

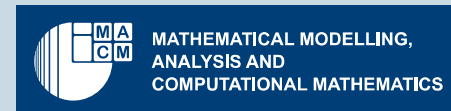
Consensus-based global optimization methods in finite and infinite dimensions

Claudia Totzeck

University of Wuppertal

9th German-Polish Conference
of Optimization Methods and Applications
Bialystok, Poland

September 22, 2026



Contents

- 1) **Non-convex global optimization and Metaheuristics**
- 2) **Consensus-based optimization: motivation and main ideas**
- 3) **Variants and current trends**

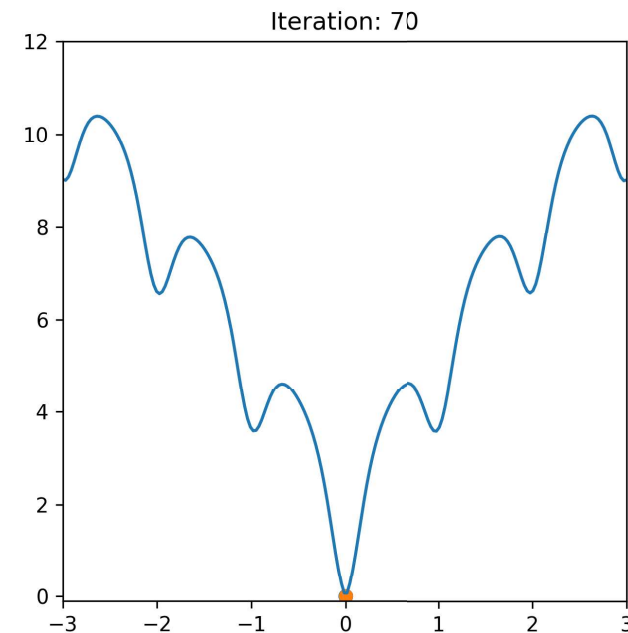
Metaheuristics

Goal: For a given continuous function $f: \mathbb{R}^d \rightarrow \mathbb{R}^+$, find

$$\bar{x} = \mathbf{argmin}_{x \in \Omega} f(x)$$

Metaheuristics:

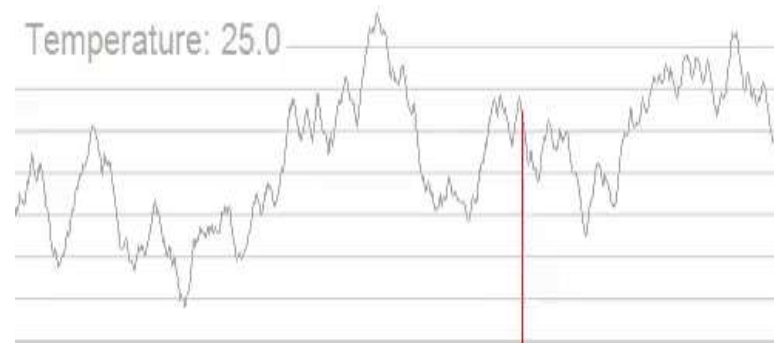
- ▶ Random search, Rastrigin (1963)
- ▶ Simplex heuristic, Nelder–Mead (1965)
- ▶ Evolutionary programming, Fogel et al. (1966)
- ▶ Metropolis–Hastings, Hastings (1970)
- ▶ Genetic algorithm, Holland (1975)
- ▶ Simulated annealing, Kirkpatrick et al. (1983)
- ▶ Ant colony, Dorigo (1992)
- ▶ Particle swarm, Kennedy et al. (1995)



Metaheuristics*

General properties:

- ✓ Allow for irregular objective functions
- No usage of gradient information
- Incorporates stochasticity
- Agent-based, collective behavior
- ✓ Resilient to finding local minimum
- ⊗ No guarantee of finding global minimum



Example: Simulated annealing (1983)

$$dX_t = -\nabla f_t(X_t)dt + (1/\beta_t)dW_t, \quad \beta_t \text{ optimal cooling schedule}$$

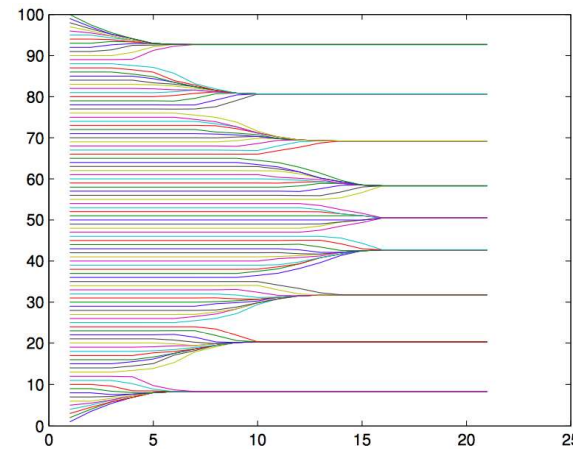
discrete versions use an acceptance rate e^{-1/β_t}

*T.S. Chiang et al. (1987), R. Holley et al. (1988, 1989)

Consensus formation[†]

Main features:

- ▶ Involves multiple agents
- ▶ Each agent may adapt other opinions
- ▶ Agents may be reluctant to change
- ▶ Individual can influence neighbors



Example: Abelson's (averaging) model (1964) $\rightsquigarrow$ Cucker–Smale

$$X_{t+1}^i = \sum_j a_{ij}(X_t^i - X_t^j) = (AX_t)^i, \quad a_{ij} \text{ rate of influence of } X_t^i \text{ over } X_t^j$$

Example: Hegselmann–Krause (bounded confidence) model (2002)

$$X_{t+1}^i = \frac{1}{|N_t^i|} \sum_{j \in N_t^i} X_t^j, \quad N_t^i = \{j \mid |X_t^i - X_t^j| \leq \epsilon\}$$

