Fast Moving Natural Evolution Strategy for High-Dimensional Problems

Masahiro Nomura, Isao Ono

Tokyo Institute of Technology

IEEE WCCI (CEC) 2022

Introduction

- Black-Box Optimization (BBO)
  - Explicit representations of objective functions are not given
  - **Only evaluation values** of solutions can be used
- Existing Method: FM-NES [Nomura and Ono, 2021]
  - One of the promising BBO methods
  - Good performance on **relatively low-dimensional problems** (~100-dimension)
- Problem of FM-NES:
  - **Difficult to apply it to high-dimensional problems**
- Goal of This Work:
  - Propose CR-FM-NES, which is **scalable and efficient in high-dimensional BBO**

FM-NES [Nomura and Ono, 2021]

1. Create $\mathcal{N}(\boldsymbol{m}^{(0)}, \mathbf{C}^{(0)} = {\sigma^{(0)}}^2 \mathbf{B}^{(0)} {\mathbf{B}^{(0)}}^{\mathrm{T}})$ and set $g = 0$

   $\sigma^{(0)}$: step size, $\mathbf{B}^{(0)}$: normalization transformation matrix

2. Generate λ solutions following $\mathcal{N}(\boldsymbol{m}^{(g)}, \mathbf{C}^{(g)})$

   $\boldsymbol{x}_i = \boldsymbol{m}^{(g)} + \sigma^{(g)} \mathbf{B}^{(g)} \boldsymbol{z}_i\,,\ \ \boldsymbol{z}_i \sim \mathcal{N}(0, \mathbf{I})$

3. Evaluate and sort the solutions, and give weights to them

4. Update the parameters $\boldsymbol{m}^{(g)}$, $\sigma^{(g)}$, $\mathbf{B}^{(g)}$

   $\boldsymbol{m}^{(g+1)} = \boldsymbol{m}^{(g)} + \eta_m \sigma^{(g)} \mathbf{B}^{(g)} \mathbf{G}_\delta \qquad \mathbf{G}_\delta = \sum_{i=1}^{\lambda} w_i \boldsymbol{z}_i$

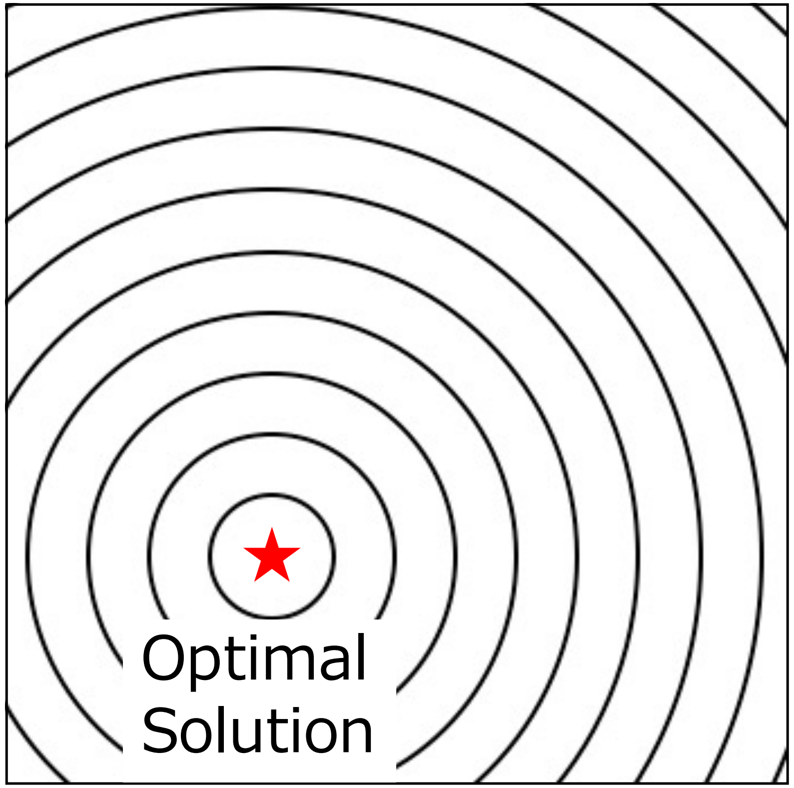
   $\sigma^{(g+1)} = \sigma^{(g)} \exp(\eta_\sigma G_\sigma / 2) \qquad \mathbf{G}_\mathrm{M} = \sum_{i=1}^{\lambda} w_i (\boldsymbol{z}_i \boldsymbol{z}_i^{\mathrm{T}} - \mathbf{I})$

   $\mathbf{B}^{(g+1)} = \mathbf{B}^{(g)} \exp(\eta_\mathrm{B} \mathbf{G}_\mathrm{B} / 2) \qquad G_\sigma = \mathrm{tr}(\mathbf{G}_\mathrm{M}) / d\,, \mathbf{G}_\mathrm{B} = \mathbf{G}_\mathrm{M} - G_\sigma \mathbf{I}$

5. Emphasis on the expansion of the distribution

6. Terminate if the criterion is met,

   otherwise $g \leftarrow g + 1$ and go to Step 2.



FM-NES [Nomura and Ono, 2021]

1. Create $\mathcal{N}(\boldsymbol{m}^{(0)}, \mathbf{C}^{(0)} = {\sigma^{(0)}}^2 \mathbf{B}^{(0)} {\mathbf{B}^{(0)}}^{\mathrm{T}})$ and set $g = 0$

   $\sigma^{(0)}$: step size, $\mathbf{B}^{(0)}$: normalization transformation matrix

2. Generate λ solutions following $\mathcal{N}(\boldsymbol{m}^{(g)}, \mathbf{C}^{(g)})$

   $$\boldsymbol{x}_i = \boldsymbol{m}^{(g)} + \sigma^{(g)} \mathbf{B}^{(g)} \boldsymbol{z}_i\,, \quad \boldsymbol{z}_i \sim \mathcal{N}(0, \mathbf{I})$$

3. Evaluate and sort the solutions, and give weights to them

4. Update the parameters $\boldsymbol{m}^{(g)}$, $\sigma^{(g)}$, $\mathbf{B}^{(g)}$

   $$\boldsymbol{m}^{(g+1)} = \boldsymbol{m}^{(g)} + \eta_m \sigma^{(g)} \mathbf{B}^{(g)} \mathbf{G}_\delta \qquad \mathbf{G}_\delta = \sum_{i=1}^{\lambda} w_i \boldsymbol{z}_i$$

