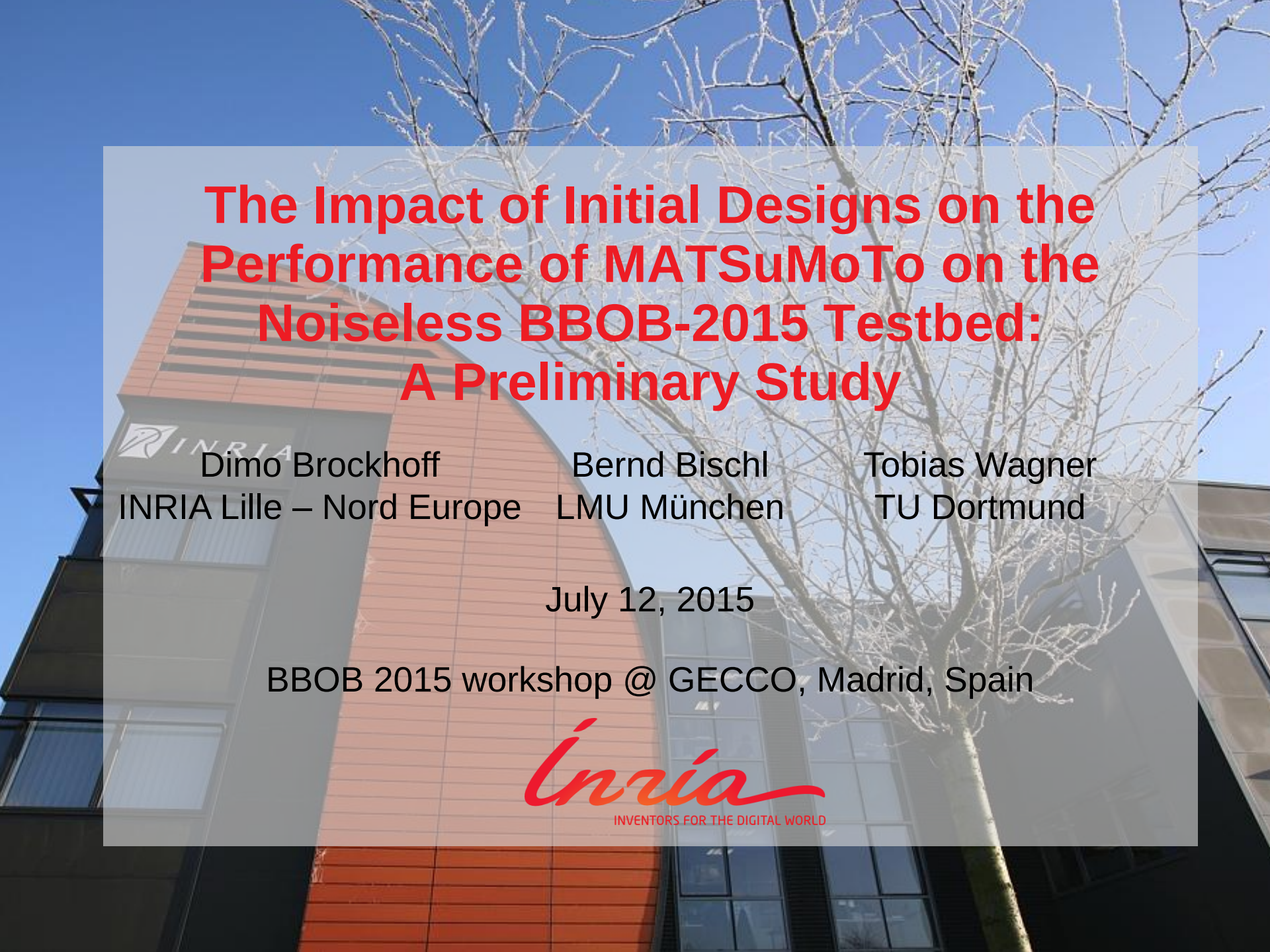


The Impact of Initial Designs on the Performance of MATSuMoTo on the Noiseless BBOB-2015 Testbed: A Preliminary Study

Dimo Brockhoff
INRIA Lille – Nord Europe

Bernd Bischl
LMU München

Tobias Wagner
TU Dortmund

July 12, 2015

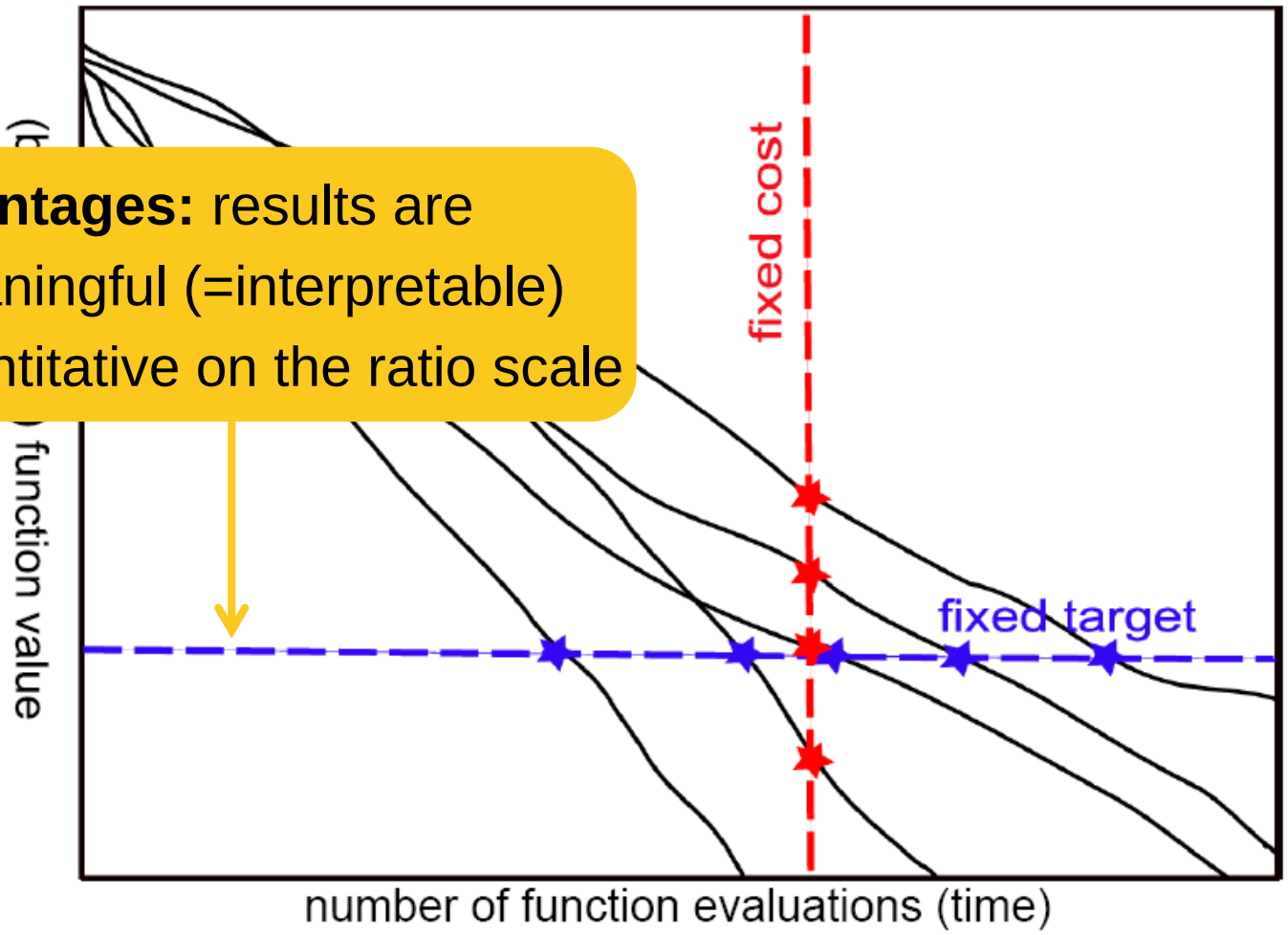
BBOB 2015 workshop @ GECCO, Madrid, Spain