[†]S. Chartterjee, E. Senata (1975, 1977), R. Hegselmann, U. Krause (2002),

Consensus-based optimization: motivation and main ideas

joint work with



José A. Carrillo
U Oxford



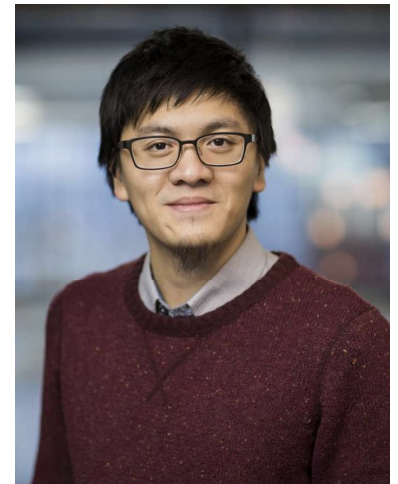
Young-Pil Choi
Yonsei U



Stephan Martin
SAP



René Pinnau
RPTU KL



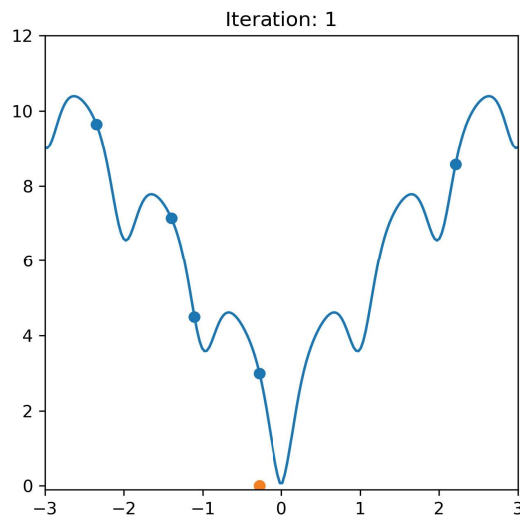
Oliver Tse
TU Eindhoven

Introduction to CBO - motivation

Let $f: \mathbb{R}^D \rightarrow \mathbb{R}^+, D \geq 1$, be a **continuous** function that admits a **unique** minimizer $\bar{x}$.

We consider $\operatorname{argmin}_{x \in \mathbb{R}^D} f(x)$.

Task: Construct an algorithm that finds the global minimizer of f !



Features we like:

- ▶ global minimum
- ▶ no gradient information
- ▶ particle scheme
- ▶ indistinguishable particles
- ▶ mean-field equation for analysis

trade-off: simplicity & efficiency

Consensus-Based Global Optimization (CBO) [Pinnau, T, Tse, Martin 2017]

Combine **swarm optimization** and **opinion dynamics** to obtain:

$$dX_t^i = -\lambda(X_t^i - v_f)dt + \sigma|X_t^i - v_f|dB_t^i, \quad i = 1, \dots, N,$$

where the **weighted average** v_f is given by

$$v_f = \frac{\sum_{i=1}^N X_t^i \omega_f^\alpha(X_t^i)}{\sum_{i=1}^N \omega_f^\alpha(X_t^i)}, \quad \text{with} \quad \omega_f^\alpha(x) = \exp(-\alpha f(x)),$$

supplemented with random initial data $\rho_0 = \text{law}(X_0^i)$. We use the notation

f objective function,

λ drift parameter,

σ diffusion parameter,

α weight parameter,

B_t^i Brownian Motion.

Note:

- v_f is our candidate for the global minimizer (Laplace principle!)
- for explicit schemes: interactions scale **linear** in N

Relationship of v_f and the Laplace principle

Theorem 1 (Approximation of the global minimizer)

Assume that $f \in \mathcal{C}_b(\mathbb{R}^d, \mathbb{R}^+)$ attains a unique global minimum at x_* and let $\rho \in \mathcal{P}^{ac}(\mathbb{R}^d)$ with $x_* \in \text{supp}(\rho)$. Then we have that $\lim_{\alpha \rightarrow \infty} v_f^\alpha = x_*$.

Proof: Let $f_* := f(x_*)$. By construction $\eta^\alpha = e^{-\alpha f(x)} \rho(x) / \|e^{-\alpha f}\|_{L^1(\rho)} \in \mathcal{P}^{ad}(\mathbb{R}^d)$. We first show that the functional

$$\mathcal{E}_\alpha(f - f_*) := \int_{\mathbb{R}^d} f d\eta^\alpha \longrightarrow 0 \quad \text{as } \alpha \rightarrow \infty.$$

Indeed, it holds

$$\begin{aligned} \frac{d}{d\alpha} \mathcal{E}_\alpha(f - f_*) &= \frac{d}{d\alpha} \int f - f_* d\eta^\alpha(x) \\ &= - \int f(x)^2 d\eta^\alpha(x) + \int f(x) d\eta^\alpha(x) \int f(y) d\eta^\alpha(y) \\ &= -\frac{1}{2} \int |f(x) - f(y)|^2 d\eta^\alpha(x) d\eta^\alpha(y) < 0 \quad (\text{strictly as } f \text{ nonconst}) \end{aligned}$$

$\mathcal{E}_\alpha(f - f_*) \geq 0$ for all α and 0 is largest lower bound $\Rightarrow \mathcal{E}_\alpha(f - f_*) \rightarrow 0$ as $\alpha \rightarrow \infty$.

Relationship of v_f and the Laplace principle

Theorem 1 (Approximation of the global minimizer)

Assume that $f \in \mathcal{C}_b(\mathbb{R}^d, \mathbb{R}^+)$ attains a unique global minimum at x_* and let $\rho \in \mathcal{P}^{ac}(\mathbb{R}^d)$ with $x_* \in \text{supp}(\rho)$. Then we have that $\lim_{\alpha \rightarrow \infty} v_f^\alpha = x_*$.