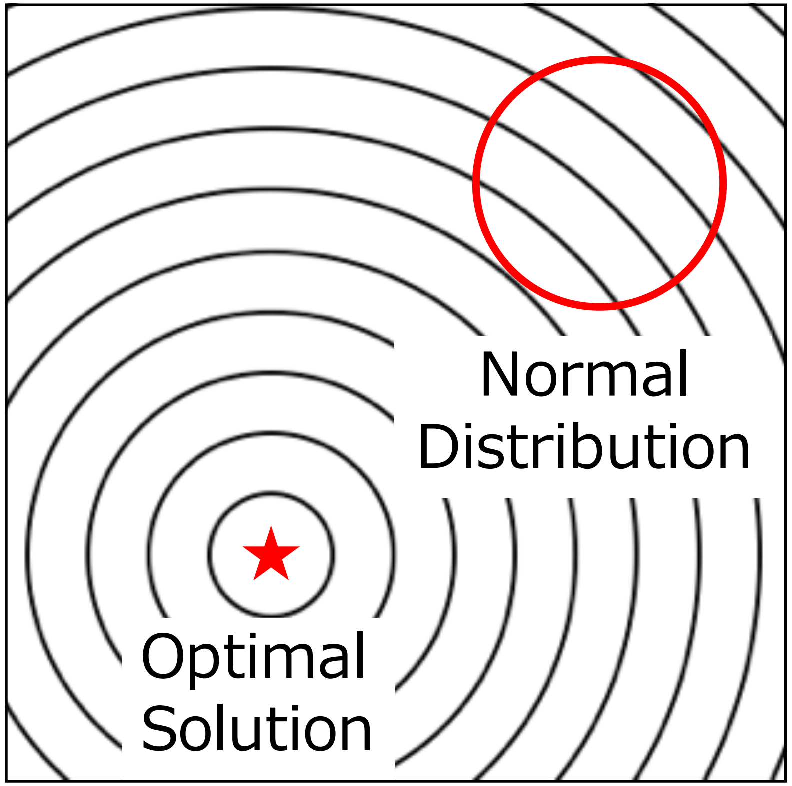
   $$\sigma^{(g+1)} = \sigma^{(g)} \exp(\eta_\sigma G_\sigma / 2) \qquad \mathbf{G}_\mathrm{M} = \sum_{i=1}^{\lambda} w_i (\boldsymbol{z}_i \boldsymbol{z}_i^{\mathrm{T}} - \mathbf{I})$$

   $$\mathbf{B}^{(g+1)} = \mathbf{B}^{(g)} \exp(\eta_\mathrm{B} \mathbf{G}_\mathrm{B} / 2) \qquad G_\sigma = \mathrm{tr}(\mathbf{G}_\mathrm{M})/d\,, \mathbf{G}_\mathrm{B} = \mathbf{G}_\mathrm{M} - G_\sigma \mathbf{I}$$

5. Emphasis on the expansion of the distribution

6. Terminate if the criterion is met,

   otherwise $g \leftarrow g + 1$ and go to Step 2.



FM-NES [Nomura and Ono, 2021]

1. Create $\mathcal{N}(\boldsymbol{m}^{(0)}, \mathbf{C}^{(0)} = {\sigma^{(0)}}^2 \mathbf{B}^{(0)} {\mathbf{B}^{(0)}}^{\mathrm{T}})$ and set $g = 0$

   $\sigma^{(0)}$: step size, $\mathbf{B}^{(0)}$: normalization transformation matrix

2. Generate λ solutions following $\mathcal{N}(\boldsymbol{m}^{(g)}, \mathbf{C}^{(g)})$

   $\boldsymbol{x}_i = \boldsymbol{m}^{(g)} + \sigma^{(g)} \mathbf{B}^{(g)} \boldsymbol{z}_i \,, \; \boldsymbol{z}_i \sim \mathcal{N}(0, \mathbf{I})$

3. Evaluate and sort the solutions, and give weights to them

4. Update the parameters $\boldsymbol{m}^{(g)}$, $\sigma^{(g)}$, $\mathbf{B}^{(g)}$

   $\boldsymbol{m}^{(g+1)} = \boldsymbol{m}^{(g)} + \eta_m \sigma^{(g)} \mathbf{B}^{(g)} \mathbf{G}_\delta \qquad \mathbf{G}_\delta = \sum_{i=1}^{\lambda} w_i \boldsymbol{z}_i$

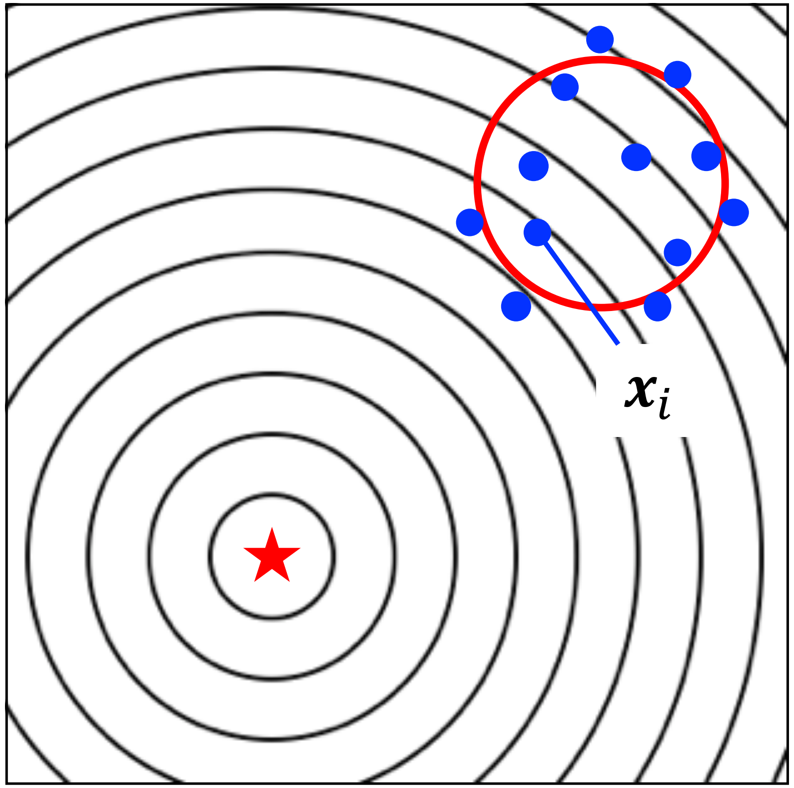
   $\sigma^{(g+1)} = \sigma^{(g)} \exp(\eta_\sigma G_\sigma / 2) \qquad \mathbf{G}_{\mathrm{M}} = \sum_{i=1}^{\lambda} w_i (\boldsymbol{z}_i \boldsymbol{z}_i^{\mathrm{T}} - \mathbf{I})$

   $\mathbf{B}^{(g+1)} = \mathbf{B}^{(g)} \exp(\eta_{\mathrm{B}} \mathbf{G}_{\mathrm{B}} / 2) \qquad G_\sigma = \mathrm{tr}(\mathbf{G}_{\mathrm{M}})/d \,, \mathbf{G}_{\mathrm{B}} = \mathbf{G}_{\mathrm{M}} - G_\sigma \mathbf{I}$

5. Emphasis on the expansion of the distribution

6. Terminate if the criterion is met,

   otherwise $g \leftarrow g + 1$ and go to Step 2.



