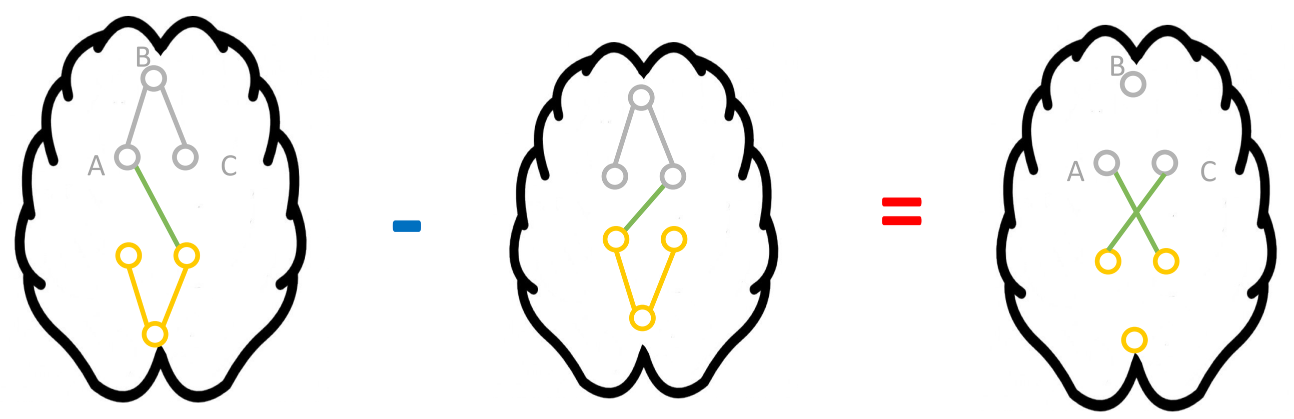
Fast and Scalable Learning of Sparse Changes in High-Dimensional Gaussian Graphical Model Structure

Beilun Wang, Arshdeep Sekhon and Yanjun Qi
Department of Computer Science, University of Virginia

