Experimental Analysis of the Tournament Size on Genetic Algorithms

Yuri Cossich Lavinas[1] Claus Aranha[1] Marcelo Ladeira[2]
[1]Tsukuba University, Graduate School of Systems and Information Engineering, Japan
[2]University of Brasilia, Computer Science Department

Introduction

Genetic Algorithm (GA) is composed of:

- crossover operator (new solutions)
- mutation operator (sustain diversity)
- selection operator (select solutions to compose new population)

One of the most used selection is the tournament operator which has one parameter, the tournament size (K).

This parameter is said to adjust the balance between the exploration and explotation.

Goal: Observe the performance of the tournament selection in a real-valued Genetic Algorithm to learn the role of K.

Tournament Selection

- K candidates are sampled randomly and the one with best fitness is selected
- Controlling the tournament size - expect to adjust the balance exploration X explotation
- The selection pressure is expected to increase with larger tournament size values
- Review of the literature shows a preference to small values for the tournament size, such as as 2 or 3
- Little experimental justification

Experiment Design

- Real-valued GA, with solution's value between [-4,4]
- Configuration: Uniform crossover (0.9), elitism (1), Gaussian mutation (0.1, mean=0, std=0, prob, attribute to mutate-0.1)
- Fitness Functions: 24 noise-free BBOB benchmark functions with 10, 20 and 40 dimensions
- Tournament sizes: 2 to 25

Discussion

- We found no significant relationship between the tournament size values and the final quality
- A better understanding of the role of the tournament size in the final quality is necessary
- We want to explore this effect on other GA operators, including self-adaptative approaches

Statistical Analysis

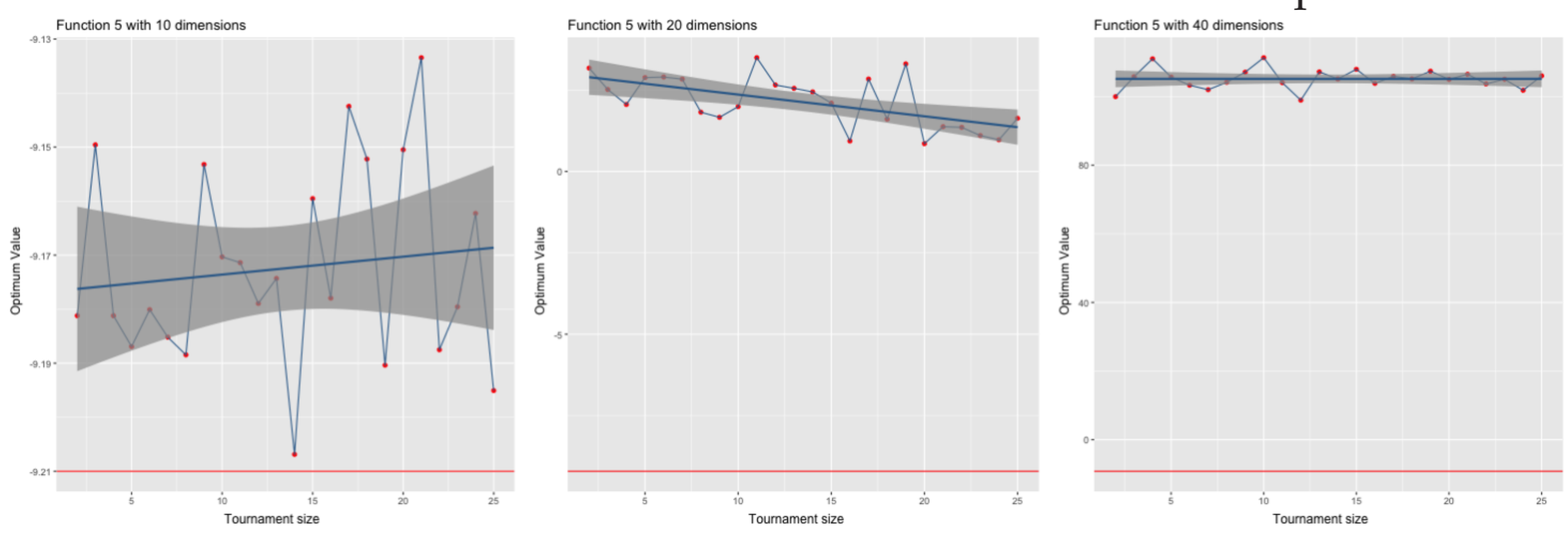
Dimension size	Chi-squared	P-value
10	31.497	0.1111
20	27.903	0.2195
40	23.288	0.444