Proof: Let $f_* := f(x_*)$. By construction $\eta^\alpha = e^{-\alpha f(x)} \rho(x) / \|e^{-\alpha f}\|_{L^1(\rho)} \in \mathcal{P}^{ad}(\mathbb{R}^d)$. We first show that the functional

$$\mathcal{E}_\alpha(f - f_*) := \int_{\mathbb{R}^d} f d\eta^\alpha \longrightarrow 0 \quad \text{as} \quad \alpha \rightarrow \infty.$$

Indeed, it holds

$$\begin{aligned} \frac{d}{d\alpha} \mathcal{E}_\alpha(f - f_*) &= \frac{d}{d\alpha} \int f - f_* d\eta^\alpha(x) \\ &= - \int f(x)^2 d\eta^\alpha(x) + \int f(x) d\eta^\alpha(x) \int f(y) d\eta^\alpha(y) \\ &= -\frac{1}{2} \int |f(x) - f(y)|^2 d\eta^\alpha(x) d\eta^\alpha(y) < 0 \quad (\text{strictly as } f \text{ nonconst}) \end{aligned}$$

$\mathcal{E}_\alpha(f - f_*) \geq 0$ for all α and 0 is largest lower bound $\Rightarrow \mathcal{E}_\alpha(f - f_*) \rightarrow 0$ as $\alpha \rightarrow \infty$.

Relationship of v_f and the Laplace principle

Theorem 1 (Approximation of the global minimizer)

Assume that $f \in \mathcal{C}_b(\mathbb{R}^d, \mathbb{R}^+)$ attains a unique global minimum at x_* and let $\rho \in \mathcal{P}^{ac}(\mathbb{R}^d)$ with $x_* \in \text{supp}(\rho)$. Then we have that $\lim_{\alpha \rightarrow \infty} v_f^\alpha = x_*$.

Proof: Let $f_* := f(x_*)$. By construction $\eta^\alpha = e^{-\alpha f(x)} \rho(x) / \|e^{-\alpha f}\|_{L^1(\rho)} \in \mathcal{P}^{ad}(\mathbb{R}^d)$. We first show that the functional

$$\mathcal{E}_\alpha(f - f_*) := \int_{\mathbb{R}^d} f d\eta^\alpha \longrightarrow 0 \quad \text{as} \quad \alpha \rightarrow \infty.$$

Indeed, it holds

$$\begin{aligned} \frac{d}{d\alpha} \mathcal{E}_\alpha(f - f_*) &= \frac{d}{d\alpha} \int f - f_* d\eta^\alpha(x) \\ &= - \int f(x)^2 d\eta^\alpha(x) + \int f(x) d\eta^\alpha(x) \int f(y) d\eta^\alpha(y) \\ &= -\frac{1}{2} \int |f(x) - f(y)|^2 d\eta^\alpha(x) d\eta^\alpha(y) < 0 \quad (\text{strictly as } f \text{ nonconst}) \end{aligned}$$

$\mathcal{E}_\alpha(f - f_*) \geq 0$ for all α and 0 is largest lower bound $\Rightarrow \mathcal{E}_\alpha(f - f_*) \rightarrow 0$ as $\alpha \rightarrow \infty$.

Relationship of v_f and the Laplace principle

Theorem 1 (Approximation of the global minimizer)

Assume that $f \in \mathcal{C}_b(\mathbb{R}^d, \mathbb{R}^+)$ attains a unique global minimum at x_* and let $\rho \in \mathcal{P}^{ac}(\mathbb{R}^d)$ with $x_* \in \text{supp}(\rho)$. Then we have that $\lim_{\alpha \rightarrow \infty} v_f^\alpha = x_*$.

Proof: Let $f_* := f(x_*)$. By construction $\eta^\alpha = e^{-\alpha f(x)} \rho(x) / \|e^{-\alpha f}\|_{L^1(\rho)} \in \mathcal{P}^{ad}(\mathbb{R}^d)$. We first show that the functional

$$\mathcal{E}_\alpha(f - f_*) := \int_{\mathbb{R}^d} f d\eta^\alpha \longrightarrow 0 \quad \text{as } \alpha \rightarrow \infty.$$

Indeed, it holds

$$\begin{aligned} \frac{d}{d\alpha} \mathcal{E}_\alpha(f - f_*) &= \frac{d}{d\alpha} \int f - f_* d\eta^\alpha(x) \\ &= - \int f(x)^2 d\eta^\alpha(x) + \int f(x) d\eta^\alpha(x) \int f(y) d\eta^\alpha(y) \\ &= -\frac{1}{2} \int |f(x) - f(y)|^2 d\eta^\alpha(x) d\eta^\alpha(y) < 0 \quad (\text{strictly as } f \text{ nonconst}) \end{aligned}$$

$\mathcal{E}_\alpha(f - f_*) \geq 0$ for all α and 0 is largest lower bound $\Rightarrow \mathcal{E}_\alpha(f - f_*) \rightarrow 0$ as $\alpha \rightarrow \infty$.

Relationship of v_f and the Laplace principle

Theorem 1 (Approximation of the global minimizer)

Assume that $f \in \mathcal{C}_b(\mathbb{R}^d, \mathbb{R}^+)$ attains a unique global minimum at x_* and let $\rho \in \mathcal{P}^{ac}(\mathbb{R}^d)$ with $x_* \in \text{supp}(\rho)$. Then we have that $\lim_{\alpha \rightarrow \infty} v_f^\alpha = x_*$.