FM-NES [Nomura and Ono, 2021]

1. Create $\mathcal{N}(\boldsymbol{m}^{(0)}, \mathbf{C}^{(0)} = {\sigma^{(0)}}^2 \mathbf{B}^{(0)} {\mathbf{B}^{(0)}}^{\mathrm{T}})$ and set $g = 0$

   $\sigma^{(0)}$: step size, $\mathbf{B}^{(0)}$: normalization transformation matrix

2. Generate λ solutions following $\mathcal{N}(\boldsymbol{m}^{(g)}, \mathbf{C}^{(g)})$

   $\boldsymbol{x}_i = \boldsymbol{m}^{(g)} + \sigma^{(g)} \mathbf{B}^{(g)} \boldsymbol{z}_i\,,\ \boldsymbol{z}_i \sim \mathcal{N}(0, \mathbf{I})$

3. Evaluate and sort the solutions, and give weights to them

4. Update the parameters $\boldsymbol{m}^{(g)}$, $\sigma^{(g)}$, $\mathbf{B}^{(g)}$

   $\boldsymbol{m}^{(g+1)} = \boldsymbol{m}^{(g)} + \eta_m \sigma^{(g)} \mathbf{B}^{(g)} \mathbf{G}_\delta \qquad \mathbf{G}_\delta = \sum_{i=1}^{\lambda} w_i \boldsymbol{z}_i$

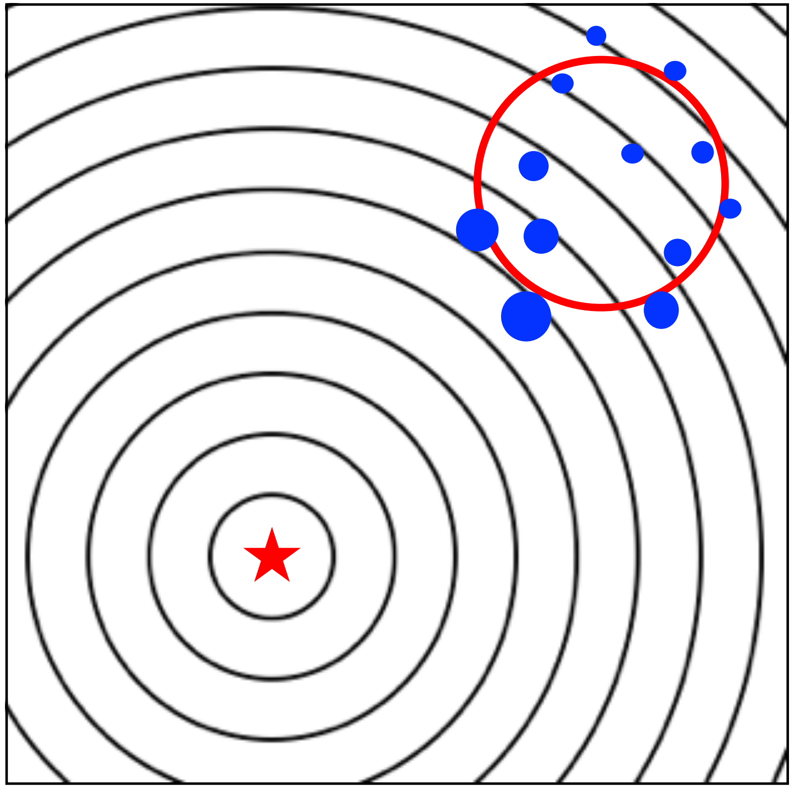
   $\sigma^{(g+1)} = \sigma^{(g)} \exp(\eta_\sigma G_\sigma / 2) \qquad \mathbf{G}_\mathrm{M} = \sum_{i=1}^{\lambda} w_i (\boldsymbol{z}_i \boldsymbol{z}_i^\mathrm{T} - \mathbf{I})$

   $\mathbf{B}^{(g+1)} = \mathbf{B}^{(g)} \exp(\eta_\mathrm{B} \mathbf{G}_\mathrm{B} / 2) \qquad G_\sigma = \mathrm{tr}(\mathbf{G}_\mathrm{M}) / d\,, \mathbf{G}_\mathrm{B} = \mathbf{G}_\mathrm{M} - G_\sigma \mathbf{I}$

5. Emphasis on the expansion of the distribution

6. Terminate if the criterion is met,

   otherwise $g \leftarrow g + 1$ and go to Step 2.



FM-NES [Nomura and Ono, 2021]

Distance weight [Fukushima et al., 2011]

- When the distribution is **moving**, it gives a large weight to a solution which has a good evalution value and **is far from the mean**.
- Otherwise, it gives a large weight to a solution which has a good evaluation value.

3. Evaluate and sort the solutions, and give **weights** to them

4. Update the parameters $\boldsymbol{m}^{(g)}$, $\sigma^{(g)}$, $\mathbf{B}^{(g)}$

$$\boldsymbol{m}^{(g+1)} = \boldsymbol{m}^{(g)} + \eta_m \sigma^{(g)} \mathbf{B}^{(g)} \mathbf{G}_\delta \qquad \mathbf{G}_\delta = \sum_{i=1}^{\lambda} \boxed{w_i} \boldsymbol{z}_i$$

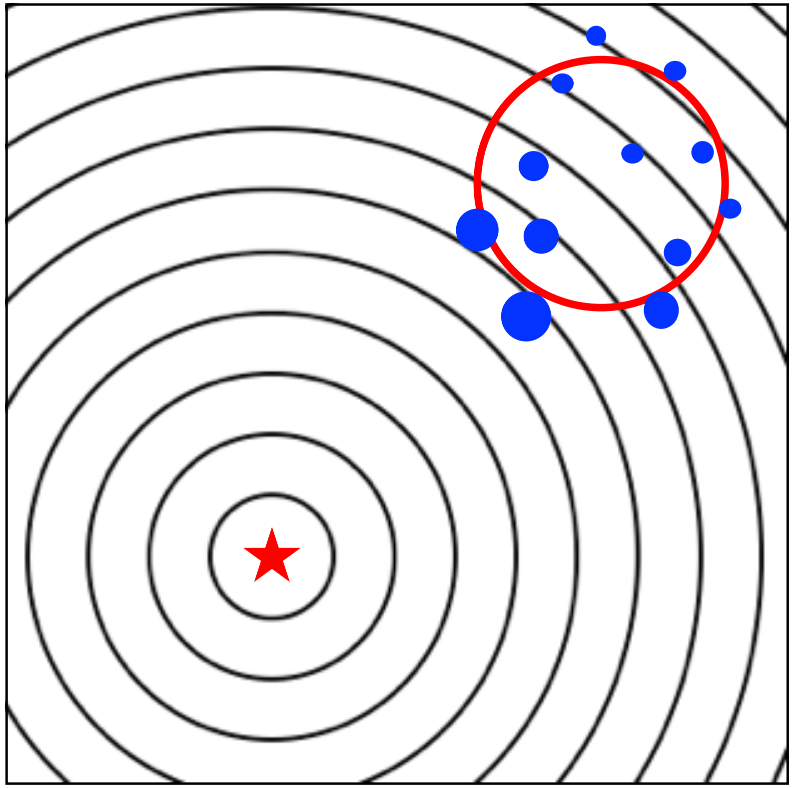
$$\sigma^{(g+1)} = \sigma^{(g)} \exp(\eta_\sigma G_\sigma / 2) \qquad \mathbf{G}_\mathrm{M} = \sum_{i=1}^{\lambda} \boxed{w_i} (\boldsymbol{z}_i \boldsymbol{z}_i^\mathrm{T} - \mathbf{I})$$

