Dynamic Interaction Network Estimation via Multivariate Mixed-Effects Models (DINE)

Samuel Ozminkowski

March 17, 2026

Northwestern University

Problem Setup

Proteomics

- Protein levels associated with
  - Alzheimer's[1,2]
  - Early-onset preeclampsia[3]
  - Late-onset preeclampsia[4]
  - Down syndrome[1]
- How do protein interactions change as disease progresses?

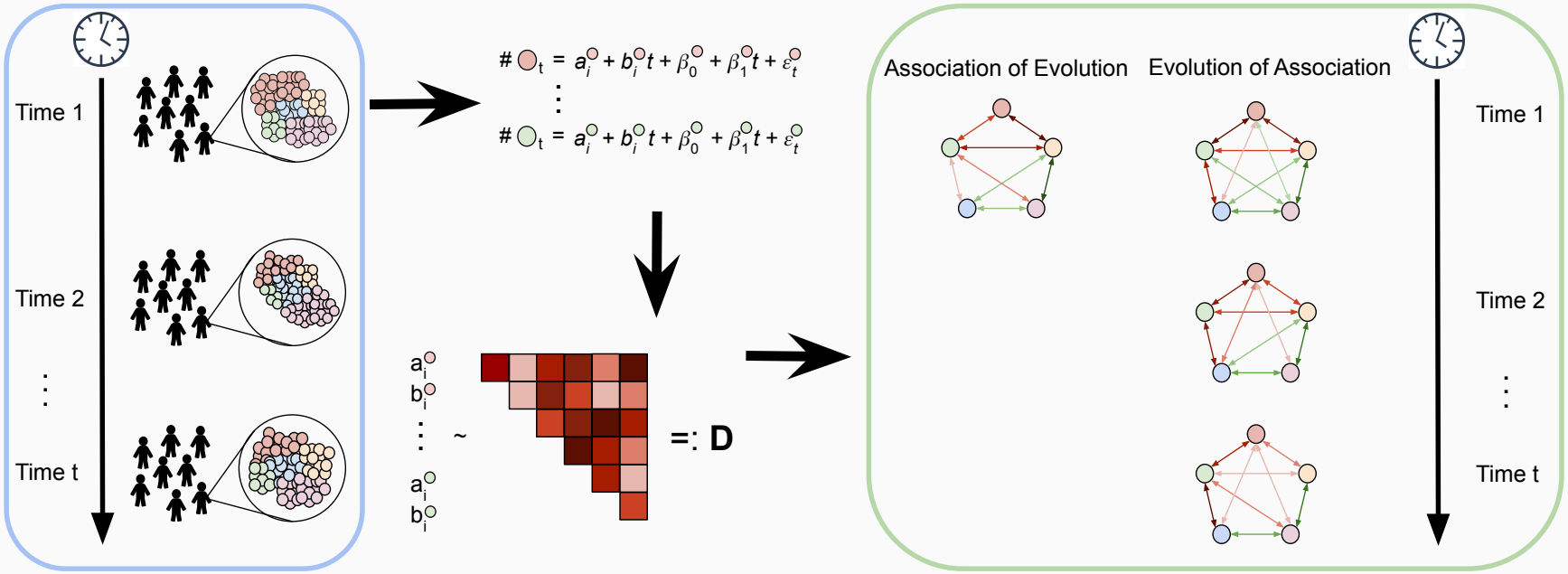
[1] Greber et al. (1999)
[2] Libiger et al. (2025)
[3] Tarca et al. (2019)
[4] Erez et al. (2025)

Proteomics[5]

- Longitudinal protein expression across multiple subjects
- Association of Evolution
  - Relationship between evolution of protein A and evolution of protein B
- Evolution of Association
  - How relationship between two proteins evolves over time

[5] Gao et al. (2017)

Joint modeling framework



$$Y \sim N(X\beta, V)$$

$$V = ZDZ^T + E$$

Joint Modeling Framework[6]

For protein $j = 1, ..., K$ at time $t = 1, .., T$ from subject $i = 1, ..., N$:

$$
\begin{aligned}
Y_{11}(t) &= \alpha_1 + a_{11} + (\beta_1 + b_{11})t + \varepsilon_{11t} \\
&\vdots \\
Y_{1j}(t) &= \alpha_j + a_{1j} + (\beta_j + b_{1j})t + \varepsilon_{1jt} \\
&\vdots \\
Y_{NK}(t) &= \alpha_K + a_{NK} + (\beta_K + b_{NK})t + \varepsilon_{NKt}
\end{aligned}
$$

The K mixed effects models are linked by the joint distribution of random factors D.

[6] Fieuws & Verbeke (2006)

Joint Modeling Framework

$$\begin{bmatrix} a_{i1} \\ b_{i1} \\ \vdots \\ a_{iK} \\ b_{iK} \end{bmatrix} \sim N(\mathbf{0}, D), \quad D = \begin{pmatrix} \sigma^2_{a_1} & \sigma_{a_1 b_1} & \cdots & \sigma_{a_1 b_K} \\ & \ddots & & \\ & & \ddots & \\ & & & \sigma^2_{b_K} \end{pmatrix}$$

$$\begin{bmatrix} \varepsilon_{i1t} \\ \vdots \\ \varepsilon_{iKt} \end{bmatrix} \sim N(\mathbf{0}, E), \quad E = \begin{pmatrix} \sigma_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sigma_K \end{pmatrix}$$

Proteomics

- Association of Evolution
  - Correlation of random intercepts (slopes)

$$\rho^{jk}_{intercept} = \frac{\sigma_{a_j a_k}}{\sqrt{\sigma^2_{a_j}}\sqrt{\sigma^2_{a_k}}} \text{ and } \rho^{jk}_{slope} = \frac{\sigma_{b_j b_k}}{\sqrt{\sigma^2_{b_j}}\sqrt{\sigma^2_{b_k}}}$$

- Evolution of Association
  - Marginal correlation

Challenges and Methods

- When $K = 2$, use R package LME4
- When $K > 2$, fit pairwise models (Fieuws & Verbeke 2006)
- Our proposed approach
  - Estimation for the K nodes simultaneously
  - Sparsity restriction on D

Covariance-adjusted thresholding[7]

- How to estimate D?
- Covariance adjusted thresholding

  (i) $\|h_\lambda(x)\| \leq \|x\|$

  (ii) $\|h_\lambda(x)\| = 0$ for $\|x\| \leq \lambda$

  (iii) $\|h_\lambda(x) - x\| \leq \lambda$
- Apply h_λ to each cell in D
- Select λ via covariance-adjusted process
  - $\lambda_{ij} = \delta\sqrt{\frac{\log p}{n}\hat{\theta}_{ij}}$
  - $\hat{\theta}_{ij} = \frac{1}{n}\sum_{k=1}^{n}\left[(X_{ki} - \bar{X}_i)(X_{kj} - \bar{X}_j) - \hat{\sigma}_{ij}\right]^2$

[7] Cai and Liu (2011)