Proof: Let $f_* := f(x_*)$. By construction $\eta^\alpha = e^{-\alpha f(x)} \rho(x) / \|e^{-\alpha f}\|_{L^1(\rho)} \in \mathcal{P}^{ad}(\mathbb{R}^d)$. We first show that the functional

$$\mathcal{E}_\alpha(f - f_*) := \int_{\mathbb{R}^d} f d\eta^\alpha \longrightarrow 0 \quad \text{as } \alpha \rightarrow \infty.$$

Indeed, it holds

$$\begin{aligned} \frac{d}{d\alpha} \mathcal{E}_\alpha(f - f_*) &= \frac{d}{d\alpha} \int f - f_* d\eta^\alpha(x) \\ &= - \int f(x)^2 d\eta^\alpha(x) + \int f(x) d\eta^\alpha(x) \int f(y) d\eta^\alpha(y) \\ &= -\frac{1}{2} \int |f(x) - f(y)|^2 d\eta^\alpha(x) d\eta^\alpha(y) < 0 \quad (\text{strictly as } f \text{ nonconst}) \end{aligned}$$

$\mathcal{E}_\alpha(f - f_*) \geq 0$ for all α and 0 is largest lower bound $\Rightarrow \mathcal{E}_\alpha(f - f_*) \rightarrow 0$ as $\alpha \rightarrow \infty$.

Relationship of v_f and the Laplace principle

Theorem 1 (Approximation of the global minimizer)

Assume that $f \in \mathcal{C}_b(\mathbb{R}^d, \mathbb{R}^+)$ attains a unique global minimum at x_* and let $\rho \in \mathcal{P}^{ac}(\mathbb{R}^d)$ with $x_* \in \text{supp}(\rho)$. Then we have that $\lim_{\alpha \rightarrow \infty} v_f^\alpha = x_*$.

Proof: Let $f_* := f(x_*)$. By construction $\eta^\alpha = e^{-\alpha f(x)} \rho(x) / \|e^{-\alpha f}\|_{L^1(\rho)} \in \mathcal{P}^{ad}(\mathbb{R}^d)$. We first show that the functional

$$\mathcal{E}_\alpha(f - f_*) := \int_{\mathbb{R}^d} f d\eta^\alpha \longrightarrow 0 \quad \text{as} \quad \alpha \rightarrow \infty.$$

Indeed, it holds

$$\begin{aligned} \frac{d}{d\alpha} \mathcal{E}_\alpha(f - f_*) &= \frac{d}{d\alpha} \int f - f_* d\eta^\alpha(x) \\ &= - \int f(x)^2 d\eta^\alpha(x) + \int f(x) d\eta^\alpha(x) \int f(y) d\eta^\alpha(y) \\ &= -\frac{1}{2} \int |f(x) - f(y)|^2 d\eta^\alpha(x) d\eta^\alpha(y) < 0 \quad (\text{strictly as } f \text{ nonconst}) \end{aligned}$$

$\mathcal{E}_\alpha(f - f_*) \geq 0$ for all α and 0 is largest lower bound $\Rightarrow \mathcal{E}_\alpha(f - f_*) \rightarrow 0$ as $\alpha \rightarrow \infty$.

Relationship of v_f and the Laplace principle

Theorem 2 (Approximation of the global minimizer)

Assume that $f \in \mathcal{C}_b(\mathbb{R}^d, \mathbb{R}^+)$ attains a unique global minimum at x_* and let $\rho \in \mathcal{P}^{ac}(\mathbb{R}^d)$ with $x_* \in \text{supp}(\rho)$. Then we have that $\lim_{\alpha \rightarrow \infty} v_f^\alpha = x_*$.

Proof (continued): We have $\mathcal{E}_\alpha(f - f_*) \rightarrow 0$ as $\alpha \rightarrow \infty$.

In the second step we show that $\eta^\alpha \rightarrow \delta_{x_*}$ in the sense of distributions.

Let $\epsilon > 0$. By Chebyshev inequality we obtain

$$\begin{aligned} \eta^\alpha(\{x \in \mathbb{R}^d : f(x) - f_* \geq \epsilon\}) &\leq \frac{1}{\epsilon} \int_{\{f-f_* \geq \epsilon\}} (f - f_*) d\eta^\alpha(x) \\ &\leq \frac{1}{\epsilon} \mathcal{E}_\alpha(f - f_*) \rightarrow 0 \quad \text{as } \alpha \rightarrow \infty. \end{aligned}$$

Hence, $\eta^\alpha \rightarrow \delta_{x_*}$ which implies $v_f \rightarrow x_*$.

Remark:

This is the reason why the weights $e^{-\alpha f(x)}$ appear so often in global optimization methods!

Relationship of v_f and the Laplace principle

Theorem 2 (Approximation of the global minimizer)

Assume that $f \in \mathcal{C}_b(\mathbb{R}^d, \mathbb{R}^+)$ attains a unique global minimum at x_* and let $\rho \in \mathcal{P}^{ac}(\mathbb{R}^d)$ with $x_* \in \text{supp}(\rho)$. Then we have that $\lim_{\alpha \rightarrow \infty} v_f^\alpha = x_*$.

Proof (continued): We have $\mathcal{E}_\alpha(f - f_*) \rightarrow 0$ as $\alpha \rightarrow \infty$.

In the second step we show that $\eta^\alpha \rightarrow \delta_{x_*}$ in the sense of distributions.

Let $\epsilon > 0$. By Chebyshev inequality we obtain

$$\begin{aligned} \eta^\alpha(\{x \in \mathbb{R}^d : f(x) - f_* \geq \epsilon\}) &\leq \frac{1}{\epsilon} \int_{\{f-f_* \geq \epsilon\}} (f - f_*) d\eta^\alpha(x) \\ &\leq \frac{1}{\epsilon} \mathcal{E}_\alpha(f - f_*) \longrightarrow 0 \quad \text{as } \alpha \rightarrow \infty. \end{aligned}$$

Hence, $\eta^\alpha \rightarrow \delta_{x_*}$ which implies $v_f \rightarrow x_*$.

Remark:

This is the reason why the weights $e^{-\alpha f(x)}$ appear so often in global optimization methods!

Mean-Field Equation

Since we use v_f , the particles are indistinguishable!

