Benchmarking the SMS-EMOA with Self-adaptation on the bbob-biobj Test Suite

Simon Wessing

Chair of Algorithm Engineering
Computer Science Department
Technische Universität Dortmund

16 July 2017

Introduction

- Evolutionary multiobjective optimization
- Continuous decision variables
- $(1+1)$-SMS-EMOA is algorithmically equivalent to single-objective $(1+1)$-EA

$\Rightarrow$ Theory about optimal step size from single-objective optimization applies

Introduction

- Evolutionary multiobjective optimization
- Continuous decision variables
- $(1+1)$-SMS-EMOA is algorithmically equivalent to single-objective $(1+1)$-EA

$\Rightarrow$ Theory about optimal step size from single-objective optimization applies

- Situation for $(\mu+1)$, $(\mu+\lambda)$ unknown
- How to define step size optimality?
- How to adapt step size if not with very sophisticated MO-CMA-ES?

Development of Control Mechanism

- Idea: use self-adaptation from single-objective optimization

Development of Control Mechanism

- Idea: use self-adaptation from single-objective optimization
- Mutation of genome: $\boldsymbol{y} = \boldsymbol{x} + \sigma\mathcal{N}(\mathbf{0}, \mathbf{I})$
- Mutation of step size: $\sigma = \tilde{\sigma} \cdot \exp(\tau\mathcal{N}(0, 1))$
- Learning parameter $\tau \propto 1/\sqrt{n}$

Development of Control Mechanism

- Idea: use self-adaptation from single-objective optimization
- Mutation of genome: $\boldsymbol{y} = \boldsymbol{x} + \sigma \mathcal{N}(\mathbf{0}, \mathbf{I})$
- Mutation of step size: $\sigma = \tilde{\sigma} \cdot \exp(\tau \mathcal{N}(0, 1))$
- Learning parameter $\tau \propto 1/\sqrt{n}$

- Not state of the art any more
- Behavior is **emergent**
- Theoretical analysis is difficult
- Application to multiobjective optimization is scarce

$\Rightarrow$ Experiment to find good parameter configurations

Experimental Setup

Factor	Type	Symbol	Levels
Number variables	observable	n	$\{2, 3, 5, 10, 20\}$
Learning param. constant	control	c	$\{2^{-2}, 2^{-1}, \mathbf{2^0}, 2^1, 2^2, 2^3\}$
Population size	control	μ	$\{10, 50\}$
Number offspring	control	λ	$\{1, \mu, 5\mu\}$
Recombination	control		{discrete, intermediate, arithmetic, none}

- Full factorial design
- 15 **unimodal** problems of BBOB-BIOBJ 2016 (only first instance)
- Budget: $10^4 n$ function evaluations
- Assessment: rank-transformed HV values **of whole EA runs**

Other Factors Held Constant

- Initial mutation strength $\sigma_{\text{init}} = 0.025$
- Repair method for bound violations: Lamarckian reflection (search space $[-100, 100]^n$, scaled to unit hypercube)
- Selection: iteratively removes worst individual, until μ reached (backward elimination)

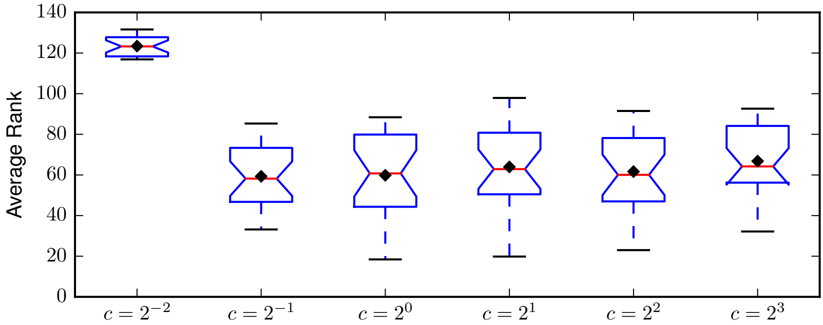
$\Rightarrow$ Might have to reconsider in the future

