

# Spectral clustering of Big Data in genetics: applications to RNA-seq data\*

Tsvetelin Zaeovski, Ognyan Kunchev, Dean Palejev, Eugenia Stoimenova  
Institute of Mathematics and Informatics, Bulgarian Academy of Sciences

`t_s_zaeovski@math.bas.bg`

## Abstract

The purpose of this research is to use a spectral clustering algorithm to improve the results available in the literature in the area of clustering of Big data in genetics. The particular application we present is to RNA-seq data, published by Bottomly et al. The data consist of two strands of mice, and the goal is to identify the genes with different overall expression levels between the two strands. There are many statistical methods which lead to clustering of genes. We apply a spectral clustering method to improve some of the existing differential expression methods.

---

\*The present research has been supported by projects I 02/19 and DH02-13 with DFNI.