$$\mathbf{B}^{(g+1)} = \mathbf{B}^{(g)} \exp(\eta_\mathrm{B} \mathbf{G}_\mathrm{B} / 2) \qquad G_\sigma = \mathrm{tr}(\mathbf{G}_\mathrm{M})/d\,, \mathbf{G}_\mathrm{B} = \mathbf{G}_\mathrm{M} - G_\sigma \mathbf{I}$$

5. Emphasis on the expansion of the distribution

6. Terminate if the criterion is met,

   otherwise $g \leftarrow g + 1$ and go to Step 2.



FM-NES [Nomura and Ono, 2021]

1. Create $\mathcal{N}(\boldsymbol{m}^{(0)}, \mathbf{C}^{(0)} = {\sigma^{(0)}}^2 \mathbf{B}^{(0)} {\mathbf{B}^{(0)}}^{\mathrm{T}})$ and set $g = 0$

    $\sigma^{(0)}$: step size, $\mathbf{B}^{(0)}$: normalization transformation matrix

2. Generate λ solutions following $\mathcal{N}(\boldsymbol{m}^{(g)}, \mathbf{C}^{(g)})$

    $\boldsymbol{x}_i = \boldsymbol{m}^{(g)} + \sigma^{(g)} \mathbf{B}^{(g)} \boldsymbol{z}_i \,, \ \boldsymbol{z}_i \sim \mathcal{N}(0, \mathbf{I})$

3. Evaluate and sort the solutions, and give weights to them

4. Update the parameters $\boldsymbol{m}^{(g)}$, $\sigma^{(g)}$, $\mathbf{B}^{(g)}$

    $\boldsymbol{m}^{(g+1)} = \boldsymbol{m}^{(g)} + \eta_m \sigma^{(g)} \mathbf{B}^{(g)} \mathbf{G}_\delta \qquad \mathbf{G}_\delta = \sum_{i=1}^{\lambda} w_i \boldsymbol{z}_i$

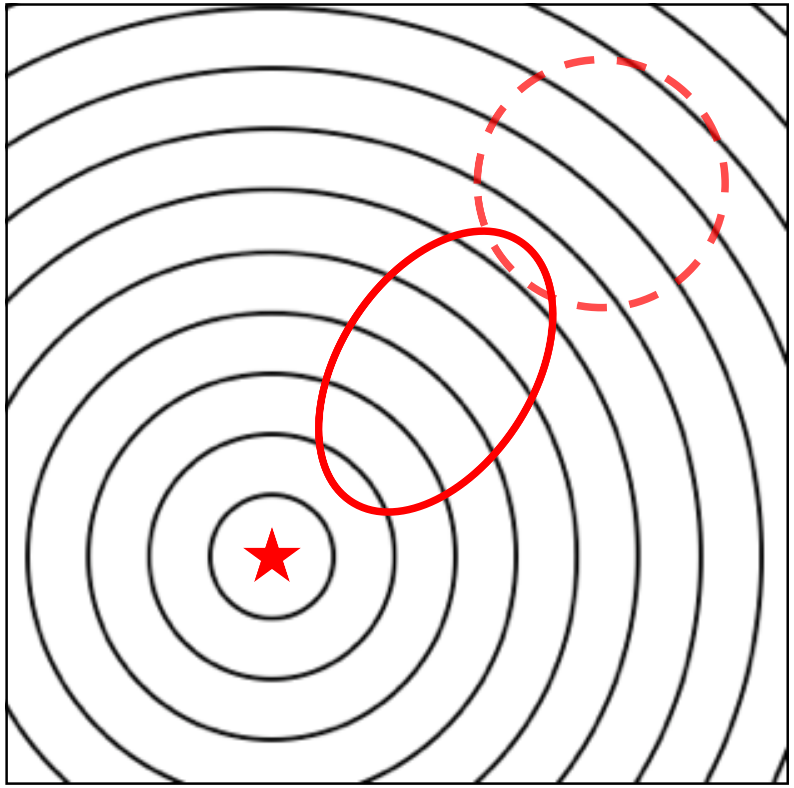
    $\sigma^{(g+1)} = \sigma^{(g)} \exp(\eta_\sigma G_\sigma / 2) \qquad \mathbf{G}_{\mathrm{M}} = \sum_{i=1}^{\lambda} w_i (\boldsymbol{z}_i \boldsymbol{z}_i^{\mathrm{T}} - \mathbf{I})$

    $\mathbf{B}^{(g+1)} = \mathbf{B}^{(g)} \exp(\eta_{\mathrm{B}} \mathbf{G}_{\mathrm{B}} / 2) \qquad G_\sigma = \mathrm{tr}(\mathbf{G}_{\mathrm{M}})/d \,, \mathbf{G}_{\mathrm{B}} = \mathbf{G}_{\mathrm{M}} - G_\sigma \mathbf{I}$

5. Emphasis on the expansion of the distribution

6. Terminate if the criterion is met,

    otherwise $g \leftarrow g + 1$ and go to Step 2.



FM-NES [Nomura and Ono, 2021]

1. Create $\mathcal{N}(\boldsymbol{m}^{(0)}, \mathbf{C}^{(0)} = {\sigma^{(0)}}^2 \mathbf{B}^{(0)} {\mathbf{B}^{(0)}}^{\mathrm{T}})$ and set $g = 0$

   $\sigma^{(0)}$: step size, $\mathbf{B}^{(0)}$: normalization transformation ma

2. Generate λ solutions following $\mathcal{N}(\boldsymbol{m}^{(g)}, \mathbf{C}^{(g)})$

   $$\boldsymbol{x}_i = \boldsymbol{m}^{(g)} + \sigma^{(g)} \mathbf{B}^{(g)} \boldsymbol{z}_i\,, \;\; \boldsymbol{z}_i \sim \mathcal{N}(0, \mathbf{I})$$

