Online Supplementary Materials for

"Cooperative Coevolution for Non-Separable Large-Scale Black-Box Optimization: Convergence Analyses and Distributed Accelerations"

(Under Reviews)

Date: 2023-11-20

https://github.com/Evolutionary-Intelligence/DCC/blob/main/SupplementaryMaterials.pdf

1. Convergence to a Pure Nash Equilibrium for Cooperative Coevolution (CC):

***Theorem* 1:** Given *any* partition $p = \{g_1, \ldots, g_m\}$ of the objective function $f(\boldsymbol{x}): \mathbb{R}^n \to \mathbb{R}$, where $1 < m \leq n$, we say that the convergence point of CC under this partition is also one *pure Nash equilibrium (PNE).*

Proof: If $\boldsymbol{x}^* = (x_1^*, x_2^*, \cdots, x_n^*) = (\boldsymbol{x}_{g_1}^*, \cdots, \boldsymbol{x}_{g_m}^*)$ is a convergence point of CC under the partition $p = \{g_1, \ldots, g_m\}$, then $\boldsymbol{x}_{g_i}^*$ is one of the global optima of the function $f\left(\boldsymbol{x}_{g_1}^*, \cdots, \boldsymbol{x}_{g_i}, \cdots, \boldsymbol{x}_{g_m}^*\right), i = 1, \cdots, m.$

If $x^* = (\boldsymbol{x}_{g_1}^*, \cdots, \boldsymbol{x}_{g_m}^*)$ is not one *pure Nash equilibrium* (w.r.t. p), according to the definition of **Pure Nash Equilibrium (PNE),** we have that $\exists i \in \{1, \cdots, m\}$, $\exists \breve{\boldsymbol{x}}_{g_i} \in \mathbb{R}^{|g_i|} \backslash \{\boldsymbol{x}_{g_i}^*\}$, such that $f\left(\boldsymbol{x}_{g_1}^*, \cdots, \breve{\boldsymbol{x}}_{g_i}, \cdots, \boldsymbol{x}_{g_m}^*\right) < f\left(\boldsymbol{x}_{g_1}^*, \cdots, \boldsymbol{x}_{g_i}^*, \cdots, \boldsymbol{x}_{g_m}^*\right)$, namely, then $\boldsymbol{x}_{g_i}^*$ is not the global optimum of $f\left(\boldsymbol{x}_{g_1}^*, \cdots, \boldsymbol{x}_{g_i}, \cdots, \boldsymbol{x}_{g_m}^*\right)$. It is a contradiction.

Note that we assume that for CC, the suboptimizer could obtain the global optimum for each subproblem given a *limited* number of function evaluations (but which could be any large number). We admit that such an assumption appears to be difficult to satisfy in practice. However, it helps to understand the convergence behavior of CC and capture the *game-theoretical* essence of CC.

2. Convergence Analyses on Four Representative Test Functions:

$$\boldsymbol{f_1(x,y) = 7x^2 + 6xy + 8y^2}$$

Since its Hessian matrix $\begin{bmatrix} 7 & 6 \\ 6 & 8 \end{bmatrix}$ is positive definite, f_1 is differentiable strictly convex, then it has a unique global optimum $(0,0)$.

Given (x_0, y_0) is a PNE, by the definition of PNE, (x_0, y_0) is the unique global optimum of differentiable strictly convex $f_1(x, y_0)$ and $f_1(x_0, y)$.

Then $\frac{\partial f_1(x,y_0)}{\partial x} = 14x + 6y_0 = 0$, we have $x = -\frac{6y_0}{14} = x_0$, $\frac{\partial f_1(x_0,y)}{\partial y} = 16y + 6x_0 = 0$, we have $y = -\frac{6x_0}{16} = y_0$. So, $(x_0, y_0) = (0,0)$.

$$\boldsymbol{f_2(x,y) = x^2 + 10^6 y^2}$$

Since its Hessian matrix $\begin{bmatrix} 1 & 0 \\ 0 & 10^6 \end{bmatrix}$ is positive definite, f_1 and its rotation variant are all strictly convex and have the unique global optimum $(0,0)$.

Like the above proof, they have only one PNE, namely global optimum $(0,0)$.

$$\boldsymbol{f_3(x,y) = 100(x^2 - y)^2 + (x-1)^2}$$

Obviously, $f_3 \geq 0$ and

$$\begin{cases} \frac{\partial f_3(x,y)}{\partial x} = 400x(x^2 - y) + 2(x-1) = 0 \\ \frac{\partial f_3(x,y)}{\partial y} = 200(y - x^2) = 0 \end{cases},$$

So, it has a unique global optimum $(1,1)$.

Suppose (x_0, y_0) is a PNE, then x_0 and y_0 are the global minima of differentiable $f_3(x, y_0)$ and $f_3(x_0, y)$, respectively. Then

$$\begin{cases} \frac{df_3(x,y_0)}{dx} = 400x^3 - 400xy_0 + 2x - 2 = 0 \\ \frac{df_3(x_0,y)}{dy} = 200y - 200x_0^2 = 0 \end{cases},$$

we have $200x^3 - 200xx_0^2 + x - 1 = 0, y = x_0^2 = y_0$.

