comme la distribution de Boltzmann (ex. Pinnau et al. (2017); Serré et al. (2025)).

Systèmes de particules

Une approche moderne de l'optimisation globale

Concevoir des systèmes de particules en interaction qui se comportent asymptotiquement comme la distribution de Boltzmann (ex. Pinnau et al. (2017); Serré et al. (2025)).

$$\mathrm{d}X_t^i = \left[\frac{1}{N}\sum_{j=1}^{N} -\kappa\left(X_t^i, X_t^j\right) k \nabla f\left(X_t^j\right) + \nabla_{X_t^j}\kappa\left(X_t^i, X_t^j\right)\right]\mathrm{d}t$$

où κ est un noyau symétrique défini positif
k est un paramètre de température

SBS

Systèmes de particules

Une approche moderne de l'optimisation globale

Concevoir des systèmes de particules en interaction qui se comportent asymptotiquement comme la distribution de Boltzmann (ex. Pinnau et al. (2017); Serré et al. (2025)).

$$\mathrm{d}X_t^i = \left[\frac{1}{N}\sum_{j=1}^{N} -\kappa\left(X_t^i, X_t^j\right)k\nabla f\left(X_t^j\right) + \nabla_{X_t^j}\kappa\left(X_t^i, X_t^j\right)\right]\mathrm{d}t$$

où κ est un noyau symétrique défini positif

k est un paramètre de température

SBS

$$\mathrm{d}X_t^i = -\lambda\left(X_t^i - v_f\right)\mathrm{d}t + \gamma\left\|X_t^i - v_f\right\|_2 \mathrm{d}B_t^i$$

où $v_f = \dfrac{\sum_{i=1}^N X_t^i e^{-\alpha f\left(X_t^i\right)}}{\sum_{i=1}^N e^{-\alpha f\left(X_t^i\right)}}$

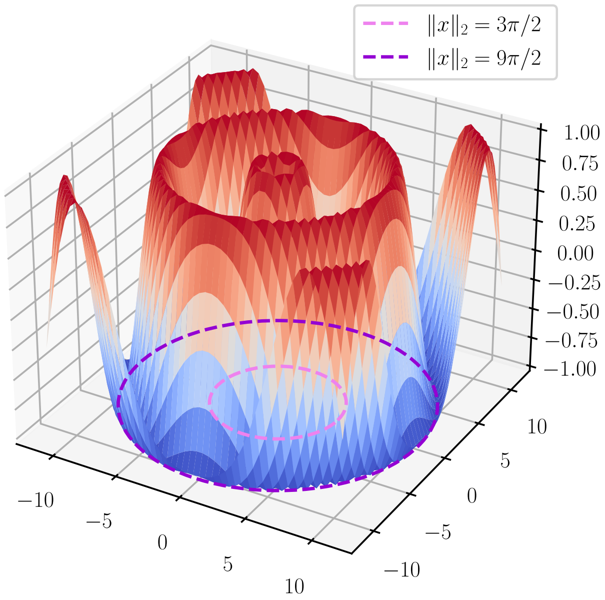
λ, γ, α sont des paramètres positifs

CBO

Systèmes de particules

Une approche moderne de l'optimisation globale

Nous pouvons visualiser le champ vectoriel associé à ces systèmes de particules :

