CEC2017 method comparison based on empirical runtime distributions

Jaroslaw Arabas and Dariusz Jagodzinski jarabas@elka.pw.edu.pl d.jagodzinski@elka.pw.edu.pl

January 7, 2018

Table of contents

1 CEC2017 Competition Evaluation criteria Results record

Performance Measures Empirical runtime distributions Example

3 Results

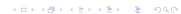
CEC2017 bound constrained case

Test functions are minimization problems defined as following:

Min
$$F_P(x)$$
, $x = [x_1, x_2, ..., x_D]^D$

D: Dimension = 10,30,50,100 **P:** Problem Number = 1,...,30 **Search range** = $[-100,100]^D$ **Budget:** f-evals = 10000*D **Initialization:** Uniform random initialization within the search space. Random seed is based on time.

Global Optimum: $F_P(x^*) = 100 * P$



CEC2017 results record

Function error value is defined as:

$$ERR_P = F_P(x^{best}) - F_P(x^*)$$

The accuracy of finding the optimum is 10^{-8} :

$$ERR_P \le 10^{-8} \implies ERR_P = 0$$

Tuning search method parameters for each problem or dimension is **not allowed.**

Runs/problem: 51

Recorded ERR: 14 error values are recorded for each problem for each run, after $(0.01, 0.02, 0.03, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0)*<math>Budget_D$.

ECDF empirical runtime distributions

Runlength-based target quality indicator corresponds to the function P value to be reached, and is defined as:

$$I_P^t = F_P(x^*) + \Delta I_P^t$$
 $F_P(x^{best}) \leq \Delta I_P^t \leq F_P(x^{worst})$ $rac{\Delta I_P^{t+1}}{\Delta I_P^t} = 10^{0.2} \; , \; t \in T_P$

The ratio between two neighboring ΔI_P target precision values is $10^{0.2}$. The largest ΔI_P value is the error value chosen such that the worst algorithm reaches the first budget step. Smallest ΔI_P value is related to the best algorithm error value on the last budget step.

Budget steps: 14 moments on which error values are reported

ECDF performance assessment: • Link





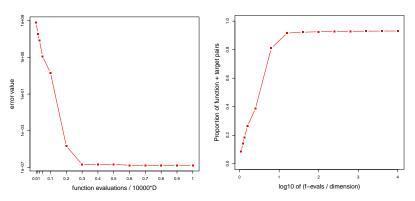
ECDF empirical runtime distributions

ECDF plot displays the proportion of problems solved within a specified budget, where the budget is given on the x-axis. The y-axis shows the fraction of target quality indicators which have been reached. The fraction is defined as:

$$P_P(b) = rac{\sum_{t=1}^{|T_P|} [ERR_P(b) \le I_P^t]}{|T_P|} \;,\; b \in \{1,..,14\}$$

The $P_P(b)$ ratio can be considered as **probability of achieving** success for a given budget b.

ECDF plot example



Convergence plot for a single run, $\;\;$ ECDF plot aggregated for 51 runs, with 14 budget steps on which error for a single problem: P=1, N=10. values have been reported.

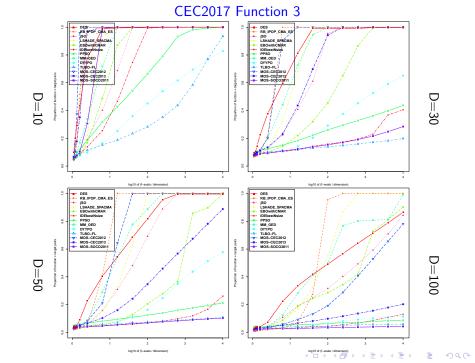
We can see in right plot, for example, that about 90 percent of the problems were solved within $10^2 * D = 10^3$ function evaluations.