Pseudocode

Input: population size μ, initial population P_0, number of offspring λ

$t \leftarrow 0$
while stopping criterion not fulfilled **do**
    $O_t \leftarrow \text{createOffspring}(P_t)$ // create λ offspring
    $\text{evaluate}(O_t)$ // calculate objective values
    $Q_t \leftarrow P_t \cup O_t$
    $\boldsymbol{r} \leftarrow \text{createReferencePoint}(Q_t)$
    while $|Q_t| > \mu$ **do**
        $\{F_1, \dots, F_w\} \leftarrow \text{nondominatedSort}(Q_t)$ // sort in fronts
        $\boldsymbol{x}^* \leftarrow \text{argmin}_{\boldsymbol{x} \in F_w}(\Delta_s(\boldsymbol{x}, F_w, \boldsymbol{r}))$ // $\boldsymbol{x}^*$ with smallest contr.
        $Q_t \leftarrow Q_t \setminus \{\boldsymbol{x}^*\}$ // remove worst individual
    end while
    $P_{t+1} \leftarrow Q_t$
    $t \leftarrow t + 1$
end while

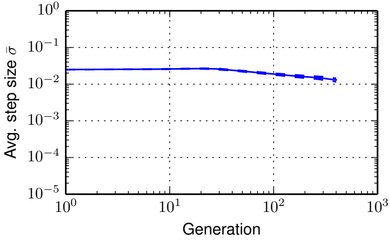
Main Effect: Learning Parameters $\tau = c/\sqrt{n}$



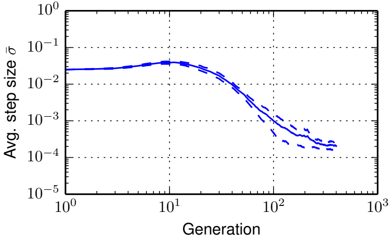
- $c = 2^{-2}$ is always the worst choice

$\Rightarrow$ Exclude $c = 2^{-2}$ from further analysis

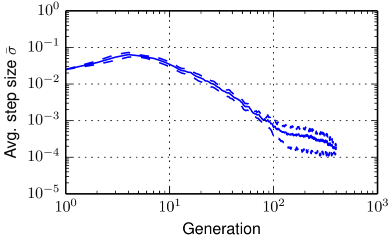
Mutation Strength vs. Generation



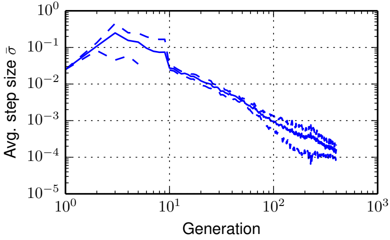
(a) $\tau = 2^{-2}/\sqrt{n}$.



(b) $\tau = 2^{0}/\sqrt{n}$.

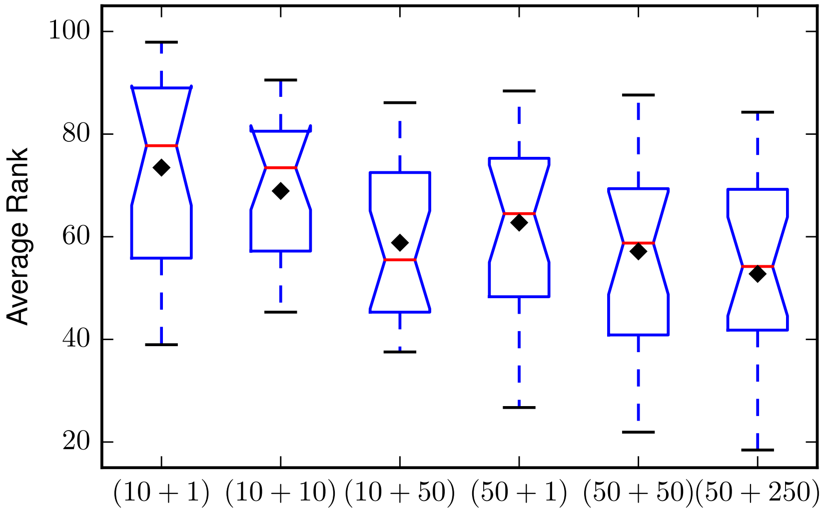


(c) $\tau = 2^{2}/\sqrt{n}$.

