

# Discussion of ‘Multiscale changepoint inference’

by K. Frick, A. Munk, and H. Sieling

Guenther Walther

Department of Statistics, 390 Serra Mall  
Stanford University, Stanford, CA 94305, USA

gwalther@stanford.edu

January 26, 2014

The multiscale test considered in the paper requires the computation of local test statistics on all intervals  $[i/n, j/n]$  for  $1 \leq i < j \leq n$ . A straightforward implementation will thus result in a  $O(n^2)$  algorithm, making computation infeasible for problems of moderate size. The authors address this by introducing methodology based on dynamic programming, which results in an improved computation time in many cases. Alternatively, it may be promising to evaluate the statistic only on a sparse approximating set of intervals. Rivera and Walther (2013) introduce such an approximating set for the closely related problem of constructing a multiscale likelihood ratio statistic for densities and intensities. The idea is that after considering an interval  $[i/n, j/n]$  with large  $j - i$ , not much is gained by also looking at, say,  $[i/n, (j + 1)/n]$ . It turns out that it is possible to construct an approximating set of intervals that is sparse enough to allow computation in  $O(n \log n)$  time, but which is still rich enough to allow optimal detection of jumps.

Another advantage in employing such an approximating set is that it considerably simplifies the theoretical treatment of the multiscale statistic. It is notoriously difficult to establish theoretical results about the null distribution of the multiscale statistic such as Theorem 1. Even if one is not interested in the limiting distribution because the critical value is obtained by simulation (an option made feasible by the  $O(n \log n)$  algorithm described above!), it is still necessary to show that the null distribution is  $O_p(1)$  to establish optimality results such as Theorem 5. The standard method of proof introduced in Dümbgen and Spokoiny (2001) requires establishing two exponential inequalities: First, one needs to establish subgaussian tails for the local statistic, which is often

straightforward. The second exponential inequality, however, concerns the change between local test statistics, and this inequality is often very difficult to derive. Rivera and Walther (2013) show that if one employs a sparse approximating set, then the  $O_p(1)$  result follows directly from the subgaussian tail property of the local statistics, which is typically easy to obtain.

It appears that sparse approximating sets can similarly be constructed for relevant multivariate problems, see Walther (2010). The advantages of these sets for computation as well as theoretical analysis suggest that these approximating sets may play an important role for these types of problems in general.

References:

C. Rivera and G. Walther: Optimal detection of a jump in the intensity of a Poisson process or in a density with likelihood ratio statistics.

*Scandinavian Journal of Statistics*. **40** (2013), 752–769.

G. Walther: Optimal and fast detection of spatial clusters with scan statistics.

*Annals of Statistics* **38** (2010), 1010–1033.