CEC2017 Function 1 25/10----RB IPOP CMA ES - RB IPOP CMA ES LSHADE SPACMA LSHADE SPACMA FROwith CMAR FROwith CMAR IDEbestNsize IDEbestNsize PPSO PPSO MM OFD MM_OED DYYPO DYYPO TLBQ-FL TLBO-FL MOS-CEC2012 MOS-CEC2012 MOS-CEC2013 MOS-SOCO2011 * MOS-CEC2013 MQS-SOCO2011 log10 of (f-evals / dimension) log10 of (f-evals / dimension) RB_IPOP_CMA_ES RB_IPOP_CMA_ES LSHADE_SPACMA LSHADE_SPACMA EBOwithCMAR EBOwithCMAR IDEbestNsize IDEbestNsize PPSO/ PPSO MM DED MM. OED DYYPO DYYPO TLBO-FL TLBO-FL MOS-CEC2012 MOS-CEC2012 MOS-CEC2013 MOS-CEC2013 MOS-SOC02011 MOS-SOC02011

log10 of (F-evals / dimension)

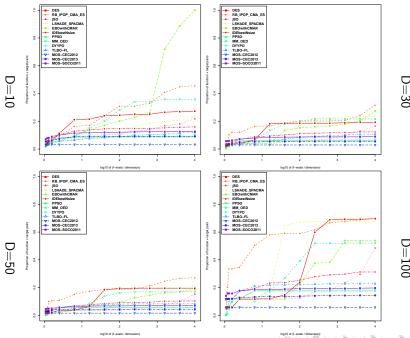
CEC2017 Function 2 RB_IPOP_CMA_ES RB IPOP CMA ES-LSHADE SPACMA LSHADE SPACMA FROwith CMAR EBOWRHCMAR IDEbestNsize IDEbestNstze PPSO PP90 MM_OED MM_OED DYYPO TLBO-FL N-BO-FL/ MOS-CEC2012 MOS-CEC2013 MOS-SOCO2011 MOS-GEC2012 MOS-CEC2013 MOS-SOCO2011 log10 of (f-evals / dimension) log10 of (f-evals / dimension) RB IPOP-CMA ES RB_IPOP_CMA_ES ISO LSHADE SPACMA USHADE SPACMA EBOWINCMAR / IDEbestNsize - IDEBestNsize PPSO / PP50 MM_OED MM OED DYYPO 7LBO-FL MOS-CEC2012 TLBO-FL MØS-CEC2012 MOS-CEC2013 MOS-SOCO2011 MOS-CEC2013 MOS-SOCO2011 9.0

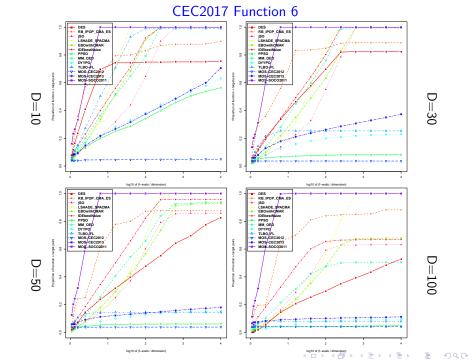
log10 of (f-evals / dimension)



CEC2017 Function 4 RB IPOP CMA ES RB IPOP CMA ES LSHADE SPACMA LSHADE SPACMA FROwith CMAR FROwith CMAR IDEbestNsize IDEbestNsize PPSO PPSO MM_OED MM_OED DYYPO DYYPÓ TLBO-FL TLBO-FL MOS-GEC2012 MOS-CEC2012 MOS-CEC2013 MOS-SOCO2011 MOS-CEC2013 MOS-SOCO2011 log10 of (f-evels / dimension) log10 of (F-evels / dimension) RB_IPOP_CMA_ES RB_IPOP_CMA_ES LSHADE_SPACMA LSHADE_SPACMA EBOwithCMAR EBOwithCMAR IDEbestNsize IDEbestNsize PPSO PPSO MM OED MM OED DYYPO DYYPO TLBO-FL TLBO-FL MOS-CEC2012 MOS-CEC2012 MOS-CEC2013 MOS-CEC2013 MOS-SOC02011 MOS-SOCO2011

log10 of (F-evals / dimension)