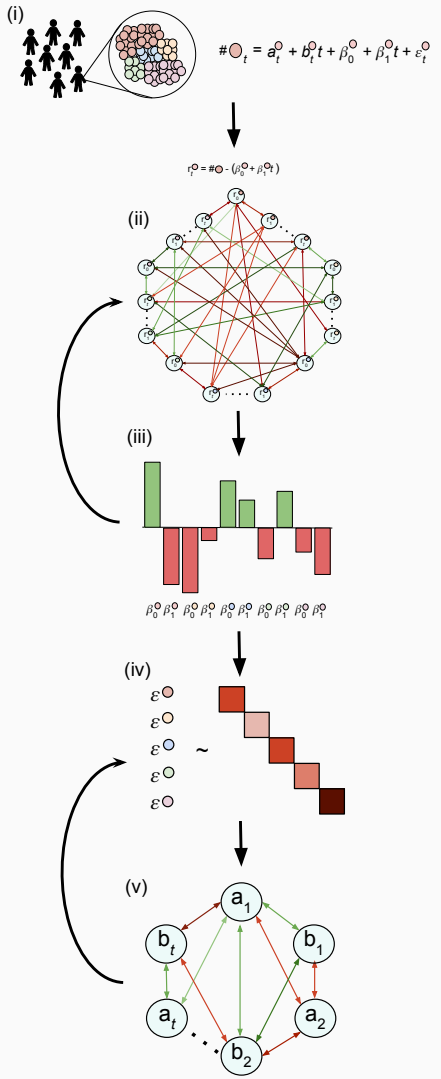
Algorithms

Algorithms

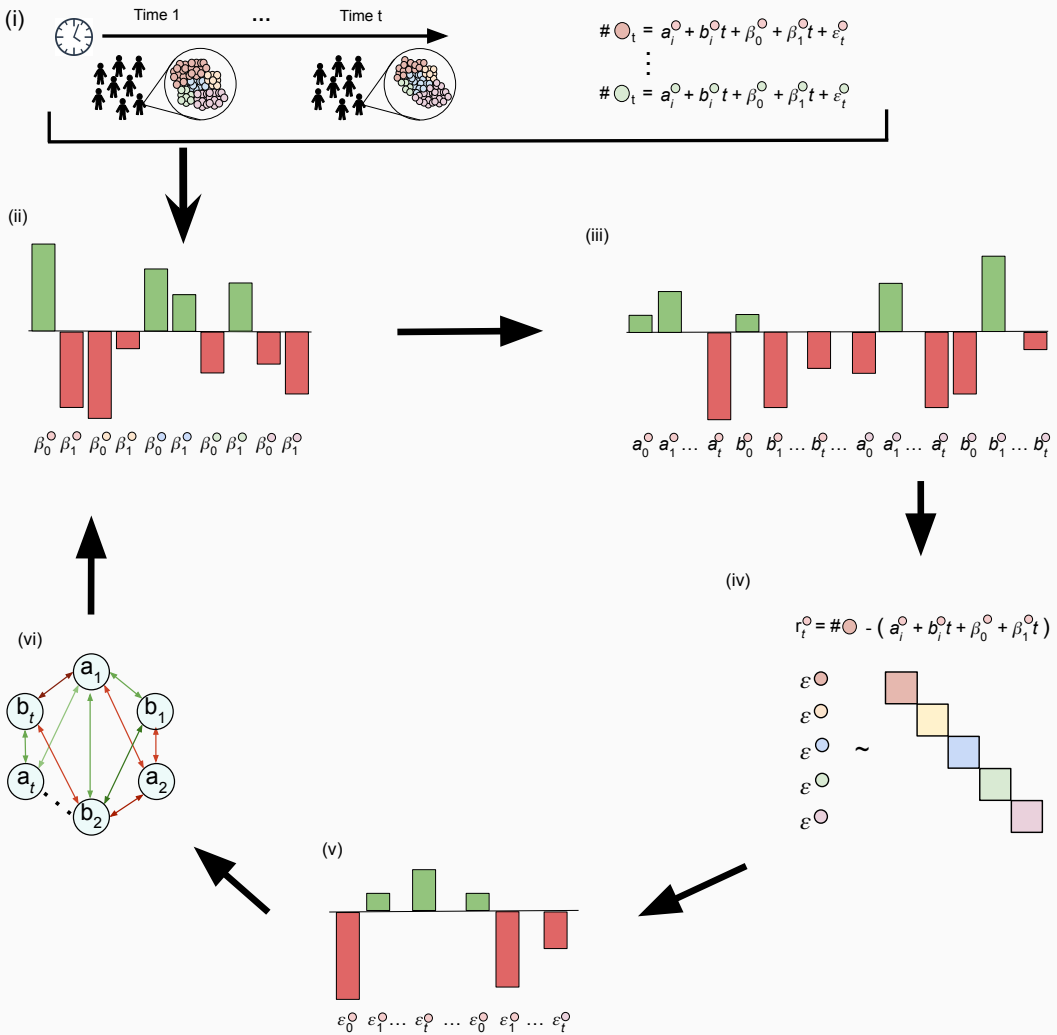
- Estimate D, E, β with Maximum Likelihood
- via Expectation-Maximization type algorithms
- Two algorithms

DINE Algorithm 1

1. Estimate β and V
2. Use V to estimate D and E



DINE Algorithm 2



Simulations

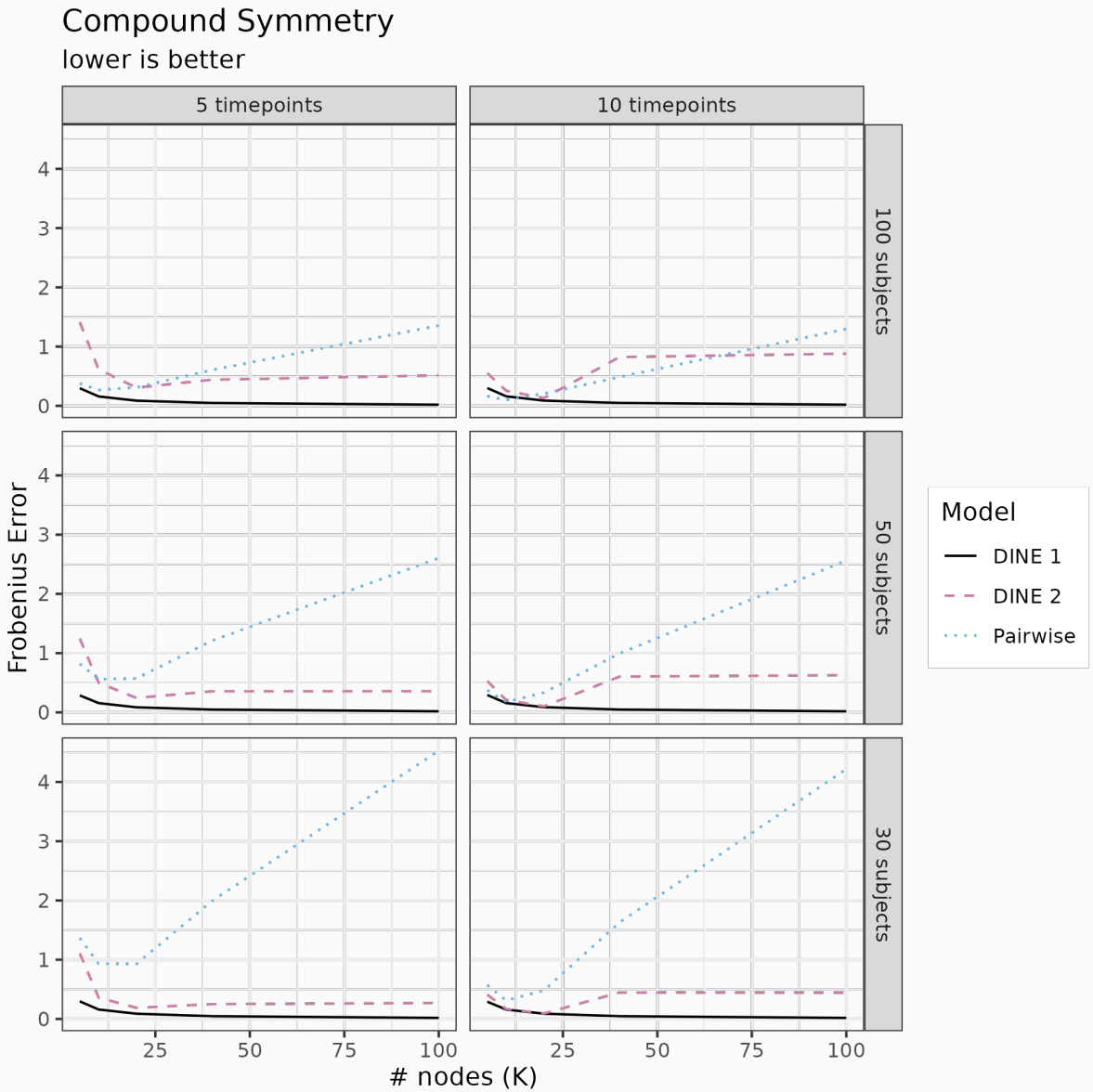
Covariance Types

- Compound symmetry (CS)
  - Constant variance, equal correlation between timepoints per protein
  - Constant cross-protein correlations

