Convergence en temps long d'une méthode d'optimisation par consensus

Victor PRISER[1] Pascal BIANCHI[1] Radu-Alexandru DRAGOMIR[1]

[1]LTCI, Télécom Paris, Institut Polytechnique de Paris, France

Résumé – Cet article étudie un algorithme à particules d'optimisation globale d'ordre 0 pour une fonction non convexe f. En particulier, l'algorithme proposé est une variante de l'algorithme de Consensus-Based Optimization (CBO). Contrairement aux autres approches existantes qui se concentrent sur une fenêtre finie, nous nous intéressons à la convergence en temps long. L'étude de cet algorithme est menée d'abord dans le cadre de la limite de champ moyen, où nous montrons la convergence en temps long vers la mesure de Dirac centrée sur le minimiseur. Dans un second temps, nous démontrons la convergence, à nombre fini de particules, en temps long, vers un ensemble de mesures se concentrant autour du minimiseur de f.

Abstract – This paper studies a global optimization particle algorithm of order 0 for a non-convex function f. In particular, the proposed algorithm is a variant of the Consensus-Based Optimization (CBO) algorithm. Unlike other existing approaches that focus on a finite window, we are interested in long-time convergence. The study of this algorithm is first conducted in the mean-field limit framework, where we show long-time convergence to the Dirac measure centered at the minimizer. In a second step, we demonstrate the long-time convergence, for a finite number of particles, to a set of measures concentrating around the minimizer of f.

1 Introduction

Nous cherchons à résoudre le problème d'optimisation globale

$$\min_{x\in\mathbb{R}^d} f(x)\,,$$

où f est une fonction potentiellement non-convexe. Nous supposons par ailleurs que le gradient de f est inconnu ou difficile à calculer. Dans ce contexte, les méthodes d'ordre 0, qui ne font appel qu'aux valeurs de la fonction, sont privilégiées.

Parmi les méthodes d'optimisation globale, on peut citer le recuit simulé (Simulated Annealing) [10], les algorithmes génétiques [6] et évolutionnistes, l'optimisation bayésienne [5], ainsi que les méthodes inspirées de comportements collectifs telles que l'optimisation par essaims de particules (*Particle Swarm Optimization*, PSO) [8].

Les méthodes PSO reposent sur l'idée fondamentale de la collaboration entre plusieurs particules, chacune possédant une information locale sur la fonction objectif, et cherchant à converger vers une solution optimale en partageant cette information avec le groupe. Bien que minimiser une fonction non-convexe soit un problème NP-complet, les algorithmes PSO sont remarquablement efficaces pour des problèmes de taille raisonnable. Ces méthodes peuvent notamment être utilisées en traitement du signal [7].

Ces méthodes sont considérées comme *heuristiques* en raison de leur manque de garanties théoriques. Ainsi, pouvoir expliquer leurs bonnes performances pratiques demeure un problème fondamental. Une des pistes consiste à en considérer une version simplifiée, l'optimisation par consensus (*Consensus-based Optimization*, CBO), qui ignore la mémoire locale des particules et l'effet d'inertie [1, 4]. Celle-ci se prête à des analyses théoriques au travers d'approximations à champ moyen et de processus de diffusion.

L'algorithme CBO considère l'évolution de n particules $(X_k^1,\dots,X_k^n)\in(\mathbb{R}^d)^n$, définies pour chaque itération $k\in\mathbb{N}$, suivant la dynamique

$$X_{k+1}^i = X_k^i - \eta(X_k^i - C_k) + \eta\epsilon_k^i\,, \tag{1}$$

où $\eta>0$ est le pas, ϵ_k^i un terme de bruit de moyenne nulle, et C_k une approximation de la meilleure position parmi toutes les particules à l'instant k. Ainsi, les particules cherchent à se rapprocher du *terme de consensus* C_k afin de minimiser l'objectif, tout en étant poussées à explorer l'espace par le bruit. C_k est typiquement donné par

$$\frac{\sum_{i=1}^n e^{-\alpha f(X_k^i)}X_k^i}{\sum_{i=1}^n e^{-\alpha f(X_k^i)}}, \tag{2}$$

avec $\alpha>0$. Cette expression a l'avantage d'être différentiable en X, et approche $\arg\min\{f(x)\,:\,x\in\{X_k^1,\dots X_k^n\}\}$ quand $\alpha\to\infty$ et quand cette dernière expression est un singleton.