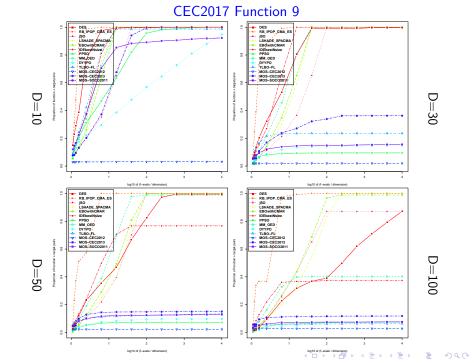


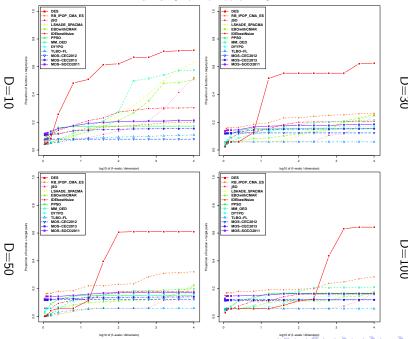
CEC2017 Function 7 RB IPOP CMA ES - RB IPOP CMA ES LSHADE SPACMA LSHADE SPACMA FROwithCMAR FROwith CMAR IDEbestNsize IDEbestNsize PPSO PPSO MM_OED MM_OED DYYPO DYYPO TLBO-FL TLBO-FL MOS-CEC2012 MOS-CEC2012 MOS-CEC2013 MOS-CEC2013 MOS-SOCO2011 MOS-SOC02011 log10 of (f-evals / dimension) log10 of (F-evels / dimension) - DES RB_IPOP_CMA_ES RB_IPOP_CMA_ES LSHADE_SPACMA LSHADE_SPACMA EBOwithCMAR EBOwithCMAR IDEbestNsize IDEbestNsize PPSO PPSO MM OED MM_OED DYYPO DYYPO TLBO-FL TLBO-FL MOS-CEC2012 MOS-CEC2012 MOS-CEC2013 MOS-CEC2013 MOS-SOCO2011 MOS-SOCO2011 7 90 00 9.0

log 10 of (I-evals / dimension)

CEC2017 Function 8 RB IPOP CMA ES RB IPOP CMA ES LSHADE SPACMA LSHADE SPACMA FROwithCMAR FROwith CMAR IDEbestNsize IDEbestNsize PPSO PPSO MM_OED MM_OED DYYPO DYYPO TLBO-FL TLBO-FL MOS-CEC2012 MOS-CEC2012 MOS-CEC2013 * MOS-CEC2013 MOS-SOC02011 MOS-SOCO2011 log10 of (f-evals / dimension) log10 of (F-evels / dimension) RB_IPOP_CMA_ES RB_IPOP_CMA_ES LSHADE_SPACMA LSHADE_SPACMA EBOwithCMAR EBOwithCMAR IDEbestNsize IDEbestNsize PPSO PPSO MM OED MM OED DYYPO DYYPO TLBO-FL TLBO-FL MOS-CEC2012 MOS-CEC2012 MOS-CEC2013 MOS-CEC2013 MOS-SOC02011 MOS-SOCO2011

