

# MBPTA-CV brief user guide, v1.0

Jaume Abella  
Barcelona Supercomputing Center (BSC)  
April 2017

## **MBPTA-CV input parameters**

### **Size of the sample you want to use:**

```
file_size <- N
```

Replace [N] by the number of measurements you want to use. Note that the input file must contain at least N measurements. If it contains more, only the first N measurements will be used.

### **Input file with execution time measurements. One value per line:**

```
input_file <- "filename"
```

Note that the file name must include the absolute path of the file. Use double quotation marks.

### **Label to appear in the plots**

```
label_plot <- "label"
```

### **Output file names for the pWCET plot and the CV plot:**

```
pWCETplotname <- 'pWCETfilename.png'
```

```
CVplotname <- 'CVfilename.png'
```

As for the input file, use absolute path of the file. Note that you must use single quotation marks.

### **Whether you want to generate the pWCET plot and the CV plot:**

```
doWCETplot <- TRUE
```

```
doCVplot <- TRUE
```

Set to TRUE or FALSE. If FALSE, the label and filenames for the plot become irrelevant.

## MBPTA-CV output example

### Successful execution:

```
[1] indep + i.d. test p-values: 0.15107017296622          0.592937271815384
[1] Number of tail values used: 50
[1] CV-value (ideally 1.0): 0.960344650212143
[1] MET: 22387
[1] pWCET    10-3,    10-6,    10-9,    10-12: 23043    24871    26698    28526
```

“Indep + i.d. test p-values” indicate the result of the independence and identically distributed tests. MBPTA-CV is configured so that p-value needs to be above 0.05 to pass the tests.

Number of tail values used indicates how many tail values have been used for the pWCET curve generation. This number is equal or larger than 50.

CV-value indicates the CV value for the tail size used for pWCET generation. It is, naturally, a value close to 1.0.

MET corresponds to the Maximum Execution Time in the input sample.

The pWCET values at exceedance probabilities  $10^{-3}$ ,  $10^{-6}$ ,  $10^{-9}$  and  $10^{-12}$  are also reported.

### Unsuccessful executions:

Four different reasons may lead to unsuccessful execution:

- 1) The independence test is failed. This is explicitly reported.
- 2) The identical distribution test is failed. This is also explicitly reported.
- 3) No convergence is achieved. This means that the tail is regarded as heavy, thus lacking enough values from the real tail. The script indicates that the sample size needs to be increased. In this case the CVplot is useful to see how many values belong to the tail (are within the exponentiality range), which indicates how much the sample size needs to be increased approximately. For instance, given that 50 tail values are wanted, if only the last 20 fall in the exponentiality range, then the sample size needs to be increased (approximately) by 2.5X at least.
- 4) Execution crashes. This may occur due to limitations in the R implementation of the solvers. For instance, if all values are identical, this may happen. The default solution is increasing the sample size.

## **Known limitations**

There are 2 known limitations of the method (not so related to the implementation itself):

### **Discretization issues:**

If data is too discrete (i.e. there are few different values in the sample that repeat many times), the best approximation can be systematically a heavy tail, thus not reaching convergence despite of the size of the sample. In fact, for very large samples, it may occur that the highest 10 values are identical, so they are regarded as a heavy tail (or a failure to converge), and further values cannot solve this issue.

This is a known limitation of EVT, so it spans beyond MBPTA-CV method. Tackling this issue is part of our ongoing research.

### **Pessimistic pWCET estimates:**

MBPTA-CV always delivers an exponential distribution. While this is an upper-bound by construction for real-time programs, some of them may be upper-bounded with light tails. Thus, using an exponential distribution in those cases is unnecessarily pessimistic. However, we lack a mechanism to prove when a light tail is a reliable upper-bound, so research on this matter is also needed if tighter estimates are wanted.