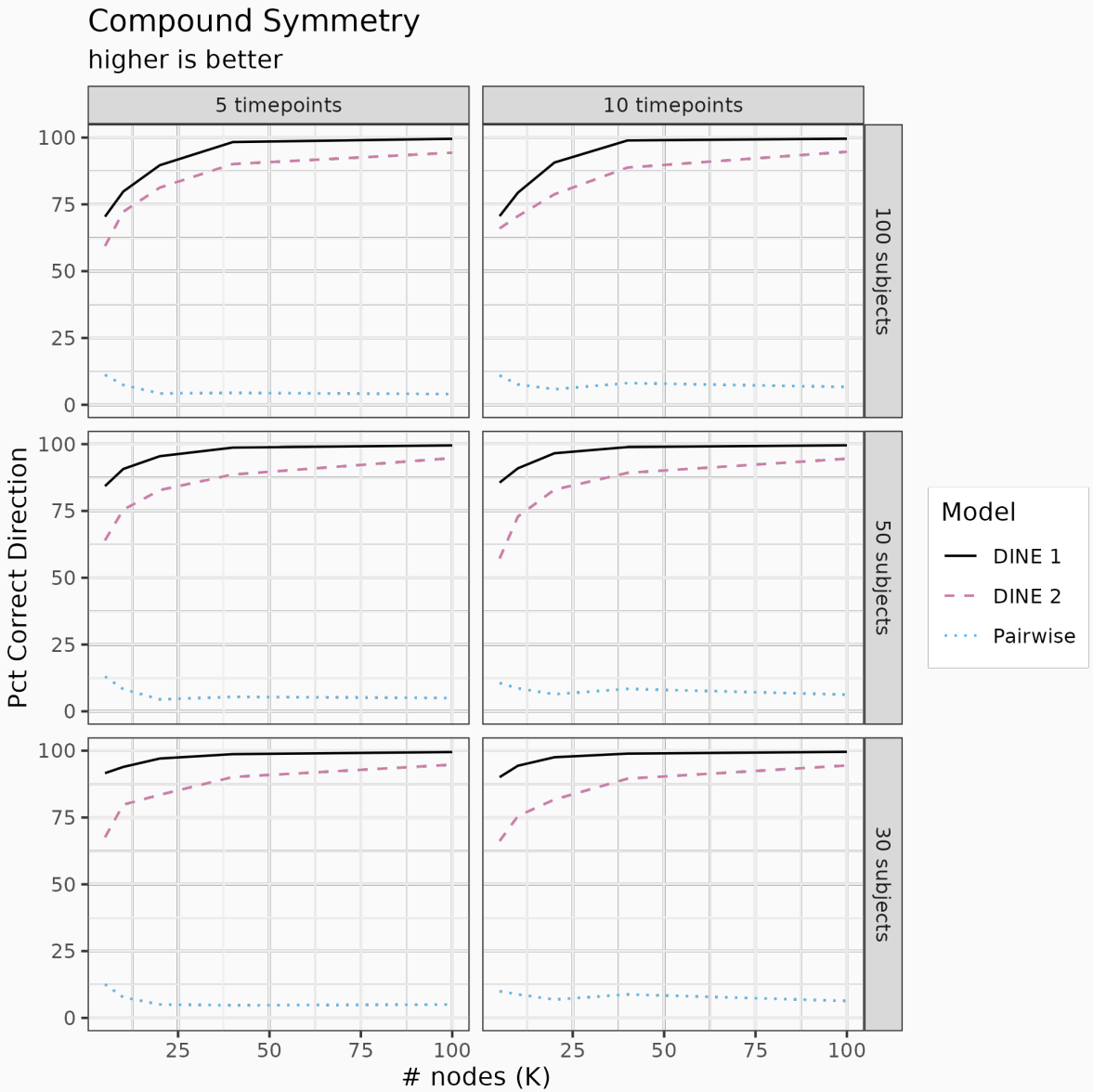
$$D = \begin{pmatrix} \sigma_{a_1}^2 & 0 & \rho_{12}\sigma_{a_1}\sigma_{a_2} & 0 & \dots & \rho_{1K}\sigma_{a_1}\sigma_{a_K} & 0 \\ & 0 & 0 & 0 & \dots & 0 & 0 \\ & & \sigma_{a_2}^2 & 0 & \dots & \rho_{2K}\sigma_{a_2}\sigma_{a_K} & 0 \\ & & & 0 & \dots & 0 & 0 \\ & & & & \ddots & & \vdots \\ & & & & & \sigma_{a_K}^2 & 0 \\ & & & & & & 0 \end{pmatrix}.$$

- $\sigma_{a_j} \sim U(2,3)$ and $\rho_{ij} \sim U(0.2,0.8) \times \text{Rademacher}(0.5)$
- $\sigma_k \sim U(5,10)$
- $\alpha_j \sim N(1,0.2)$ and $\beta_j \sim N(0,0.2)$
- $\mathbf{Y} \sim N(\boldsymbol{\alpha} + \boldsymbol{\beta}[\mathbf{t}], ZDZ^T + \Sigma)$

Normalized Frobenius Error



Directional Recovery

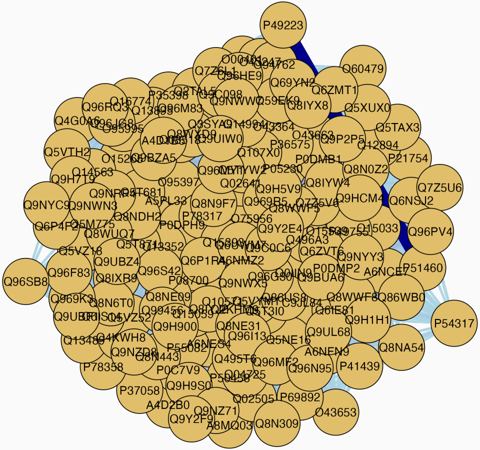


CARDIA Study

CARDIA

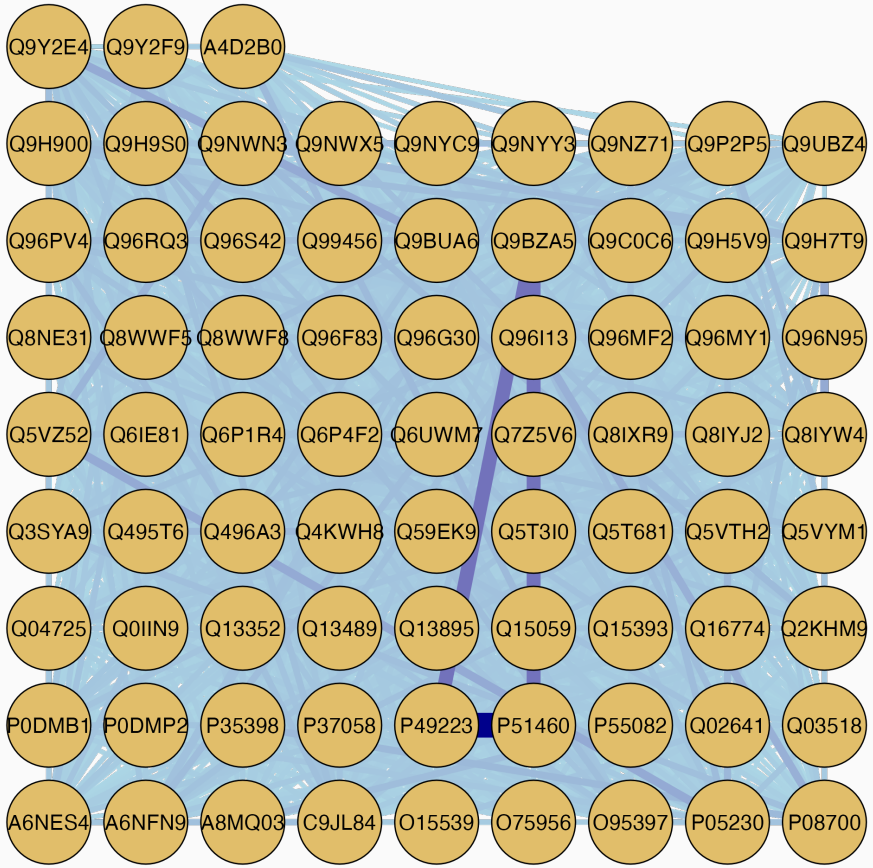
- Long-running study of cardiovascular health (launched in 1985!)
- Four centers: AL, IL, MN, CA
- Protein expression values for 5346 proteins across 15 years (15, 20, 25, 30)
  - 2529 subjects ($>$ 1 timepoint)
- Preprocessing: 150 most variable proteins

Results


P49223
Q60479
Q6ZMT1
Q5XUX0
Q5TAX3
P21754
Q7Z5U6
Q6NSJ2
Q96PV4
P51460
Q86WB0
P54317
Q8NA54
P41439
O43653
Q8N309
Q69YN2
Q8IYX8
Q9P2P5
Q8N0Z2
Q9HCM4
Q9NYY3
Q9H1H1
Q9UL68
Q96N95
P69892
Q96SB8
Q9NYC9
Q5VTH2
Q96F83
P78358
P37058
A4D2B0
Q9Y2F9
Q9NZ71
Q02505
Q9H900
Q8NE31
Q9NWX5
Q96S42
Q8IXR9
Q9UBZ4
Q8WUQ7
Q8NDH2
P78317
Q8N9F7
O75956
Q9Y2E4
Q6P1RA
P0DMP2
Q9BUA6
Q8WWF8
Q6ZVT6
Q496A3
O15033
A6NCE7
Q9H5V9
Q8IYW4
P0DMB1
P36575
Q107X0
O43364
Q9NWW0
Q59EK9
Q9UIW0
Q96M83
O95397
Q9BZA5
Q14563
Q9H7T9
Q9NWN3
Q5M775
Q5VZ18
Q8N6T0
Q4KWH8
Q9NZD8
Q8N443
P0C7V9
Q9H9S0
Q96MF2
Q5NE16
A6NEN9
Q495T6
Q96I13
P08700
Q8NE09
Q02641
P0DPH9