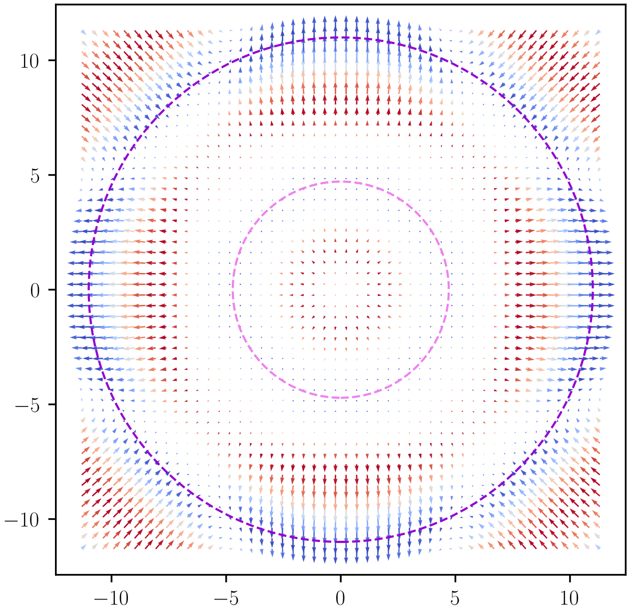


$x : \mathbb{R}^2 \mapsto \sin(\|x\|_2)$

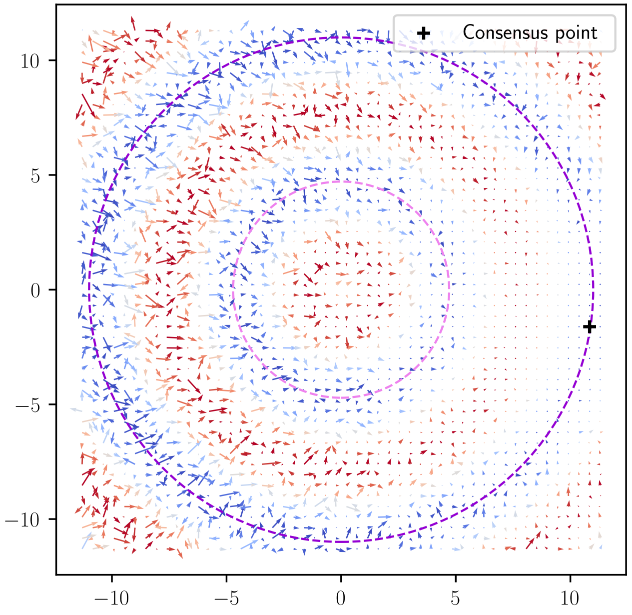
Systèmes de particules

Une approche moderne de l'optimisation globale

Nous pouvons visualiser le champ vectoriel associé à ces systèmes de particules :



