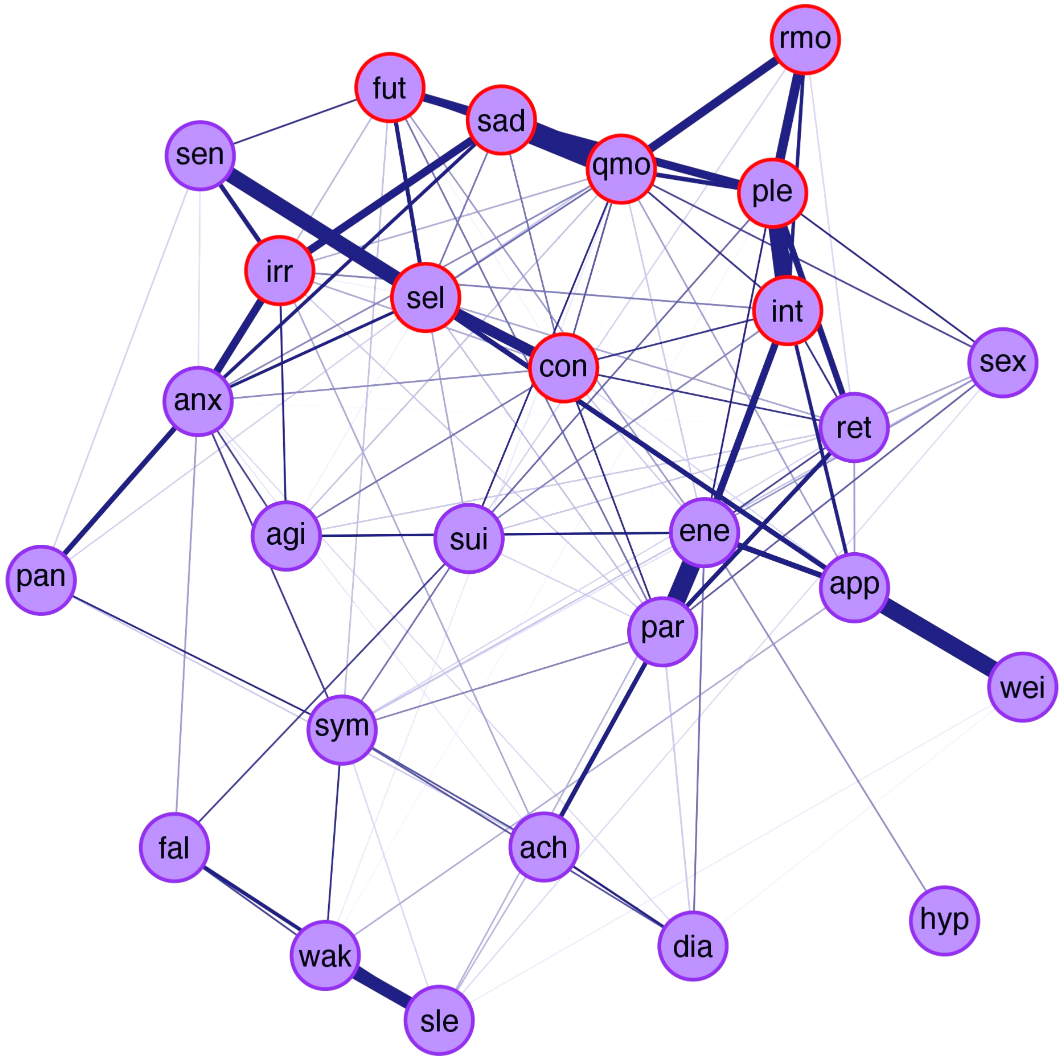
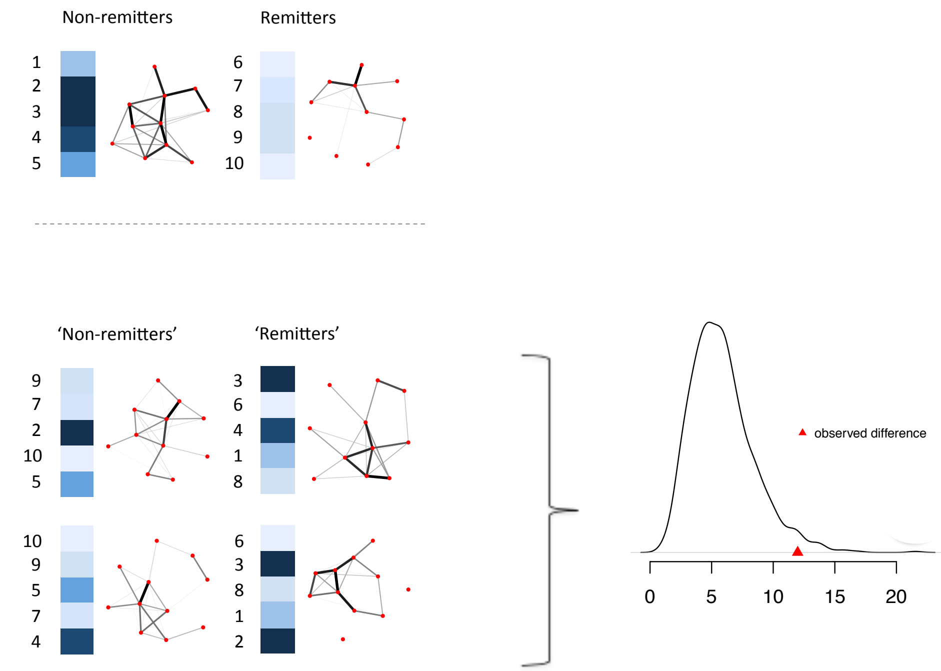
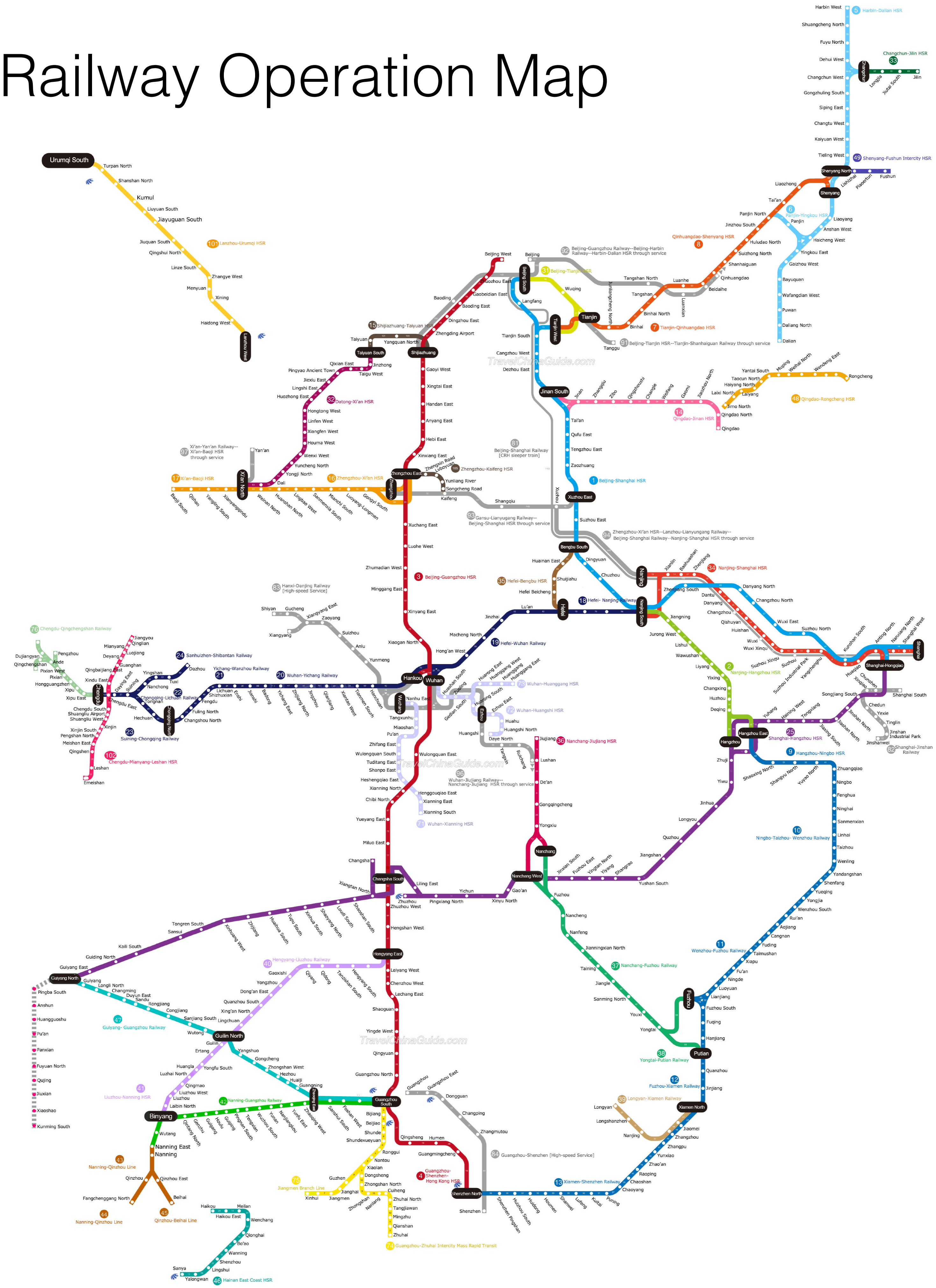