INVENTORS FOR THE DIGITAL WORLD

Reminder: Target-Based Runlengths @ BBOB

Measuring Performance from Convergence Graphs

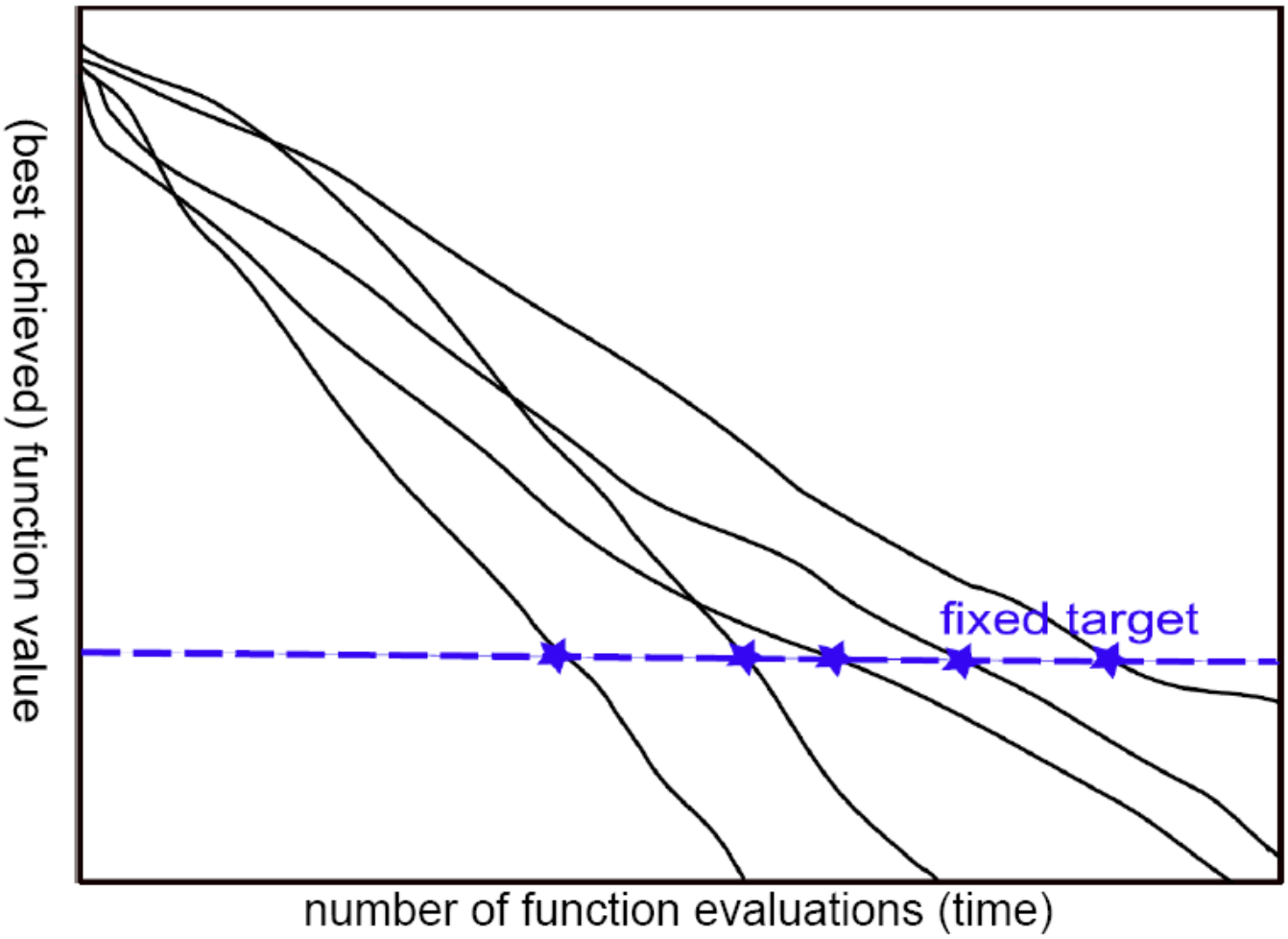
fixed-cost versus fixed-target



Advantages: results are
- meaningful (=interpretable)
- quantitative on the ratio scale

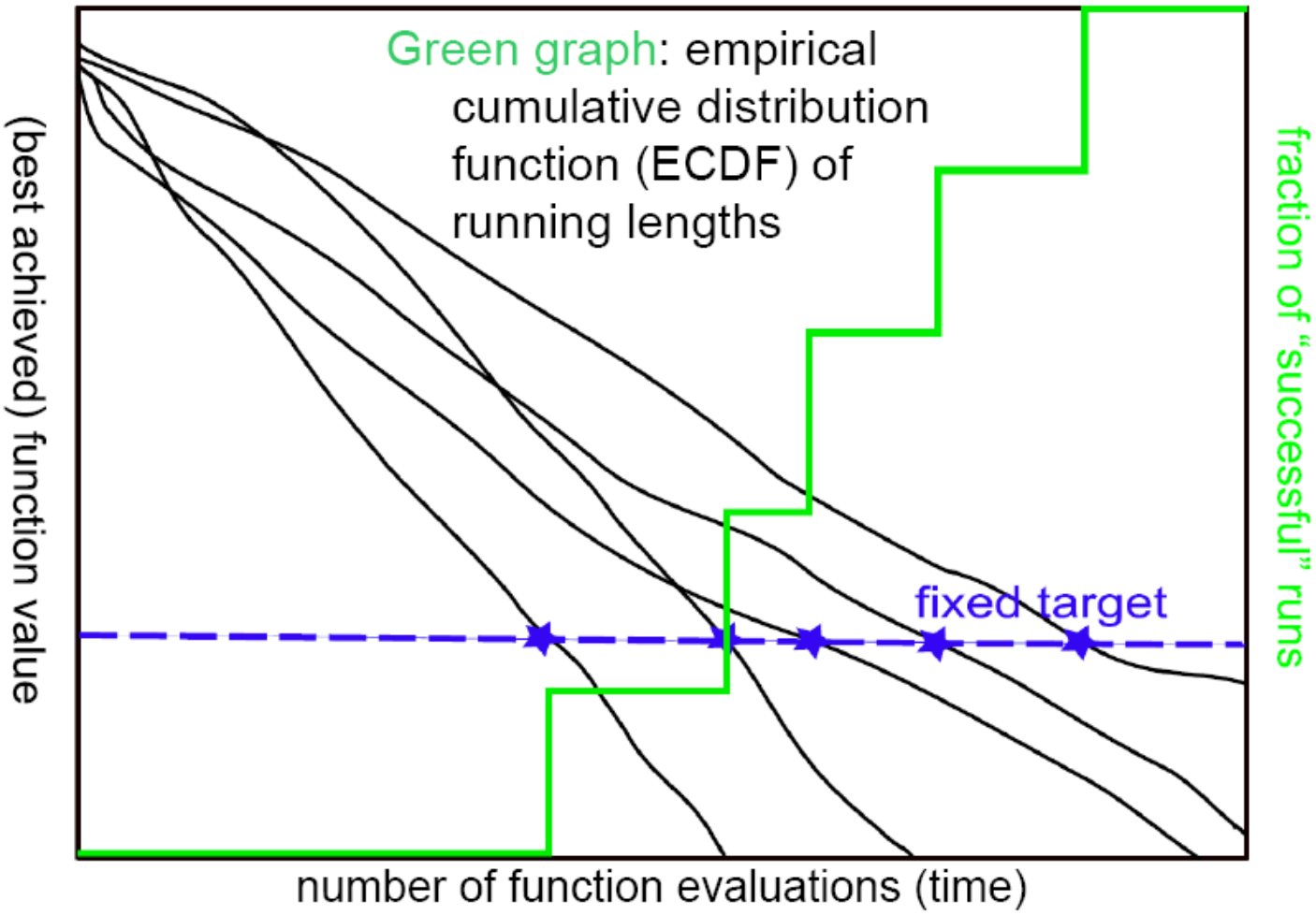
Reminder: Target-Based Runlengths @ BBOB