L'objectif est alors de savoir sous quelles conditions les particules convergent vers le minimiseur global x^* de la fonction (en supposant que celui-ci soit unique), lorsque $(k,n)\to(\infty,\infty)$, et à quelle vitesse. Dans cette optique, le choix du terme de bruit est particulièrement crucial.

Etat de l'art. Dans les travaux précédents [1, 3], le bruit est choisi comme

$$\epsilon_k^i = \sigma\|X_k^i - C_k\|\xi_i^k,$$

où $\sigma>0$ et ξ_i^k est une variable gaussienne centrée réduite. Ainsi, la variance dépend de la position de la particule. Ce choix a pour but de favoriser l'émergence d'un consensus. c'est à dire la convergence vers une mesure de Dirac.

L'étude de cet algorithme a conduit à des résultats de convergence dans le régime ($n=\infty$, $k\to\infty$). Ce régime est également appelé *limite de champ moyen*. Les auteurs de [1] démontrent la convergence de la loi de chaque particule vers une mesure de Dirac centrée en un point situé à une distance

$\mathcal{O}\left(\alpha^{-1/2}\right)$ du minimiseur. Toutefois, le résultat repose sur une condition forte concernant la mesure initiale. Dans [3], cette hypothèses restrictive est levée, et l'étude est étendue au cas où le nombre de particules est fini, dans le régime où $(n,k) \to (\infty,\infty)$ avec $n \gtrsim \exp(\exp(k))$, ce qui est une contrainte forte sur le nombre de particules.

Contributions. Un défaut majeur des travaux précédents est que ceux-ci ne montrent pas la consistance de l'algorithme, c'est à dire la convergence vers une mesure de Dirac centrée en x^*. Une des raisons est que CBO étudié dans [1, 3] admet une infinité d'états d'équilibres, correspondant aux configurations où toutes les particules coïncident en un même point quelconque de $\mathbb{R}^d$.

Pour y remédier, nous proposons une variante de CBO où la variance du bruit est de la forme :

$$\epsilon_i^k = \sigma_k \xi_i^k,$$

avec ξ_i^k gaussiennee centrée réduite et une variance $\sigma_k > 0$ qui peut être constante ou décroissante au cours du temps. Nous faisons également varier le paramètre α_k dans le terme de consensus 2, afin de garantir la convergence exacte vers x^*.

Nos résultats principaux sont les suivants.

- **Régime à champs moyen, temps continu (Th. 1) :** nous étudions la version à temps continu l'algorithme avec $n = \infty$. En choisissant $\alpha_t \to \infty$ à la vitesse de $\log\log(t+1)$, et $\sigma_t = \sqrt{\frac{2\gamma}{\alpha_t}}$, nous montrons la convergence de loi des particules vers δ_{x_*} quand $t \to \infty$.
- **Nombre de particules fini, temps discret (Th. 2) :** avec α_k et σ_k constants, nous montrons que la loi des particules (X_k^i) converge, dans n'importe quel régime $(n,k,\eta) \to (\infty,\infty,0)$, vers un ensemble de gaussiennes centrées dans une boule de rayon $\mathcal{O}(\alpha^{-1/2})$ autour de x_* et de variance $\mathcal{O}(\alpha^{-1})$. On ne démontre pas la consistance comme décrit plus haut, mais notre résultat est nouveau par rapport à celui de [3]. En effet, il permet, pour un n fixé, d'estimer la distance entre la loi d'une particule X_k^i et δ_{x_*} à tout instant k. En revanche, le résultat de [3] ne fournit cette estimation que pour $k < \log\log(n)$.