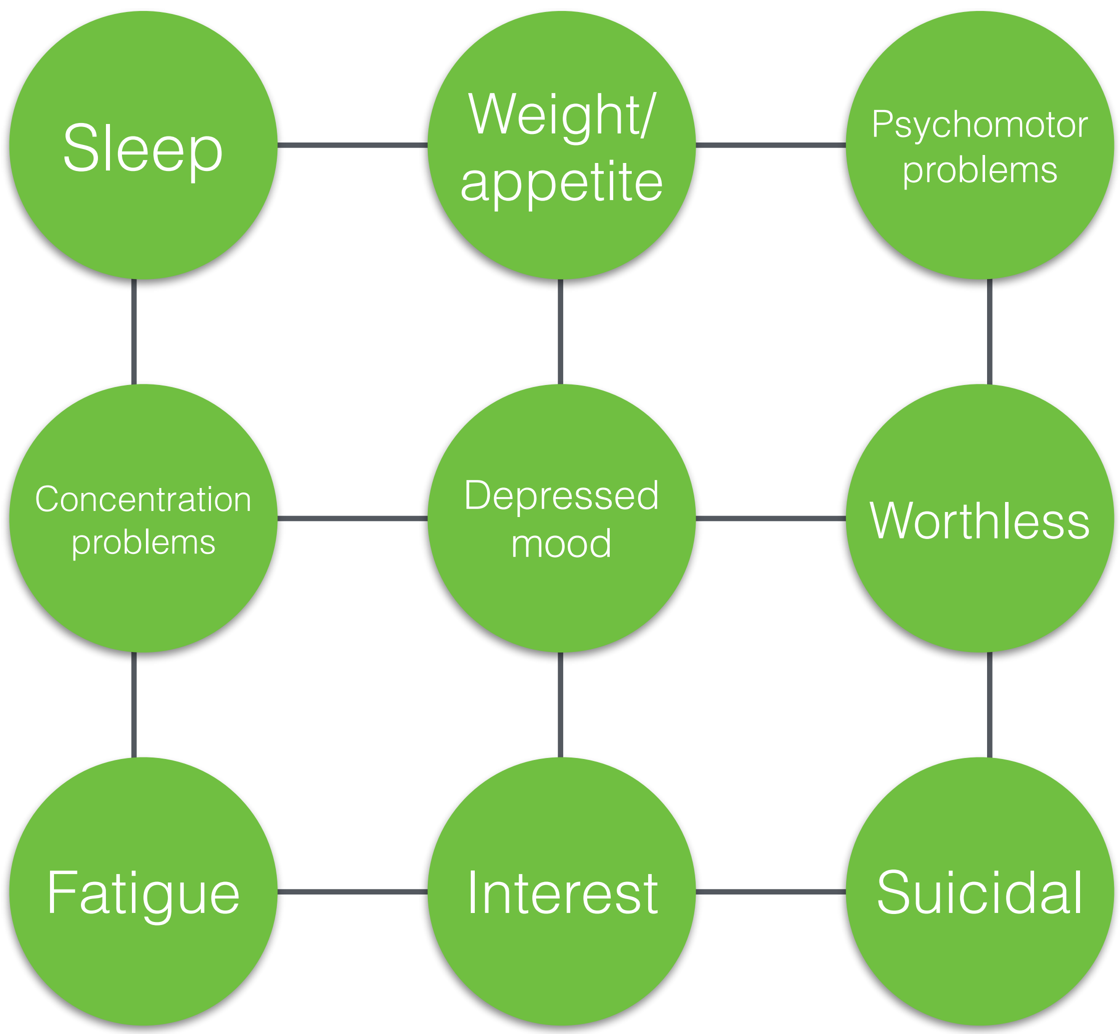
Constructing networks from binary data

Claudia van Borkulo
University Medical Center Groningen - University of Amsterdam
Promotors: R.A. Schoevers, D. Borsboom