Results

- P49223: Kunitz-type protease inhibitor 3
- P51460: Insulin-like 3 (Leydig insulin-like peptide)
- Q9BZA5: Putative gamma-taxilin 2



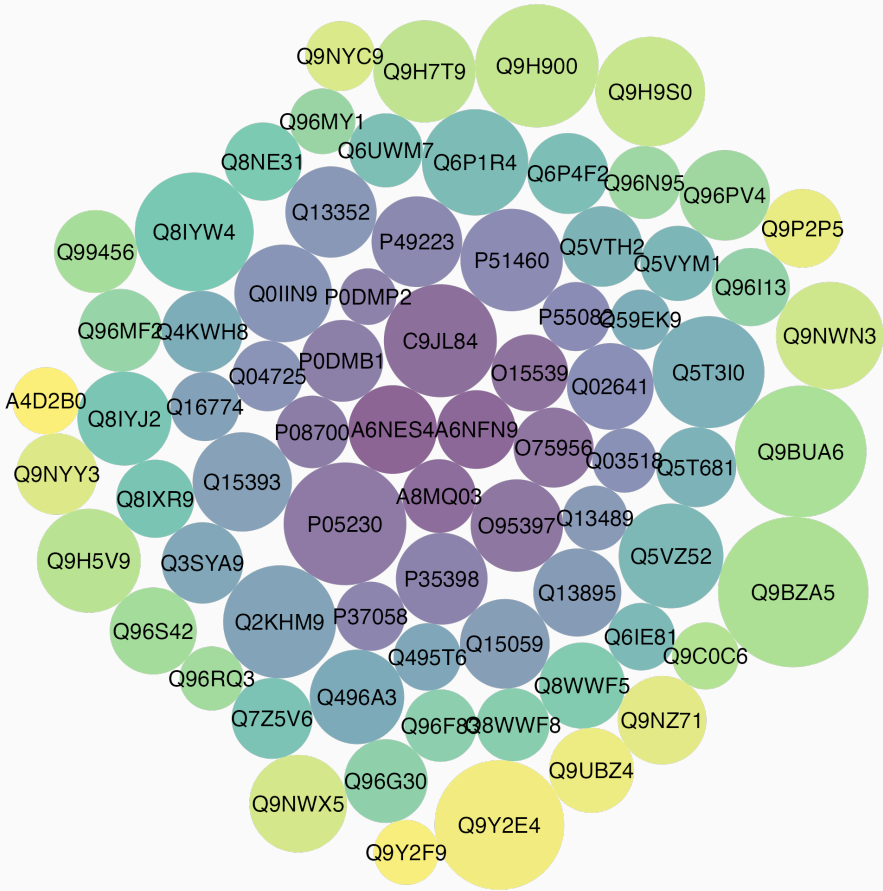
Results

Top Proteins (total edge weight)

1. Q9Y2E4: DIP2 homolog C
2. **Q9BZA5: Putative gamma-taxilin 2**
3. Q9BUA6: Myosin regulatory light chain 10
4. Q0IIN9: ZNF252P antisense gene protein 1
5. Q5T3I0: G patch domain-containing protein 4

33 **P51460: Insulin-like 3 (Leydig insulin-like peptide)**

52 **P49223: Kunitz-type protease inhibitor 3**



Results

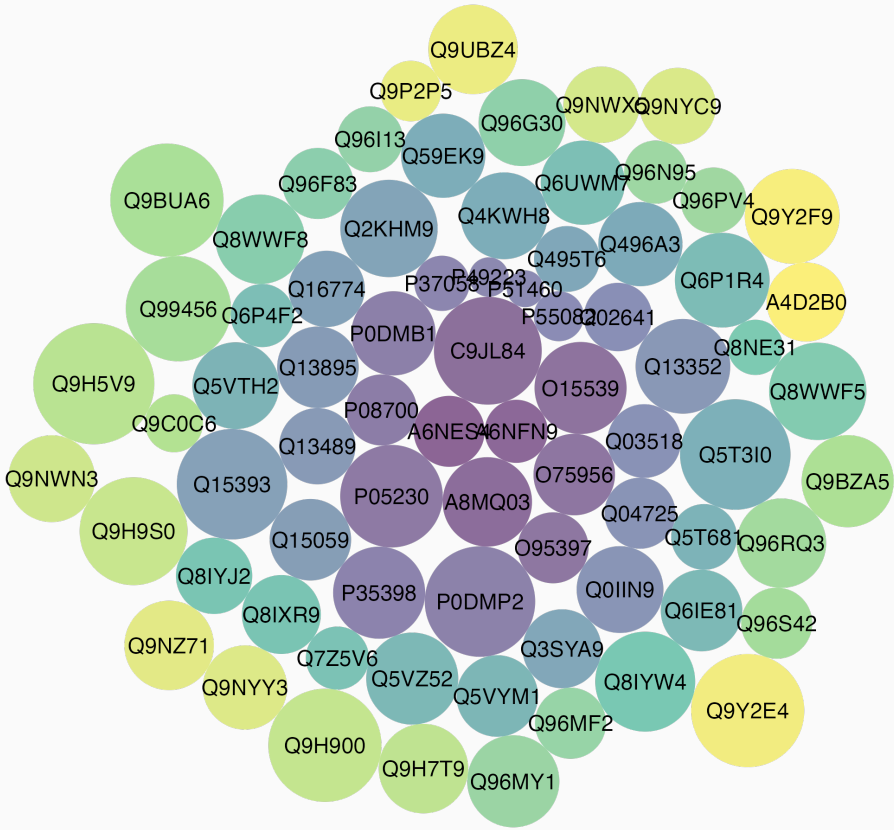
Top Proteins (degree)

1. Q9Y2E4: DIP2 homolog C (98)
2. P0DMP2: SLIT-ROBO Rho GTPase-activating protein 2B (97)
3. Q9H5V9: STING ER exit protein (97)
4. O75956: CDK2-associated protein 2 (96)
5. C9JL84: HERV-H LTR-associating protein 1(95)

25 **Q9BZA5: Putative gamma-taxilin 2** (85)

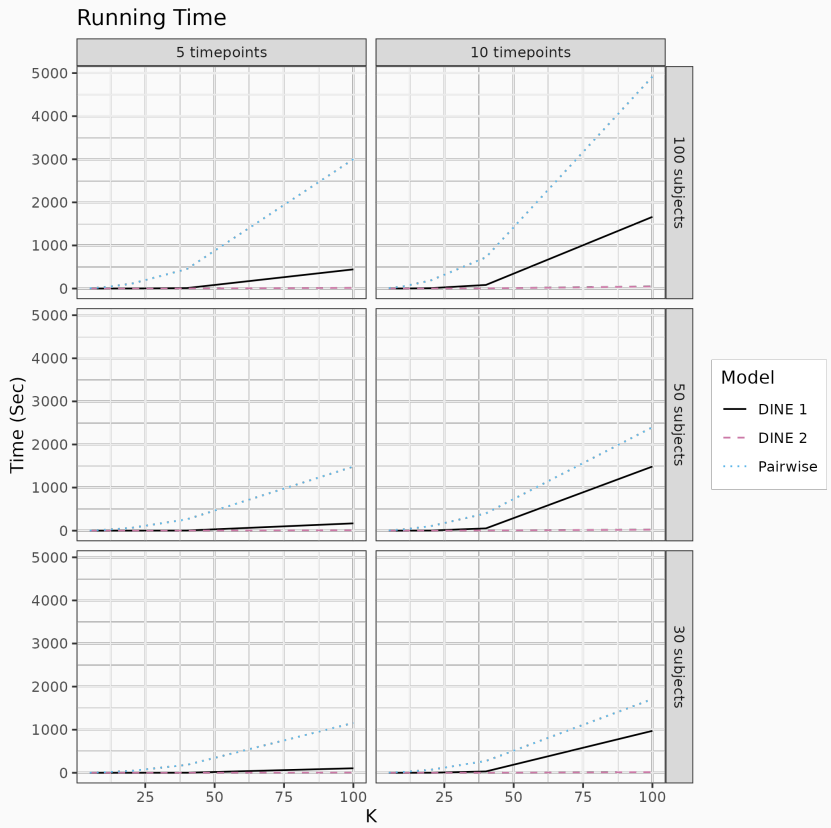
139 **P51460: Insulin-like 3** (45)

147 **P49223: Kunitz-type protease inhibitor 3** (36)



Advantages of DINE methods

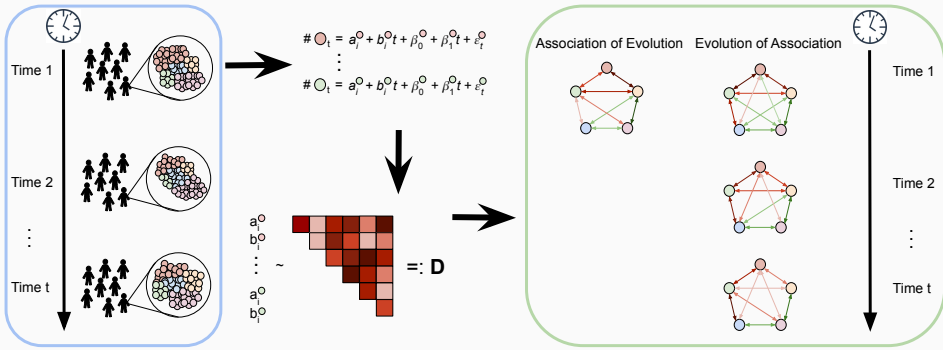
- Simultaneous estimation
- Computational speed
- Estimation accuracy
- Not just for proteins!