Mean-Field system [Huang and Qiu, 2021]

As $N \rightarrow \infty$ the SDE system turns into the McKean non-linear Process

$$d\bar{X}_t = -\lambda(\bar{X}_t - v_f)dt + \sqrt{2}\sigma|\bar{X}_t - v_f|dB_t,$$

with initial data $\text{law}(\bar{X}_0) = \rho_0$. Here the weighted average is given by

$$v_f = \frac{1}{\int_{\mathbb{R}^d} \omega_f^\alpha d\rho_t} \int_{\mathbb{R}^d} x \omega_f^\alpha d\rho_t, \quad \rho_t = \text{law}(\bar{X}_t),$$

which can be expressed equivalently by the **non-local, non-linear** PDE

$$\partial_t \rho_t = \Delta(\kappa \rho_t) + \nabla \cdot (\mu \rho_t),$$

with $\kappa = \sigma^2|x - v_f|^2$ and $\mu = -\lambda(x - v_f)$.

Theory is mostly on the PDE level!

Well-posedness of the PDE

Assumptions:

1. $f: \mathbb{R}^d \rightarrow \mathbb{R}$ is bounded from below with $\underline{f} := \inf f$.
2. There exist constants L_f and $c_u > 0$ such that

$$\begin{cases} |f(x) - f(y)| \leq L_f(|x| + |y|)|x - y| & \text{for all } x, y \in \mathbb{R}^d, \\ f(x) - \underline{f} \leq c_u(1 + |x|^2) & \text{for all } x \in \mathbb{R}^d. \end{cases}$$

Theorem 3 (Existence result for the PDE)

Let f satisfy the above Assumption and have at least quadratic growth at infinity, and $\rho_0 \in \mathcal{P}_4(\mathbb{R}^d)$. Then there exists a unique nonlinear process $\bar{X} \in \mathcal{C}([0, T], \mathbb{R}^d)$, $T > 0$, satisfying

$$d\bar{X}_t = -\lambda(\bar{X}_t - v_f[\rho_t]) dt + \sigma|\bar{X}_t - v_f[\rho_t]| dW_t, \quad \rho_t = \text{law}(\bar{X}_t),$$

in the strong sense, and $\rho \in \mathcal{C}([0, T], \mathcal{P}_2(\mathbb{R}^d))$ satisfies the corresponding PDE (in the weak sense) with $\lim_{t \rightarrow 0} \rho_t = \rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$.

Remark: proof is based on a Schauder fixed point argument. Freeze v_f and iterate. Difficulty is compactness which we obtain by an embedding argument.

Convergence Result [Carrillo, Choi, T, Tse 2018]

Goal

We aim to prove that $\delta_{\hat{x}}$ is the invariant measure of ρ_t with $\hat{x} \in B_\epsilon(x_*)$ for $1 \gg \epsilon > 0$. In other words, that ρ_t concentrates close to x_* as $t \rightarrow \infty$.

Notation: We denote the expectation and variance of ρ_t by

$$E(\rho_t) := \int x d\rho_t \quad \text{and} \quad V(\rho_t) := \frac{1}{2} \int |x - E(\rho_t)|^2 d\rho_t.$$

Strategy of the proof:

- (i) show concentration at $\hat{x}$ for some $\hat{x} \in \mathbb{R}^d$:
compute evolution of $V(\rho_t)$ and $E(\rho_t)$ and use Chebyshev's inequality
- (ii) show $\hat{x} \in B_\epsilon(x_*)$.
similar idea to Laplace principle

Assumption: In the following we assume that $f \in \mathcal{C}^2(\mathbb{R}^d)$ satisfies additionally

1. $\inf f > 0$.
2. $\|\nabla^2 f\|_\infty \leq c_f$ and there exists constants $c_0, c_1 > 0$, such that

$$\Delta f \leq c_0 + c_1 |\nabla f|^2 \quad \text{in } \mathbb{R}^d.$$

Formation of consensus

Altogether, we have

$$\frac{d}{dt} V(\rho_t) \leq -(2\lambda - d\sigma^2 e^{-\alpha \underline{f}} / \|e^{-\alpha f}\|_{L^1(\rho_t)}) V(\rho_t)$$

bounded objective functions

For f bounded, we could estimate the denominator and obtain

$$\frac{d}{dt} V(\rho_t) \leq -(2\lambda - d\sigma^2 e^{-\alpha(\underline{f} - \bar{f})}) V(\rho_t)$$

but this is of no help as we want to send $\alpha \rightarrow \infty$ (Laplace principle) !

With some more work (skipped here) we obtain for λ, σ such that $2\lambda \|\omega_f^\alpha\|_{L^1(\rho_0)}^2 - V(\rho_0) - 2d\sigma^2 \|\omega_f^\alpha\|_{L^1(\rho_0)} e^{-\alpha \underline{f}} \geq 0$ the existence of $q > 0$ and

$$\frac{d}{dt} V(\rho_t) \leq V(\rho_0) e^{-qt}.$$

Remark: This yields concentration as $t \rightarrow \infty$ by Gronwall's Lemma. Note that the condition for λ, σ is independent of α !

Evolution of the expectation

With a similar idea we compute the evolution

$$\frac{d}{dt}E(\rho_t) = -\lambda \int |x - v_f| d\rho_t.$$

which yields

$$\int_0^t \left| \frac{d}{ds}E(\rho_s) \right| ds \leq \frac{4\lambda}{q\|\omega_f^\alpha\|_{L^1(\rho_0)}} e^{-\alpha f} V(\rho_0),$$

in particular, $\frac{d}{dt}E(\rho_t) \in L^1(0, \infty)$.

