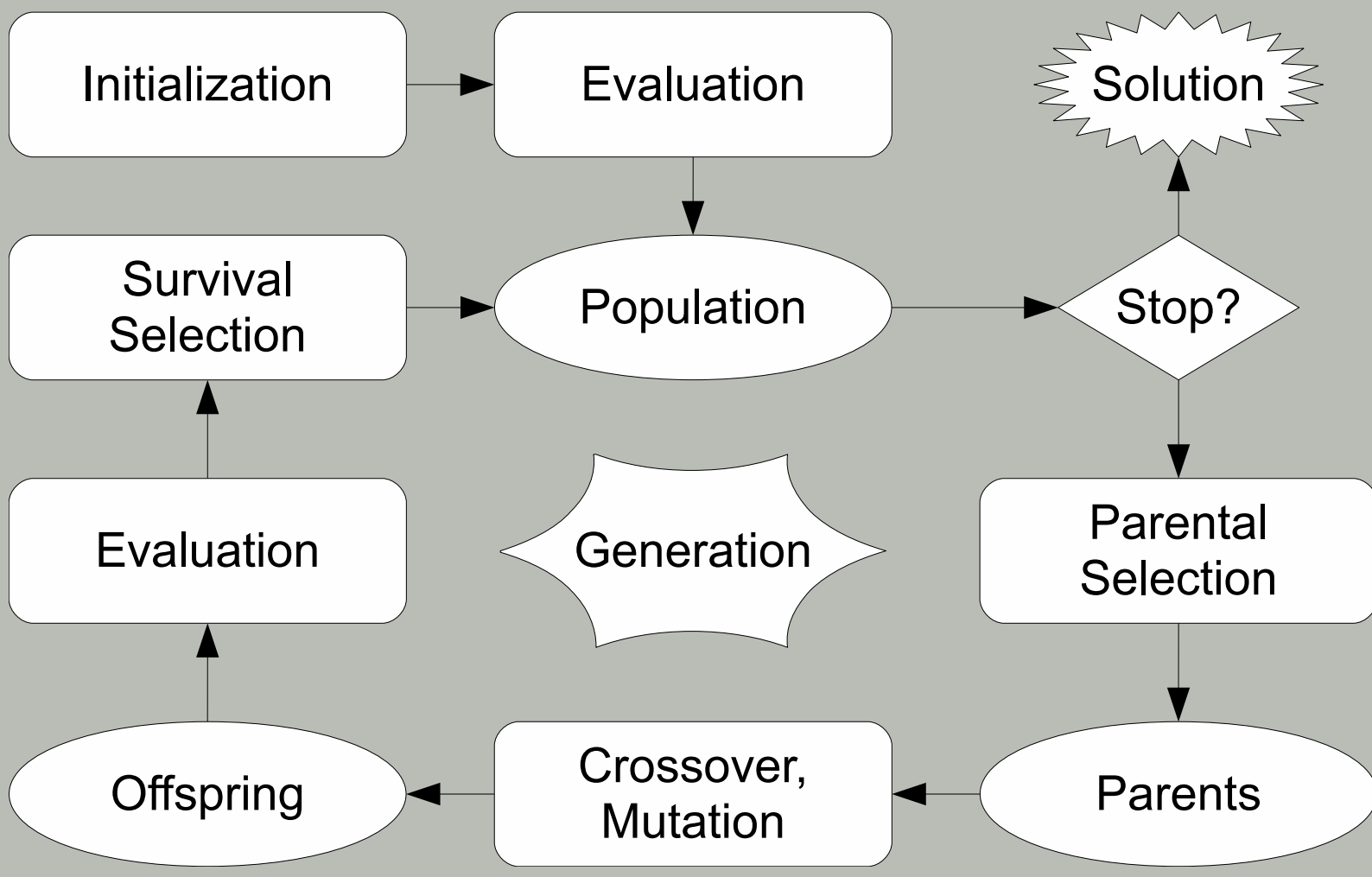


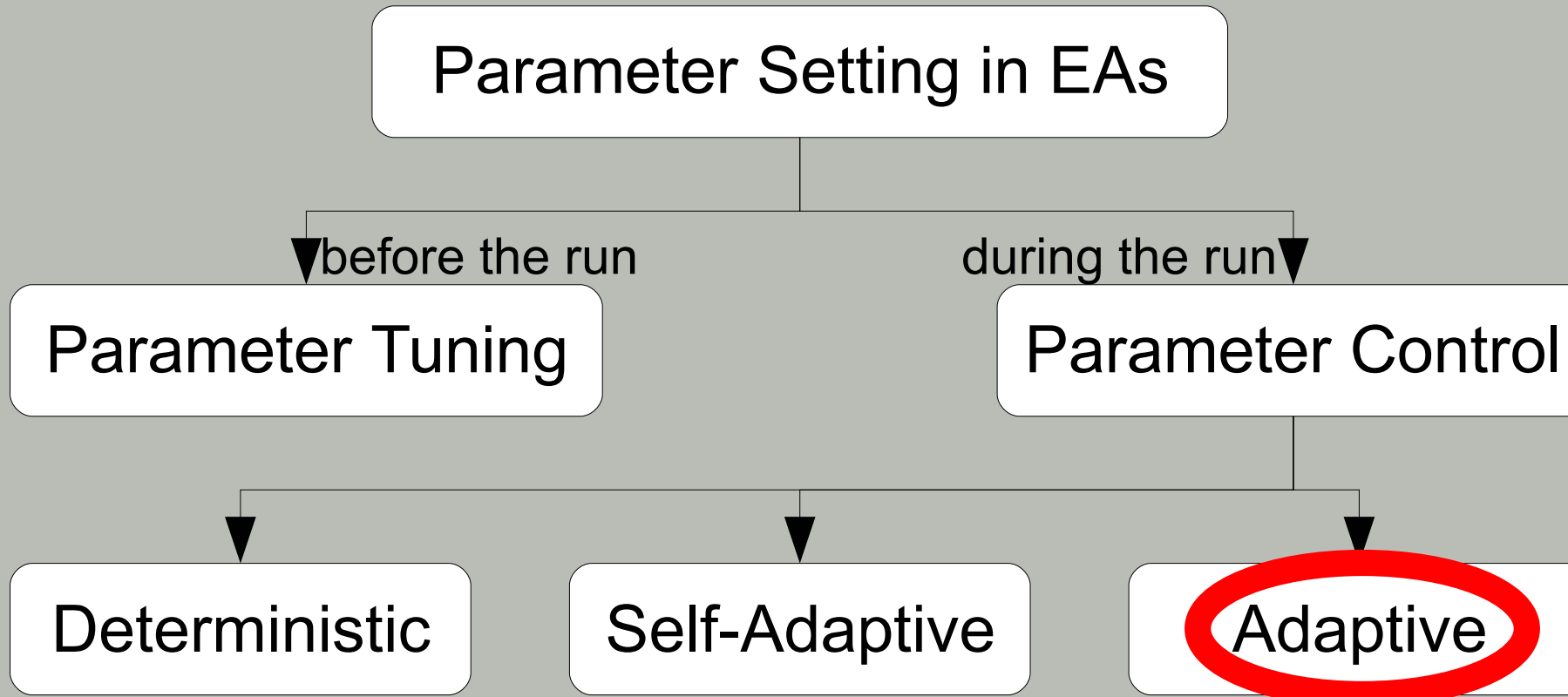
# Comparison-based Adaptive Strategy Selection with Bandits in Differential Evolution

Álvaro Fialho<sup>1</sup>, Raymond Ros<sup>2</sup>, Marc Schoenauer<sup>1,2</sup> and Michèle Sebag<sup>1,2</sup>