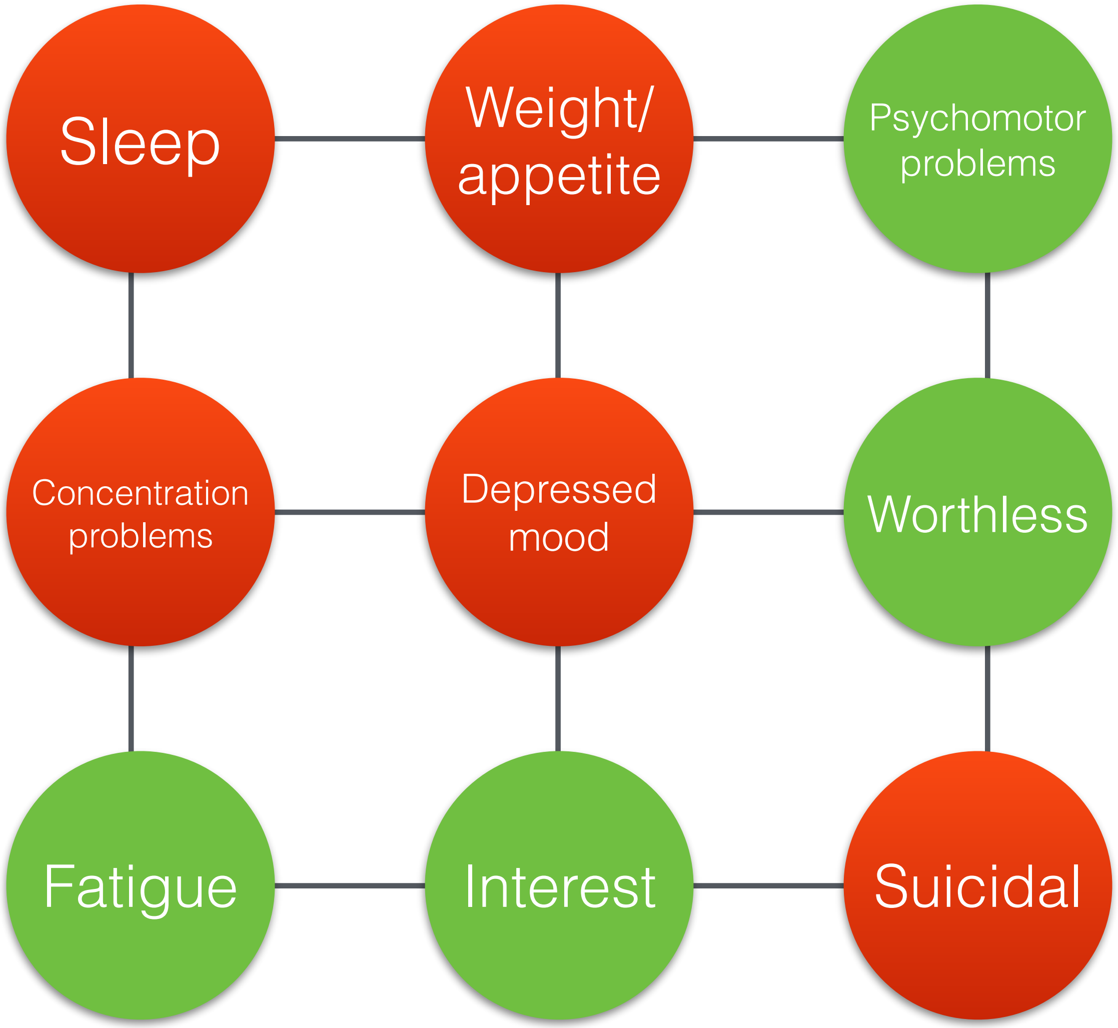
China Speed Railway Operation Map



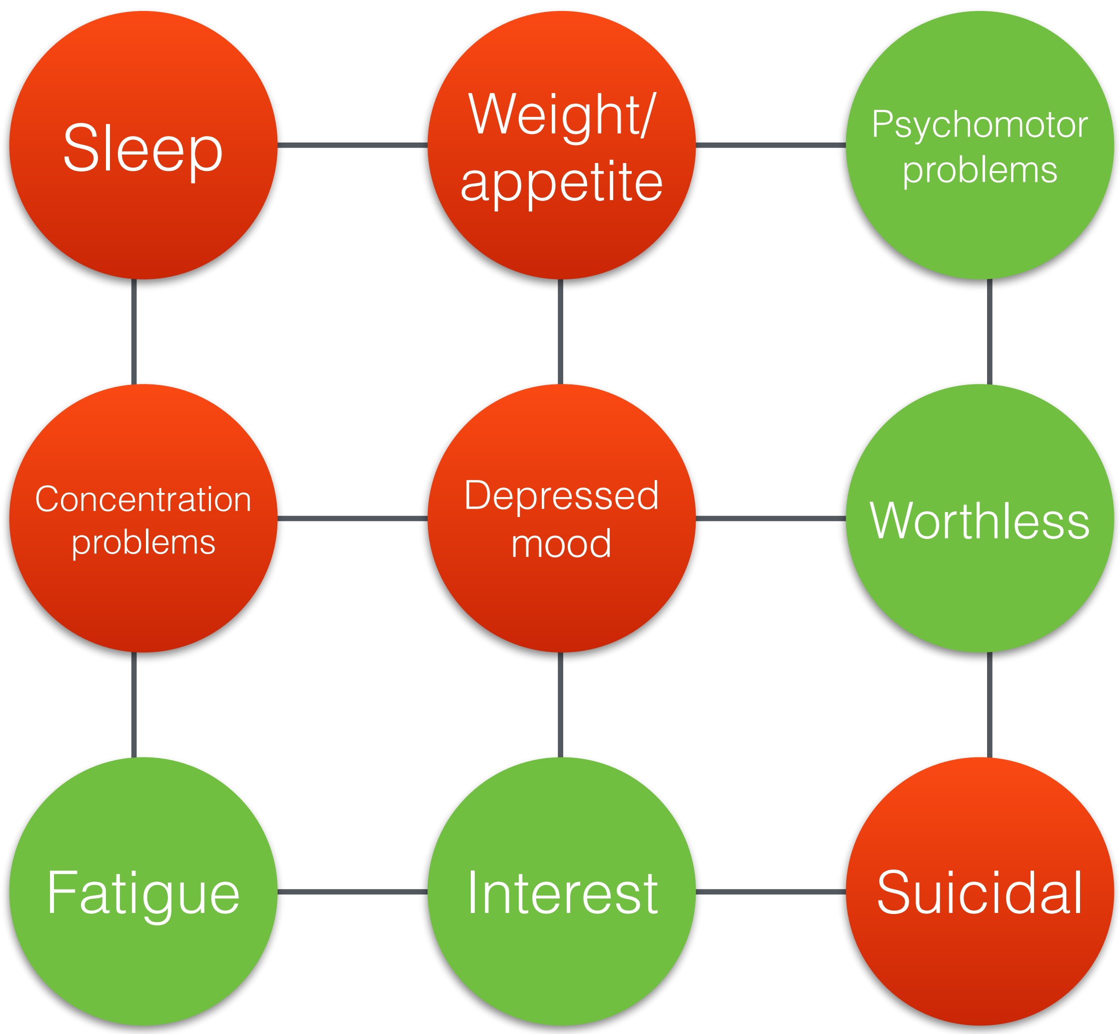
Psychopathology

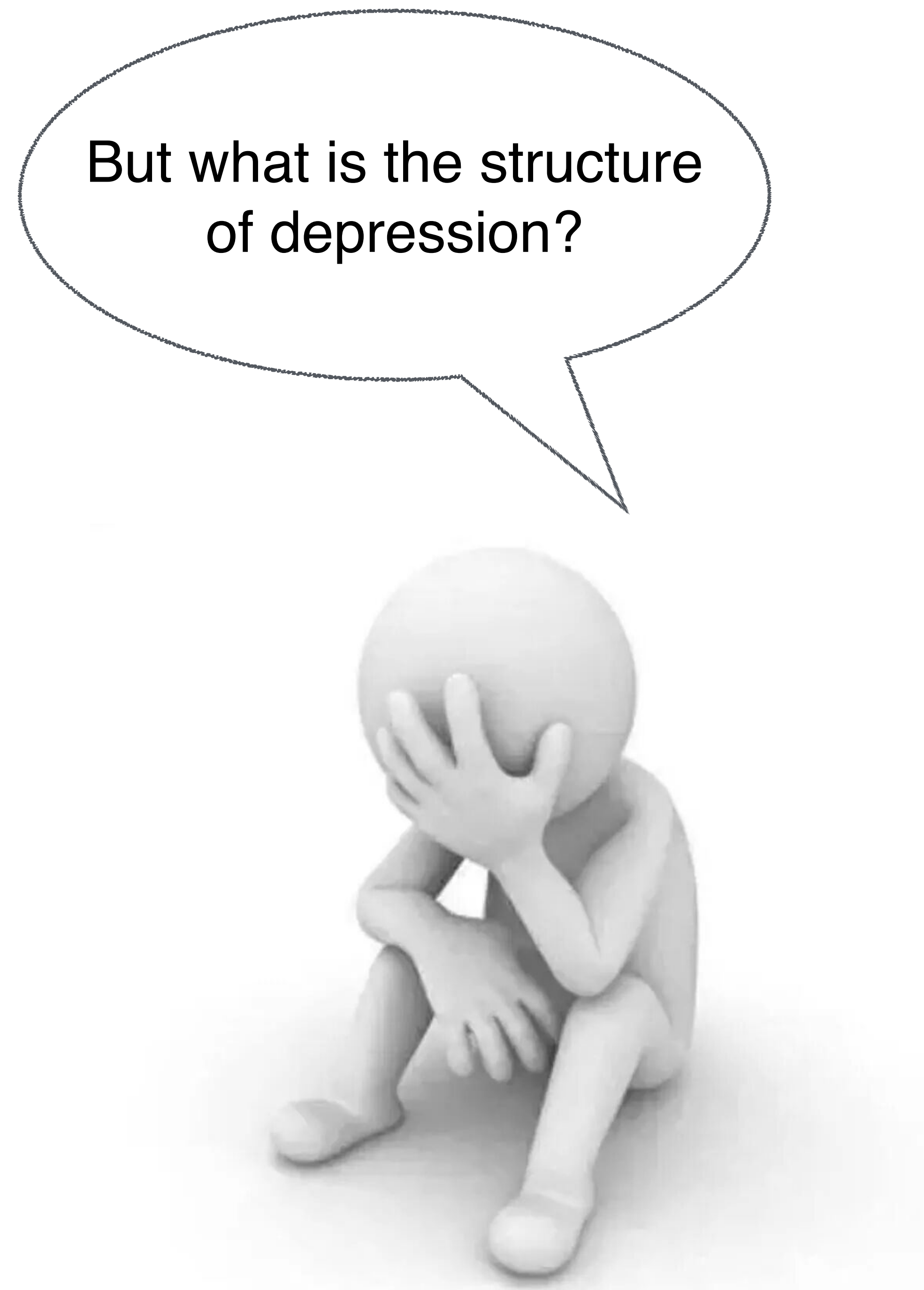


Psychopathology



Psychopathology

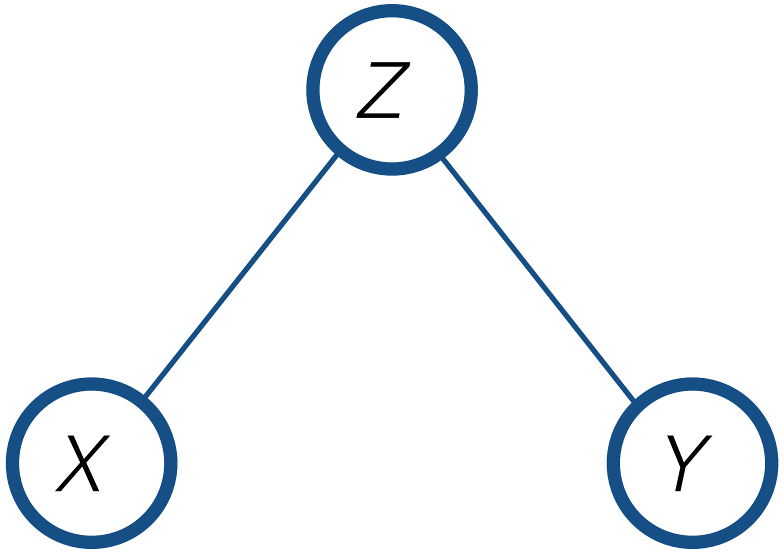




Gaussian data

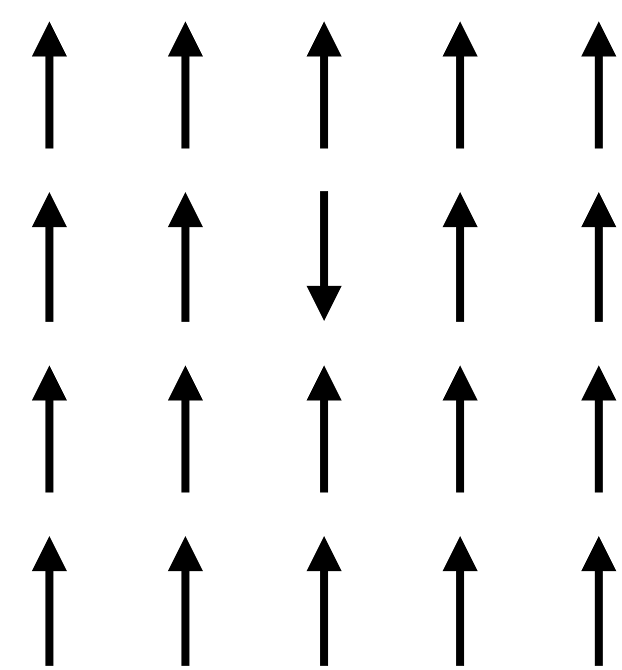
Inverse covariance matrix

- representation of a Markov Random Field (undirected network)