This allows us to conclude that there exists $\hat{x} \in \mathbb{R}$ with

$$\hat{x} = E(\rho_0) + \int_0^\infty \frac{d}{dt}E(\rho_t) dt = \lim_{t \rightarrow \infty} E(\rho_t).$$

Moreover, since $|E(\rho_t) - v_f[\rho_t]| \rightarrow 0$ for $t \rightarrow \infty$, we get $\lim_{t \rightarrow \infty} v_f[\rho_t] = \hat{x}$.

Remark: This prove the convergence to some *arbitrary* point $\hat{x} \in \mathbb{R}$.

Convergence to minimizer

Note: We are left to show the convergence to the global minimum.

The main idea is very similar to that of the proof $v_f^\alpha \rightarrow x_*$ as $\alpha \rightarrow \infty$. Indeed one shows that

$$\left| \int e^{-\alpha f} d\rho_t - \int_{B_\epsilon(x_*)} e^{-\alpha f} d\rho_t \right| \leq c(V(\rho_t) + |E(\rho_t) - \hat{x}|^2)$$

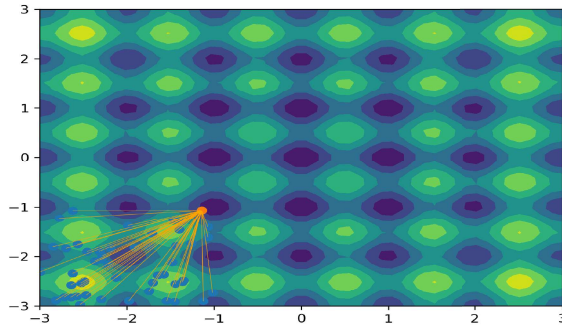
and uses Chebyshev's inequality to see that for α large enough the concentration point is close to the global minimizer of f .

Theorem 4 (Uniform Consensus)

Let $\rho_0 \in \mathcal{P}_4(\mathbb{R}^D)$ and $f \in \text{Lip}_{loc}(\mathbb{R}^D)$, $\inf f > 0$ and all of the above assumptions hold.

Then there exist constants $\alpha \ll 1$ and $\lambda, \sigma > 0$ such that we obtain uniform consensus for ρ_t as $t \rightarrow \infty$ arbitrary close to the global minimizer.

Purpose-driven particle systems (PdIPS) on github



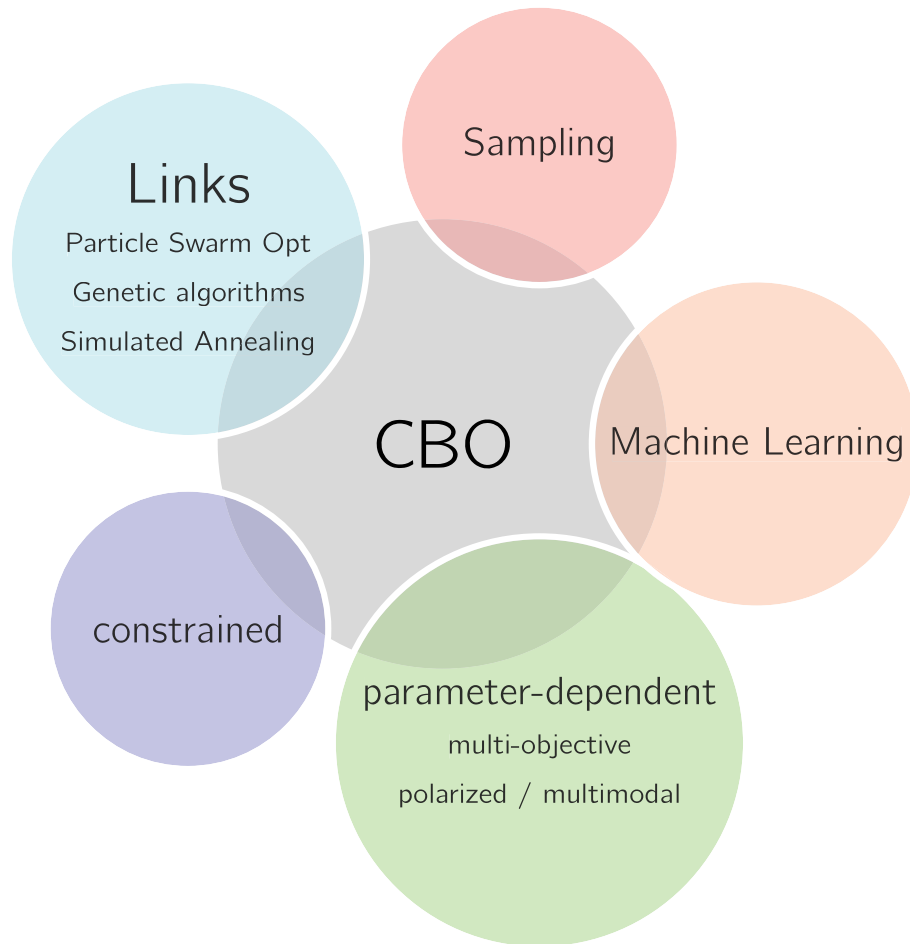
CBX github repository: <https://github.com/PdIPS>

Example: Visualizing CBX dynamics

Other **preimplemented examples**:

- ▶ NN training with CBO
- ▶ Consensus-based sampling
- ▶ Polarized CBO
- ▶ Custom noise variants

Overview of the variants



Contributing researchers:

- ▶ Massimo Fornasier
- ▶ Lorenzo Pareschi
- ▶ Michael Herty
- ▶ Hui Huang
- ▶ Franca Hoffmann
- ▶ Shi Jin
- ▶ Andrew Stuart
- ▶ Marie-Therese Wolfram
- ▶ and many more ...!

Overarching goal:

rigorous convergence analysis

Towards infinite dimensional problems (joint work with M. Khatab)

Extend CBO to global optimization in Sobolev spaces:

$$\min_{u \in \mathcal{U}} f(u),$$



where $\mathcal{U}$ is a Sobolev space with elements that can be represented by a Gaussian process (GP) and f is nonconvex with unique minimizer in $\mathcal{U}$.