3. Evaluate and sort the solutions, and give weights to them

4. Update the parameters $\boldsymbol{m}^{(g)}$, $\sigma^{(g)}$, $\mathbf{B}^{(g)}$

   $$\boldsymbol{m}^{(g+1)} = \boldsymbol{m}^{(g)} + \eta_m \sigma^{(g)} \mathbf{B}^{(g)} \mathbf{G}_\delta \qquad \mathbf{G}_\delta = \textstyle\sum_{i=1}^{\lambda} w_i \boldsymbol{z}_i$$

   $$\sigma^{(g+1)} = \sigma^{(g)} \exp(\boxed{\eta_\sigma} G_\sigma / 2) \qquad \mathbf{G}_{\mathrm{M}} = \textstyle\sum_{i=1}^{\lambda} w_i (\boldsymbol{z}_i \boldsymbol{z}_i^{\mathrm{T}} - \mathbf{I})$$

   $$\mathbf{B}^{(g+1)} = \mathbf{B}^{(g)} \exp(\boxed{\eta_{\mathrm{B}}} \mathbf{G}_{\mathrm{B}} / 2) \qquad G_\sigma = \mathrm{tr}(\mathbf{G}_{\mathrm{M}})/d\,, \mathbf{G}_{\mathrm{B}} = \mathbf{G}_{\mathrm{M}} - G_\sigma \mathbf{I}$$

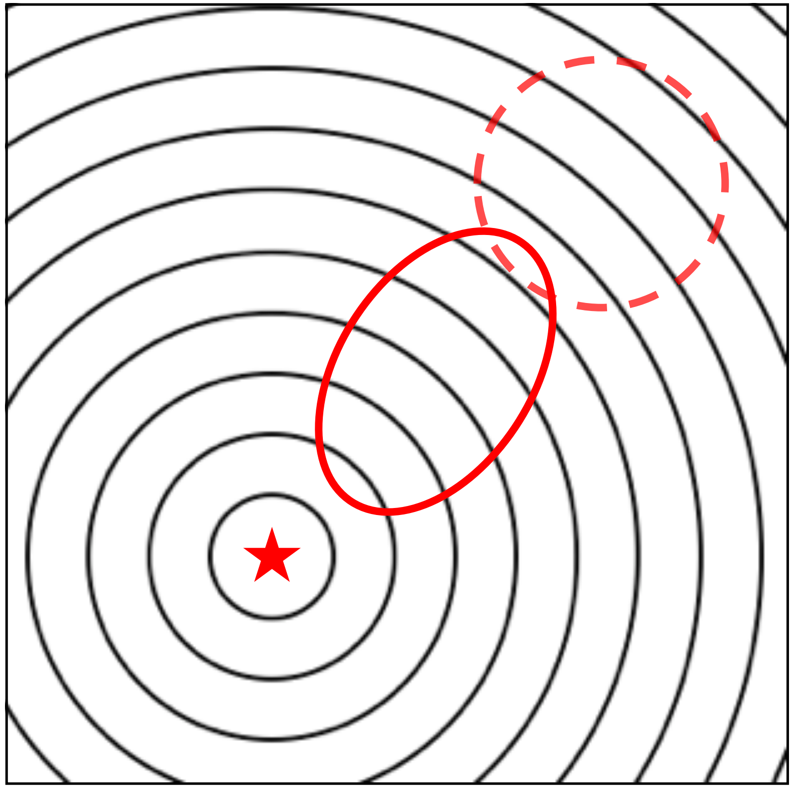
5. Emphasis on the expansion of the distribution

6. Terminate if the criterion is met,

   otherwise $g \leftarrow g + 1$ and go to Step 2.

Switching the learning rates if the distribution is moving

- Set **higher** learning rates $\eta_\sigma, \eta_{\mathrm{B}}$



FM-NES [Nomura and Ono, 2021]

1. Create $\mathcal{N}(\boldsymbol{m}^{(0)}, \mathbf{C}^{(0)} = {\sigma^{(0)}}^2 \mathbf{B}^{(0)} {\mathbf{B}^{(0)}}^{\mathrm{T}})$ and set $g = 0$

   $\sigma^{(0)}$: step size, $\mathbf{B}^{(0)}$: normalization transformation matrix

2. Generate λ solutions following $\mathcal{N}(\boldsymbol{m}^{(g)}, \mathbf{C}^{(g)})$

   $\boldsymbol{x}_i = \boldsymbol{m}^{(g)} + \sigma^{(g)} \mathbf{B}^{(g)} \boldsymbol{z}_i \,, \; \boldsymbol{z}_i \sim \mathcal{N}(0, \mathbf{I})$

3. Evaluate and sort the solutions, and give weights to them

4. Update the parameters $\boldsymbol{m}^{(g)}$, $\sigma^{(g)}$, $\mathbf{B}^{(g)}$

   $\boldsymbol{m}^{(g+1)} = \boldsymbol{m}^{(g)} + \eta_m \sigma^{(g)} \mathbf{B}^{(g)} \mathbf{G}_\delta \qquad \mathbf{G}_\delta = \sum_{i=1}^{\lambda} w_i \boldsymbol{z}_i$

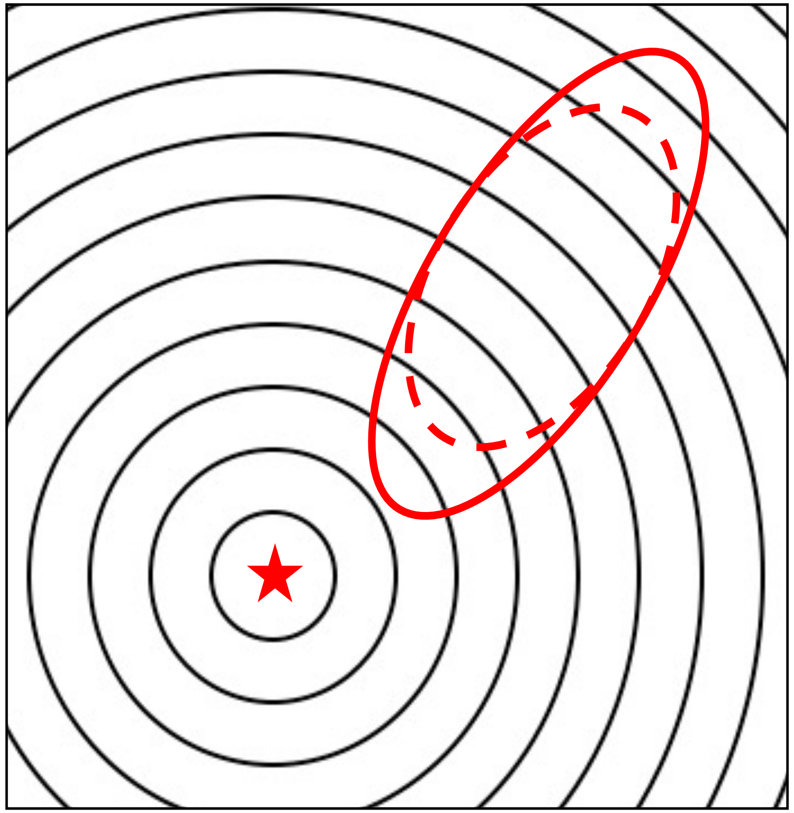
   $\sigma^{(g+1)} = \sigma^{(g)} \exp(\eta_\sigma G_\sigma / 2) \qquad \mathbf{G}_{\mathrm{M}} = \sum_{i=1}^{\lambda} w_i (\boldsymbol{z}_i \boldsymbol{z}_i^{\mathrm{T}} - \mathbf{I})$

   $\mathbf{B}^{(g+1)} = \mathbf{B}^{(g)} \exp(\eta_{\mathrm{B}} \mathbf{G}_{\mathrm{B}} / 2) \qquad G_\sigma = \mathrm{tr}(\mathbf{G}_{\mathrm{M}})/d \,, \mathbf{G}_{\mathrm{B}} = \mathbf{G}_{\mathrm{M}} - G_\sigma \mathbf{I}$

5. Emphasis on the expansion of the distribution

6. Terminate if the criterion is met,

   otherwise $g \leftarrow g + 1$ and go to Step 2.