Notations Soit $\langle\cdot,\cdot\rangle$ le produit scalaire euclidien de $\mathbb{R}^d$ et $\|\cdot\|$ sa norme associée.

La notation $B(x,r)$ désigne la boule ouverte de rayon r centrée en x.

Pour deux nombres réels a, b, on note $a \wedge b := \min(a,b)$ et $a \vee b := \max(a,b)$.

Soit $p \geq 1$. On définit l'espace $\mathcal{P}_p(\mathbb{R}^d)$ comme l'ensemble des mesures de probabilité vérifiant $\int \|x\|^p \, d\mu(x) < \infty$. Cet espace est muni de la distance de Wasserstein W_p.

Pour un ensemble $A \subset \mathcal{P}_p(\mathbb{R}^d)$ et $\mu \in \mathcal{P}_p(\mathbb{R}^d)$, on définit $W_p(\mu, A) := \inf_{\nu\in A} W_p(\mu,\nu)$.

Dans la suite, les variables aléatoires sont définies dans un ensemble de probabilité $(\Omega, \mathbb{P}, \mathcal{F})$.

2 Algorithme

Soit $\alpha > 0$. L'algorithme CBO s'initialise avec $n \in \mathbb{N}^*$ particules $(X_0^1, \dots, X_0^n) \in (\mathbb{R}^d)^n$. À l'étape $k \in \mathbb{N}$ de l'algorithme, les particules $(X_k^1, \dots, X_k^n)$ interagissent à travers le *point de consensus* $C_\alpha(\mu_k^n)$, où, pour une mesure μ,

$$C_\alpha(\mu) := \frac{\int x e^{-\alpha f(x)} \, \mathrm{d}\mu(x)}{\int e^{-\alpha f(x)} \, \mathrm{d}\mu(x)},$$

et où la *mesure empirique* est définie par $\mu_k^n = \frac{1}{n}\sum_{i=1}^n \delta_{X_k^i}$. Le point de consensus est une approximation de la meilleure particule à l'instant k. En effet, lorsque $\alpha \to \infty$, on a $C_\alpha(\mu_k^n) \approx \arg\min\{f(X_k^i)\}$: c'est le principe de Laplace.

Notre algorithme s'écrit, avec une suite α_k fixée,

$$X_{k+1}^i = X_k^i + \eta(\mathrm{clip}_R(C_{\alpha_k}(\mu_k^n)) - X_k^i) + \sqrt{2\frac{\gamma}{\alpha_k}\eta}\xi_{k+1}^i\,, \quad (3)$$

avec $R > 0$, $(\xi_k^i)_{k\in\mathbb{N}^*, i\leq n}$ sont des variables gaussiennes i.i.d. centrées et réduites. et on définit la fonction de troncature :

$$\mathrm{clip}_R(x) := (R \wedge \|x\|)\,\frac{x}{\|x\|}.$$

Cette fonction de troncature est également un ajout par rapport à l'algorithme CBO original. Elle joue un rôle purement technique en garantissant la stabilité de l'algorithme, c'est-à-dire le bornage uniforme dans le temps des moments des particules.

3 La limite de champ moyen

3.1 Résultat principal

Soit $R, \gamma > 0$. Lorsque $n \to \infty$ et $\eta \to 0$, l'algorithme (3) copie la dynamique d'une équation différentielle stochastique (EDS) non linéaire. Ce lien est établi par la propagation du chaos [2]. Dans le cas où α est une constante, ce lien est explicité dans la section 4. L'EDS associée à l'algorithme (3) s'écrit ci-dessous pour une mesure initiale $\nu \in \mathcal{P}_2(\mathbb{R}^d)$:

$$dX_t = (\mathrm{clip}_R(C_{\alpha_t}(\rho_t)) - X_t)\,dt + \sqrt{\frac{2\gamma}{\alpha_t}}\,dB_t, \quad \rho_0 = \nu\,, \quad (4)$$

où ρ_t désigne la loi de X_t à l'instant t, $(B_t)_{t\geq 0}$ est un mouvement brownien sur $\mathbb{R}^d$, et $\alpha : \mathbb{R}_+ \to \mathbb{R}_+^*$ est une fonction continue.