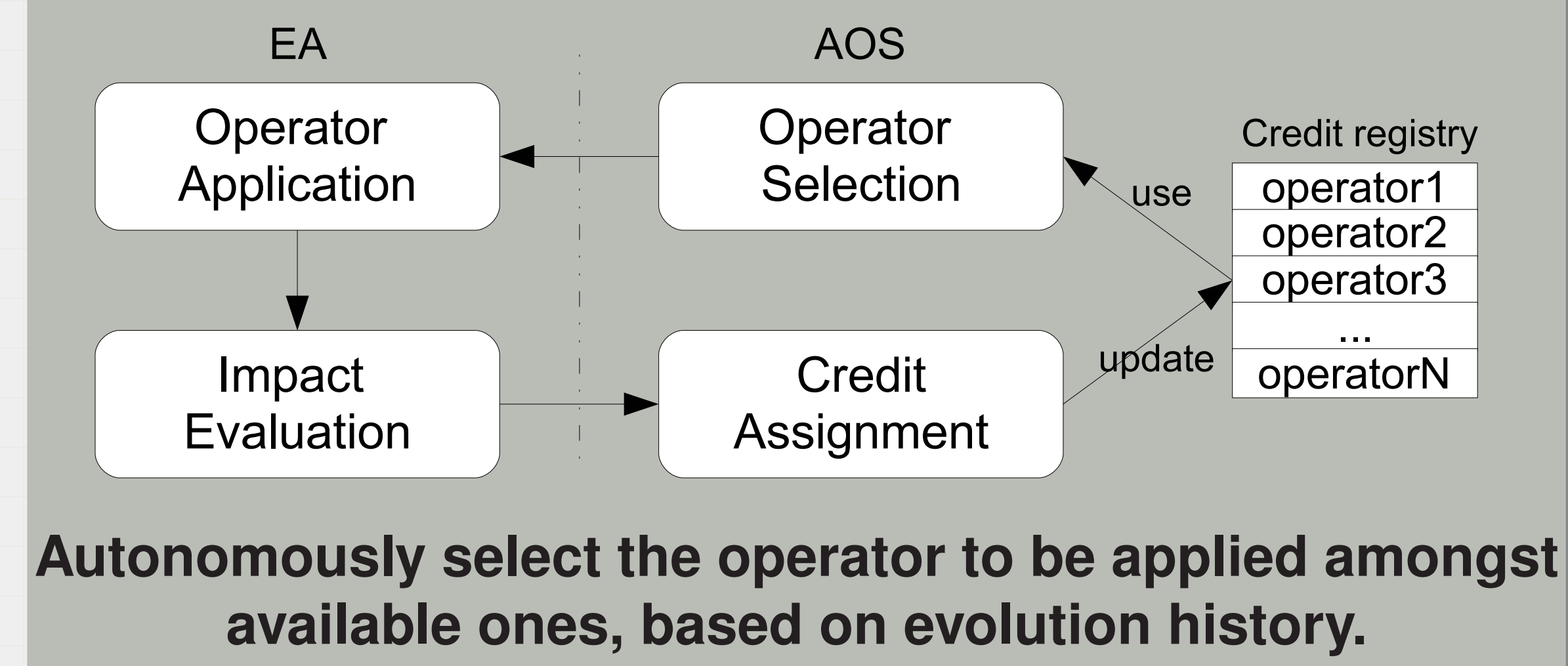
## Evolutionary Algorithms



## Parameter Setting in EAs



## Adaptive Operator Selection



## Credit Assignment:

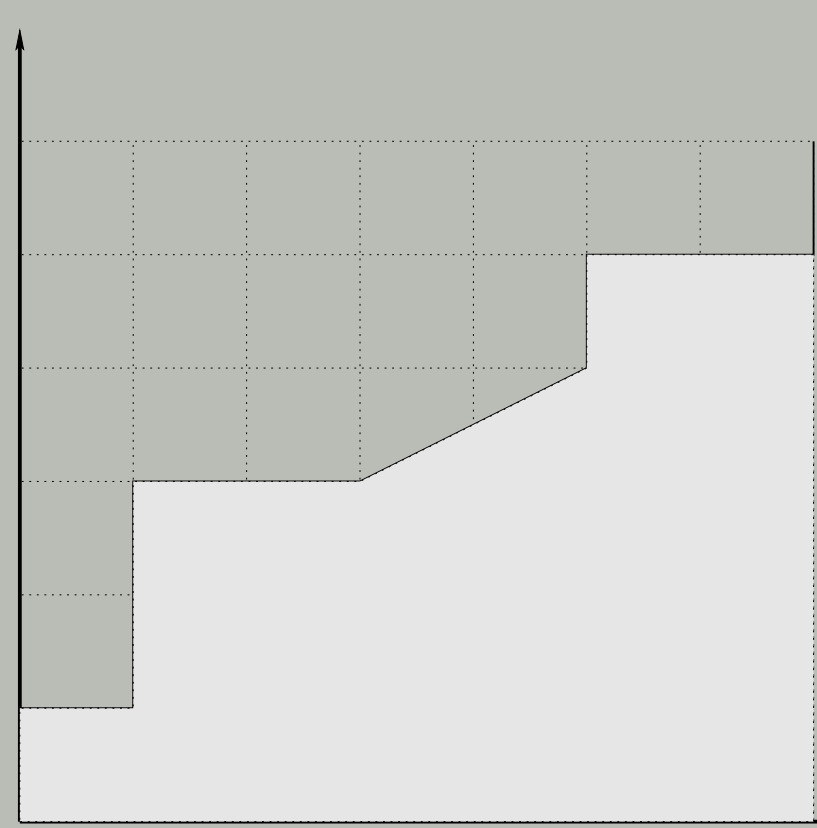
## Area Under the Curve

### Receiver Operating Characteristic curve

- ML: compare binary classifiers
- AOS: compare 1 vs. all operators

### Rank-based Credit Assignment

- Credits received by operators are ranked
- Position  $r$  is assigned value  $D^r(W - r)$
- Size of the segment = assigned rank-value



ROC curve with  $D = 1$

## Operator Selection:

## Multi-Armed Bandit

Yet another *Exploration vs. Exploitation* dilemma.



Multi-Armed Bandits  
At time  $t$ , gambler plays arm  $j$

$$r_{j,t} = \begin{cases} 1 & \text{with prob} = p_j \\ 0 & \text{with prob} = 1 - p_j \end{cases}$$

### Original UCB algorithm [Auer 2002]

- Be optimistic in face of the unknown.
- At time  $t$ , choose arm  $j$  maximizing:

$$\hat{q}_{j,t} = \hat{r}_{j,t} + S \sqrt{\frac{2 \log \sum_k n_{k,t}}{n_{j,t}}}, \text{ where } \begin{cases} \hat{r}_{j,t} & \text{empirical reward for arm } j \\ n_{j,t} & \text{number of times arm } j \text{ was played} \end{cases}$$

### AUC-UCB algorithm

- AUC is already an aggregation
- directly use AUC in UCB equation

$$\hat{q}_{j,t} = AUC_{j,t} + S \sqrt{\frac{2 \log \sum_k n_{k,t}}{n_{j,t}}}$$

## Previous Bandit-based AOS

Operator Selection: Dynamic-MAB = UCB + Page-Hinkley  
Credit Assignment: Extreme stat. on Fitness Improvements

### Issues

- Original MAB: very slow adaptation (AOS is dynamic)
  - AUC incorporates the behavior of **all operators**
  - Rewarding **one operator** affects **all AUCs**
- Extreme D-MAB: highly problem-dependent
- AUC is **rank-based**