Key idea:

CBO agents and noise are sampled from GPs that are tailored to the problem at hand.

Example [Henderson 2024]:

If the domain $\mathcal{D}$ is sufficiently smooth, samples of GP with Matérn kernel of order ν lie in $H^m(\mathcal{D})$ for $m \in [0, \nu) \cap \mathbb{N}$. Typical values in ML: $\nu = 1.5$ or 2.5 .

Agents sampled from GP

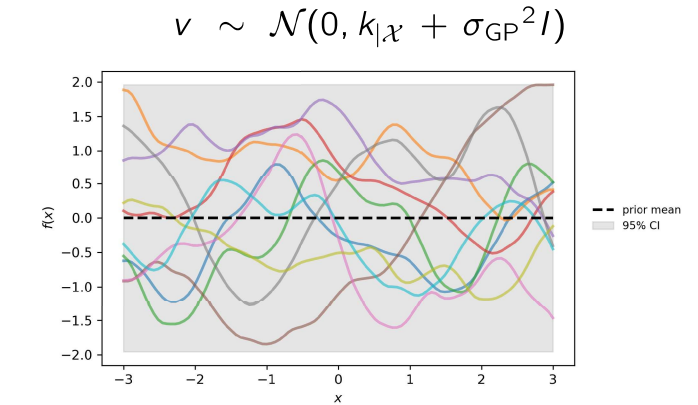
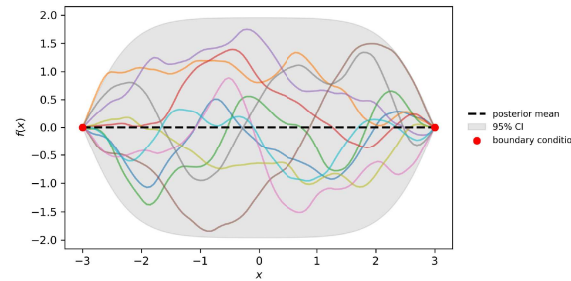
A GP is a distribution over functions:

$$f \sim \text{GP}(m(x), k(x, x')),$$

where m is the mean function and k the covariance kernel.

Data like initial / boundary values or constraints can be included via Bayes rule:

$$p(y_1|y_2) = \frac{p(y_1, y_2)}{p(y_2)}$$



marginalization property: $U^i \sim \mathcal{N}(BC^{-1}x_{\text{train}}, A - BC^{-1}B^\top),$

where $A := k_{|\mathcal{X}} + \sigma_{\text{GP}}^2 I \in \mathbb{R}^{P \times P}$, $B = k_{|\mathcal{X}, x_{\text{train}}} \in \mathbb{R}^{P \times d_{\text{data}}}$, $C := k_{|x_{\text{train}}} \in \mathbb{R}^{d_{\text{data}} \times d_{\text{data}}}$.

CBO with Gaussian processes

Time-discrete CBO in $\mathcal{U}$

$$U^i \leftarrow U^i - \lambda(U^i - v_f^\alpha(U))\tau + \sqrt{2\tau}\|U^i - v_f^\alpha(U)\|_{\mathcal{U}}\xi^i, \\ U_0 \in \text{GP}_c(m, k, \theta), \quad \xi^i \in \text{GP}_0(m, k, \theta).$$

with $\text{GP}_c(m, k, \theta)$ and $\text{GP}_0(m, k, \theta)$ tailored to the problem.

Example

$$\frac{d}{dt}u = g(u), \quad u(t_0) = u_0.$$

with cost functional $f(u) = \|\frac{d}{dt}u - g(u)\|_{L^2(0,T)}^2$ and state space

$$\mathcal{U} = \{u \in H^1(0, T): u(t_0) = u_0 \neq 0\}.$$

Then $U_0^i \sim \text{GP}_c$ satisfies $U_0^i(t_0) = u_0$ and $\xi^i \sim \text{GP}_0$ satisfies $\xi^i(t_0) = 0$, $i = 1, \dots, N$.

CBO Algorithm in Function Spaces

Require: number of CBO agents N ; constraint posterior $\text{GP}_c(m, k, \theta)$; homogeneous posterior $\text{GP}_0(m, k, \theta)$; time horizon T , time step size τ ; weight parameter α , drift parameter λ ;

```
1:  $J \leftarrow \text{int}(T/\tau)$ 
2:  $U_0 = (U^i)_i \leftarrow N$  independent samples  $U^i \sim \text{GP}_c(m, k, \theta)$ 
3: for  $j = 0, \dots, J - 1$  do
4:   evaluate  $v_f^\alpha(U_j)$ 
5:    $\xi = (\xi^i)_i \leftarrow N$  independent samples  $\xi^i \sim \text{GP}_0(m, k, \theta)$ 
6:   compute the weighed mean  $v_f^\alpha(U_j)$ 
7:   for  $i = 1, \dots, N$  do
8:      $U_{j+1}^i \leftarrow U_j^i - \lambda(U_j^i - v_f^\alpha(U_j))\tau + \sqrt{2\tau}\|U_j^i - v_f^\alpha(U_j)\|_{\mathcal{U}} \xi_j^i$ 
9:   end for
10: end for
11: return  $v_f^\alpha(U_J), U_J$ 
```

Nonlinear Optimal Control with ODE constraint

$$\min J(x, v, d) \quad \text{subject to}$$

State constraint modelling the interaction of N sheep and M dogs

$$\frac{d}{dt}x_i = v_i, \quad x_i(0) = x_{0,i} \in \mathbb{R}^2, i = 1, \dots, N$$

$$\frac{d}{dt}v_i = -\alpha v_i - \frac{1}{N_S} \sum_{j=1}^{N_S} K_s(x_j - x_i) - K_d(d - x_i), \quad v_i(0) = v_{0,i},$$

$$\frac{d}{dt}d = u, \quad d(0) = d_0 \in \mathbb{R}^{2M},$$

K_s, K_d : gradients of regularized Morse potentials $\rightsquigarrow$ **nonlinear, nonlocal**