- Zero entry: conditional independence between two variables (given all other variables)
- $X \perp\!\!\!\perp Y \mid Z$: X and Y are not connected



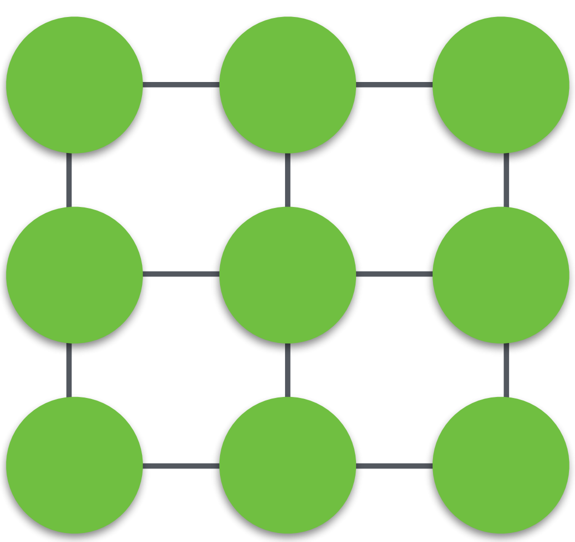
Binary data

- Discrete MRF are intractable
- We propose a computationally efficient method for binary data:
  - Ising model
  - l1-regularized logistic regression