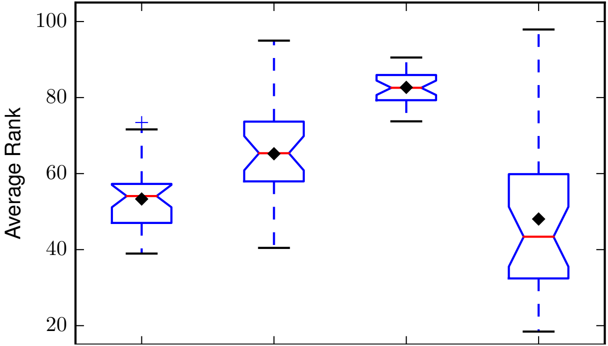


(d) $\tau = 2^{3}/\sqrt{n}$.

Main Effect: Selection Variants



Main and Interaction Effects: Recombination & Selection

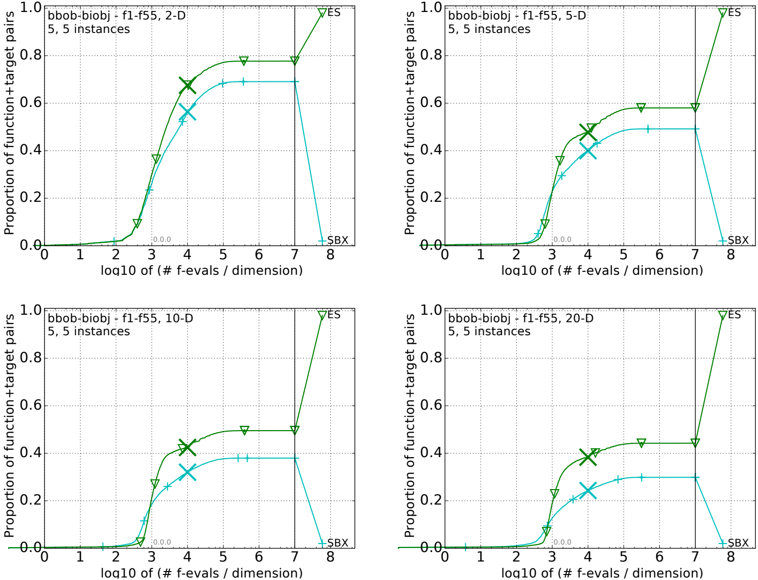


	arithmetic	discrete	intermediate	none
$(10+1)$	46.97	85.43	82.53	78.95
$(10+10)$	51.29	72.55	83.48	68.34
$(10+50)$	47.69	62.90	82.25	42.50
$(50+1)$	61.93	63.21	84.93	40.95
$(50+50)$	58.23	55.88	84.06	30.43
$(50+250)$	53.77	51.34	78.82	27.14

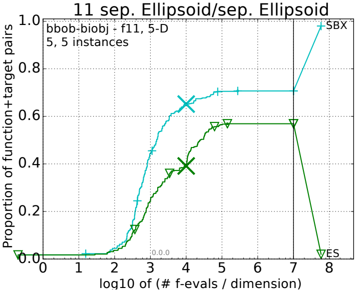
Interaction Effect: Learning Parameter vs. Recombination

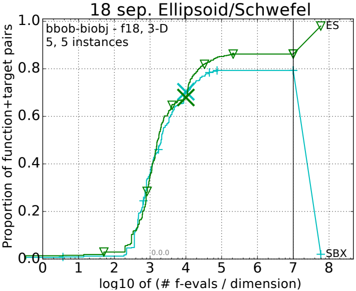
	arithmetic	discrete	intermediate	none
$2^{-1}/\sqrt{n}$	49.96	66.60	79.90	40.82
$2^{0}/\sqrt{n}$	57.01	53.97	83.87	44.49
$2^{1}/\sqrt{n}$	55.65	65.43	82.33	52.42
$2^{2}/\sqrt{n}$	48.70	66.57	80.38	50.98
$2^{3}/\sqrt{n}$	55.25	73.53	86.90	51.54