log10 of (f-evals / dimension

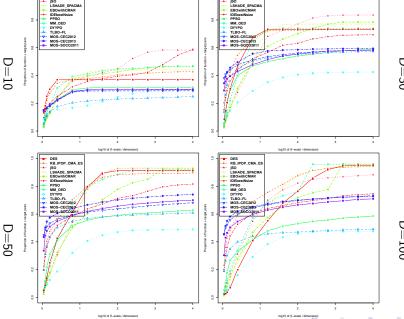




CEC2017 Function 11 RB IPOP CMA ES RB IPOP CMA ES LSHADE SPACMA LSHADE SPACMA FROwithCMAR FROwith CMAR IDEbestNsize PPSO IDEbestNsize PPSO MM_OED MM OFD DYYPO DYYPO TLBO-FL TLBO-FL MOS-CEC2012 MOS-CEC2012 MOS-CEC2013 MOS-CEC2013 MOS-SOC02011 MOS-SOCO2011 log10 of (f-evals / dimension) log10 of (f-evals / dimension) RB_IPOP_CMA_ES RB_IPOP_CMA_ES LSHADE_SPACMA LSHADE_SPACMA EBOwithCMAR EBOwithCMAR IDEbestNsize IDEbestNsize PPSO PPSO MM OED MM OED DYYPO DYYPO TLBO-FL TLBO-FL MOS-CE02012 MOS-CEC2012 MOS-CEC2013 MOS-CEC2013 MOS-80C02011 MOS-SOCO201



CEC2017 Function 12 RB IPOP CMA ES LSHADE SPACMA FROwithCMAR IDEbestNsize PPSO MM OFD DYYPO TLBO-FL MOS-CEC2012 MOS-CEC2013 MOS-SQC02011



RB IPOP CMA ES

CEC2017 Function 13 RB IPOP CMA ES RB IPOP CMA ES LSHADE SPACMA LSHADE SPACMA FROwithCMAR - FROwithCMAR IDEbestNsize PPS0 IDEbestNsize PPSO MM_OED MM OFD DYYPO DYYPO TLBO-FL TLBO-FL MOS-CEC2012 MOS-CEC2012 MOS-CEC2013 MOS-CEC2013 MOS-SOC02011 log10 of (f-evals / dimension) log10 of (F-evels / dimension) - DES RB_IPOP_CMA_ES ··· RB_IPOP_CMA_ES LSHADE_SPACMA LSHADE_SPACMA EBOwithCMAR EBOwithCMAR IDEbestNsize - IDEbestNsize PPSO=:=::6:= PPSO MM OED MM OED DYYPO TLBO-FE MOS-CEC2012 TLBO-FL MOS-CEC2012 MOS-CEC2013 MOS-CEC2013 MOS-50C02811 MOS-SOCO2011 970

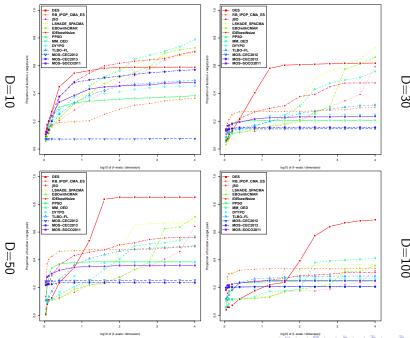
log10 of (F-evals / dimension)

CEC2017 Function 14 RB IPOP CMA ES RB IPOP CMA ES LSHADE SPACMA LSHADE SPACMA FROwithCMAR FROwith CMAR IDEbestNsize IDEbestNsize PPSO PPSO MM OFD MM OFD DYYPO DYYPO TLBO-FL TLBO-FL MOS-CEC2012 MOS-CEC2012 MOS-CEC2013 NOS-CEC2013 MOS-SOC02011 MOS-SOCO2011 log10 of (f-evels / dimension) log10 of (f-evals / dimension) RB_IPOP_CMA_ES ··· RB_IPOP_CMA_ES - ISO LSHADE_SPACMA LSHADE_SPACMA EBOwithCMAR/ EBOwithCMAR IDEbestNsize IDEbestNsize PPSO PPSO MM OED MM OED DYYPO TLBO-FL DYYPO TLBO-FL MOS-CEC2012 MOS-CEC2012 MOS-CEC2013 MOS-CEC2013 MGS/SGC02011 MOS-SOC02011 80 9.0