Ising model

$$\mathbb{P}_\Theta\left(x_j \middle| x_{\setminus j}\right) = \frac{\exp\left[\tau_j x_j + x_j \sum_{k \in V_{\setminus j}} \beta_{jk} x_k\right]}{1 + \exp\left[\tau_j + \sum_{k \in V_{\setminus j}} \beta_{jk} x_k\right]}$$

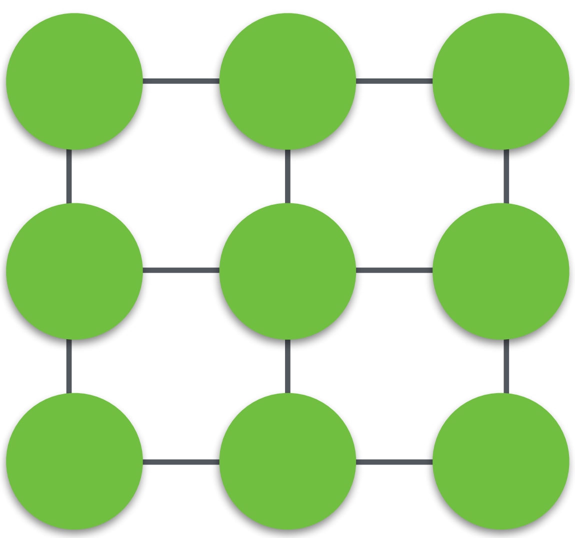


- x = (x1, x2, …,xn)
- x_j = 0 or 1
- τ_j: node parameter (threshold)
- β_{jk}: pairwise interaction parameter

Ising model

Conditional probability

$$\mathbb{P}_{\Theta}(x_j | x_{\setminus j}) = \frac{\exp\left[\tau_j x_j + x_j \sum_{k \in V_{\setminus j}} \beta_{jk} x_k\right]}{1 + \exp\left[\tau_j + \sum_{k \in V_{\setminus j}} \beta_{jk} x_k\right]}$$

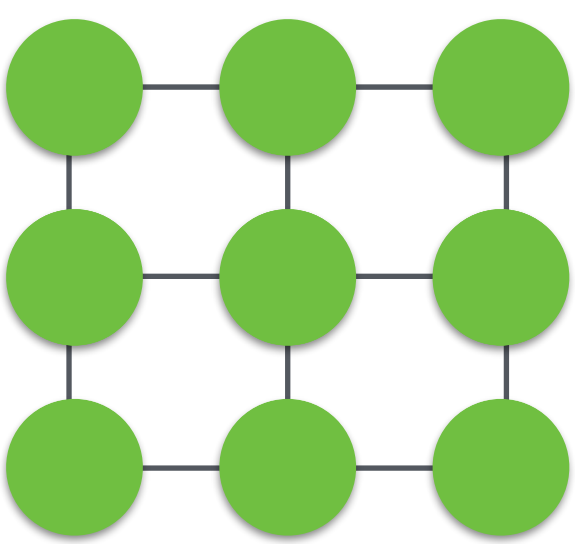


- x = (x1, x2, …,xn)
- x_j = 0 or 1
- τ_j: node parameter (threshold)
- β_{jk}: pairwise interaction parameter

Ising model

Conditional probability

$$\mathbb{P}_\Theta\left(x_j \middle| x_{\setminus j}\right) = \frac{\exp\left[\tau_j x_j + x_j \sum_{k \in V_{\setminus j}} \beta_{jk} x_k\right]}{1 + \exp\left[\tau_j + \sum_{k \in V_{\setminus j}} \beta_{jk} x_k\right]}$$



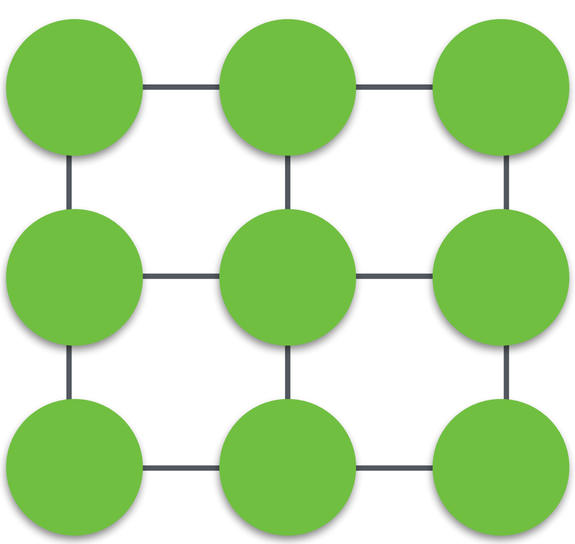
- $x = (x1, x2, \ldots, xn)$
- $x_j = 0$ or 1
- τ_j: node parameter (threshold)
- β_{jk}: pairwise interaction parameter

Ising model

Conditional probability

$$\mathbb{P}_\Theta(x_j | x_{\setminus j}) = \frac{\exp\left[\tau_j x_j + x_j \sum_{k \in V_{\setminus j}} \beta_{jk} x_k\right]}{1 + \exp\left[\tau_j + \sum_{k \in V_{\setminus j}} \beta_{jk} x_k\right]}$$

$\tau_1 + \beta_{12} X_2 + \beta_{13} X_3 + \ldots$



- x = (x1, x2, …,xn)
- x_j = 0 or 1
- τ_j: node parameter (threshold)
- β_{jk}: pairwise interaction parameter

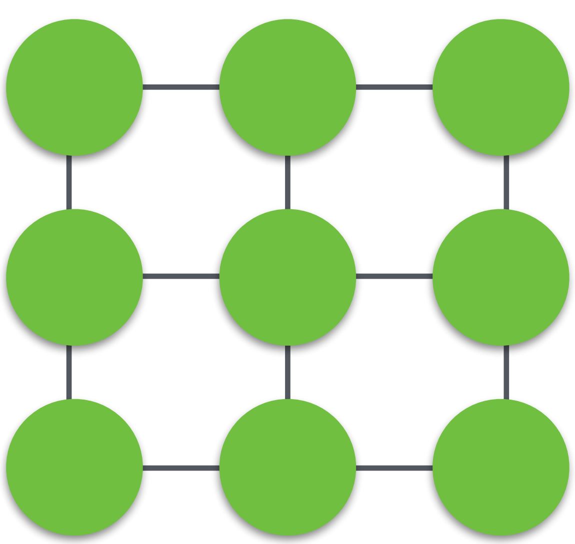
Ising model

Autonomous disposition of x_j

Conditional probability

$$\mathbb{P}_{\Theta}(x_j | x_{\setminus j}) = \frac{\exp\left[\tau_j x_j + x_j \sum_{k \in V_{\setminus j}} \beta_{jk} x_k\right]}{1 + \exp\left[\tau_j + \sum_{k \in V_{\setminus j}} \beta_{jk} x_k\right]}$$