Comparison with $(50 + 250)$ SBX on bbob-biobj 2016



Comparison with $(50 + 250)$ SBX on bbob-biobj 2016





- SBX is better/competitive on separable problems

Discussion

- Self-adaptive step size adaptation works in both directions (increasing/decreasing)
- Best configuration for budget of $10^4 n$:
  - No recombination
  - $\tau = 2^0/\sqrt{n}$
  - $(50 + 250)$-selection
- Surprisingly similar to single-objective case
- Only arithmetic and no recombination seem to be worth investigating further

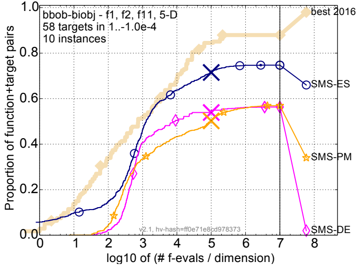
Application to bbob-biobj 2017

Modifications to previous experiments:

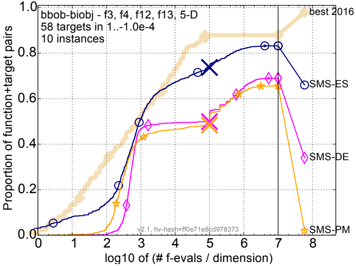
- Initialization in $[0.475, 0.525]^n$ (normalized), corresponding to $[-5, 5]^n$ in original problem space
- Budget of $10^5 n$
- Comparison to $(\mu + 1)$-SMS-EMOA from bbob-biobj 2016
  - DE variation
  - SBX/PM variation

Some Results 5-D

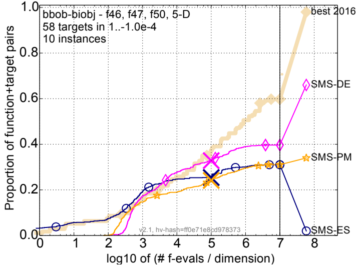
separable-separable



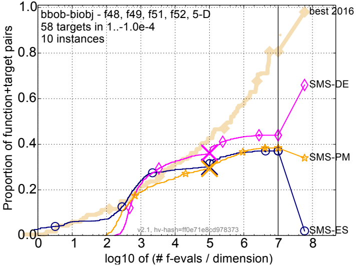
separable-moderate



multimodal-multimodal

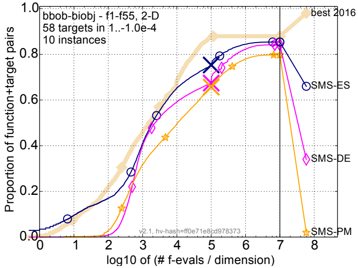


multimodal-weakstructure

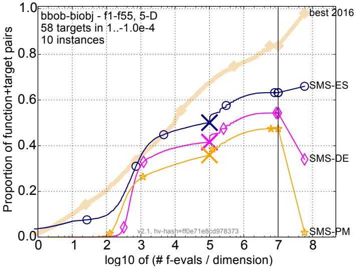


All 55 Functions

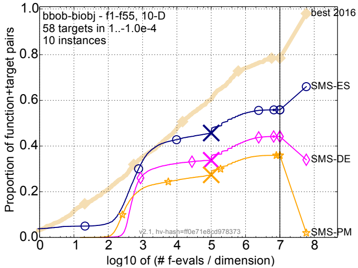
2-D



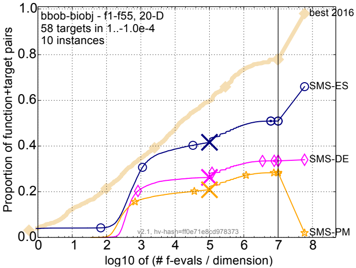
5-D



10-D



20-D



Conclusions and Outlook

Conclusions:

- Self-adaptive variation better than SBX in all tested dimensions, also on multimodal problems
- But not better than DE on multimodal problems
- Not a good anytime algorithm
- Restarts?

Outlook:

- Separate step size for each decision variable?
- Exploit knowledge that dominated solutions need higher mutation strength?
- More sophisticated recombination variants?
- Does variation interact with backward/forward greedy selection?