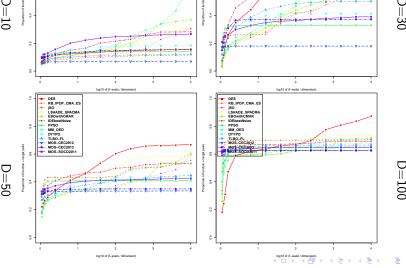
log10 of (F-evals / dimension)

CEC2017 Function 15 RB IPOP CMA ES RB IPOP CMA ES LSHADE SPACMA LSHADE SPACMA FROwithCMAR - FROwithCMAR IDEbestNsize PPS0 IDEbestNsize PPSO MM_OED MM OFD DYYPO DYYPO TLBO-FL TLBO-FL MOS-CEC2012 MOS-CEC2012 MOS-CEC2013 MOS-05C2013 MOS-80C02011 MOS-SOC02011 log10 of (f-evals / dimension) log10 of (f-evals / dimension) RB_IPOP_CMA_ES ··· RB_IPOP_CMA_ES LSHADE_SPACMA LSHADE_SPACMA **EBOwithCMAR** EBOwithCMAR IDEbestNsize PPSO MM_OED MM OED DYYPO DYYPO / MOS-DEC2012 MOS-DEC2012 MOS-DEC2013 MOS-SOC02011 MOS-CEC2012 MOS-CEC2013 MOS-SOC02011 1 9 ·

log10 of (F-evals / dimension)



CEC2017 Function 17 RB IPOP CMA ES LSHADE SPACMA FROwithCMAR IDEbestNsize PPSO MM OFD DYYPO TLBO-FL MOS-CEC2012 * MOS-CEC2013 MOS-SOCO2011 log10 of (F-evels / dimension) RB_IPOP_CMA_ES LSHADE_SPACMA EBOwithCMAR IDEbestNsize PPSO MM OED DYYPO TLBO-FL MOS-CEC2012



RB IPOP CMA ES

LSHADE SPACMA

FROwithCMAR

MOS-CEC2012

MOS-CEC2013

MOS-SOC02011

IDEbestNsize

PPSO

MM_OED

DYYPO TLBO-FL

CEC2017 Function 18 RB IPOP CMA ES RB IPOP CMA ES LSHADE SPACMA LSHADE SPACMA FROwithCMAR - FROwithCMAR IDEbestNsize PPS0 IDEbestNsize PPSO MM_OED MM OFD DYYPO DYYPO TLBO-FL TLBO-FL MOS-CEC2012 MOS-@EC2012 MOS-CEC2013 MO3-CEC2013 MOS-SOCO2011 MOS-SOC02011 log10 of (f-evals / dimension) log10 of (f-evals / dimension) RB_IPOP_CMA_ES ··· RB_IPOP_CMA_ES LSHADE_SPACMA LSHADE_SPACMA EBOwithCMAR EBOwithCMAR IDEbestNsize IDEbestNsize PPSO MM OED MM OED DYYPO DYYPO TLBO-FL TLBO-FL MOS-CEC2012 MOS-CEC2012 MOS-GEC2013 MOS-CEC2013 MOS-SOCO2011 MOS-SOCO201

log10 of (F-evals / dimension)

CEC2017 Function 19 RB IPOP CMA ES RB IPOP CMA ES LSHADE SPACMA LSHADE SPACMA FROwithCMAR FROwith CMAR IDEbestNsize IDEbestNsize PPSO PPSO MM OFD MM_OEO DYYPO DYYPO TLBO-FL TLBO-FL MOS-CEC2012 MOS-CEC2012 MOS-CEC2013 MOS-CEC2013 MOS-30C02011 MOS-SOC02011 log10 of (f-evals / dimension) log 10 of (F-evals / dimension) ··· RB_IPOP_CMA_ES RB_IPOP_CMA_ES LSHADE_SPACMA LSHADE_SPACMA EBOwithCMAR **EBOwithCMAR** IDEbestNsize PPSO PRSD MM_OED MM OED DYYPO DYYPO TLBO-FL TLBO-FL MOS-CEC2012 MOS-CEC2013 MOS-SOCO201

log10 of (F-evals / dimension)