$\tau_1 + \beta_{12}X_2 + \beta_{13}X_3 + \ldots$



- x = (x1, x2, …,xn)
- x_j = 0 or 1
- τ_j: node parameter (threshold)
- β_{jk}: pairwise interaction parameter

Ising model

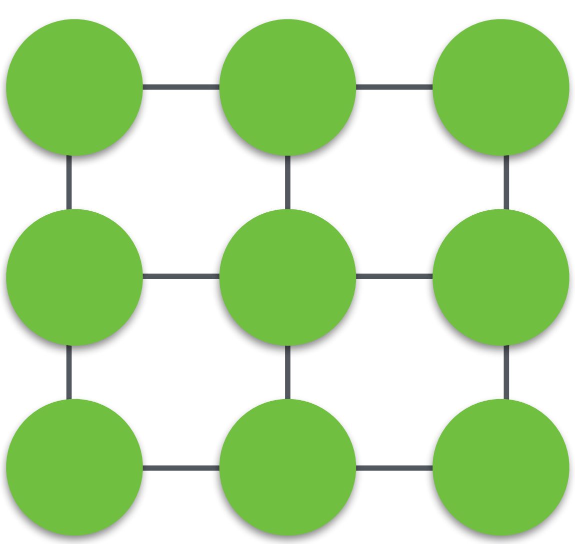
Interaction strength between x_j and x_k

Autonomous disposition of x_j

Conditional probability

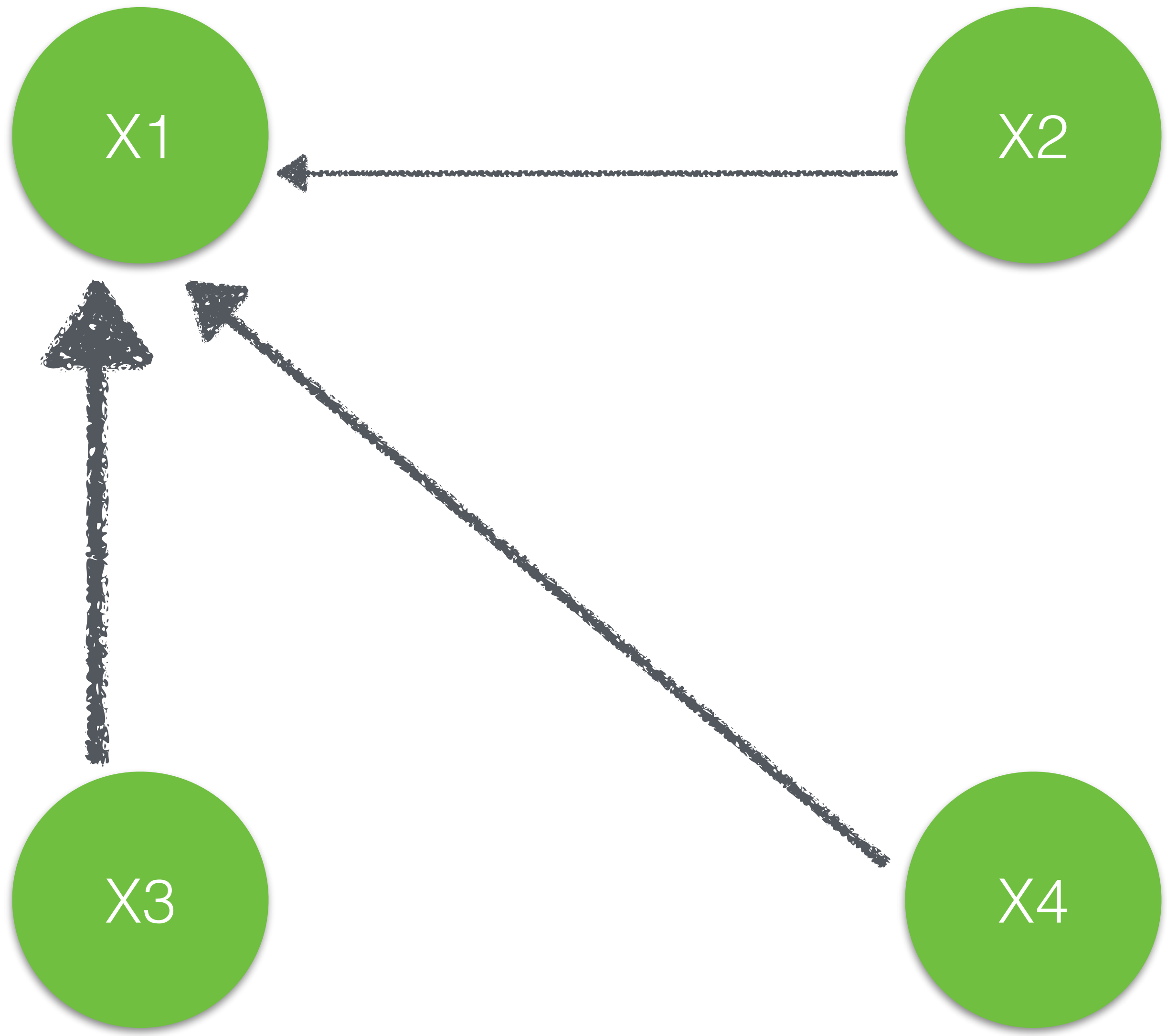
$$\mathbb{P}_\Theta(x_j|x_{\setminus j}) = \frac{\exp\left[\tau_j x_j + x_j \sum_{k\in V_{\setminus j}} \beta_{jk} x_k\right]}{1+\exp\left[\tau_j + \sum_{k \in V_{\setminus j}} \beta_{jk} x_k\right]}$$

