


# A fast algorithm to compute a curve of confidence upper bounds for the False Discovery Proportion using a reference family with a forest structure

Guillermo Durand <sup>1</sup> Laboratoire de Mathématiques d'Orsay, Université Paris-Saclay

Date published: 2024-06-30 Last modified: 2024-06-30

## Abstract

This paper presents a new algorithm (and an additional trick) that allows to compute fastly an entire curve of post hoc bounds for the False Discovery Proportion when the underlying bound  $V_{\mathfrak{R}}^*$  construction is based on a reference family  $\mathfrak{R}$  with a forest structure à la [Durand et al. \(2020\)](#). By an entire curve, we mean the values  $V_{\mathfrak{R}}^*(S_1), \dots, V_{\mathfrak{R}}^*(S_m)$  computed on a path of increasing selection sets  $S_1 \subsetneq \dots \subsetneq S_m, |S_t| = t$ . The new algorithm leverages the fact that going from  $S_t$  to  $S_{t+1}$  is done by adding only one hypothesis.

**Keywords:** multiple testing, algorithmic, post hoc inference, false discovery proportion, confidence bound

1	<b>Contents</b>	
2	<b>1 Introduction</b>	2
3	<b>2 Notation and reference family methodology</b>	2
4	2.1 Multiple testing notation . . . . .	2
5	2.2 Post hoc bounds with reference families . . . . .	2
6	2.3 Deterministic regions with a forest structure . . . . .	2
7	<b>3 New algorithms</b>	2
8	3.1 Pruning the forest . . . . .	2
9	3.2 Fast algorithm to compute a curve of confidence bounds on a path of selection sets	2
10	<b>4 Numerical experiments</b>	2
11	<b>5 Conclusion</b>	2
12	<b>References</b>	2
13	<b>Session information</b>	2

<sup>1</sup>Corresponding author: [guillermo.durand@universite-paris-saclay.fr](mailto:guillermo.durand@universite-paris-saclay.fr)

14	<b>1 Introduction</b>
15	<b>2 Notation and reference family methodology</b>
16	<b>2.1 Multiple testing notation</b>
17	<b>2.2 Post hoc bounds with reference families</b>
18	<b>2.3 Deterministic regions with a forest structure</b>
19	<b>3 New algorithms</b>
20	<b>3.1 Pruning the forest</b>
21	<b>3.2 Fast algorithm to compute a curve of confidence bounds on a path of selection sets</b>
22	
23	<b>4 Numerical experiments</b>
24	<b>5 Conclusion</b>

## 25 References

26 Guillermo Durand, Gilles Blanchard, Pierre Neuvial, and Etienne Roquain. Post hoc false positive  
 27 control for structured hypotheses. *Scand. J. Stat.*, 47(4):1114–1148, 2020. ISSN 0303-6898. doi:  
 28 10.1111/sjos.12453. URL <https://doi.org/10.1111/sjos.12453>.

## 29 Session information

```

30 R version 4.4.0 (2024-04-24)
31 Platform: x86_64-pc-linux-gnu
32 Running under: Ubuntu 22.04.4 LTS
33
34 Matrix products: default
35 BLAS: /usr/lib/x86_64-linux-gnu/openblas-pthread/libblas.so.3
36 LAPACK: /usr/lib/x86_64-linux-gnu/openblas-pthread/libopenblas-r0.3.20.so; LAPACK version 3.10.0
37
38 locale:
39  [1] LC_CTYPE=C.UTF-8      LC_NUMERIC=C           LC_TIME=C.UTF-8
40  [4] LC_COLLATE=C.UTF-8    LC_MONETARY=C.UTF-8    LC_MESSAGES=C.UTF-8
41  [7] LC_PAPER=C.UTF-8      LC_NAME=C               LC_ADDRESS=C
42  [10] LC_TELEPHONE=C        LC_MEASUREMENT=C.UTF-8 LC_IDENTIFICATION=C
43
44 time zone: UTC
45 tzcode source: system (glibc)
46
47 attached base packages:
48 [1] stats      graphics  grDevices datasets  utils      methods    base
49
50 loaded via a namespace (and not attached):
51 [1] compiler_4.4.0    fastmap_1.1.1      cli_3.6.2          htmltools_0.5.8.1

```

```
52  [5] tools_4.4.0      yaml_2.3.8      rmarkdown_2.26  knitr_1.46
53  [9] jsonlite_1.8.8   xfun_0.43       digest_0.6.35   rlang_1.1.3
54 [13] renv_1.0.7       evaluate_0.23
```