Empirical Cumulative Distribution with a given target value



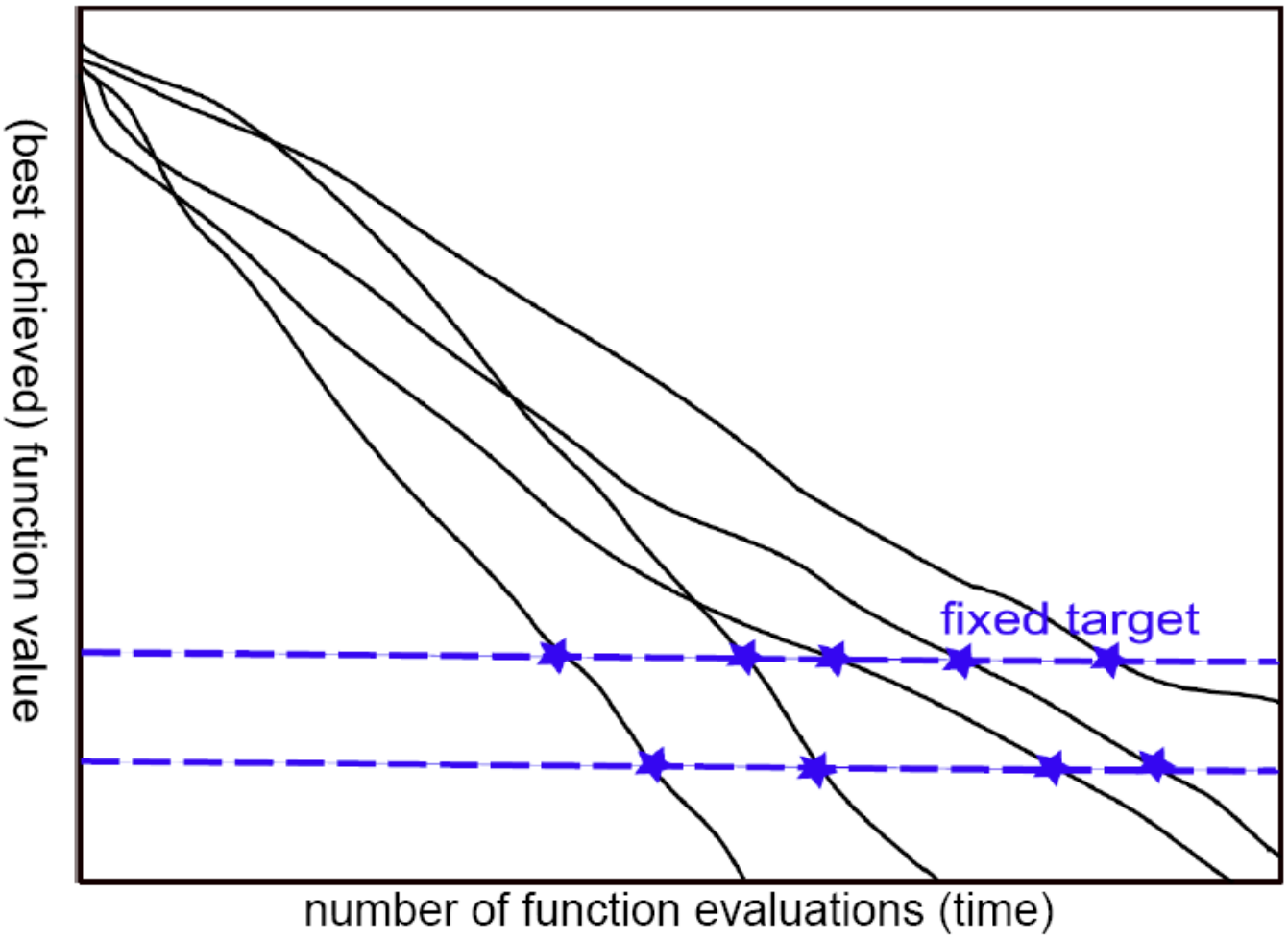
Reminder: Target-Based Runlengths @ BBOB

Empirical Cumulative Distribution with a given target value



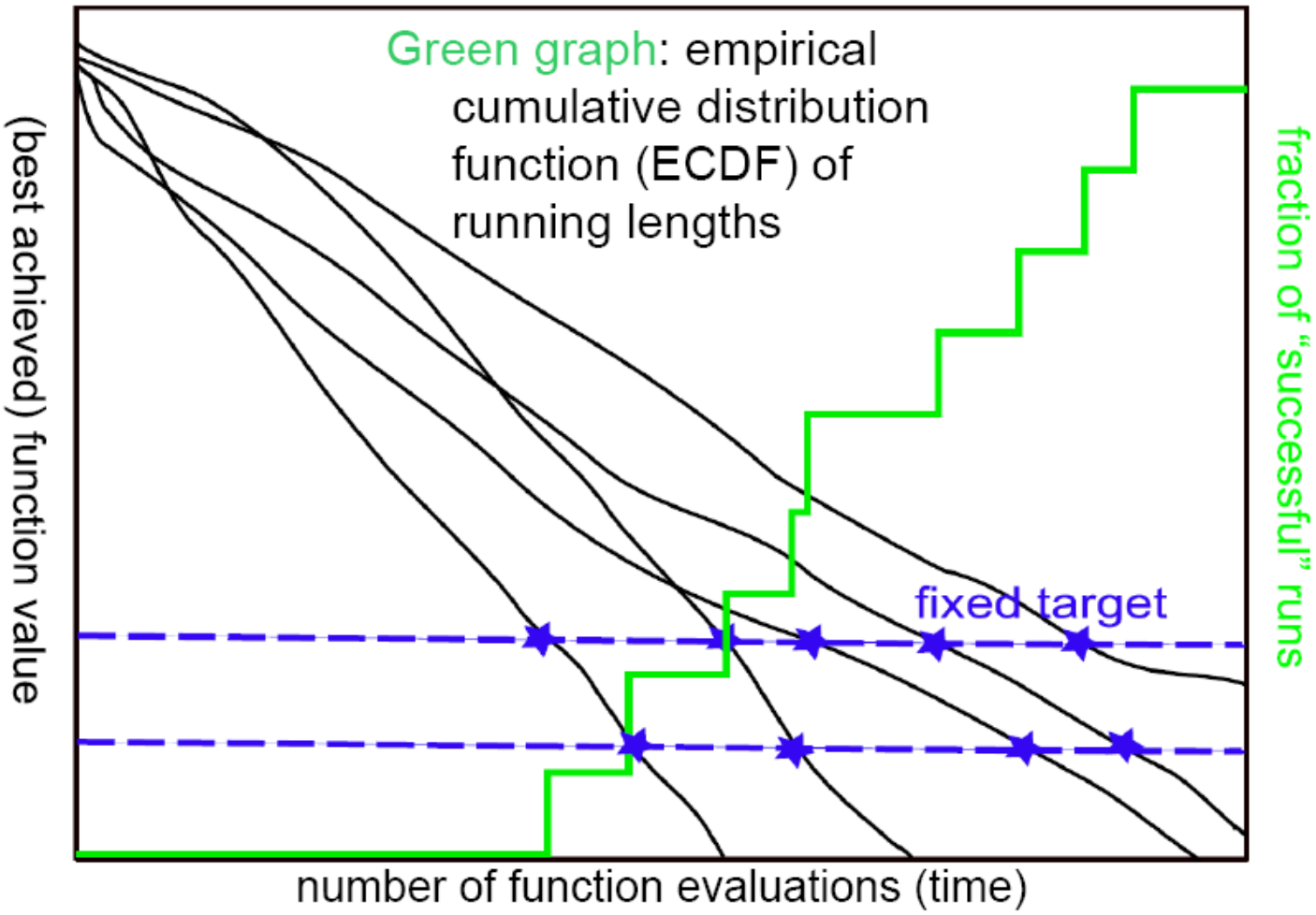
Reminder: Target-Based Runlengths @ BBOB

