

# Comparison of SiMLR and RGCCA: summary of all experiments

Brian B. Avants et al.

2020-10-15

Table 1: Summary of RGCCA-SGCCA-SiMLR comparison results. RGCCA = regularized generalized canonical correlation analysis; SGCCA = sparse generalized canonical correlation analysis; Sim = similarity-driven multivariate linear reconstruction (SiMLR); Reg = regression; CCA = absolute canonical covariance; ICA = ICA mixing method; SVD = SVD mixing method.

Best results are highlighted in cadet blue; worst in antiquewhite. SiMLR with the absolute canonical covariance similarity measurement and SVD (SimCCASVD) as a mixing method performs best overall. SiMLR with the regression energy and ICA mixing method (SimRegICA) outperforms SGCCA most consistently across sparseness levels, provides closely competitive performance overall, and is highlighted in pink. MSE for the matrix completion error on SNP embeddings is measured by the

Frobenius norm between the reconstructed and the true SNP embeddings. The PING examples are exploratory analyses described in the supplementary information as we cannot directly share the data. The "n comp" description in the PING table refers to the number of significant components related to either anxiety or depression.

study	RGCCA	SGCCA	SimCCAICA	SimCCASVD	SimRegICA	SimRegSVD	metric
Signal-Sens.	0.35+/-0.18	0.45+/-0.17	0.5+/-0.15	0.51+/-0.14	0.49+/-0.13	0.49+/-0.14	R-squared
Noise-Sens.	0.09	0.16	0.09	0.06	0.07	0.1	R-squared
Mixomics	N/A	0.62 +/- 0.01	0.64 +/- 0.03	0.65 +/- 0.03	0.65 +/- 0.04	0.61 +/- 0.03	Concordance
brainAge	N/A	2+/-1.5	1.6+/-1.2	1.4+/-1.2	1.6+/-1.3	1.7+/-1.2	MAE
PING-Anx	N/A	1 comp.	N/A	3 comp.	5 comp.	N/A	Inferential
PING-Dep	N/A	1 comp.	N/A	1 comp.	5 comp.	N/A	Inferential