Questions?

- Acknowledgements
  - Hongmei Jiang
  - Lifang Hou
- Try it for yourself!

github.com/samozm/DINE



References

1. Fieuws, S. & Verbeke, G. **Pairwise Fitting of Mixed Models for the Joint Modeling of Multivariate Longitudinal Profiles.** *Biometrics* **62,** 424–431. ISSN: 0006-341X (June 2006).
2. Greber, S., Lubec, G., Cairns, N. & Fountoulakis, M. **Decreased Levels of Synaptosomal Associated Protein 25 in the Brain of Patients with Down Syndrome and Alzheimer's Disease.** *Electrophoresis* **20,** 928–934 (1999).
3. Libiger, O. *et al.* **Longitudinal CSF Proteomics Identifies NPTX2 as a Prognostic Biomarker of Alzheimer's Disease.** *Alzheimer's & Dementia* **17,** 1976–1987. ISSN: 1552-5260. https://alz-journals.onlinelibrary.wiley.com/doi/full/10.1002/alz.12353 (2025) (Dec. 2021).

References

4. Tarca, A. L. *et al.* **The Prediction of Early Preeclampsia: Results from a Longitudinal Proteomics Study.** *PLOS ONE* **14,** e0217273. ISSN: 1932-6203. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0217273 (2025) (June 4, 2019).

5. Erez, O. *et al.* **The Prediction of Late-Onset Preeclampsia: Results from a Longitudinal Proteomics Study.** *PLOS ONE* **12,** e0181468. ISSN: 1932-6203. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0181468 (2025) (July 24, 2017).

References

6. Gao, F. *et al.* **Estimating Correlation between Multivariate Longitudinal Data in the Presence of Heterogeneity.** *BMC Medical Research Methodology* **17,** 124. ISSN: 1471-2288 (Dec. 2017).

7. Cai, T. & Liu, W. **Adaptive Thresholding for Sparse Covariance Matrix Estimation.** *Journal of the American Statistical Association* **106,** 672–684. ISSN: 0162-1459. https://doi.org/10.1198/jasa.2011.tm10560 (2024) (June 1, 2011).

Other Simulations

Unstructured D

- No structure on the matrix
- $\sigma_{a_j} \sim U(2.5, 3)$
- $\sigma_{b_j} \sim U(2, 2.5)$
- $\sigma_{a_j b_k} \sim U(0.2, 0.6) \times$ Rademacher(0.5)
- $\sigma_k \sim U(5, 10)$
- $\alpha_j \sim N(1, 0.2)$ and $\beta_j \sim N(0, 0.2)$
- $\mathbf{Y} \sim N(\boldsymbol{\alpha} + \boldsymbol{\beta}[\mathbf{t}], ZDZ^T + \Sigma)$

AR(1)

- Within-variable covariance matrix is

$$\Sigma_v = \begin{pmatrix} \sigma_v^2 & \rho_v\sigma_v & \rho_v^2\sigma_v & \dots & \rho_v^{T-1}\sigma_v \\ & \sigma_v^2 & \rho_v\sigma_v & \dots & \rho_v^{T-2}\sigma_v \\ & & \sigma_v^2 & \dots & \rho_v^{T-3}\sigma_v^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ & & & & \sigma_v^2 \end{pmatrix}.$$

- Between-variable covariance matrix is

$$\Sigma_{vw} = \begin{pmatrix} \rho_{vw}\sigma_v\sigma_w & \rho_v\rho_w\rho_{vw}\sigma_v\sigma_w & \dots & \rho_v^{T-1}\rho_w^{T-1}\rho_{vw}\sigma_v\sigma_w \\ & \rho_{vw}\sigma_v\sigma_w & \dots & \rho_v^{T-2}\rho_w^{T-2}\rho_{vw}\sigma_v\sigma_w \\ & & \ddots & \vdots \\ & & & \rho_{vw}\sigma_v\sigma_w \end{pmatrix}$$

Toeplitz

- within-variable covariance matrix (Σ_v) and between-variable covariance matrix (Σ_{vw}) are

$$\Sigma_v = \begin{pmatrix} \Sigma_v^0 & \mathbf{0} \\ \mathbf{0} & \Sigma_v^1 \end{pmatrix} \qquad \Sigma_{vw} = \begin{pmatrix} \Sigma_{vw}^0 & \mathbf{0} \\ \mathbf{0} & \Sigma_{vw}^1 \end{pmatrix}$$

$$\Sigma_v^0 = \begin{pmatrix} \sigma_v^2 & \rho_v\sigma_v & \frac{\rho_v}{2}\sigma_v & \cdots & \frac{\rho_v}{t^\star}\sigma_v \\ & \sigma_v^2 & \rho_v\sigma_v & \cdots & \frac{\rho_v}{t^\star-1}\sigma_v \\ & & \sigma_v^2 & \cdots & \frac{\rho_v}{t^\star-2}\sigma_v \\ & & & \ddots & \vdots \\ & & & & \sigma_v^2 \end{pmatrix}$$

Toeplitz

$$\Sigma_v^1 = \begin{pmatrix} \sigma_v^2 & 0 & \dots & 0 \\ & \sigma_v^2 & 0 & 0 \\ & & \ddots & 0 \\ & & & \sigma_v^2 \end{pmatrix}$$

$$\Sigma_{vw}^0 = \rho_{vw}\sigma_v\sigma_w \begin{pmatrix} 1 & \rho_v\rho_w & \cdots & \frac{\rho_v}{t^\star-1}\frac{\rho_w}{t^\star-1} \\ & 1 & \cdots & \frac{\rho_v}{t^\star-2}\frac{\rho_w}{t^\star-1} \\ & & \ddots & \vdots \\ & & & 1 \end{pmatrix}$$

$$\Sigma_{vw}^1 = \rho_{vw}\sigma_v\sigma_w \begin{pmatrix} 1 & 0 & \dots & 0 \\ & 1 & 0 & 0 \\ & & \ddots & 0 \\ & & & 1 \end{pmatrix}$$

AR(1), Toeplitz, Overall Unstructured

- First $K^\star = \left[\frac{K}{2}\right]$ nodes connected
- $\mu_i \sim U(0, 10)$, $\sigma_v \sim U(0.8, 1.2)$
- $\rho_v \sim U(0.2, 0.8) \times \text{Rademacher}(0.5)$
- $\rho_{vw} \sim U(0.2, 0.8) \times \text{Rademacher}(0.5)$.
- Draw Y from one of two models
  1. $\mathbf{Y}_i \sim N(\mu_i, \mathbf{R}_i)$
  2. $\mathbf{Y}_i = \mu + \frac{F_i U_k}{\sqrt{10}}$ where $F_i F_i^T = R_i$ and the components of U_k are drawn from $\Gamma(10, 1)$

Algorithms

Joint Modeling Framework[6]

For protein $j = 1, ..., K$ at time $t = 1, .., T$:

$$Y_{1j}(t) = \alpha_j + a_{1j} + (\beta_j + b_{1j})t + \varepsilon_{1jt}$$

$$\vdots$$

$$Y_{Nj}(t) = \alpha_j + a_{Nj} + (\beta_j + b_{Nj})t + \varepsilon_{Njt}$$

[6] Fieuws & Verbeke (2006)

Joint Modeling Framework

$$\begin{bmatrix} a_{i1} \\ b_{i1} \\ \vdots \\ a_{iK} \\ b_{iK} \end{bmatrix} \sim N(\mathbf{0}, D), \quad D = \begin{pmatrix} \sigma^2_{a_1} & \sigma_{a_1 b_1} & \cdots & \sigma_{a_1 b_K} \\ & \ddots & & \\ & & \ddots & \\ & & & \sigma^2_{b_K} \end{pmatrix}$$

$$\begin{bmatrix} \varepsilon_{i1t} \\ \vdots \\ \varepsilon_{iKt} \end{bmatrix} \sim N(\mathbf{0}, E), \quad E = \begin{pmatrix} \sigma_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sigma_K \end{pmatrix}$$

Joint Modeling Framework

$$Y \sim N(X\beta, V)$$
$$V = ZDZ^T + E$$

Algorithm 1

Algorithm 1 Estimation based on full V

Output: D is the random-effect variance-covariance matrix

Output: E is the diagonal error variance matrix

Require: K is number of nodes, T is number of timepoints, N is number of subjects

Require: y is the data, vectorized in a nested order, with observations arranged so that all timepoints for node 1 for subject 1 appear first, then all timepoints for node 2 for subject 1 appear, and this pattern repeats for each subject

Require: Z is the mixed-effects design matrix.