$\tau_1 + \beta_{12}X_2 + \beta_{13}X_3 + \ldots$



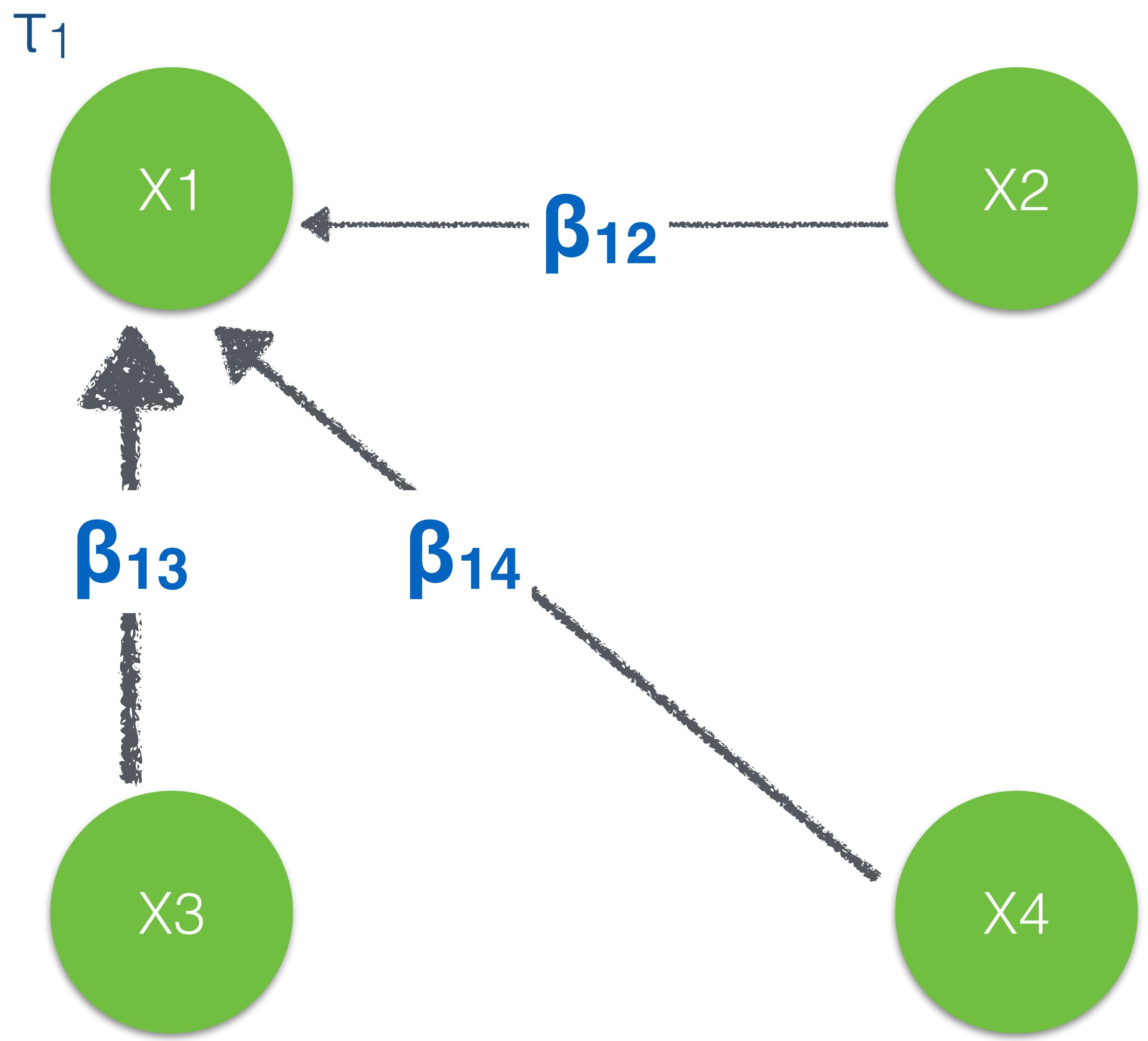
- $x = (x1, x2, \ldots, xn)$
- x_j = 0 or 1
- τ_j: node parameter (threshold)
- β_{jk}: pairwise interaction parameter

