Learning Latent Variable Gaussian Graphical Models
Zhaoshi Meng\textsuperscript{1}, Brian Eriksson\textsuperscript{2}, Alfred Hero\textsuperscript{1} \hspace{1cm} \textsuperscript{1} University of Michigan \hspace{0.5cm} \textsuperscript{2} Technicolor Research Center

**Problem Statement**
Many applications where the number of covariates($p$) is much larger than the number of observations ($n$), i.e., a “large $p$, small $n$” regime

**Gaussian Graphical Model (GGM)**
GGM characterizes distribution via conditional dependency relations

- **Key Limitation:** Real-world data does not fit well to sparse GGM

**Main Assumptions**

**Assumption 1 (Restricted Fisher Eigenvalue).** The Fisher information of the true model is sufficiently curved (i.e., lower bounded by a quadratic function) along restricted sparse and low-rank directions, respectively.

Implication: Parameter consistency and error rate can be established \cite{Neyshabur2012}.

**Assumption 2 (Structural Fisher Incoherence).** The maximum eigenvalue of the projected Fisher information w.r.t. sparse and low-rank subspace pairs is upper bounded.

Implication: Parameter consistency for “dirty” superposition model \cite{Yang2013}.

**Assumption 3 (Effective Rank).** The effective rank of true marginal covariance matrix increases more slowly than $p$ \cite{Lounici2012}.

$$r_{\text{eff}}(\Sigma) := \text{tr}(\Sigma)/\|\Sigma\|_2$$

(Right: Validation of Assumption 3. Effective ranks of randomly generated LVGGMs with various configurations.)

**Main Result**
Theorem. Suppose Assumptions 1-3 hold for the true marginal precision matrix $\Theta$. Assume $n \geq O(\log^* p)$ and the regularization parameters satisfy

$$\lambda = 160C_1p^{3/2}\sqrt{\log p}/n$$

$$\mu = C_2p^{3/2}\sqrt{r_{\text{eff}} \log p}/n.$$ 

Then with high probability approaching one, we have error rate

$$\|\hat{\Theta} - \Theta\|^2_F \leq 6\lambda \sqrt{n \log p / n} + 2\sqrt{r_{\text{eff}} \cdot r \cdot \log(2p)}.$$ 

**Our Contributions**
- High-dimensional parameter estimation error bounds for LVGGM under mild conditions:
  - Restricted Fisher Eigenvalue
  - Structural Fisher Incoherence
  - Empirically, the covariance often has low effective rank which leads to loosening of convergence conditions ($n \geq O(\log^* p)$) and improved error rate (similar to sparse GGM)

*Our theory predicts required sample size and regularization parameters to ensure LVGGM convergence*

**Acknowledgement** This work is in part supported by ARO grant W911NF-11-1-0391.

(2) Hsieh et al., BIC & QIC: Sparse inverse covariance estimation for a million variables. NIPS 2013.