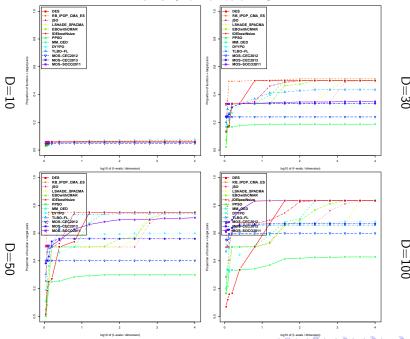
CEC2017 Function 20 RB IPOP CMA ES RB IPOP CMA ES LSHADE SPACMA LSHADE SPACMA FROwithCMAR FROwithCMAR IDEbestNsize IDEbestNsize PPSO PPSO MM_OED MM OFD DYYPO DYYPO TLBO-FL TLBO-FL MOS-CEC2012 MOS-CEC2012 MOS-CEC2013 * MOS-CEC2013 MOS-SOC02011 MOS-SOCO2011 log10 of (f-evals / dimension) log10 of (F-evels / dimension) RB_IPOP_CMA_ES RB_IPOP_CMA_ES LSHADE_SPACMA LSHADE_SPACMA EBOwithCMAR EBOwithCMAR IDEbestNsize IDEbestNsize PPSO PPSO MM OED MM OED DYYPO DYYPO TLBO-FL TLBO-FL MOS-CEC2012 MOS-CEC2012 MOS-CEC2013 MOS-CEC2013 MOS-SOC02011 MOS-SOC02011

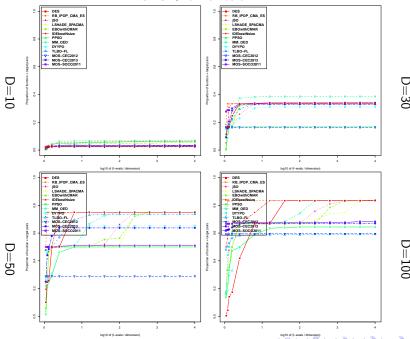
log10 of (F-evals / dimension)

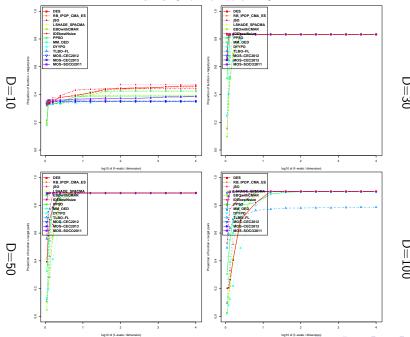
CEC2017 Function 21 RB IPOP CMA ES RB IPOP CMA ES LSHADE SPACMA LSHADE SPACMA FROwithCMAR FROwithCMAR IDEbestNsize IDEbestNsize PPSO PPSO MM OFD MM OFD DYYPO DYYPO TI BO.FI TLBO-FL MOS-CEC2012 MOS-CEC2012 MOS-CEC2013 MOS-CEC2013 MOS-SQC02011 D = 10log10 of (f-evals / dimension) log10 of (F-evels / dimension) RB_IPOP_CMA_ES RB_IPOP_CMA_ES LSHADE_SPACMA LSHADE_SPACMA EBOwithCMAR EBOwithCMAR IDEbestNsize IDEbestNsize PPSO PPSO MM OED MM OED DYYPO DYYPO TLBO-FL TLBO-FL MOS-CEC2012 MOS-CEC2012 MOS-CEC2013 MOS-CEC2013 MOS-SOC02011 MOS-SOC02011 2