SBS $(t \triangleq 0)$

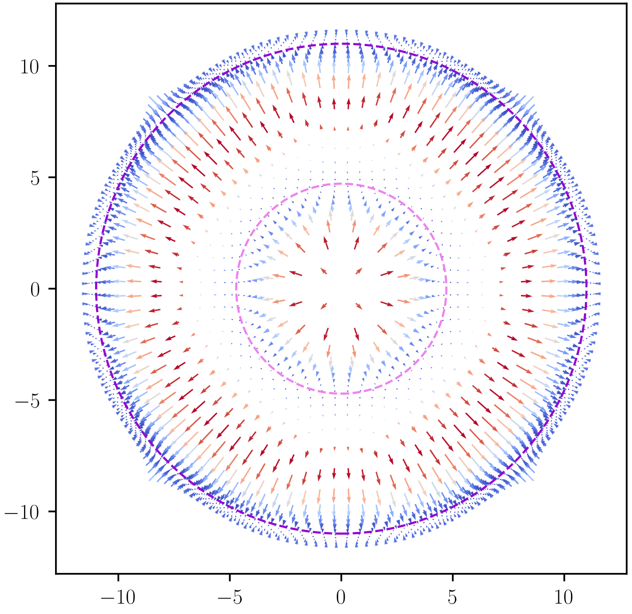


CBO $(t \triangleq 0)$

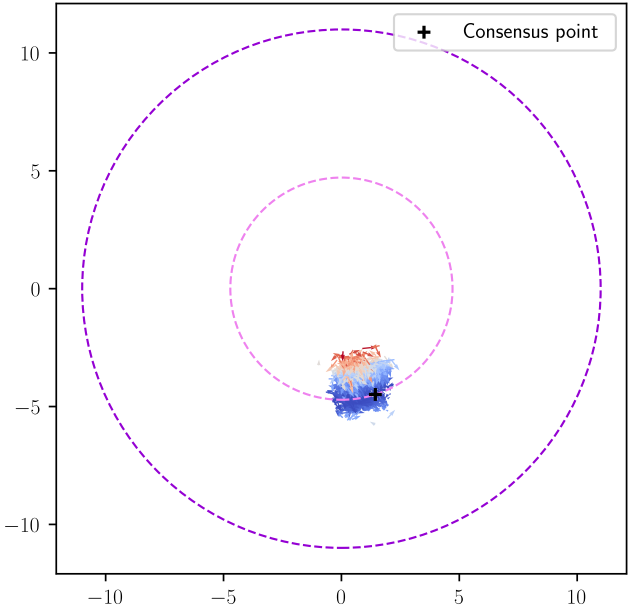
Systèmes de particules

Une approche moderne de l'optimisation globale

Nous pouvons visualiser le champ vectoriel associé à ces systèmes de particules :



SBS $(t \triangleq 700)$



CBO $(t \triangleq 250)$

Systèmes de particules

Abstraction commune

« L'abstraction […] ne consiste pas à aller vers le plus compliqué, mais bien au contraire à aller vers le plus simple. »

Claire Voisin, *Faire des mathématiques*

Systèmes de particules

Abstraction commune

« L'abstraction [...] ne consiste pas à aller vers le plus compliqué, mais bien au contraire à aller vers le plus simple. »

Claire Voisin, *Faire des mathématiques*

Systèmes d'interaction de McKean-Vlasov :

$$\mathrm{d}X_t^i = \boldsymbol{b}(X_t^i, \hat{\mu}_t)\,\mathrm{d}t + \boldsymbol{\sigma}(X_t^i, \hat{\mu}_t)\,\mathrm{d}B_t^i$$

où

- $t \in \overline{\mathbb{R}}_+$, $i \in [\![1, N]\!]$;
- $\hat{\mu}_t$ est la mesure empirique des particules au temps t ;
- B_t^i sont des mouvements browniens indépendants ;
- $\boldsymbol{b}$ est le *drift* et $\boldsymbol{\sigma}$ est la *diffusion*.

Travailler avec des assistants de preuve

Les assistants de preuve sont des logiciels pour écrire et prouver des énoncés mathématiques.
(ex. Rocq, Lean, Isabelle/HOL)

Dans quel but ?

- garantir la correction des résultats mathématiques ;
- améliorer la transparence et la reproductibilité ;
- récemment, pour construire des systèmes d’IA avec raisonnement formel (ex. Harmonic, Numina).

Travailler avec des assistants de preuve

Les assistants de preuve sont des logiciels pour écrire et prouver des énoncés mathématiques.
(ex. Rocq, Lean, Isabelle/HOL)

Dans quel but ?

- garantir la correction des résultats mathématiques ;
- améliorer la transparence et la reproductibilité ;
- récemment, pour construire des systèmes d'IA avec raisonnement formel (ex. Harmonic, Numina).

Formalisation de résultats et définitions autour de l'optimisation globale en utilisant L∃∀N.

L∃∀N

Un tutoriel

L∃∀N (avec Mathlib), est un assistant de preuve puissant pour écrire et prouver des énoncés mathématiques :

L∃∀N

Un tutoriel

L∃∀N (avec Mathlib), est un assistant de preuve puissant pour écrire et prouver des énoncés mathématiques :

L∃∀N

```
import Mathlib

example : StrictMono (fun (x : ℝ) ↦ x + 1) := by
```

LƎ∀N

Un tutoriel

LƎ∀N (avec Mathlib), est un assistant de preuve puissant pour écrire et prouver des énoncés mathématiques :

LƎ∀N

```
import Mathlib

example : StrictMono (fun (x : ℝ) ↦ x + 1) := by
  intro x y hxy
  /-
   x y : ℝ
   hxy : x < y
   ⊢ x + 1 < y + 1
  -/
```

LEAN

Un tutoriel

LEAN (avec Mathlib), est un assistant de preuve puissant pour écrire et prouver des énoncés mathématiques :