Since x_0 is one of solutions of equation $200x^3 - 200xx_0^2 + x - 1 = 0$, we conclude that $x_0 = 1,\ y_0 = 1$.

$$\boldsymbol{f_4(x,y) = |x-y| - \min(x,y) = \begin{cases} x - 2y, x > y \\ -x, x = y \\ y - 2x, x < y \end{cases}}$$

For any (x_0, y_0), the function $f_4(x, y_0)$ and $f_4(x_0, y)$ obtain their global minima at (y_0, y_0) and (x_0, x_0), respectively. So, the set $\{(x,y) | x = y\}$ is the set of PNEs.

3. Convergence Analysis on A Function (with Loss of Gradients):

Corollary 3: For the *Schwefel's Problem 2.21*, $min\,(f(\boldsymbol{x})) = min\,(\underset{i=1,\ldots,n}{max}\,(|\boldsymbol{x_i}|))$, defined on an open set[1] $\Omega \subseteq \mathbb{R}^n$, the set of global pure Nash equilibria w.r.t. any partition $p = \{g_1, \ldots, g_m\}$ is $\{\boldsymbol{x}, \underset{j\in g_1}{max}|\boldsymbol{x}_j| = \cdots = \underset{j\in g_m}{max}|\boldsymbol{x}_j|\}$. There is a unique strict global Nash equilibrium $\boldsymbol{x} = (0, \ldots, 0)$ w.r.t. any partition set, which equals the global optimum, and vice versa.

Proof: Given $\boldsymbol{x} = (\boldsymbol{x}_{g_1}, \cdots, \boldsymbol{x}_{g_m})$ is a PNE w.r.t. any partition $p = \{g_1, \ldots, g_m\}$ of $f(\boldsymbol{x}) = \underset{i=1,\ldots,n}{max}\,(|\boldsymbol{x_i}|)$, owing to the definition of PNE, we have $\boldsymbol{x}_{g_i} \in \mathbb{R}^{|g_i|}$ for each $g_i \in p, \forall i \in \{1, \ldots, m\}$, satisfies

$$f(\boldsymbol{x}_{g_i}, \boldsymbol{x}_{\neq g_i}) \leq f(\boldsymbol{x}_{\tilde{g_i}}, \boldsymbol{x}_{\neq g_i}),\ \forall \boldsymbol{x}_{\tilde{g_i}} \in \mathbb{R}^{|g_i|} \backslash \{\boldsymbol{x}_{g_i}\},$$

namely, for any $i \in \{1, \cdots, m\}$, $\boldsymbol{x}_{\tilde{g_i}} \in \mathbb{R}^{|g_i|} \backslash \{\boldsymbol{x}_{g_i}\}$, $max\{\underset{j\in g_i}{max}|\boldsymbol{x}_j|, \underset{j\in\neq g_i}{max}|\boldsymbol{x}_j|\} \leq max\{\underset{j\in g_i}{max}|\boldsymbol{x}_{\tilde{g_i}}|, \underset{j\in\neq g_i}{max}|\boldsymbol{x}_j|\}$, only if, $max\{\underset{j\in g_i}{max}|\boldsymbol{x}_j|, \underset{j\in\neq g_i}{max}|\boldsymbol{x}_j|\} \leq min\{\underset{j\in g_i}{max}|\boldsymbol{x}_{\tilde{g_i}}|, \underset{j\in\neq g_i}{max}|\boldsymbol{x}_j|\}$, only if,

$$\begin{cases} \underset{j\in g_i}{max}|\boldsymbol{x}_j| \leq \min\left\{\underset{j\in g_i}{max}|\boldsymbol{x}_{\tilde{g_i}}|\right\} \\ \underset{j\in g_i}{max}|\boldsymbol{x}_j| \leq \underset{j\in\neq g_i}{max}|\boldsymbol{x}_j| \end{cases},$$

owing to the definition of PNE,

$$\underset{j\in g_i}{max}|\boldsymbol{x}_j| \leq min\left\{\underset{j\in g_i}{max}|\boldsymbol{x}_{\tilde{g_i}}|\right\},$$

so, we only need for any $i \in \{1, \cdots, m\}$,

$$\underset{j\in g_i}{max}|\boldsymbol{x}_j| \leq \underset{j\in\neq g_i}{max}|\boldsymbol{x}_j|,$$

we have $\underset{j\in g_1}{max}|\boldsymbol{x}_j| = \cdots = \underset{j\in g_m}{max}|\boldsymbol{x}_j|$.

Since $f(\boldsymbol{x}) \geq 0$, $\boldsymbol{x_0} = (0, \cdots, 0)$ is a PNE, and $f(x_0) = f(0, \cdots, 0) = 0$, $\boldsymbol{x_0}$ is the unique global optimum, thus it is a unique strict global Nash equilibrium, and vice versa.

[1]Here it is *implicitly* assumed that there is (at least) one global optimum in this open set.