

# Integrative Analysis of Multi-Omic Data via Sparse Multiple Co-Inertia Analysis

Qi Long

*University of Pennsylvania*

*E-mail: qlong@upenn.edu*

**Abstract:** Multiple co-inertia analysis (mCIA) is a multivariate analysis method that can assess relationships and trends in multiple datasets. Recently it has been used for an integrative analysis of multiple highdimensional -omics datasets. However, the estimated loading vectors from the existing mCIA method are non-sparse, which presents challenges for interpreting analysis results. We propose two new mCIA methods: 1) a sparse mCIA (smCIA) method that produces sparse loading estimates and 2) a structured sparse mCIA (ssmCIA) method that further enables the incorporation of structural information among variables such as those from functional genomics. The two proposed methods achieve simultaneous model estimation and feature selection and yield analysis results that are more interpretable than the existing mCIA. Our extensive simulation studies demonstrate the superior performance of the smCIA and the ssmCIA methods compared to the existing mCIA. We also apply our methods to the integrative analysis of transcriptomics data and proteomics data from a cancer study.