Require: X is the fixed-effects design matrix.

Algorithm 1 - EM Process 1

$V \leftarrow \frac{1}{KT} I_{KT}$

$\beta \leftarrow (X^T X)^{-1} X^T y$

while β and V are not converged **do**

$$V \leftarrow \begin{bmatrix} D & & \\ & \ddots & \\ & & D \end{bmatrix}$$

$\beta \leftarrow (X^T V^{-1} X)^{-1} X^T V^{-1} y$

$r \leftarrow y - X\beta$

$R \leftarrow \text{re-arranged}(r)$ { vector r is re-arranged to be an N x KT matrix}

$D \leftarrow cov(R)$

end while

$D \leftarrow \text{threshold}(R)$ { composition-adjusted thresholding of R}

Algorithm 1 - EM Process 2

while $D^\star$ and Ω are not converged **do**

$D^\star \leftarrow (Z^T Z)^{-1} Z^T (V - \Omega) Z (Z^T Z)^{-1}$

if $D^\star \preceq 0$ **then**

$D^\star \leftarrow D^\star +$ offset

end if

$\Omega \leftarrow \mathrm{diag}(V) - \mathrm{diag}(Z D^\star Z^T)$

end while

$D^\star \leftarrow \mathrm{threshold}(D^\star)$ { threshold $D^\star$ so it has the same proportion of 0s as D}

for $k \in \{1, ..., K\}$ **do**

$E^{(k,k)} \leftarrow \mathrm{mean}(\Omega^{((k-1)T+1,(k-1)T+1)}, ..., \Omega^{(kT,kT)})$

end for

Algorithm 2

Algorithm 2 Estimation based on D and E

while D not converged **do**

$\hat{\beta} = (X^T V^{-1} X)^{-1} X^T V^{-1} y$ for $V = ZDZ^T + E$

$\hat{\boldsymbol{b}} = \Lambda_D(\Lambda_D^T Z^T E^{-T} E^{-1} Z \Lambda_D + I_k)^{-1} \Lambda_D^T Z^T E^{-T} E^{-1}(y - X\beta)$

$\boldsymbol{r_E} = \boldsymbol{y} - \boldsymbol{X}\hat{\beta} - \boldsymbol{Z}\hat{\boldsymbol{b}}$.

$\boldsymbol{R}_{Ei} = \boldsymbol{r}_{Ei}$ ith row of $\boldsymbol{R_E}$ so $\boldsymbol{R_E} \in \mathbb{R}^{nt \times k}$

$E_{ii} = \mathrm{var}(r_i)$

$M := \Lambda_D^{-1}(Z^T Z)^{-1} Z^T$

$\hat{\boldsymbol{\varepsilon}} | \boldsymbol{Y} = \hat{\varepsilon} = \Lambda_{E_{nt}}(\Lambda_{E_{nt}}^T M^T M \Lambda_{E_{nt}} + I_k)^{-1} \Lambda_{E_{nt}}^T M^T M(y - X\beta)$

$\boldsymbol{r} = \boldsymbol{y} - \boldsymbol{X}\hat{\beta} - \hat{\boldsymbol{\varepsilon}}$.

$\boldsymbol{R}_i = (Z_i^T Z_i)^{-1} Z_i^T \boldsymbol{r}_i$ ith row of $\boldsymbol{R}$ so $\boldsymbol{R} \in \mathbb{R}^{n \times 2k}$

$\boldsymbol{D} = \mathit{threshold}(\boldsymbol{R})$ { composition-adjusted thresholding of R}