FM-NES [Nomura and Ono, 2021]

1. Create $\mathcal{N}(\boldsymbol{m}^{(0)}, \mathbf{C}^{(0)} = {\sigma^{(0)}}^2 \mathbf{B}^{(0)} {\mathbf{B}^{(0)}}^{\mathrm{T}})$ and set $g = 0$

   $\sigma^{(0)}$: step size, $\mathbf{B}^{(0)}$: normalization transformation matrix

2. Generate λ solutions following $\mathcal{N}(\boldsymbol{m}^{(g)}, \mathbf{C}^{(g)})$

   $\boldsymbol{x}_i = \boldsymbol{m}^{(g)} + \sigma^{(g)} \mathbf{B}^{(g)} \mathbf{z}_i \,, \ \mathbf{z}_i \sim \mathcal{N}(0, \mathbf{I})$

3. Evaluate and sort the solutions, and give weights to them

4. Update the parameters $\boldsymbol{m}^{(g)}$, $\sigma^{(g)}$, $\mathbf{B}^{(g)}$

   $\boldsymbol{m}^{(g+1)} = \boldsymbol{m}^{(g)} + \eta_m \sigma^{(g)} \mathbf{B}^{(g)} \mathbf{G}_\delta \qquad \mathbf{G}_\delta = \sum_{i=1}^{\lambda} w_i \mathbf{z}_i$

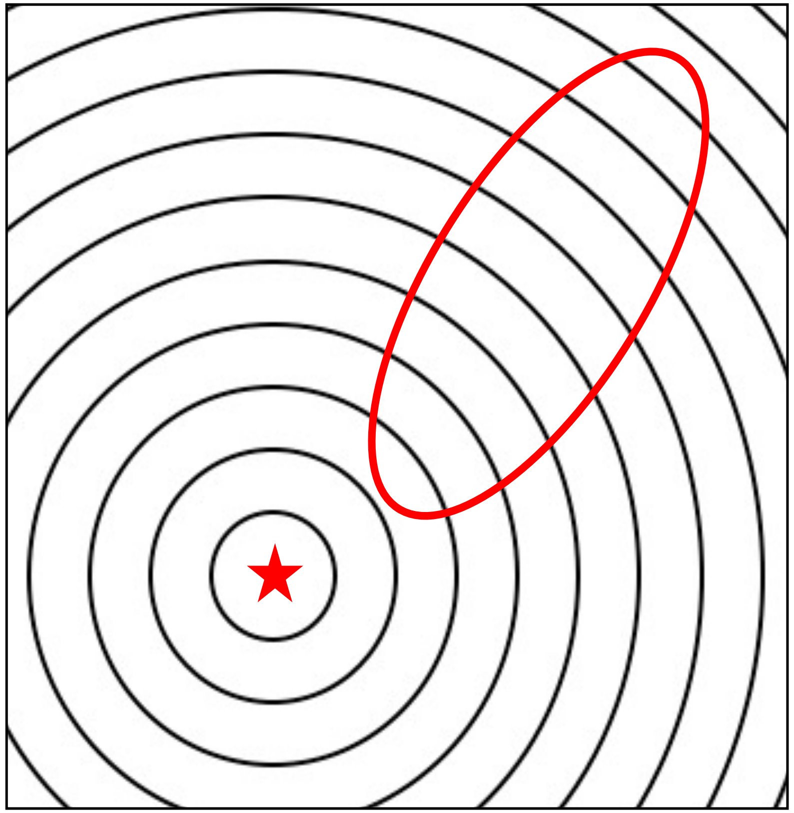
   $\sigma^{(g+1)} = \sigma^{(g)} \exp(\eta_\sigma G_\sigma / 2) \qquad \mathbf{G}_\mathrm{M} = \sum_{i=1}^{\lambda} w_i (\mathbf{z}_i \mathbf{z}_i^{\mathrm{T}} - \mathbf{I})$

   $\mathbf{B}^{(g+1)} = \mathbf{B}^{(g)} \exp(\eta_\mathrm{B} \mathbf{G}_\mathrm{B} / 2) \qquad G_\sigma = \mathrm{tr}(\mathbf{G}_\mathrm{M}) / d \,, \mathbf{G}_\mathrm{B} = \mathbf{G}_\mathrm{M} - G_\sigma \mathbf{I}$

5. Emphasis on the expansion of the distribution

6. Terminate if the criterion is met,

   otherwise $g \leftarrow g + 1$ and go to Step 2.



Problem of FM-NES

Difficult to apply FM-NES to high-dimensional problems

Main reasons:

- **Time complexity:** $\boldsymbol{\mathcal{O}(d^3)}$ … eigenvalue decomposition
- **Space complexity:** $\boldsymbol{\mathcal{O}(d^2)}$ … preserving matrix

d: the number of dimension

Cost-Reduction FM-NES (CR-FM-NES)

CR-FM-NES restricts representation of covariance matrix

$$\mathbf{C} = \sigma^2 \mathbf{D}\left(\mathbf{I} + \boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}\right)\mathbf{D},$$

$\mathbf{D} \in \mathbb{R}^{d\times d}$: diagonal matrix

$\boldsymbol{v} \in \mathbb{R}^{d}$: d-dimensional vector

$\sigma \in \mathbb{R}$: step size (scalar)

- This representation is originally used in VD-CMA [Akimoto et al., 2014]

The number of parameters in covariance matrix is reduced

from $\frac{d(d+1)}{2} + 1 \in \Theta(d^2)$ to $2d + 1 \in \Theta(d)$

Main Point: CR-FM-NES has linear space complexity

Parameter Update in Linear Time Complexity

$\boldsymbol{v}$ and $\mathbf{D}$ Update:

- The update is performed by the estimated natural gradient
- This is the same as VD-CMA and takes linear time complexity

σ Update:

- The update is also performed by the estimated natural gradient
- This is similar to FM-NES and takes linear time complexity

$$\sigma^{(g+1)} = \sigma^{(g)} \exp(\eta_\sigma G_\sigma / 2),$$
$$G_\sigma = \frac{\mathrm{Tr}\left(\sum_{i=1}^{\lambda} w_i \left(\boldsymbol{z}_i \boldsymbol{z}_i^{\mathrm{T}} - \mathbf{I}\right)\right)}{d}.$$

Main Point: CR-FM-NES has linear time complexity

Comparison: CR-FM-NES and VD-CMA

The difference come from the **difference between FM-NES and CMA-ES**

In FM-NES,

- Step-size adaptation is based on **natural gradient**
- Several operations added to **improve performance**:
  - Adaptive weighting strategy (Distance weight)
  - Adaptive learning rates based on estimation accuracy of natural gradient

Main Point: CR-FM-NES inherits the efficiency of FM-NES

Experiments

Research Questions (RQs)

- **RQ1. How is the performance of CR-FM-NES, compared to FM-NES?**

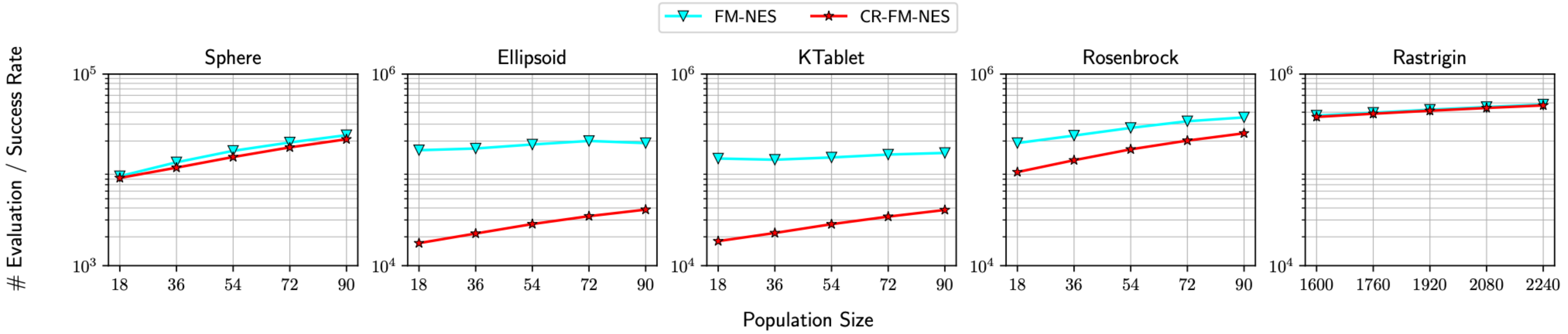
- **RQ2. Is CR-FM-NES more efficient than baseline methods for high-dimensional problems?**

Experiments: Benchmark Problems

Name	Definition
Sphere	$f_{\mathrm{sph}}(\mathbf{x}) = \sum_{i=1}^{d} \mathbf{x}_i^2$
k-Tablet	$f_{k\mathrm{tab}}(\mathbf{x}) = \sum_{i=1}^{k} \mathbf{x}_i^2 + \sum_{i=k+1}^{d} (100\mathbf{x}_i)^2$
Ellipsoid	$f_{\mathrm{ell}}(\mathbf{x}) = \sum_{i=1}^{d} (1000^{\frac{i-1}{d-1}} \mathbf{x}_i)^2$
Rosenbrock	$f_{\mathrm{rosen}}(\mathbf{x}) = \sum_{i=1}^{d-1} 100(\mathbf{x}_{i+1} - \mathbf{x}_i^2)^2 + (\mathbf{x}_i - 1)^2$
Rastrigin	$f_{\mathrm{rast}}(\mathbf{x}) = 10d + \sum_{i=1}^{d} (\mathbf{x}_i^2 - 10\cos(2\pi\mathbf{x}_i))$

The initial parameters are set to locate in the outside of the optimum.

RQ1: Performance Comparison (d=80)



For Ellipsoid, Ktablet and Rosenbrock, CR-FM-NES has achieved a **significant speedup**