Cette équation est également appelée une équation de *McKean-Vlasov* en raison de sa dépendance à la mesure ρ_t.

L'hypothèse suivante assure la bonne définition de l'EDS (4) :

Hypothèse 1 *f est continûment différentiable et α est continue. De plus, il existe un coefficient $p > 0$ tel que*

$$\sup_{x\in\mathbb{R}^d} \frac{\|\nabla f(x)\|}{1 + \|x\|^p} < \infty\,.$$

En s'inspirant de [1], nous pouvons garantir l'existence d'une solution de l'EDS. En remarquant qu'il s'agit d'un processus d'Ornstein-Uhlenbeck, on peut montrer que celui-ci se décompose simplement en la somme de deux termes : un terme lié à la mesure initiale ρ_0, qui est oublié à une vitesse exponentielle, et une gaussienne dont la moyenne x_t satisfait une équation différentielle ordinaire (EDO).

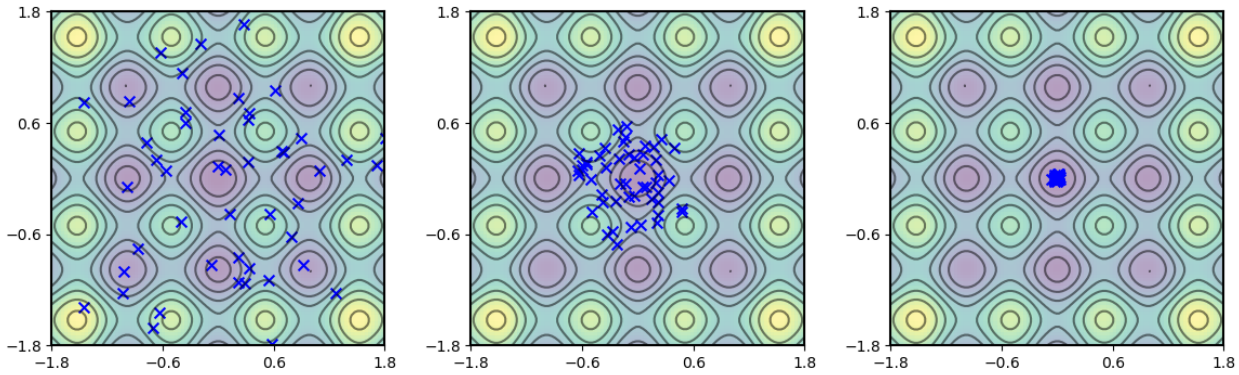

FIGURE 1 : Démonstration de CBO pour $N = 50$ particules sur une fonction en dimension 2.

Proposition 1 *Sous l'hypothèse 1, pour toute distribution initiale $\nu \in \mathcal{P}_2(\mathbb{R}^d)$ de X_0 et pour tout $T > 0$, l'EDS* (4) *admet une unique solution forte sur $[0, T]$. On note $\rho_t \in \mathcal{P}_2(\mathbb{R}^d)$ la loi de la solution forte X_t.*

De plus, pour tout $t \geq 0$, il existe un vecteur Gaussien centré Z_{α_t} de covariance $\mathrm{Cov}(Z_{\alpha_t}) = \frac{\gamma}{\alpha_t} I_d$, indépendant de X_0, tel que :

$$X_t = (X_0 - x_0)e^{-t} + x_t + \sigma(\alpha)_t Z_{\alpha_t}\,, \tag{5}$$

où

$$\sigma(\alpha)_t := \sqrt{\mathrm{e}^{-2t}\alpha_t \int_0^t \frac{\mathrm{e}^{2s}}{\alpha_s} ds}\,,$$

et où $t \mapsto x_t$ est l'unique solution de l'équation différentielle ordinaire (EDO) :

$$\dot{x}_t = \mathrm{clip}_R(C_{\alpha_t}(\rho_t)) - x_t\,, \tag{6}$$

avec condition initiale $x_0 \in \mathbb{R}^d$.