## Black-Box Optimization Benchmarking within DE

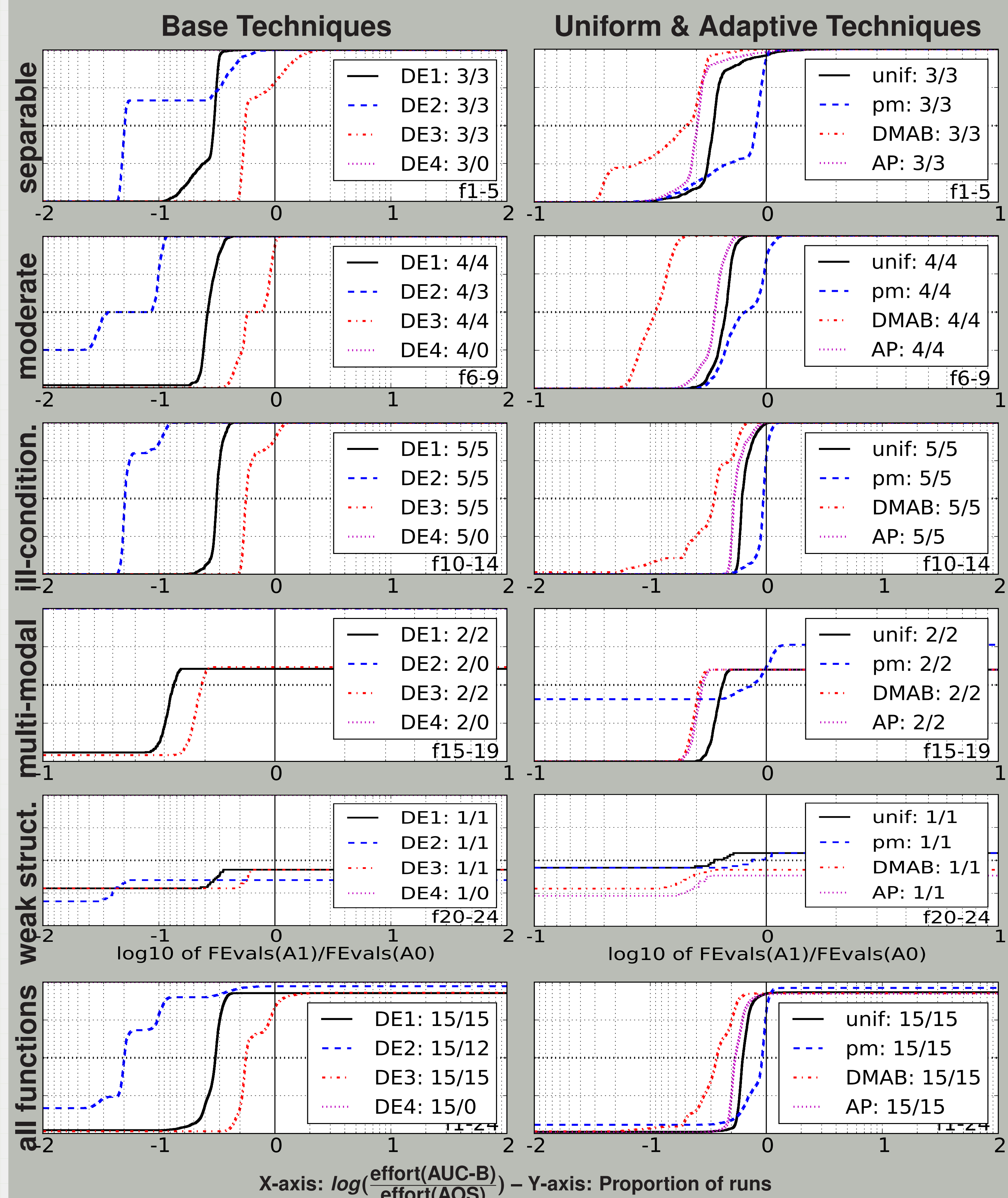
### Standard Differential Evolution algorithm

$NP = 100 \cdot DIM$ ;  $CR = 1.0$ ;  $F = 0.5$ ; 4 mutation strategies:

- “rand/1”:  $v_i = x_{r_1} + F \cdot (x_{r_2} - x_{r_3})$
- “rand/2”:  $v_i = x_{r_1} + F \cdot (x_{r_2} - x_{r_3}) + F \cdot (x_{r_4} - x_{r_5})$
- “rand-to-best/2”:  $v_i = x_{r_1} + F \cdot (x_{best} - x_{r_1}) + F \cdot (x_{r_2} - x_{r_3}) + F \cdot (x_{r_4} - x_{r_5})$
- “current-to-rand/1”:  $v_i = x_i + F \cdot (x_{r_1} - x_i) + F \cdot (x_{r_2} - x_{r_3})$

**BBOB**: 24 single-objective continuous functions, 5 classes

## Comparative Results



## Discussion

- Rank over  $F$ , instead of  $\Delta F$ : **comparison-based** AOS method
- Very **robust** w.r.t. its hyper-parameters ( $C.5D.5W50$ )
- Better** than fixed, naive and known adaptive approaches.
- Poor facing multi-modality → consider **diversity**