Control variable: dog velocity $u \in L^2(0, T; \mathbb{R}^{2m})$

Sufficient regularity $\Rightarrow$ existence of **control-to-state map** $\mathcal{S}: u \rightarrow y = (x, v)$

Nonlinear Optimal Control with ODE constraint

Cost functional

$$J(x, v, d, u) = \int_0^T \sigma_1 (V(t) - V_0)^2 + \sigma_2 \|E(t) - x_{\text{des}}\|^2 + \sigma_3 \|u(t)\|^2 dt,$$

where

$$E(t) = \frac{1}{N_S} \sum_{i=1}^{N_S} x_i(t), \quad V(t) = \frac{1}{N_S} \sum_{i=1}^{N_S} \|x_i(t) - E(t)\|^2.$$

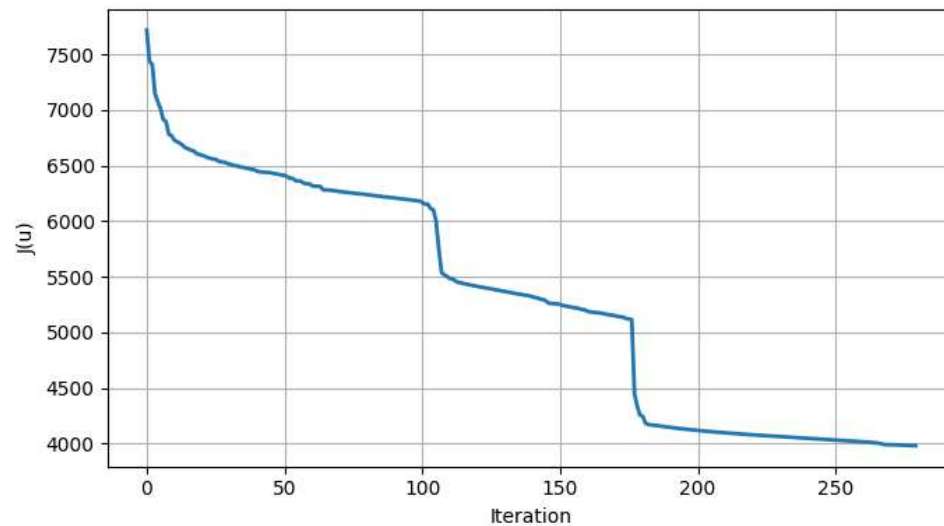
Reduced costfunctional

$$\min_{u \in \mathcal{U}} f(u) := J(\mathcal{S}(u), u), \quad \mathcal{U} = \{u \sim \text{GP}_{\text{Matérn}(0.5)} : u(0) = 0\}$$

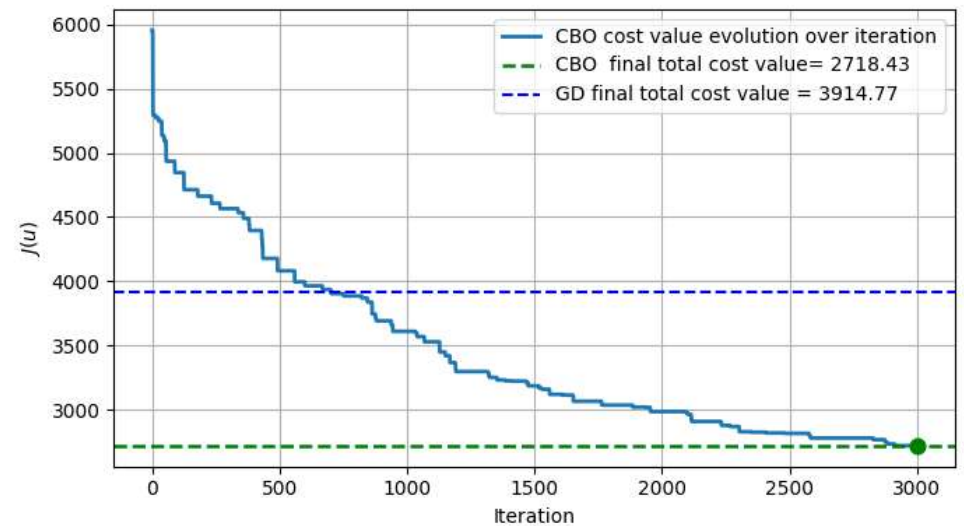
Note: in Burger et al. (2019) this was solved with gradient descent

Sheep & Dogs: Gradient Descent vs CBO

We compare the performance of the global CBO with a gradient descent (GD) method.



Cost evolution for the gradient descent method.



Cost evolution of one CBO run.

Outlook: From here we aim towards global optimization in infinite dimensions!

That's it!

Main references:

- ▶ A consensus-based model for global optimization and its mean-field limit, Pinnau, T, Tse, Martin, M3AS, 2017.
- ▶ An analytical framework for consensus-based global optimization method, Carrillo, Choi, T, Tse, M3AS, 2018.
- ▶ CBX: Python and Julia Packages for Consensus-Based Interacting Particle Methods, Bailo, Barbaro, Gomes, Riedl, Roith, T, Vaes, JOSS, 2024.
- ▶ M. Khatab, C. Totzeck A Consensus-based optimization algorithm using Gaussian processes for global optimization problems in Sobolev spaces, preprint arXiv:2603.15337, 2026.
- ▶ ... please feel free to reach out for more!

Online resources:

- ▶ CBX github repository:
<https://github.com/PdIPS>
- ▶ Mailing list:
cbx_news@lists.uni-wuppertal.de

Thanks!