On cite désormais les hypothèses nécessaire pour établir notre premier théorème.

Hypothèse 2 *Supposons que f est continûment différentiable jusque à l'ordre 3. Les conditions suivantes sur f sont vérifiées :*

1. *f admet un minimiseur unique x_* et $f(x_*) = 0$.*
2. *Il existe $r > 0$ tel que la Hessienne H_f de f est Lipschitzienne sur $B(x_*, r)$ et $H_f(x_*)$ est définie positive.*
3. *Il existe $\kappa > 0$ et $\beta \geq 1$, tels que $f(x) \geq f(x_*) + \frac{\kappa}{2}\|x - x_*\|^\beta$.*

Cette hypothèse garantit que f est fortement convexe autour de son minimiseur x_* et qu'il croît suffisamment en dehors de ce minimiseur pour assurer que les points de consensus restent dans un voisinage de x_*.

Finalement, nous introduisons l'hypothèse sur le taux de croissance du paramètre α_t.

Hypothèse 3 *La fonction $t \mapsto \alpha_t$ est continue et diverge vers l'infini à la vitesse de* $\log\log(t + 1)$.

Cette hypothèse garantit que, lorsque t est grand, X_t, défini dans la proposition 4, aura le comportement de $x_t + Z_{\alpha_t}$. On obtient le théorème principal de cette section.

Théorème 1 *Soient les hypothèses 1, 2 et 3 vérifiées. Soit $\delta > 0$ une constante dépendant de f. On suppose $R \geq \|x_*\| + \delta$. Soit $\nu \in \mathcal{P}_2(\mathbb{R}^d)$. Soit $\gamma > \frac{8(R \vee W_2(\nu, \delta_{x_*}))}{\delta^{\beta-1}\kappa}$. Soit $(\rho_t)_{t \in \mathbb{R}_+}$ la séquence de mesures définie dans la proposition 1 avec pour mesure initiale ν. Alors,*

$$W_2(\rho_t, \delta_{x_*}) \leq \frac{C^{\mathrm{conv}}}{\sqrt{\alpha_t}},$$

où la constante C^{conv} dépend de f, γ, R, et ν.

3.2 Esquisse de preuve

Dans la décomposition 5, la condition initiale étant oubliée à vitesse exponentielle, nous nous concentrons sur la partie gaussienne $x_t + \sigma(\alpha)_t Z_{\alpha_t}$, dont la variance tend vers 0.

Le principal enjeu est de montrer que la moyenne x_t, qui vérifie l'EDO (6), converge vers x^*.

L'opérateur proximal On définit l'opérateur proximal qui, à une fonction convexe g et un réel $x \in \mathbb{R}^d$, associe :

$$\mathrm{prox}_g(x) := \arg\min_{y \in \mathbb{R}^d} \{g(y) + \frac{\|x - y\|^2}{2}\}.$$

On remarque que, dans notre cas, f n'est pas convexe, et son opérateur proximal n'est donc pas nécessairement bien défini. Cependant, grâce à l'hypothèse 2, on vérifie que, pour γ suffisamment grand par rapport à $\|x - x_*\|$, on a

$$\arg\min_{y \in \mathbb{R}^d} \{f(y) + \frac{\|x - y\|^2}{2\gamma}\} \in B(x_*, \delta),$$

où δ est une constante telle que f est strictement convexe sur la boule $B(x_*, \delta)$. On peut donc définir une fonction strictement convexe et semi-continue inférieurement (s.c.i.) $\bar{f}$ telle que $\bar{f}$ coïncide avec f sur $B(x_*, \delta)$ et le précédent $\arg\min$ est égal à $\mathrm{prox}_{\gamma \bar{f}}(x)$ pour $x \in B(x_*, K)$ où γ est suffisamment grand par rapport à K.