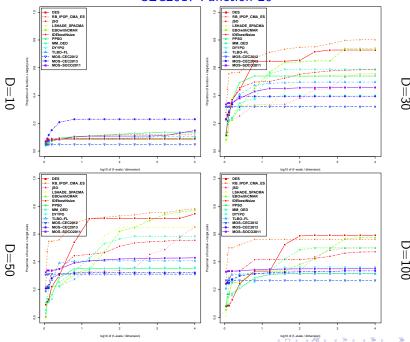
log10 of (F-evals / dimension)

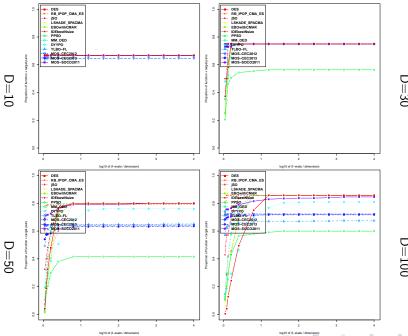
CEC2017 Function 22 RB IPOP CMA ES - RB IPOP CMA ES LSHADE SPACMA FROwithCMAR EBOwith CMAR IDEbestNsize ID EbestNsize PPSO PPSO MM OFD MM_OED DYYPO DYYPO TLBO-FL TLBO-FL MOS-CEC2012 MOS-CEC2012 MOS-CEC2013 MOS-CEC2013 MOS-SOCO2011 MOS-SOC02011 log10 of (f-evals / dimension) log10 of (f-evals / dimension) ··· RB_IPOP_CMA_ES RB_IPOP_CMA_ES LSHADE_SPACMA EBOWITHCMAR IDEbestNsize LSHADE_SPACMA EBOwithCMAR IDEbestNsize PPSO MM_OED **PPSO** MM OED DYYPO DYYPO TUBO-FL TLBO-FL MOS-CEC2012 MOS-CEC2012 MOS-CEC2013 MOS-CEC2013 MOS-SOC02011 MOS-SOC02011 80 9.0

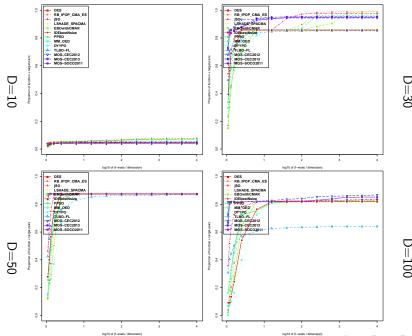


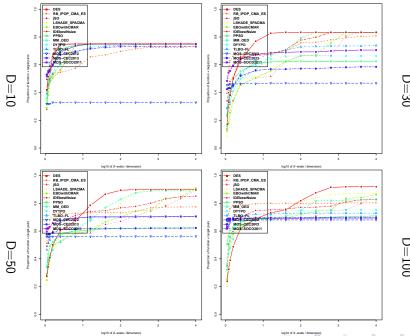












CEC2017 Function 30 - RB IPOP CMA ES RB IPOP CMA ES LSHADE SPACMA LSHADE SPACMA EBOwithCMAR FROwith CMAR IDEbestNsize IDEbestNsize PPSD PPSO MM OED MM_OED DYYPO DYWPO-TLBO-FL TLBO-EL MOS-CEC2012 MOS-CEC2012 MOS-CEC2013 MOS-SOCO2011 MOS CEC2013 MOS SOCO2011 log10 of (f-evals / dimension) log10 of (F-evels / dimension) ··· RB_IPOP_CMA_ES RB_IPOP_CMA_ES LSHADE_SPACMA LSHADE_SPACMA EBOwithCMAR EBOwithCMAR IDEbestNsize IDEbestNsize PPSO MM OED MM OED DYYPO DYYPO TLBO-FL TLBO-FL MOS-CEC2012 MQS-CEC2012 MOS-CEC2013 MOS-SOCO201 MOS-CEC2013 MOS-SOC02011 9.0

log10 of (F-evals / dimension)