Empirical Cumulative Distribution with two given target values

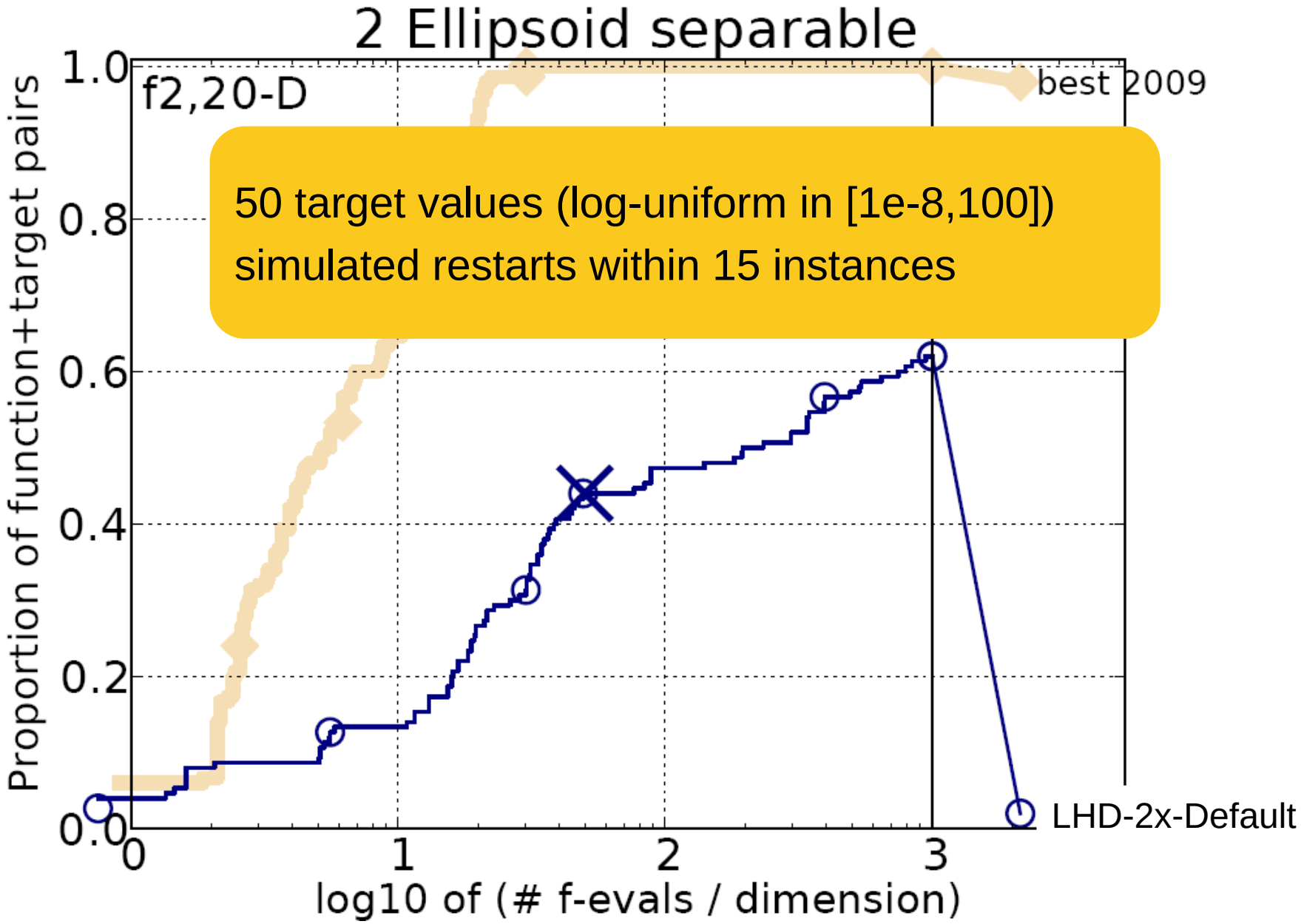


Reminder: Target-Based Runlengths @ BBOB

Empirical Cumulative Distribution with two given target values



Resulting Data Profile For a Real Algorithm



Expensive Optimization

#funevals in practical applications often restricted

~10...1000D

e.g. due to

- expensive simulations
- physical evaluations

Hence, benchmarking wrt. difficult target precisions meaningless

Expensive BBOB setting:

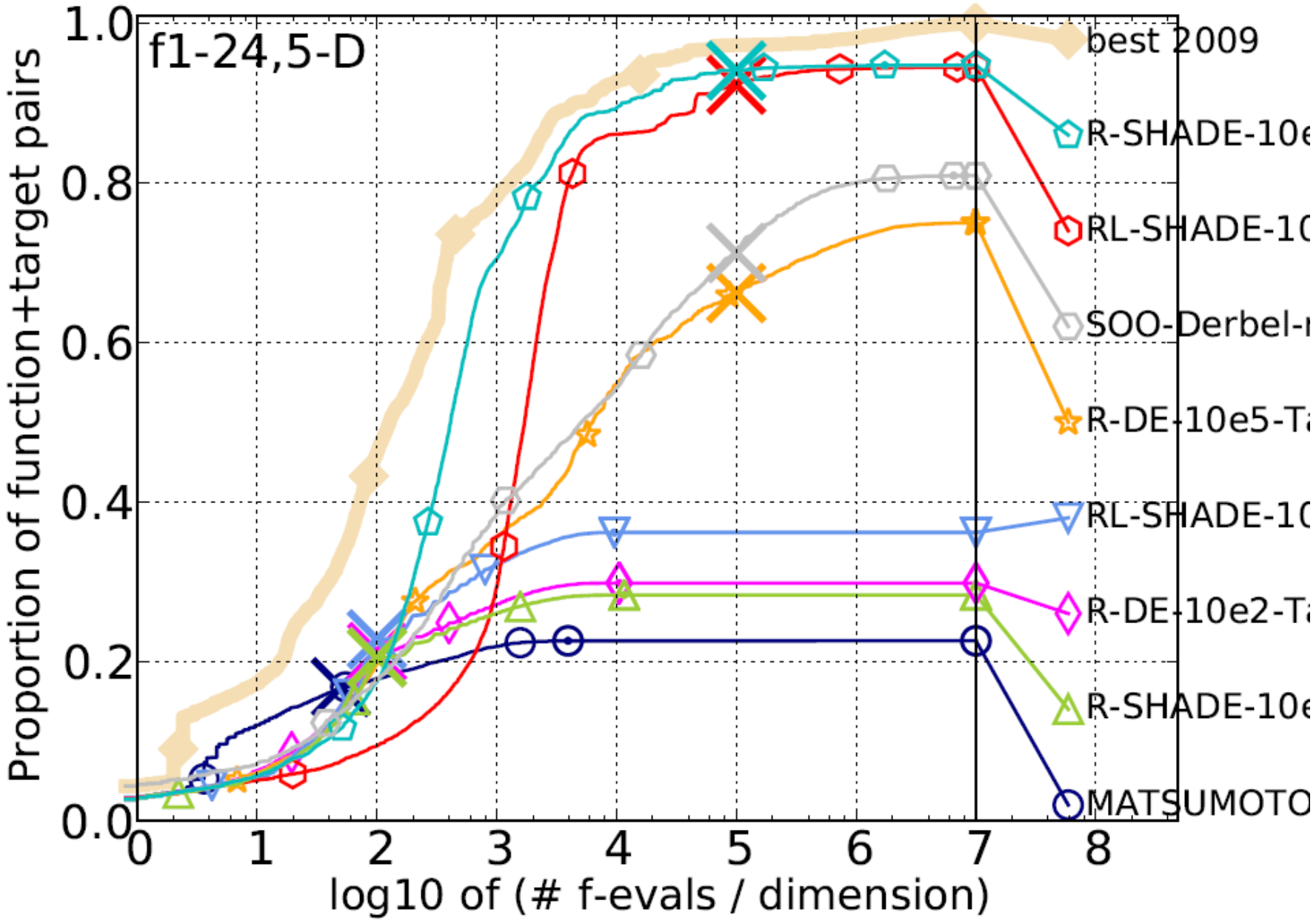
targets relative to the best algorithm submitted to BBOB-2009

Expensive BBOB Setting

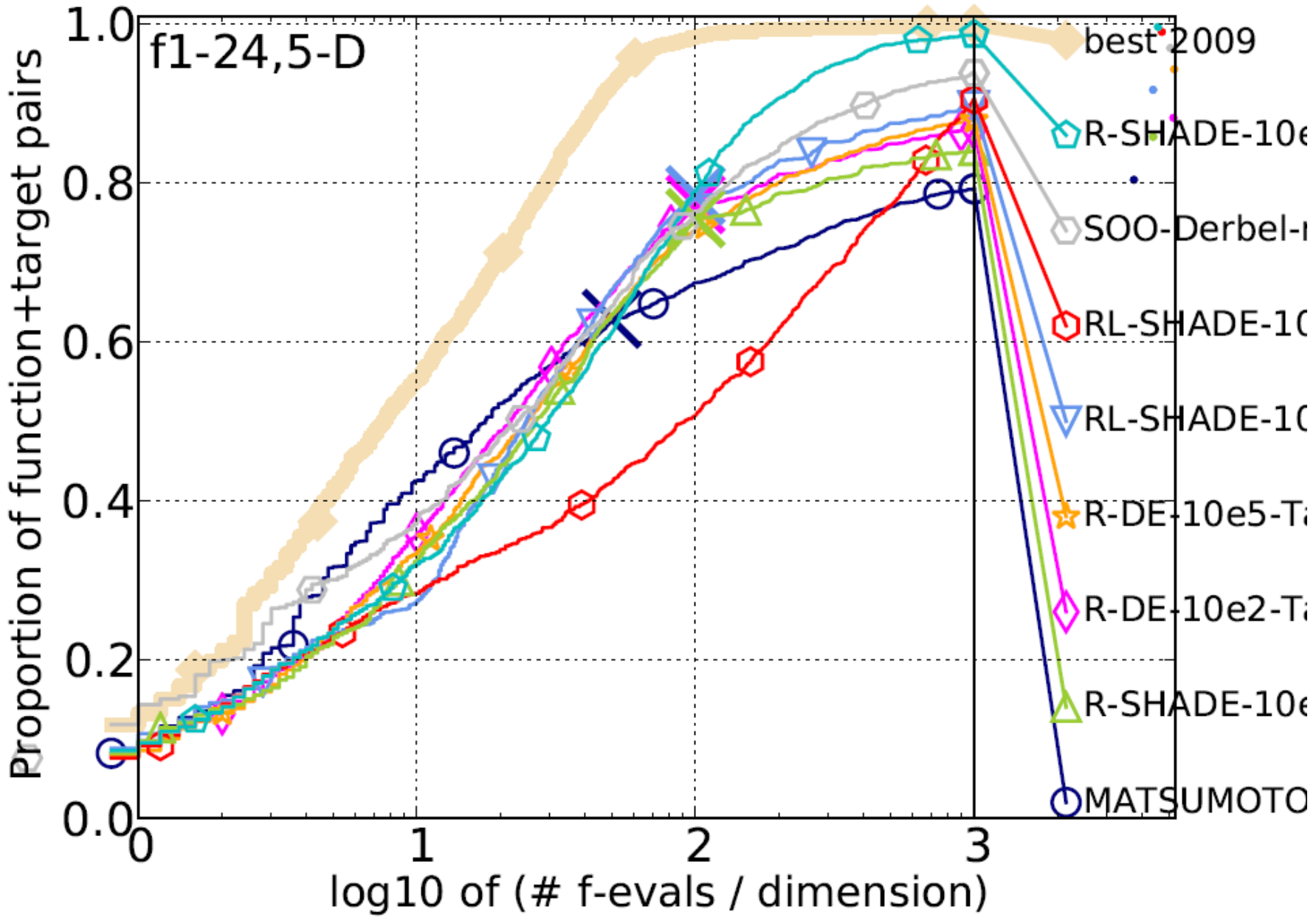
- ∀ test functions:
  - ∀ instances:
    - ∀ budgets in 0.5...50D:

use target just not reached by best algo from BBOB-2009 (out of 31 algorithms) as reference target

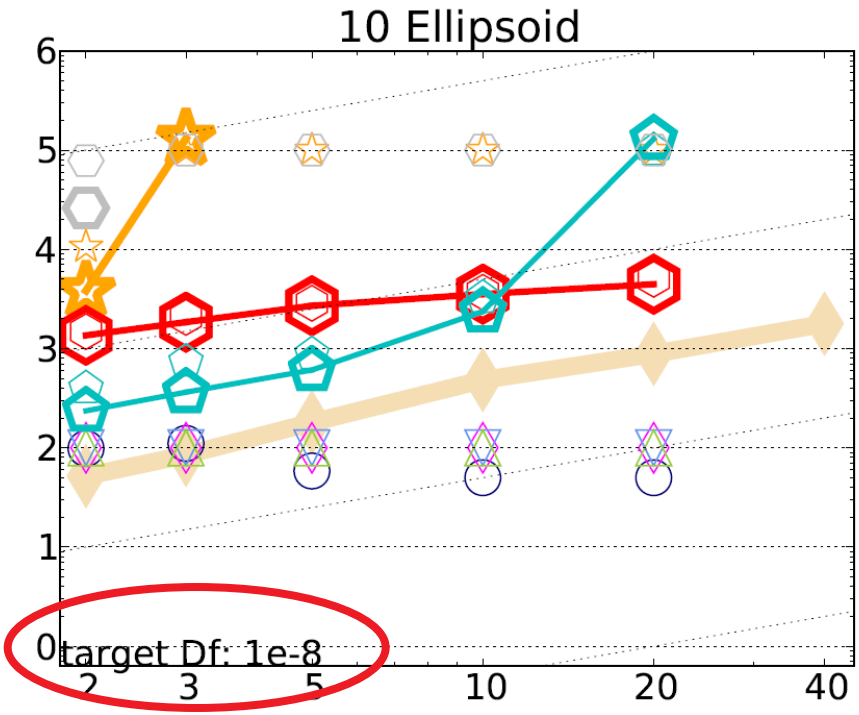
BBOB: Standard Setting

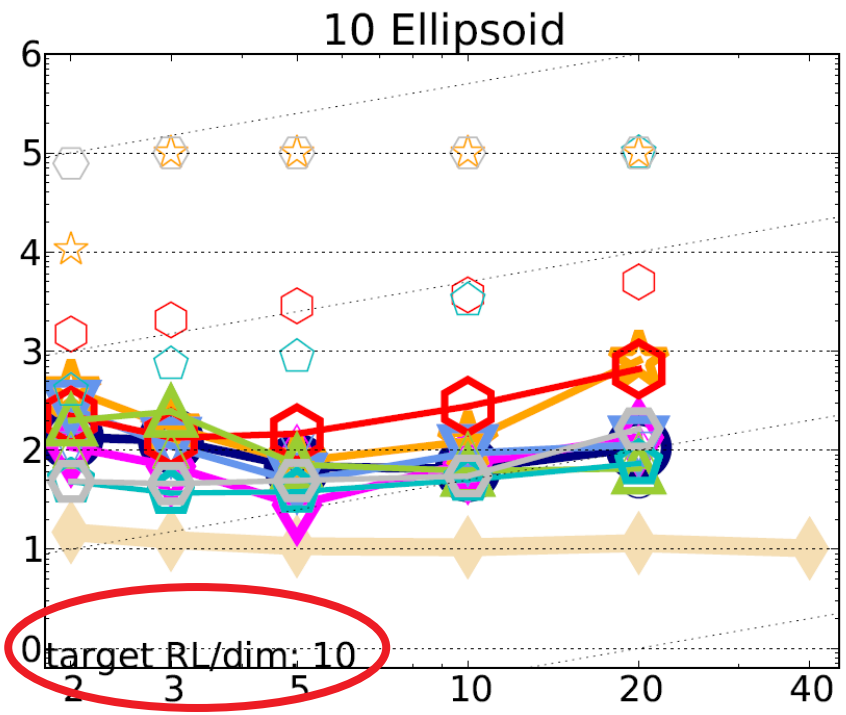


BBOB: Expensive Setting



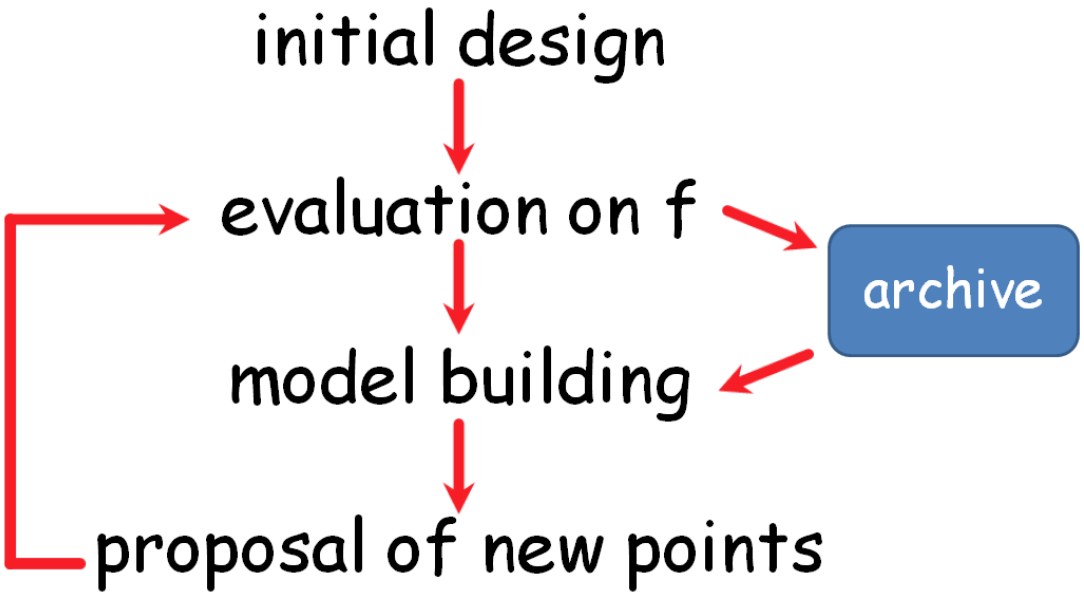
Expensive Vs. Non-Expensive Setting





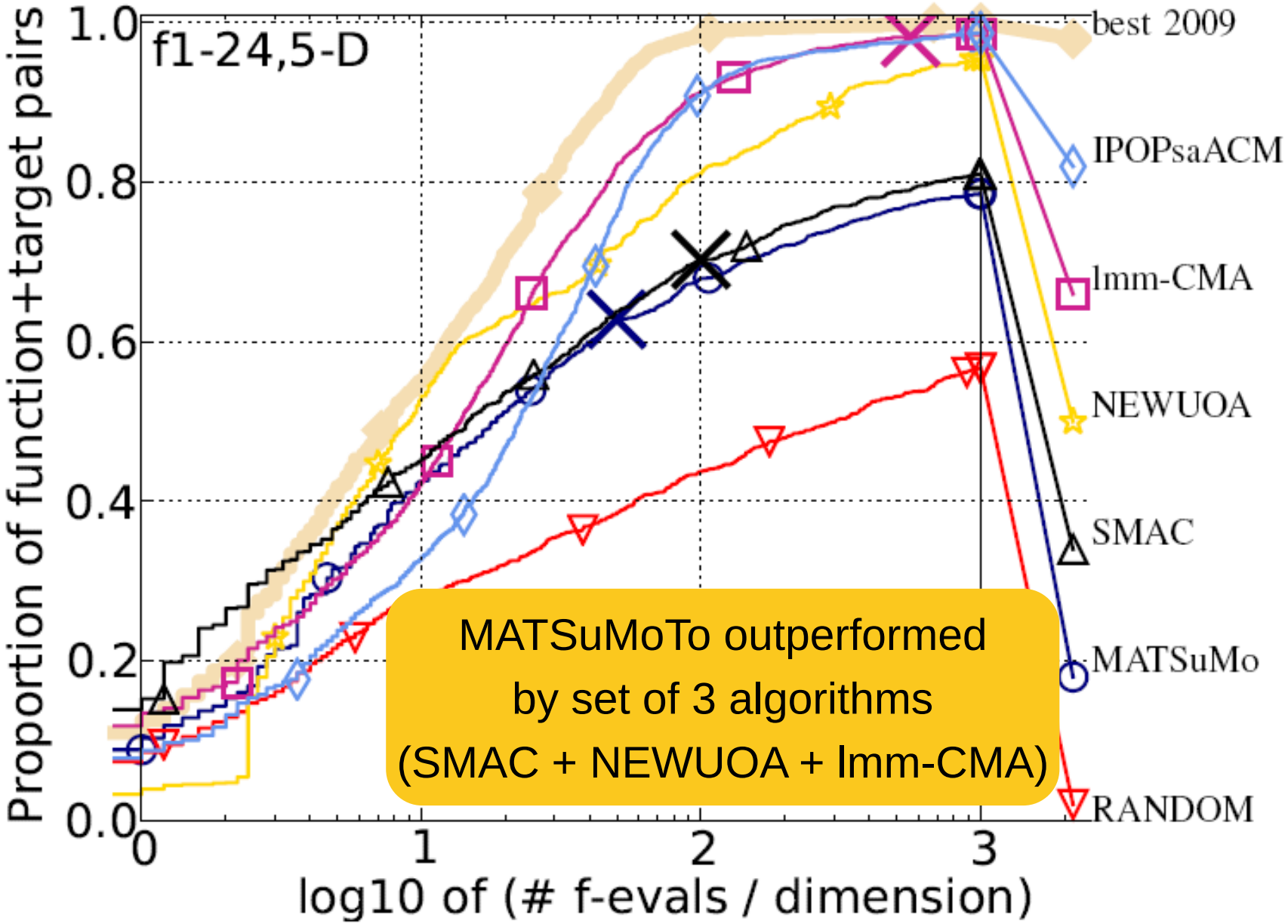
MATSuMoTo:

MATLAB Surrogate Model Toolbox by Juliane Müller



several options for each part

MATSuMoTo Default (CEC 2015 Data)



Goal of BBOB 2015 paper:

- start to study MATSuMoTo's many parameters
- impact of size and type of initial design
- paper title contains "preliminary" because not all implemented initial designs have been tested yet

