## Fast Methods for Dimensionality Reduction Based on the Generative Topographic Mapping

## M. Griebel<sup>1</sup> and <u>A. Hullmann<sup>1</sup></u>

**Abstract:** Most high-dimensional data exhibit some correlation structure, which means that they are not distributed uniformly in the data space and have a lower intrinsic dimension. The Principal Component Analysis (PCA) is the best-known method to find low-dimensional parametrisations of high-dimensional data. As it fails to capture non-linear dependencies, several more complex methods have been invented.

The Generative Topographic Mapping (GTM) is a method for non-linear dimensionality reduction, which has been published in 1998 by Bishop, Williams and Svensen. We will show how a continuous formulation of the GTM allows for the use of efficient quadrature methods and discretisations.

Instead of the usual full grid discretisation of the mapping from the latent space to the data space, we use sparse grids to break the 'curse of dimensionality' to some extent. Alternatively, we restrict the discretisation to low-order ANOVA-terms, which removes any exponential dependence of the runtime complexity on the embedding dimension. From a statistical perspective, this model exploits the independence between different groups of data space dimensions, yielding a method that is more flexible than the PCA but significantly faster that the original GTM. As a third possibility, we discuss the application of the representer theorem on the discretisation of our mapping.

As it has been shown, the Eucledian metric may not be optimal in high dimensions, which is related to the 'concentration of measure'-effect. We describe a GTM based on the *p*-Minkowski-Norm, and show that its practicality depends on the noise the data exhibit. Furthermore, we will use the GTM not only for dimensionality reduction, but also for classification problems. This allows for demonstrative experiments that study effects of the intrinsic dimension of high-dimensional data.

<sup>&</sup>lt;sup>1</sup> Institute for Numerical Simulation Wegelerstr. 6, D-53115 Bonn, Germany {griebel, hullmann}@ins.uni-bonn.de