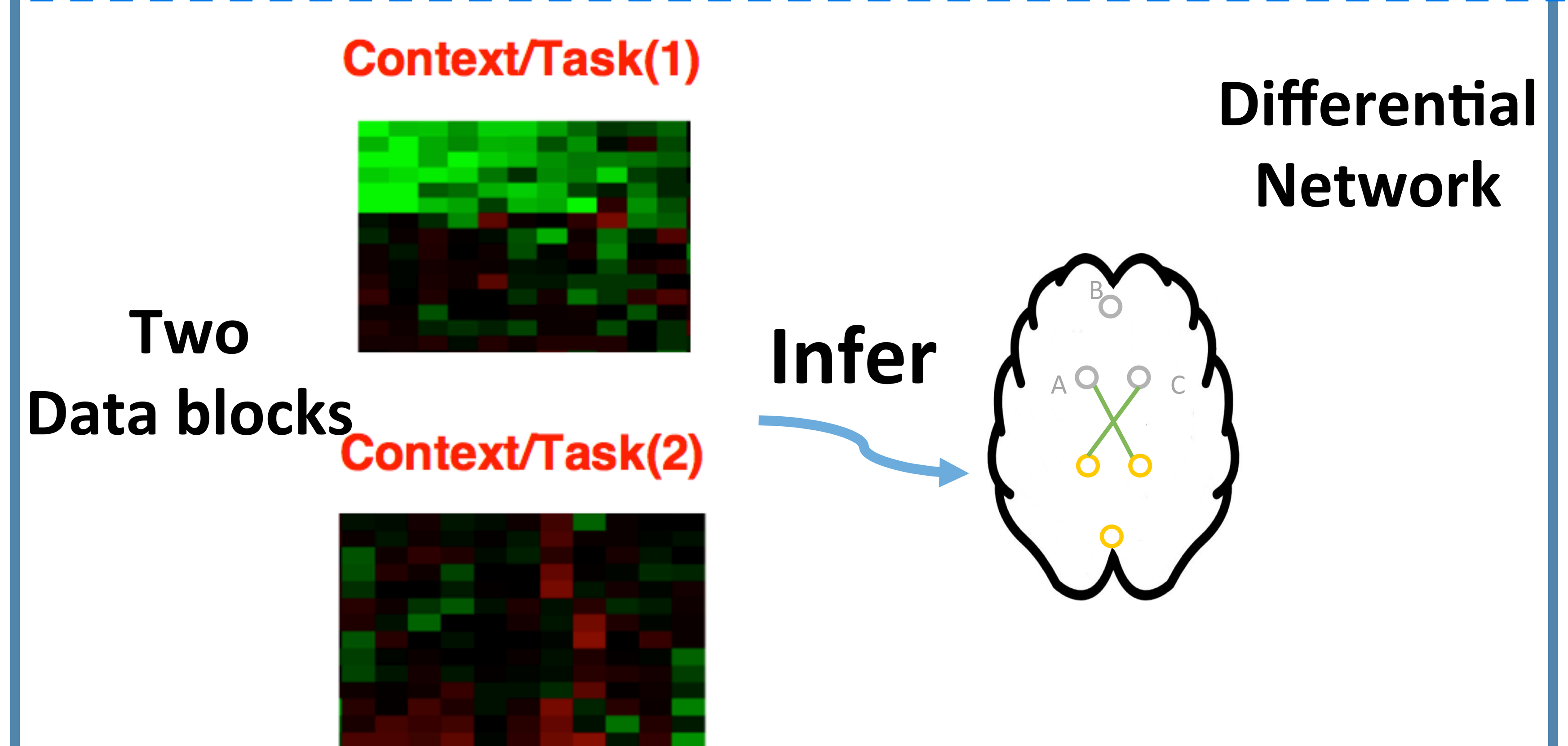
Introduction

Differential Network

Two dense graph, the changes between them is sparse



$$\Delta = \Omega_d - \Omega_c$$



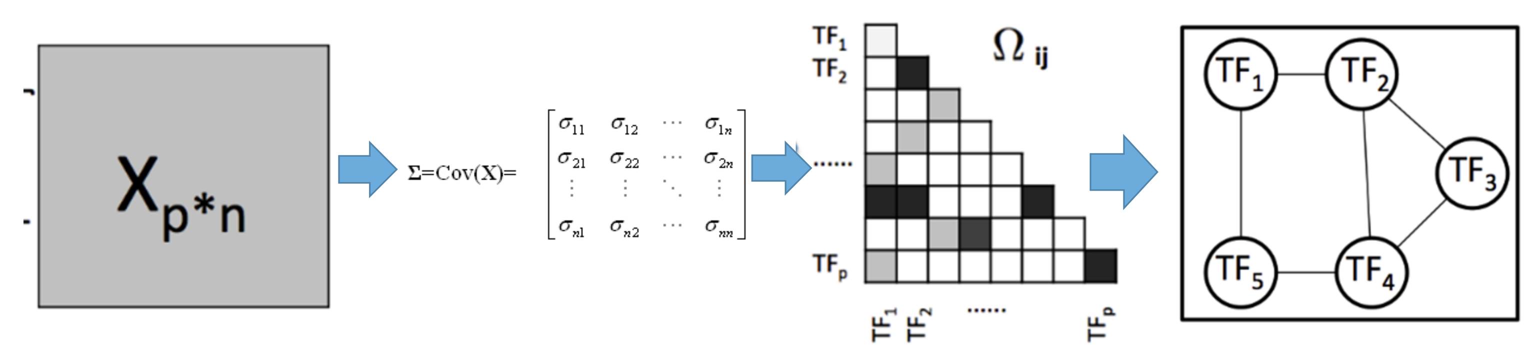
Background: Number of feature (p) are increasing.

The past decade has seen a revolution in collecting largescale heterogeneous data from many scientific fields. For instance, genomic technologies have delivered fast and accurate molecular profiling data across many cellular contexts (e.g., cell lines or stages) from national projects like ENCODE

	Previous	Now
p	Yeast data: 326 genes	Encode project: more than 20000 genes

Background: sGGM to derive Conditional Independence Graph from data.

Sparse Gaussian Graphical Model is solved by the following three steps: (1) Calculate the sample covariance matrix; (2) estimate the sparse inverse of covariance matrix; (3) extract sparsity pattern in the inverse of covariance matrix. The solution of the second step includes: gLasso, neighborhood selection or Elementary Estimator.



DIFFEE: DIFFerential network identification via an Elementary Estimator

We propose a novel approach, **DIFFEE** for fast and scalable learning of sparse changes in high-dimensional Gaussian Graphical Model structure . As the first study of differential network using the Elementary Estimator framework, our work has three major Advantages: (1) Closed-form solution; (2) Fast and scalable computation; (3) Achieves sharp convergence rate.

Elementary Estimator:

$$\underset{\theta}{\operatorname{argmin}} ||\theta||_1$$
$$\text{Subject to: } ||\theta - \mathcal{B}^*(\widehat{\phi})||_\infty \le \lambda_n$$

$$[T_v(\widehat{\Sigma}_d)]^{-1} - [T_v(\widehat{\Sigma}_c)]^{-1}$$

DIFFEE:

$$\underset{\Delta}{\operatorname{argmin}} ||\Delta||_1$$
$$\text{Subject to: } ||\Delta - \left([T_v(\widehat{\Sigma}_d)]^{-1} - [T_v(\widehat{\Sigma}_c)]^{-1}\right)||_\infty \le \lambda_n$$

Computational complexity Comparison:

DIFFEE	FusedGLasso	Density Ratio	Diff-CLIME
$O(p^3)$	$O(T * p^3)$	$O((n_c + p^2)^3)$	$O(p^8)$

Theoretical Analysis

DIFFEE:

Convergence rate

$$||\widehat{\Delta} - \Delta^*||_\infty \le \frac{16\kappa_1 a}{\kappa_2}\sqrt{\frac{\log p}{\min(n_c, n_d)}}$$
$$||\widehat{\Delta} - \Delta^*||_F \le \frac{32\kappa_1 a}{\kappa_2}\sqrt{\frac{k \log p}{\min(n_c, n_d)}}$$
$$||\widehat{\Delta} - \Delta^*||_1 \le \frac{64\kappa_1 a}{\kappa_2}k\sqrt{\frac{\log p}{\min(n_c, n_d)}}$$

Algorithm:

DIFFEE can be solved with a closed-form solution.

Algorithm 1 DIFFEE
input Two data matrices $\mathbf{X}_c$ and $\mathbf{X}_d$.
input Hyper-parameter: λ_n and v
output Δ
1: Compute $[T_v(\widehat{\Sigma}_c)]^{-1}$ and $[T_v(\widehat{\Sigma}_d)]^{-1}$ from $\widehat{\Sigma}_c$ and $\widehat{\Sigma}_d$.
2: Compute $\Delta = S_{\lambda_n}([T_v(\widehat{\Sigma}_d)]^{-1} - [T_v(\widehat{\Sigma}_c)]^{-1})$
output Δ

Experiment Evaluation