Scientific Questions

Q1: What is the effect of having larger initial designs, such as two times or ten times the default value of 2(DIM+1) function evaluations?

Q2: What is the effect of replacing the LHS with simple (uniform pseudo-)random sampling?

Experimental Setting

MATSuMoTo default

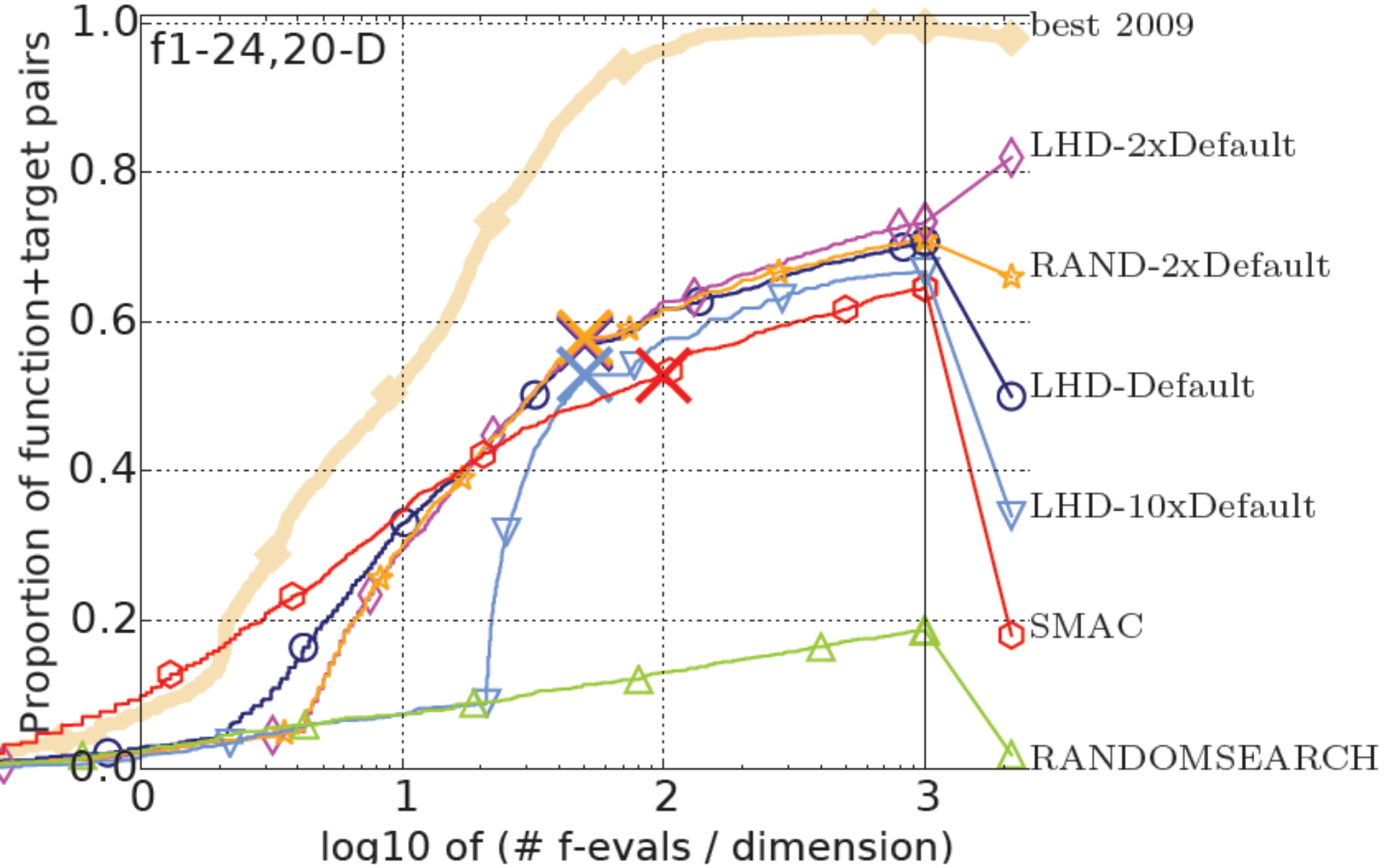
- Latin Hypercube Sampling for 2(DIM+1) funevals
- cubic radial basis functions as surrogate model
- infill criterion: (small) random perturbation around the model's minimum (exploitation) or, with a certain probability, uniformly at random in the whole variable domain (exploration).