Principe de Laplace [9]. Pour une fonction h suffisamment régulière, ce principe établit que

$$\frac{\int x \mathrm{e}^{-\alpha h(x)} dx}{\int \mathrm{e}^{-\alpha h(x)} dx} \xrightarrow{\alpha \to \infty} \arg\min_{y \in \mathbb{R}^d} h(y)\,.$$

En utilisant le paragraphe précédent, cela veut dire que pour γ suffisemment grand par rapport à K et $x \in B(x_*, K)$,

$C_{\alpha_t}(x + Z_{\alpha_t}) \simeq \mathrm{prox}_{\gamma \bar{f}}(x)$, où pour une v.a. $X : \Omega \to \mathbb{R}^d$, $C_{\alpha_t}(X) := C_{\alpha_t}(\mathcal{L}(X))$ où $\mathcal{L}(X)$ désigne la loi de X et Z_{α_t} est défini dans Prop. 1. Grâce à l'hypothèse 3, on obtient que

$$C_{\alpha_t}(X_t) \simeq C_{\alpha_t}(x_t + Z_{\alpha_t}) \simeq \mathrm{prox}_{\gamma \bar{f}}(x_t)\,,$$

Le principe de Laplace peut être quantifié grâce à [9] et permet d'obtenir une évaluation de l'approximation précédente.

Convergence dans la limite de champ moyen L'objectif est d'établir la convergence de $(x_t)_{t \geq 0}$, défini dans la proposition 1. Pour cela, il suffit de remarquer que x_t satisfait l'équation différentielle suivante :

$$\dot{x}_t = \mathrm{prox}_{\gamma \bar{f}}(x_t) - x_t + r_t,$$

où l'on part de $x_0 \in \mathbb{R}^d$, et où le terme de reste est donné par $r_t = C_\alpha(X_t) - \mathrm{prox}_{\gamma \bar{f}}(x_t)$. D'après le paragraphe précédent, ce terme de reste converge vers 0 lorsque $t \to \infty$. En utilisant la non-expansivité de l'opérateur proximal, on peut montrer que la fonction $t \mapsto \|x_t - x_*\|$ converge vers 0. On en déduit alors le résultat du théorème 1.

4 Étude du système avec un nombre fini de particules

Soit $\eta > 0$. L'objectif de cette section est d'étudier l'algorithme (4) avec un pas η et $\alpha_k = \alpha > 0$. L'étude de la convergence de cet algorithme passe par la propagation du chaos détaillée dans [2].

On définit :

$$\mathcal{N}_{x_*}^{C,\alpha} := \{\mathcal{N}(x, \frac{\gamma}{\alpha} I_d) \,:\, x \in B(x_*, \left(\frac{C}{\sqrt{\alpha}}\right))\}\,,$$

où, pour tout $x \in \mathbb{R}^d$, $\mathcal{N}(x, \frac{\gamma}{\alpha} I_d)$ désigne la loi d'un vecteur gaussien centré en x et de matrice de covariance $\frac{\gamma}{\alpha} I_d$.