LEAN

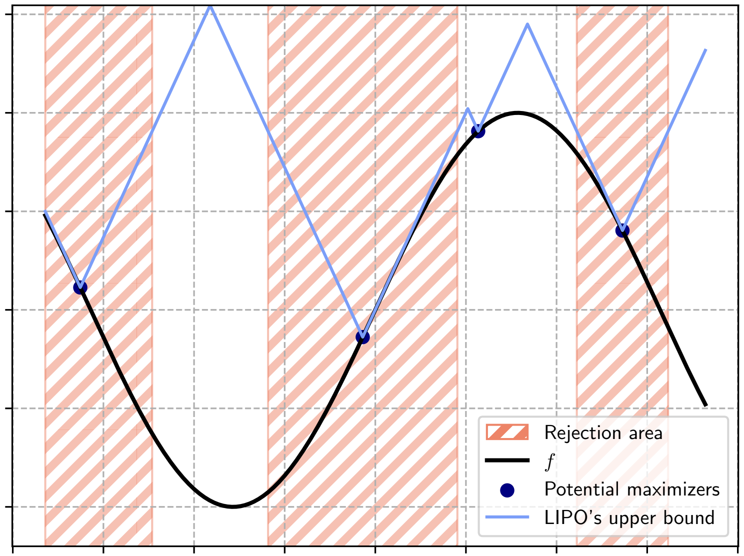
```
import Mathlib

example : StrictMono (fun (x : ℝ) ↦ x + 1) := by
  intro x y hxy
  /-
   x y : ℝ
   hxy : x < y
   ⊢ x + 1 < y + 1
  -/
  exact (add_lt_add_iff_right 1).mpr hxy
  /-
   - add_lt_add_iff_right (a : ℝ) : x + a < y + a ↔ x < y
  -/
```

LƎ∀N

En pratique – LIPO

Étant donnée une fonction de Lipschitz, LIPO construit itérativement une borne supérieure pour identifier les régions où la fonction **ne peut pas** être maximisée (Malherbe & Vayatis, 2017).



Borne supérieure de LIPO après 4 itérations

Des résultats théoriques complexes sur les algorithmes d'optimisation globale peuvent être formalisés en LΞ∀N.

Par exemple, nous avons formalisé le résultat suivant sur l'algorithme LIPO (Serré et al., 2024) :

Théorème – *Probabilité de rejet de LIPO (Serré et al., 2024)*

Pour toute fonction κ-Lipschitz f et tout $x \in \Omega$, soit $\mathscr{R}(x, t)$ l'événement de rejet de x au temps t. Alors, on a :

$$\mathbb{P}(\mathscr{R}(x, t+1)) \leq \frac{t\pi^{d/2}\Delta^d}{\kappa^d \Gamma(d/2+1)\lambda(\Omega)},$$

où $\Delta \triangleq \max_{x\in\Omega} f(x) - \min_{x\in\Omega} f(x)$.

Merci pour votre attention !

Bibliographie

Hwang, C.-R. (1980). Laplace's method revisited: weak convergence of probability measures. *The Annals of Probability*.

Kirkpatrick, S., Gelatt Jr, C. D., & Vecchi, M. P. (1983). Optimization by Simulated Annealing. *Science*.

Malherbe, C., & Vayatis, N. (2017). Global optimization of Lipschitz functions. *International conference on machine learning*, 2314-2323.

Pinnau, R., Totzeck, C., Tse, O., & Martin, S. (2017). A Consensus-Based Model for Global Optimization and Its Mean-Field Limit. *Mathematical Models and Methods in Applied Sciences*.

Serré, G., Beja-Battais, P., Chirrane, S., Kalogeratos, A., & Vayatis, N. (2024). LIPO+: Frugal Global Optimization for Lipschitz Functions. *Proceedings of the 13th Hellenic Conference on Artificial Intelligence*.

Serré, G., Kalogeratos, A., & Vayatis, N. (2025). Stein Boltzmann Sampling: A Variational Approach for Global Optimization. *Proceedings of The 28th International Conference on Artificial Intelligence and Statistics*.