

## **Numerical separation in HPLC-UV-DAD with Independent Component Analysis (ICA) using high order statistics for the automated identification of peaks.**

***P. Lebrun*<sup>1</sup>, *B. Debrus*<sup>1</sup>, *B. Boulanger*<sup>2</sup>, *G. Caliaro*, *A. Ceccato*<sup>3</sup>, *Ph. Hubert*<sup>1</sup>**

A major concern to achieve the development of analytical methods is the automated detection and identification of compounds contained in complex samples. This step is critical and often time-consuming, more especially when Designs Of Experiments (DOE) are used to generate chromatograms. The use of DOE leads to maximize the changes of the analytical conditions in order to explore an experimental domain. This often gives very different and uncontrolled chromatograms. In this case, peak picking and peak tracking can be somewhat complicated. Numerical methods such as the Independent Component Analysis (ICA) give the ability to overcome these difficulties. ICA is a statistical tool which allows separating complex and overlapping signals. The main idea is to suppose that the observed signals are resulting from a sum of several independent phenomenons, known as sources. There are two main objectives: first, to separate with ICA a matrix signal in a certain unknown number of sources. We develop an original statistical method to find this number automatically. Second, to sort the sources to identify the noise sources from the peaks and other relevant artifacts ones such as front or end of gradient. We propose to use high order statistics such as kurtosis and skewness of the distributions of the sources to select relevant sources from the complete set of computed sources [1]. Clustering method *k*-means is used to automate this statistical analysis. This new approach is very promising and has been applied successfully to real samples. Moreover, this latter does not require expensive equipments such as mass spectrometer to identify all compounds of a complex sample.

[1]A. Delorme, T. Sejnowsky, S. Makeig, "Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis", 2006, *NeuroImage* 34 (2007), 1443-1449