$\Lambda_D = \mathit{Cholesky}(D + I_{kt})$

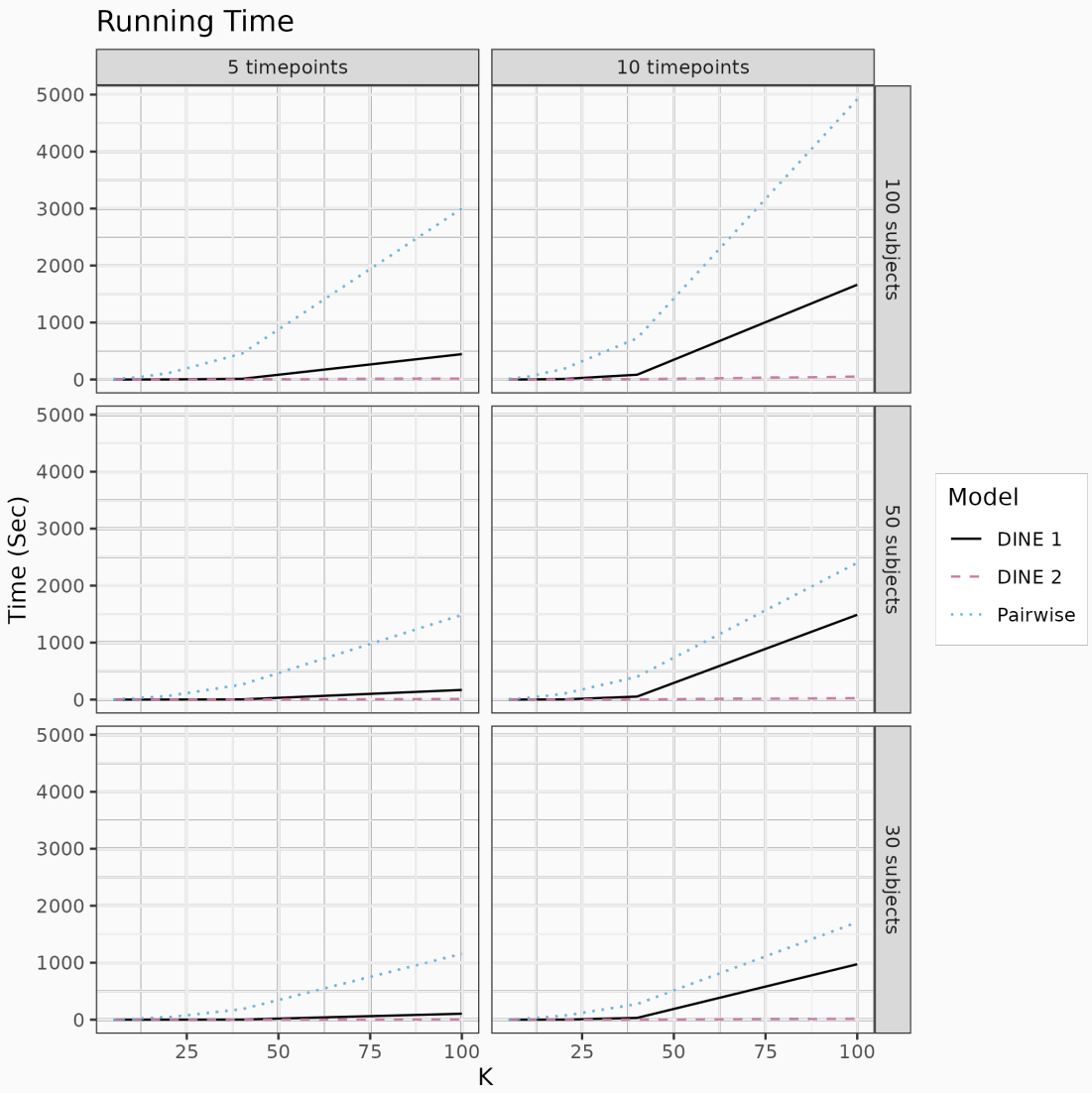
end while

Algorithm 2

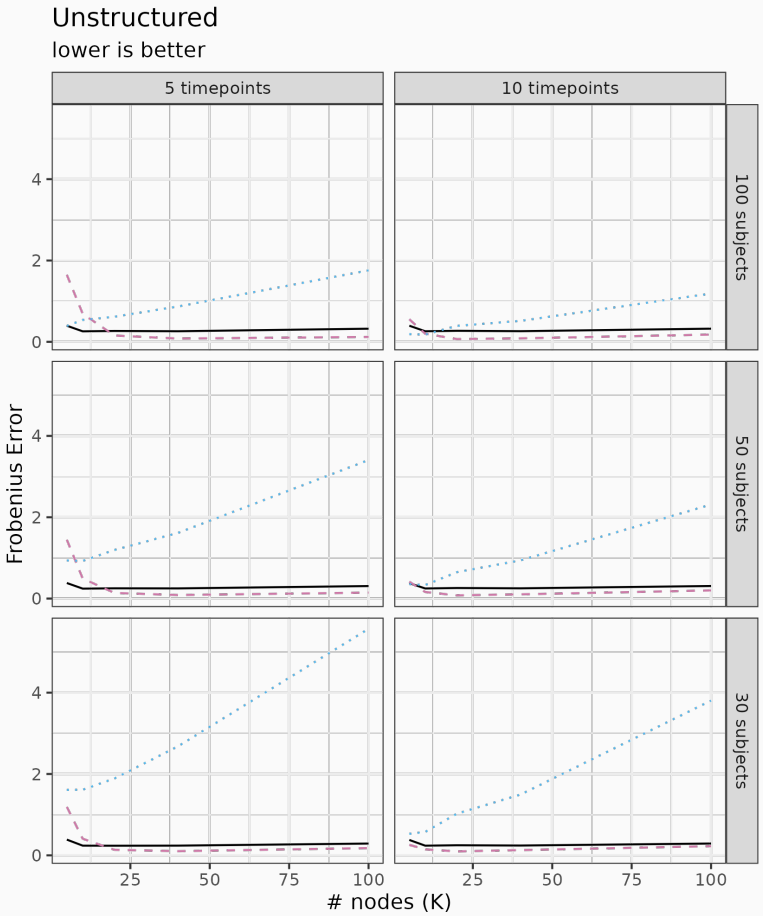
$$\hat{\sigma}^2 = \frac{1}{nkt}\|\Lambda_V^{-1}(y - X\beta)\|^2$$
$$D = \hat{\sigma}^2 \Lambda_D \Lambda_D^T$$
$$E = \hat{\sigma}^2 \Lambda_E \Lambda_E^T$$