Friedman Test on the effect of K given all functions

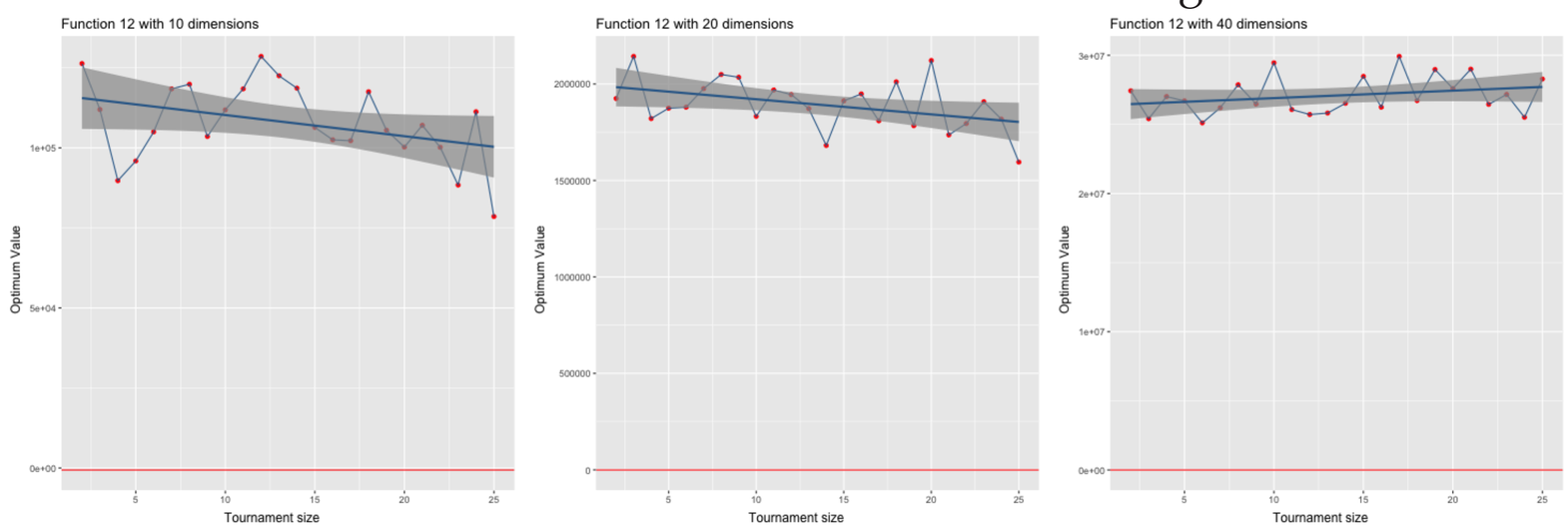
Observed Impact of the Tournament Size by Functions

The Figures below show that **changing the tournament size** does not lead to significant better final values found by the GA. The gray shaded area represents the 95% confidence level interval (predictions from a linear model) for each scenario. For all Figures, the mean of 40 repetitions is shown as bullets and the bars represent the standard deviation. All Figures show from left to right the results for 10, 20, 40 Dimensions. The red lines in the bottom of the Figures show the target value for each function.

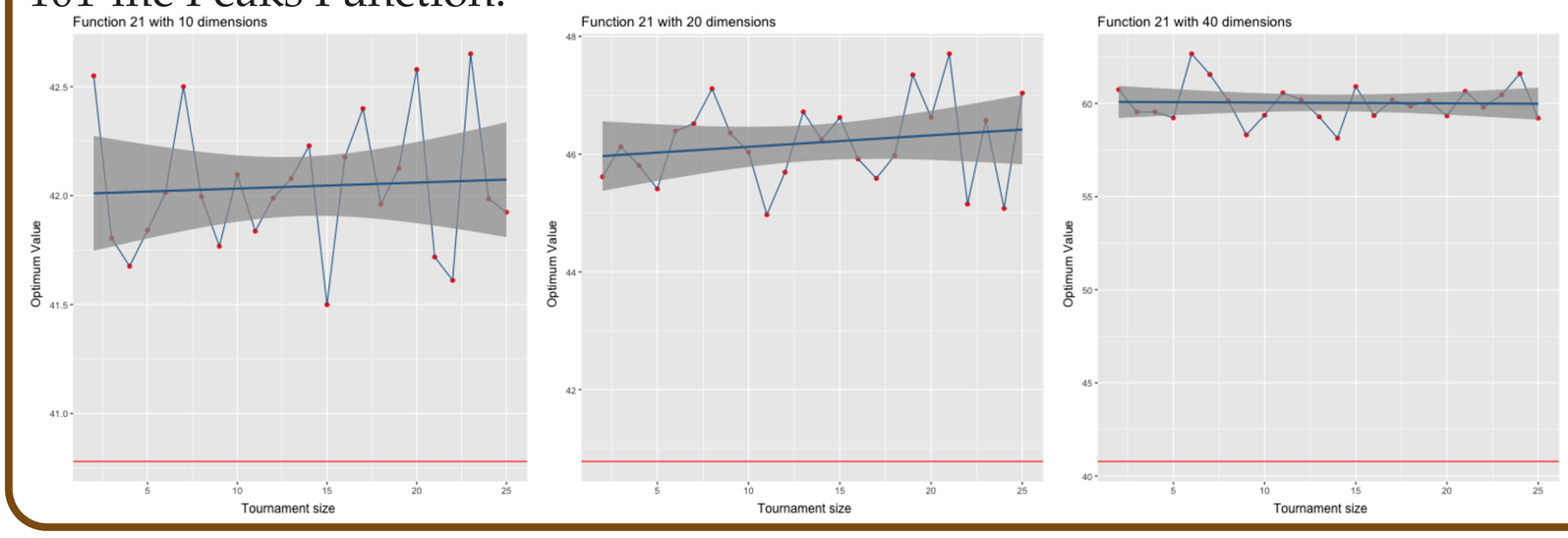
Performance on different tournament size for the Linear Slope Function.



Performance on different tournament size for the Bent Cigar Function.



Performance on different tournament size for the Gallagher's Gaussian 101-me Peaks Function.



Preliminary self-adaptative tournament size

The Figures below preliminary results of the self-adaptative tournament size when **half of the evaluations are completed**. The gray shaded area represents the 95% confidence level interval for predictions from a linear model for each scenario showed and demonstrate that no value of the tournament size has consistently better results. For all Figures, the mean of 40 repetitions is shown as bullets and the bars represent the standard deviation. The red lines in the bottom of the Figures show the target value for each function.

Performance for the Linear Slope Function,Bent Cigar Function with 40-D.