

## IN SEARCH OF NON-GAUSSIAN COMPONENTS OF A HIGH-DIMENSIONAL DISTRIBUTION

MOTOAKI KAWANABE

In high dimensional data analysis, finding non-Gaussian components of data is an important preprocessing step for efficient information processing. This article proposes a new linear method to identify the non-Gaussian subspace within a very general semi-parametric framework. Our proposed method called NGCA (Non-Gaussian Component Analysis) is essentially based on the fact that we can construct a linear operator which, to any arbitrary nonlinear (smooth) function, associates a vector which belongs to the low dimensional non-Gaussian target subspace up to an estimation error. Furthermore, we show that this procedure is consistent in the sense that the estimation error tends to zero at a parametric rate. By applying the operator to a family of different nonlinear functions, one obtains a family of different vectors lying in a vicinity of the target space. PCA is then applied to this family to estimate the non-Gaussian subspace.

A numerical study demonstrates the usefulness of our method.

This is a joint work with G.Blanchard, M.Sugiyama, V.Spokoiny and K.-R.Mueller