Overall 4 Algorithm Variants

algorithm name	initial design	length of initial design phase
LHD-Default	LHS	default
LHD-2xDefault	LHS	2x default
RAND-2xDefault	random	2x default
LHD-10xDefault	LHS	10x default

50xDIM funevals in total, initial sampling in [-5,5]

All Functions



Scientific Question 1

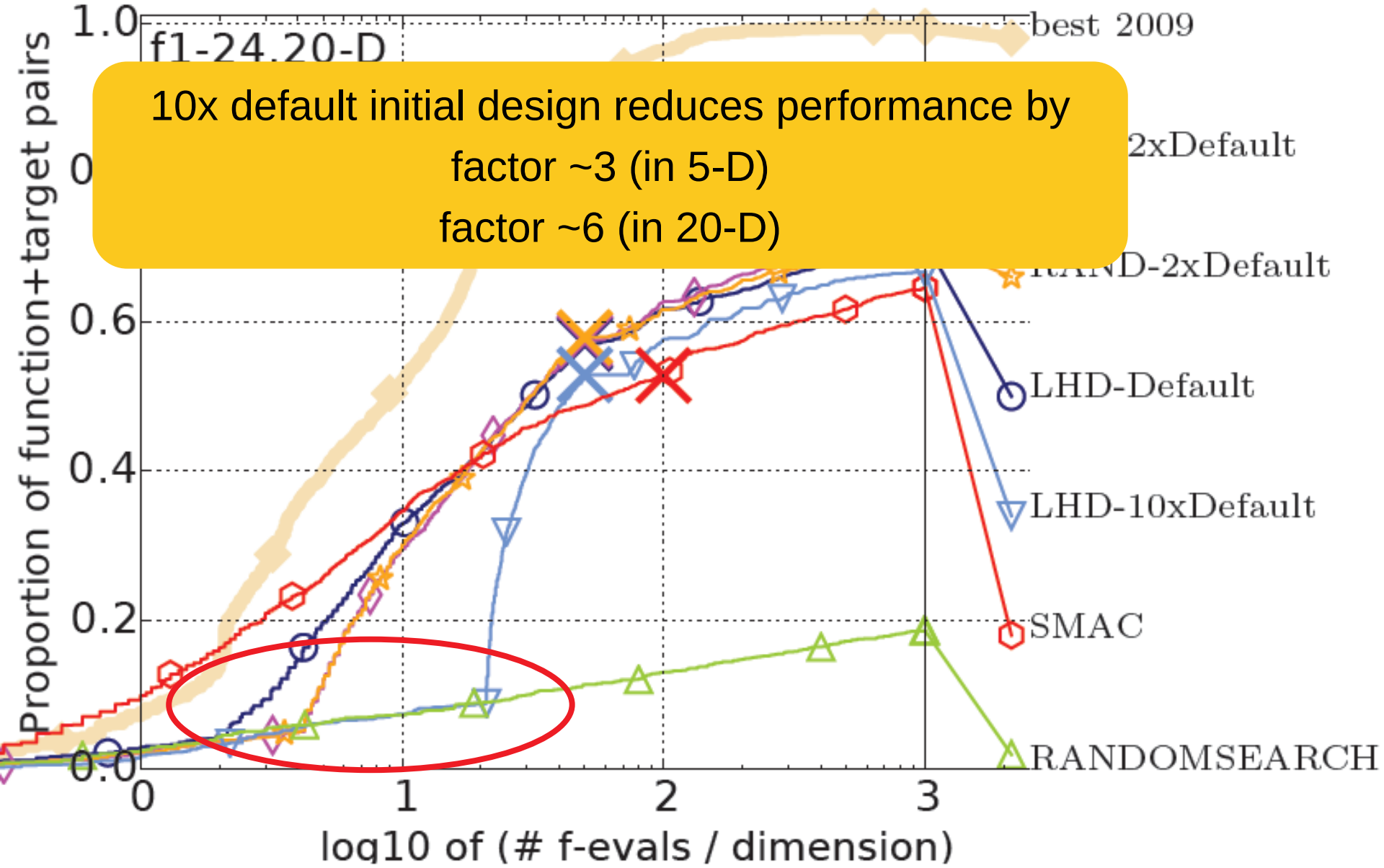
Q1: What is the effect of having larger initial designs, such as two times or ten times the default value of 2(DIM+1) function evaluations?

2 main observations:

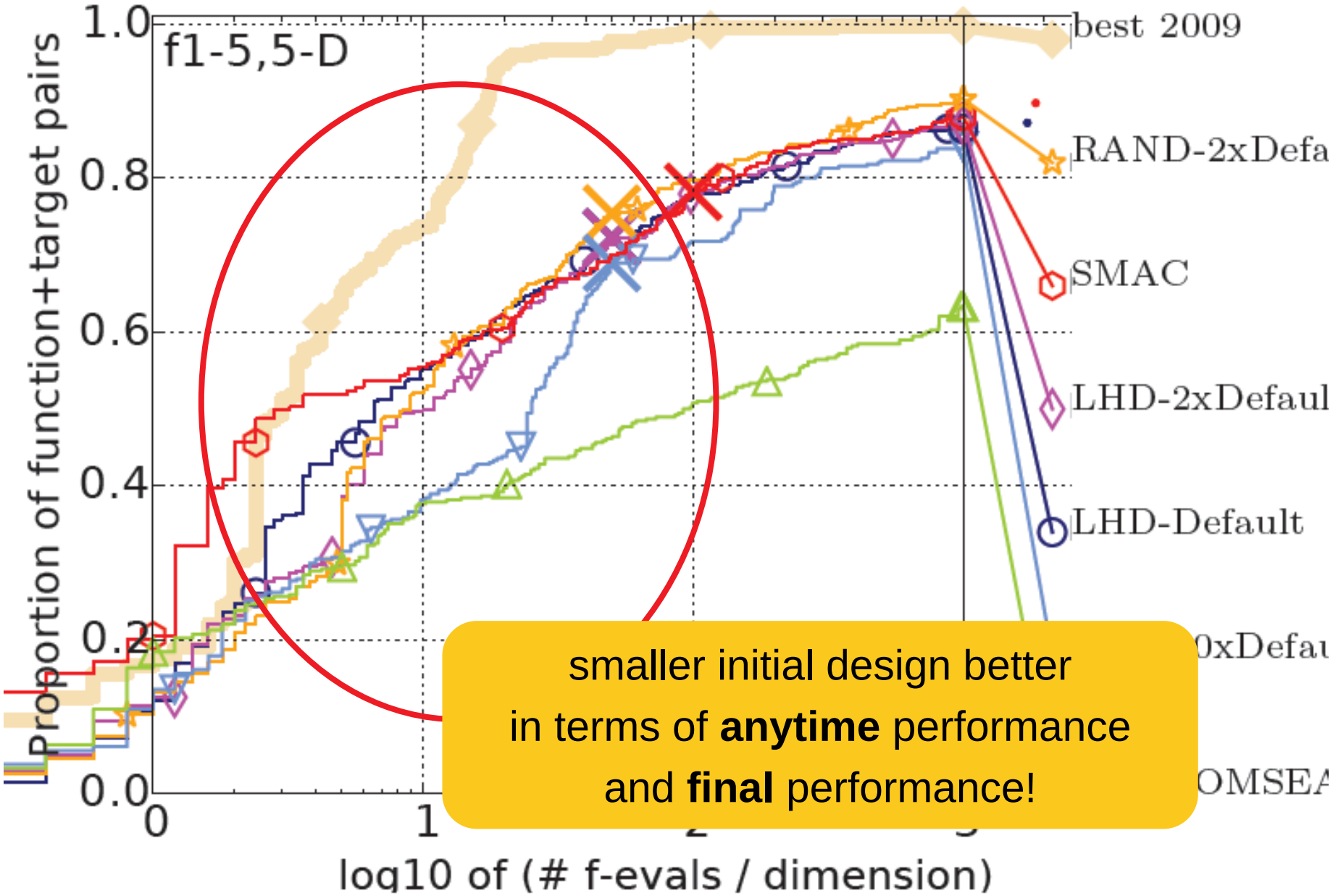
❶ initial design follows RANDOMSEARCH up to first evaluation of surrogate model

