

# Review on Manuscript COST-D-19-00031

## 1 Summary of the contribution:

The authors present a simulation study to evaluate various projection pursuit indexes for detecting specific bivariate patterns in the data. Moreover, two data sets from physics are examined using PPI.

## 2 Major remarks

The simulation study targets three particular structures the pipe, the sine, and the spiral. Some motivation and justification for selecting these structures should be added.

Second, the simulation study uses six-dimensional data of which four dimensions contribute nuisance and the other two the intended structure. Implicitly, this is clear, but the description does not make this explicit and uses the general notion of  $p$  dimensions instead. Please make this more explicit and also provide some justification that this set-up is adequate for targeting the problem at hand.

Thirdly, the simulations refer to sequences of projections (the horizontal axis in plots in Figures 2, 3, and 4. As I understand these sequences are a consequence of optimising the corresponding projection pursuit index. How can you then actually control the two different scenarios referred to in section 3.2 and 3.3? Or, do you extract a specified sequence of projections? Moreover, how do you make sure that you have the same sequence length for each trial? Please clarify!

The optimisation methods are further discussed in section 3.4 and depend on the structure present in the data. What is the value of these investigations for more realistic settings in which the structure is not known.

Finally, the relationship between the simulation study and the examples using the astrophysics data remains unclear. How are the physics application slinked to the simulation study and vice versa? Currently, the two aspects seem to me rather independent and could actually be presented in separate manuscripts. Moreover, since some rather important aspects of the simulation study have been put into an appendix. For the COST audience certainly the details of the simulation study are more relevant than the astrophysics application.

### 3 Minor remarks

When introducing the various PPIs some terms and notations are used which are not properly introduced, e.g. alpha hull,  $c_{convex}$ . Check which ones actually need to be properly defined as they are needed for further steps and for which ones you could just refer to the literature avoiding unnecessary details.

In Figure 3, each plot shows a vertical blue line. Please explain what this line represents.

p. 14, l. 26: **optimist** should be **optimise**