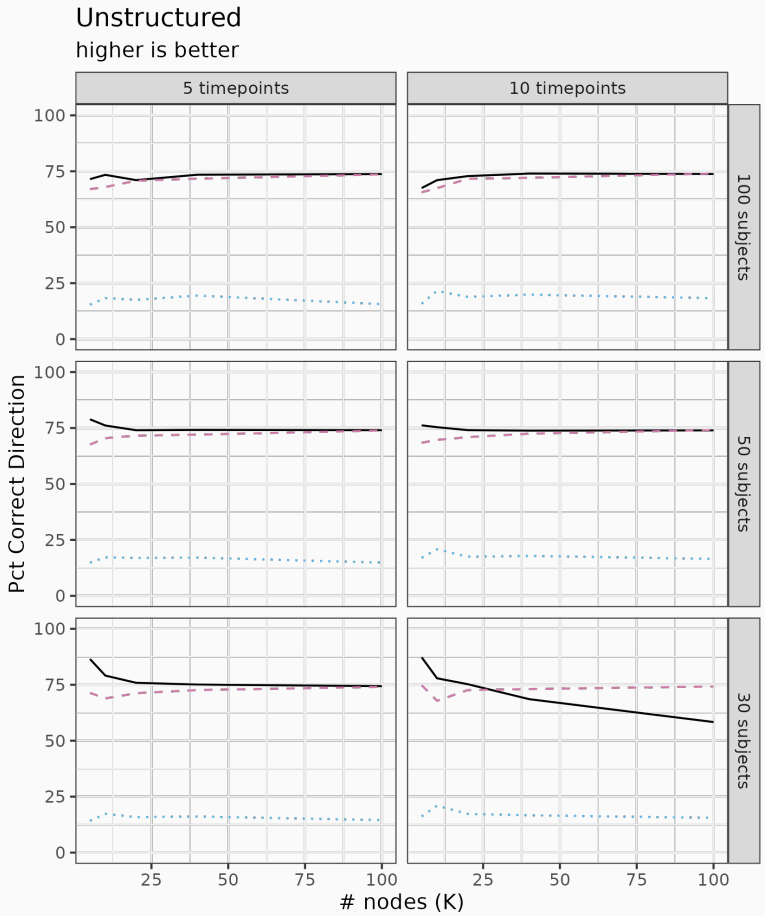
Additional Results

Computational Time

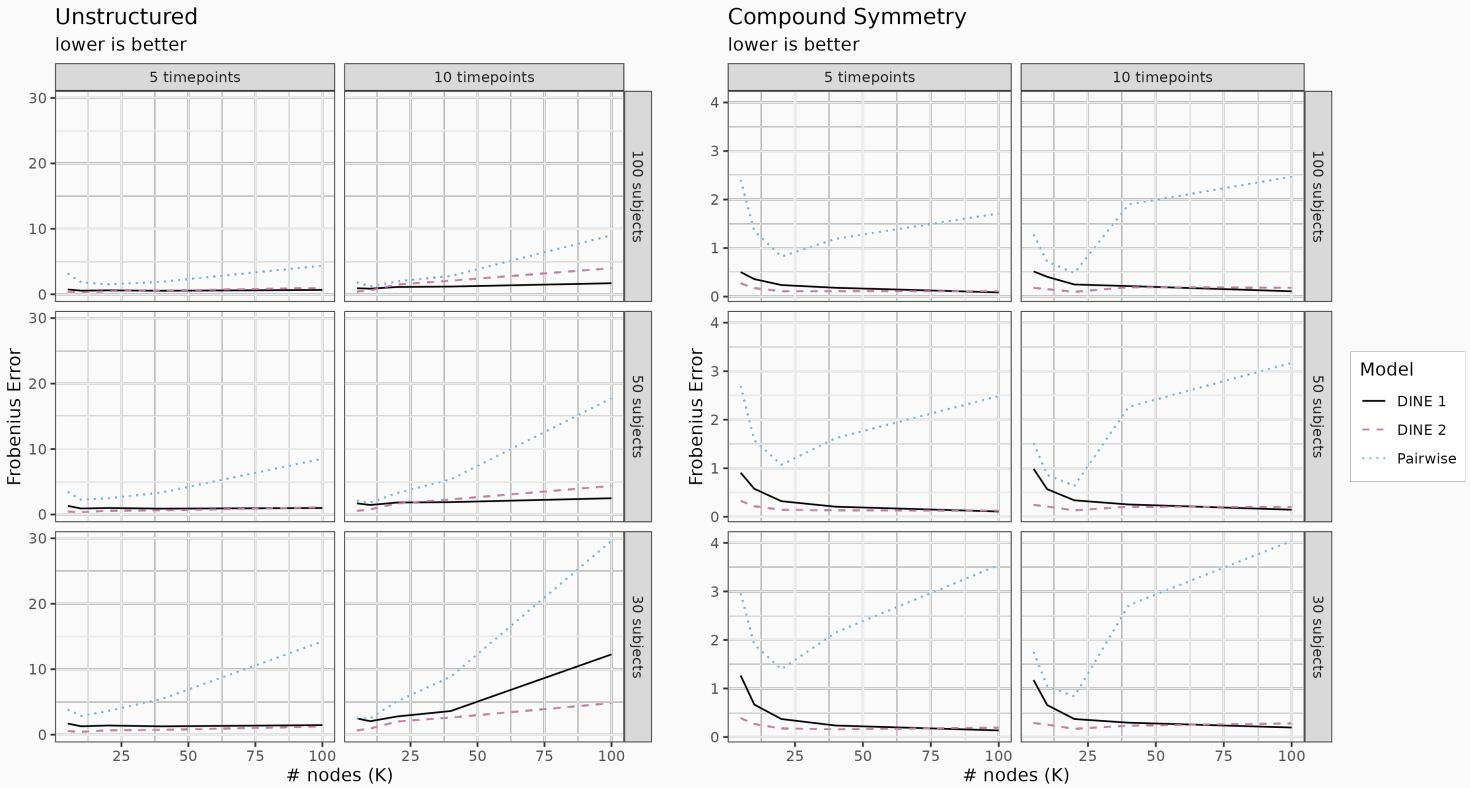


Unstructured D

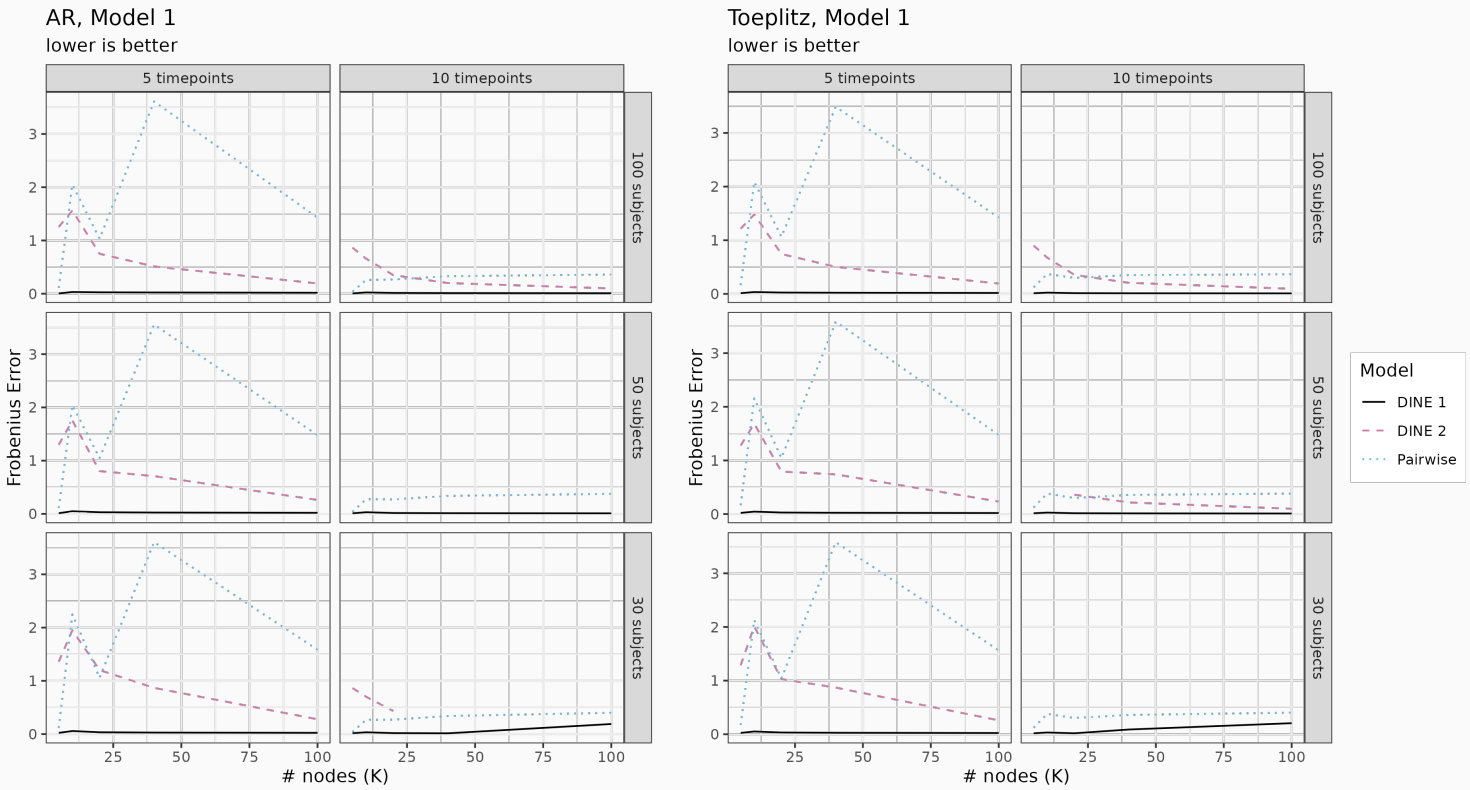




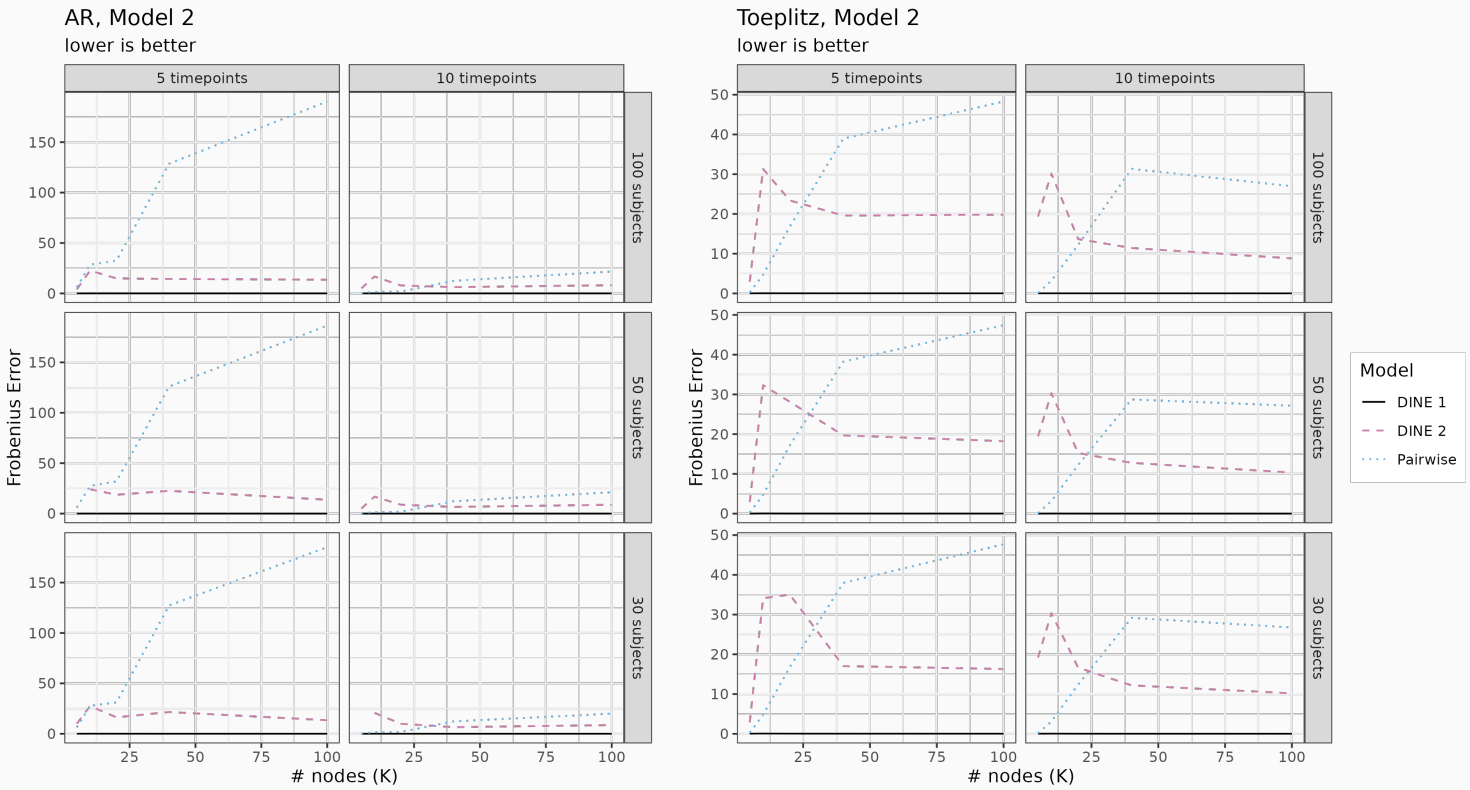
Overall Covariance



Overall Covariance



Overall Covariance



	$\boldsymbol{\beta}^{0.9}$	...	$\boldsymbol{\beta}^{0.1}$
Sim 1	$\beta_1^{1.0} - \beta_1^{0.9}$	...	$\beta_1^{1.0} - \beta_1^{0.1}$
	$\vdots$		$\vdots$
Sim 500	$\beta_{500}^{1.0} - \beta_{500}^{0.9}$	...	$\beta_{500}^{1.0} - \beta_{500}^{0.1}$

Then we have 500 estimates of each of $\beta^{0.9}$, $\beta^{0.8}$, etc. from which to estimate intervals. But the way we are currently doing this is we generate one full dataset for which we calculate $\beta^{1.0}$. So we would have

	$\boldsymbol{\beta}^{0.9}$	...	$\boldsymbol{\beta}^{0.1}$
Sim 1	$\beta^{1.0} - \beta_1^{0.9}$	...	$\beta^{1.0} - \beta_1^{0.1}$
	$\vdots$		$\vdots$
Sim 500	$\beta^{1.0} - \beta_{500}^{0.9}$	...	$\beta^{1.0} - \beta_{500}^{0.1}$

So should we actually be generating a new dataset for every simulation? Or am I misunderstanding how to do the pairing?