❷ smaller initial designs seem beneficial

❶ Initial Design Follows RANDOMSEARCH

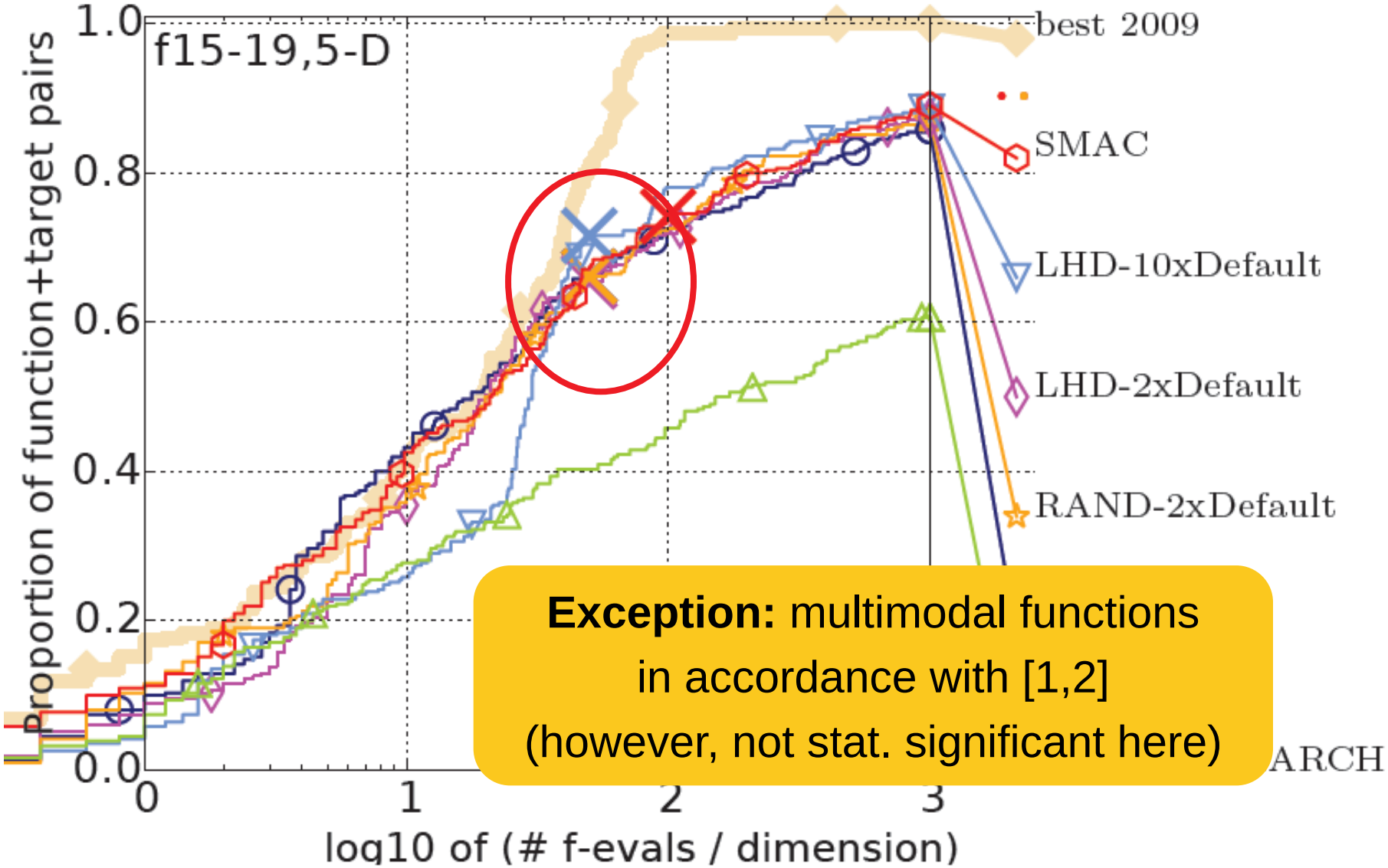


❷ Smaller Initial Designs Seem Beneficial



smaller initial design better
in terms of **anytime** performance
and **final** performance!

❷ Smaller Initial Designs Seem Beneficial



Exception: multimodal functions in accordance with [1,2] (however, not stat. significant here)

[1] T. Bartz-Beielstein and M. Preuss. Considerations of budget allocation for sequential parameter optimization (SPO). In Proc. Workshop on Empirical Methods for the Analysis of Algorithms (EMAA 2006), pages 35-40, 2006.

[2] F. Hutter, H. Hoos, and K. Leyton-Brown. An Evaluation of Sequential Model-Based Optimization for Expensive Blackbox Functions. In GECCO workshop on Black-Box Optimization Benchmarking (BBOB'2013), pages 1209-1216. ACM Press, 2013.

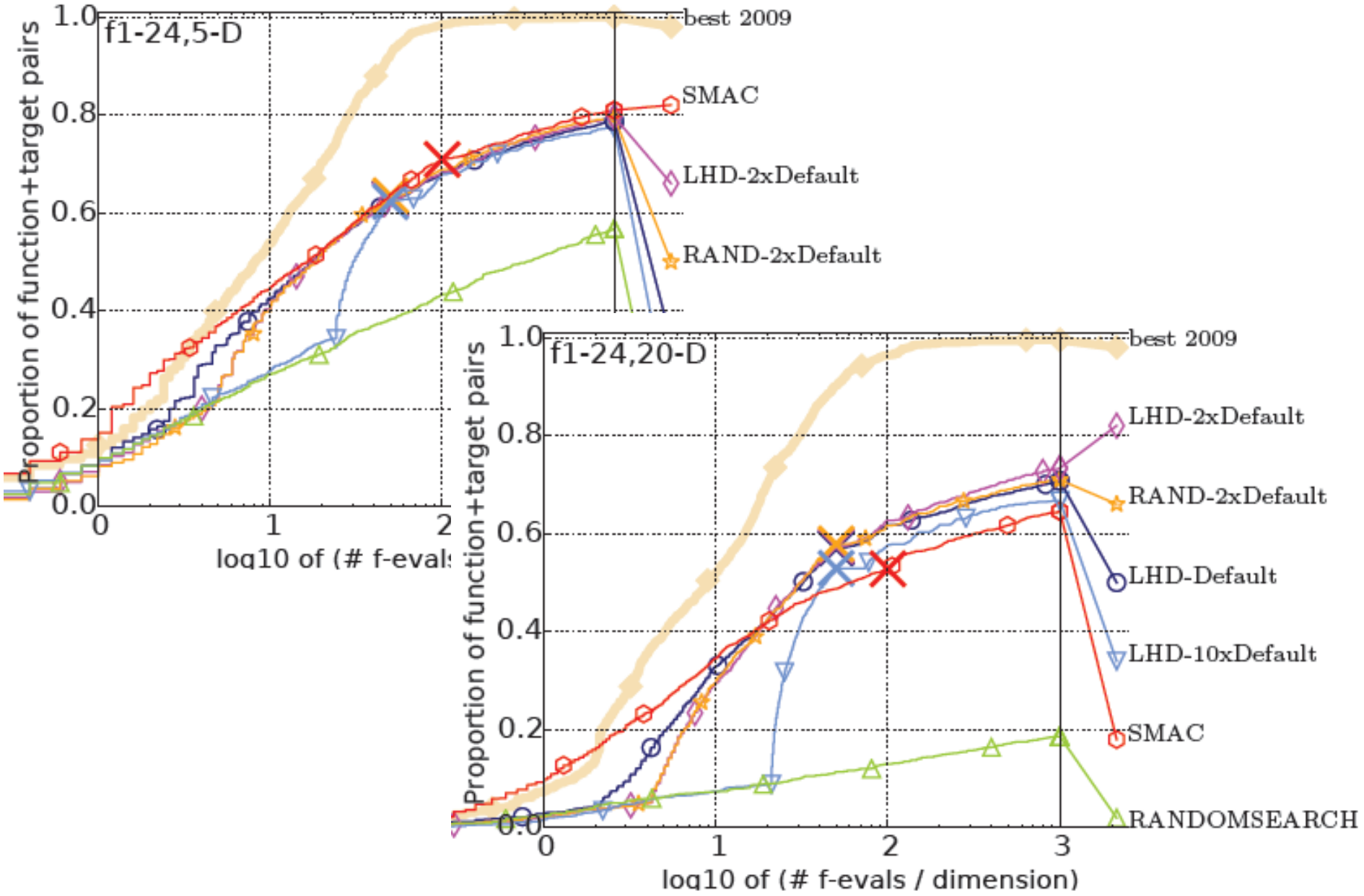
Scientific Question 2

Q2: What is the effect of replacing the LHS with simple (uniform pseudo-)random sampling?

Observation:

no (stat. significant) difference between Latin Hypercube Sampling and Random Sampling on the BBOB test functions (for 2x default initial design length)

No Difference Between LHS and Random Design



Conclusions

- First deeper investigations of the MATSuMoTo library for expensive optimization
- Impact of initial design
  - size
  - Latin Hypercube Sampling (LHS) vs. Random (RAND)
- Findings
  - smaller initial design seems better (exception: multimodality)
  - no difference between LHS and RAND
  - default MATSuMoTo variant best among tested ones
- To be done:
  - investigate other initial designs & other parameters
  - investigate how much better the results become with shorter initial design phase

questions?