Alors, on obtient notre second théorème (dont la justification est omise ici par manque de place) :

Théorème 2 *Supposons que $(X_0^1, \ldots, X_0^n)$ sont des variables aléatoires i.i.d. de loi $\nu \in \mathcal{P}_2(\mathbb{R}^d)$. Soient les hypothèses 1 et 2. On choisit R comme dans le théorème 1. Soit γ_0 une constante qui dépend de ν, f, R. On choisit $\gamma > \gamma_0$. Soient C, α_0 des constantes qui dépendent de R, γ et f. On choisit $\alpha > \alpha_0$. Il existe k^{pop} tel que pour tout $i \leq n$ et $k \in \mathbb{N}$, on obtient*

$$W_2(\mathcal{L}(X_k^i), \mathcal{N}_{x_*}^{C,\alpha}) \leq \frac{C_0}{2^{\frac{k}{k^{\mathrm{pop}}}}} + \left(\frac{C_{k^{\mathrm{pop}}}^{\mathrm{PDC}}}{\sqrt{n}} + C_{k^{\mathrm{pop}}}^{\mathrm{Euler}} \sqrt{\eta}\right)\,,$$

où $C_0 > 0$ est une constante exponentielle en α qui dépend de ν, γ, f, R. De plus, $C_{k^{\mathrm{PDC}}}^{\mathrm{Euler}}$ et $C_{k^{\mathrm{pop}}}^{\mathrm{Euler}}$ sont des constantes exponentielles en k^{pop} et α.

On remarque que, pour n fixé, grâce à ce théorème, les points d'accumulation de $\mathcal{L}\left(X_k^i\right)_k$ convergent dans la limite $(n, \eta) \to (\infty, 0)$ vers l'ensemble $\mathcal{N}_{x_*}^{C,\alpha}$.

5 Perspectives

L'objectif est de proposer un algorithme dont la loi de chaque particule converge, dans le triple régime $(k, n, \eta) \to (\infty, \infty, 0)$, vers la mesure δ_{x_*}. Cela ne peut être garanti que si α_k diverge vers l'infini. Dans cet article, nous considérons un $\alpha_k = \alpha$ constant dans le régime $n < \infty$. Une perspective intéressante, qui améliorerait considérablement les résultats de ce travail, serait d'étudier le cas où α_k est variable.

Remerciements Le premier auteur remercie la Chaire DSAIDIS, et le troisième auteur remercie la Chaire HI ! PARIS pour le financement de ce travail.

Références

[1] J. A CARRILLO, Y.-P. CHOI, C. TOTZECK et O. TSE : An analytical framework for consensus-based global optimization method. *Mathematical Models and Methods in Applied Sciences*, 28(06):1037–1066, 2018.

[2] L.-P. CHAINTRON et A. DIEZ : Propagation of chaos : A review of models, methods and applications. i. models and methods. *Kinetic and Related Models*, 15(6):895, 2022.

[3] M. FORNASIER, T. KLOCK et K. RIEDL : Consensus-based optimization methods converge globally. *SIAM Journal on Optimization*, 34(3):2973–3004, septembre 2024.

[4] S. GRASSI et L. PARESCHI : From particle swarm optimization to consensus based optimization : stochastic modeling and mean-field limit. *Math. Models Methods Appl. Sci.*, 31(8):1625–1657, 2021.

[5] W. K. HASTINGS : Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

[6] J. H. HOLLAND : *Adaptation in natural and artificial systems. An introductory analysis with applications to biology, control, and artificial intelligence*. University of Michigan Press, Ann Arbor, Mich., 1975.

[7] T. INCE, S. KIRANYAZ et M. GABBOUJ : A generic and robust system for automated patient-specific classification of ecg signals. *IEEE Transactions on Biomedical Engineering*, 56(5):1415–1426, 2009.

[8] J. KENNEDY et R. EBERHART : Particle swarm optimization. *In Proceedings of ICNN'95 - International Conference on Neural Networks*, volume 4, pages 1942–1948 vol.4, 1995.

[9] W. D KIRWIN : Higher asymptotics of laplace's approximation. *Asymptotic Analysis*, 70(3-4):231–248, 2010.

[10] M. PELLETIER : Weak convergence rates for stochastic approximation with application to multiple targets and simulated annealing. *Annals of Applied Probability*, pages 10–44, 1998.