This is because the number of parameters in CR-FM-NES is $\Theta(d)$

→ **enables to set the learning rates much higher than those in FM-NES**

RQ1: Computational Time

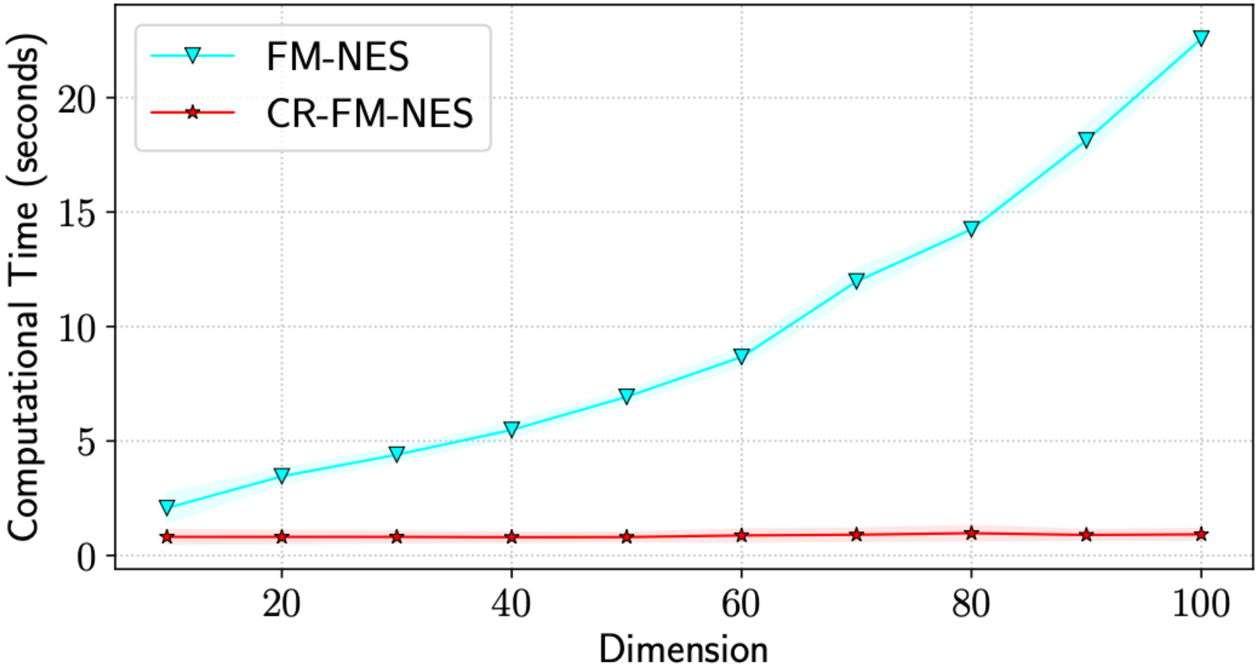


Figure: Computational time required to execute 1,000 iterations

The computational time of FM-NES increases rapidly,

while **that of CR-FM-NES is fairly small**

RQ2: Comparison in High-Dimension

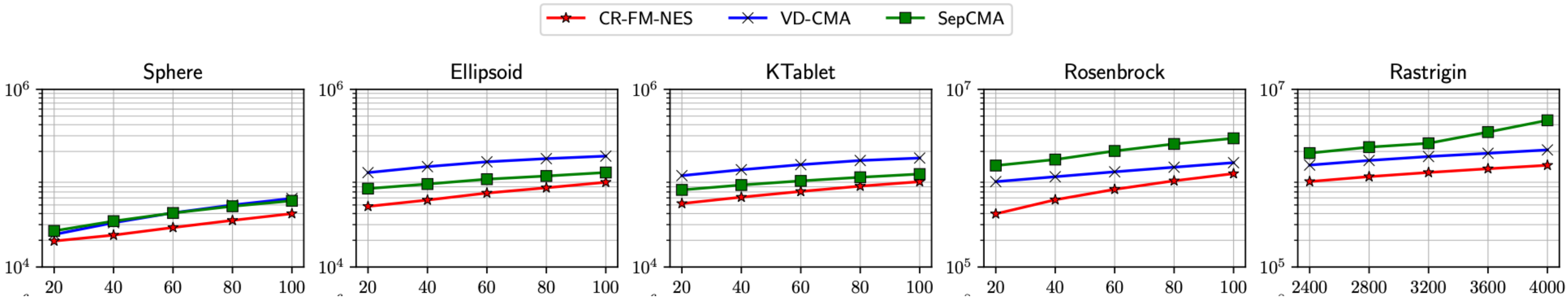
Compared Methods:

- VD-CMA [Akimoto et al., 2014]
- Sep-CMA [Ros and Hansen, 2008]

Settings:

- dimension $d = 200, 600, 1000$
- varying the population size

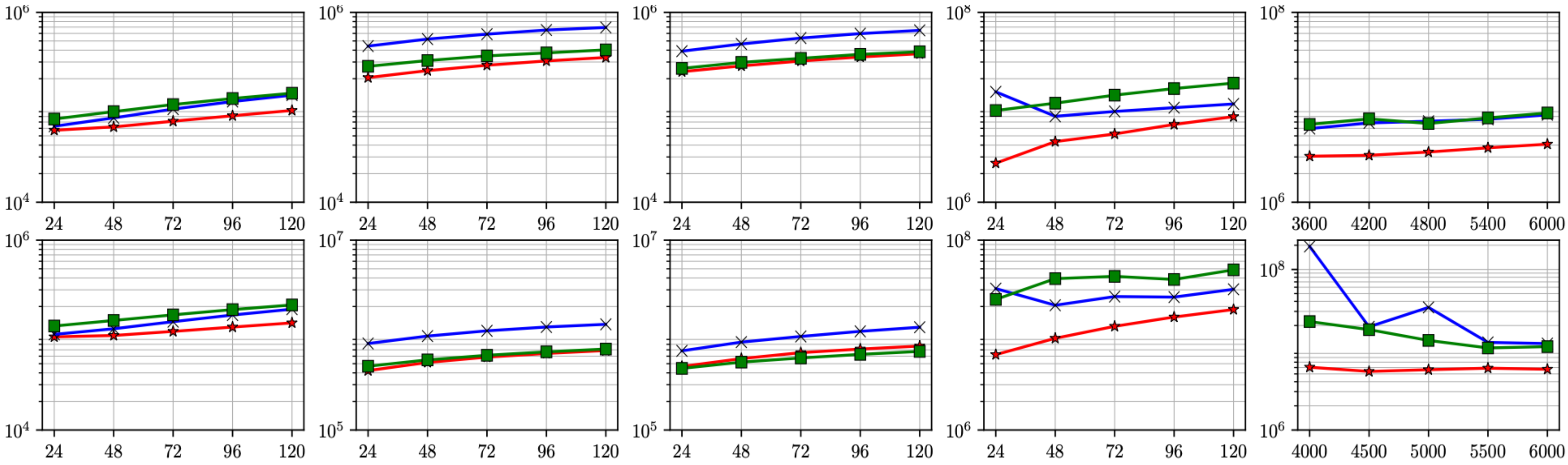
RQ2: Comparison in High-Dimension (d=200)



- CR-FM-NES outperforms VD-CMA in all the functions
- Even in the problems where inverse Hessian can be approx. by diagonal, CR-FM-NES is slightly more efficient or competitive to Sep-CMA

CR-FM-NES is more efficient or at least competitive

RQ2: Comparison in High-Dimension (d=600,1000)



The results in d=600 and 1,000 are similar to that in d=200

Main Point: CR-FM-NES is quite promising on high-dimensional BBO

Summary

- **Problem :**
  - High-dimensional black-box optimization (BBO)
- **Our approach : CR-FM-NES**
  - Scalable version of FM-NES
  - Linear time and space complexity via restricted representation of covariance matrix
- **Evaluation :**
  - 200, 600, 1,000-dimensional benchmark problems

Last Message

- Please use CR-FM-NES when to solve high-dim. BBO
- Python implementation is available at github:

  https://github.com/nomuramasahir0/crfmnes

  - You can use CR-FM-NES via **pip install crfmnes**
  - C# and C++ versions are also available

Thank you for listening!

References

- [Akimoto et al., 2014] Y. Akimoto, A. Auger, and N. Hansen, “Comparison-Based Natural Gradient Optimization in High Dimension,” in Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation. ACM, 2014, pp. 373–380.
- [Fukushima et al., 2011] N. Fukushima, Y. Nagata, S. Kobayashi, and I. Ono, “Proposal of distance-weighted exponential natural evolution strategies,” in 2011 IEEE Congress of Evolutionary Computation (CEC). IEEE, 2011, pp.164–171.
- [Glasmachers et al., 2010] T. Glasmachers, T. Schaul, S. Yi, D. Wierstra, and J. Schmidhuber, “Exponential Natural Evolution Strategies,” in Proceedings of the 12th annual conference on Genetic and evolutionary computation. ACM, 2010, pp.393–400.
- [Hansen and Ostermeier, 2001] N. Hansen and A. Ostermeier, “Completely Derandomized Self- Adaptation in Evolution Strategies,” Evolutionary computation, 2001, vol. 9, no. 2, pp. 159–195.
- [Nomura and Ono, 2021] M. Nomura and I. Ono, “Natural evolution strategy for unconstrained and implicitly constrained problems with ridge structure,” in 2021 IEEE Symposium Series on Computational Intelligence (SSCI). IEEE, 2021, pp. 1–7.
- [Ros and Hansen, 2008] R. Ros and N. Hansen, “A Simple Modification in CMA-ES Achieving Linear Time and Space Complexity,” in International Conference on Parallel Problem Solving from Nature. Springer, 2008, pp. 296–305.