

Supplementary Material A

Distribution of the optimal number of components.

We performed 1000 rearrangements of the rows of the dataset used to build figure 1 (Supplementary Data, Dataset1.xlsx).

For each rearrangement, the value of the optimal number of component (ONC) was determined by looking for the first local minimum of $p_{CV-ANOVA}$ when NC is incremented.

The ONC distribution is represented in figure S1. We observe that for the 1000 rearrangements, values of ONC spread from 1 to 6. This corresponds to a Q2 that spreads from -0.8 to 0.38.

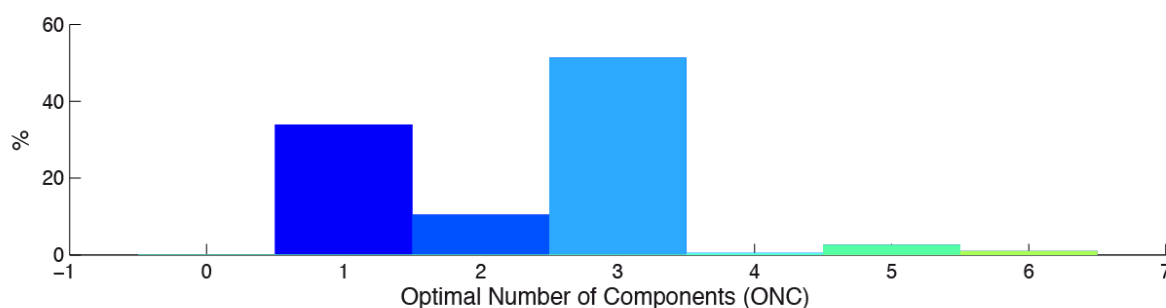


Figure S1: Variation of ONC when 1000 random permutations of the dataset lines are performed. Each color is associated to a specific value of NSC.