Basic idea

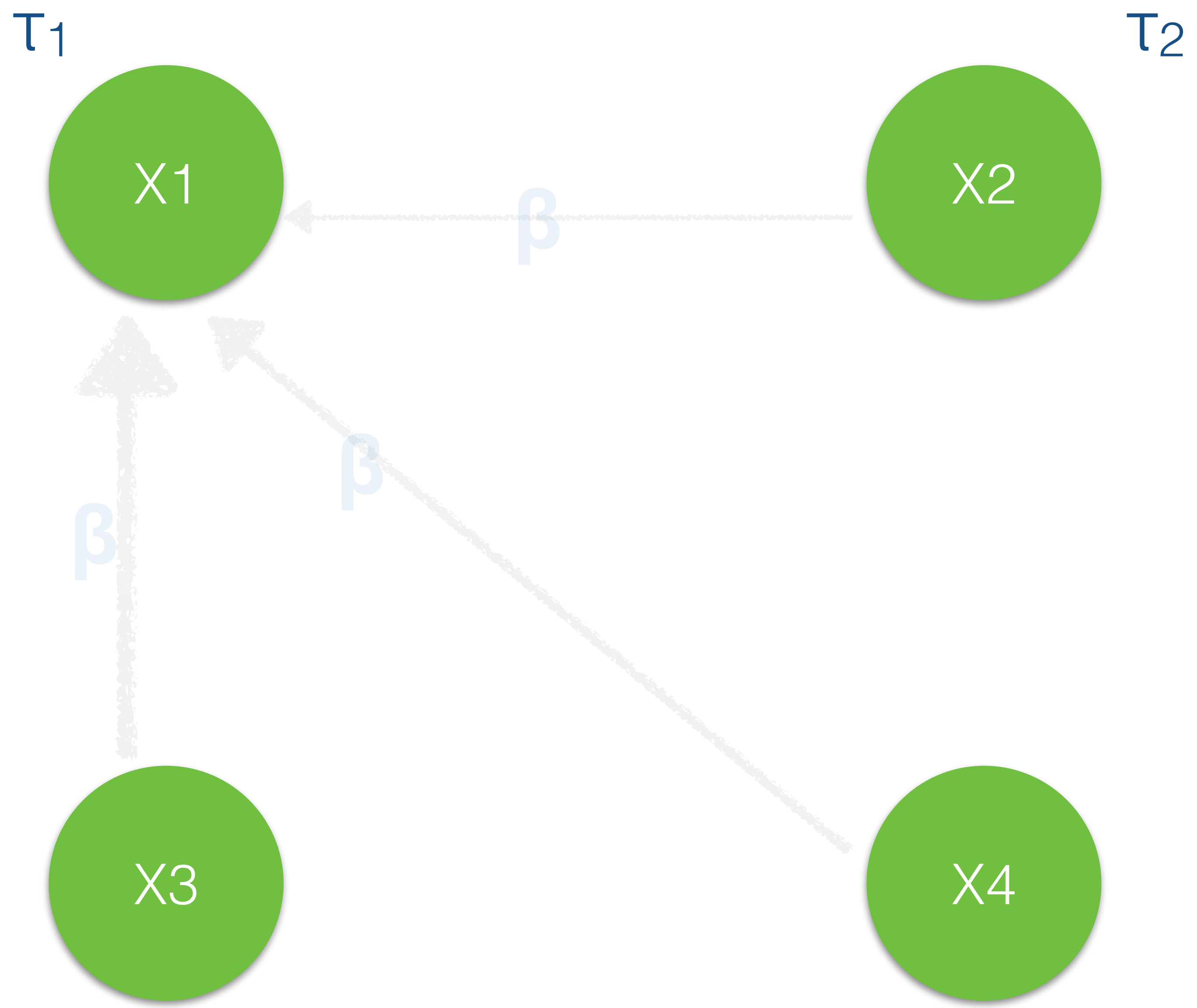


Perform regression of X_1 on all other variables

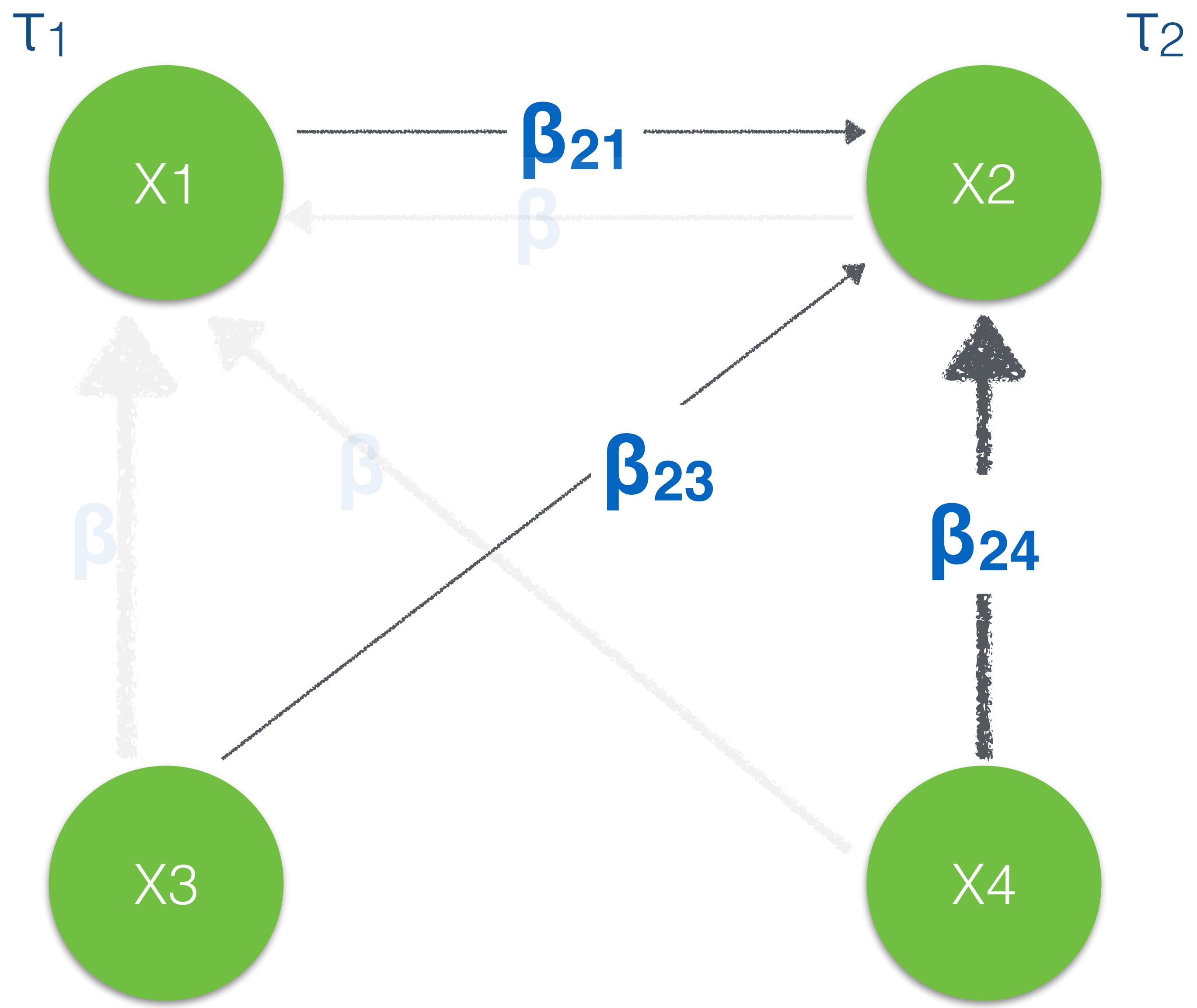
Basic idea



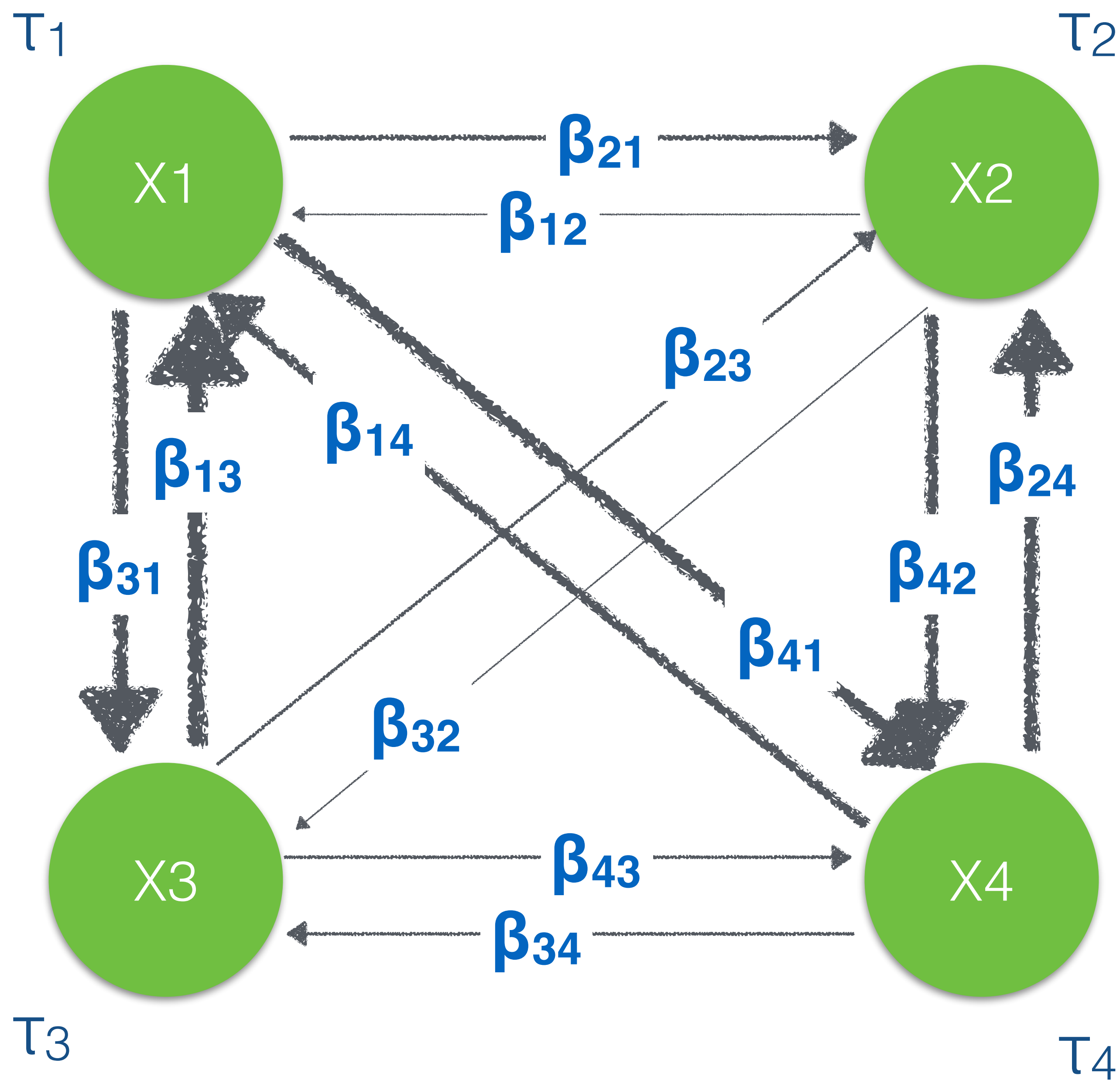
Perform regression of X_1 on all other variables

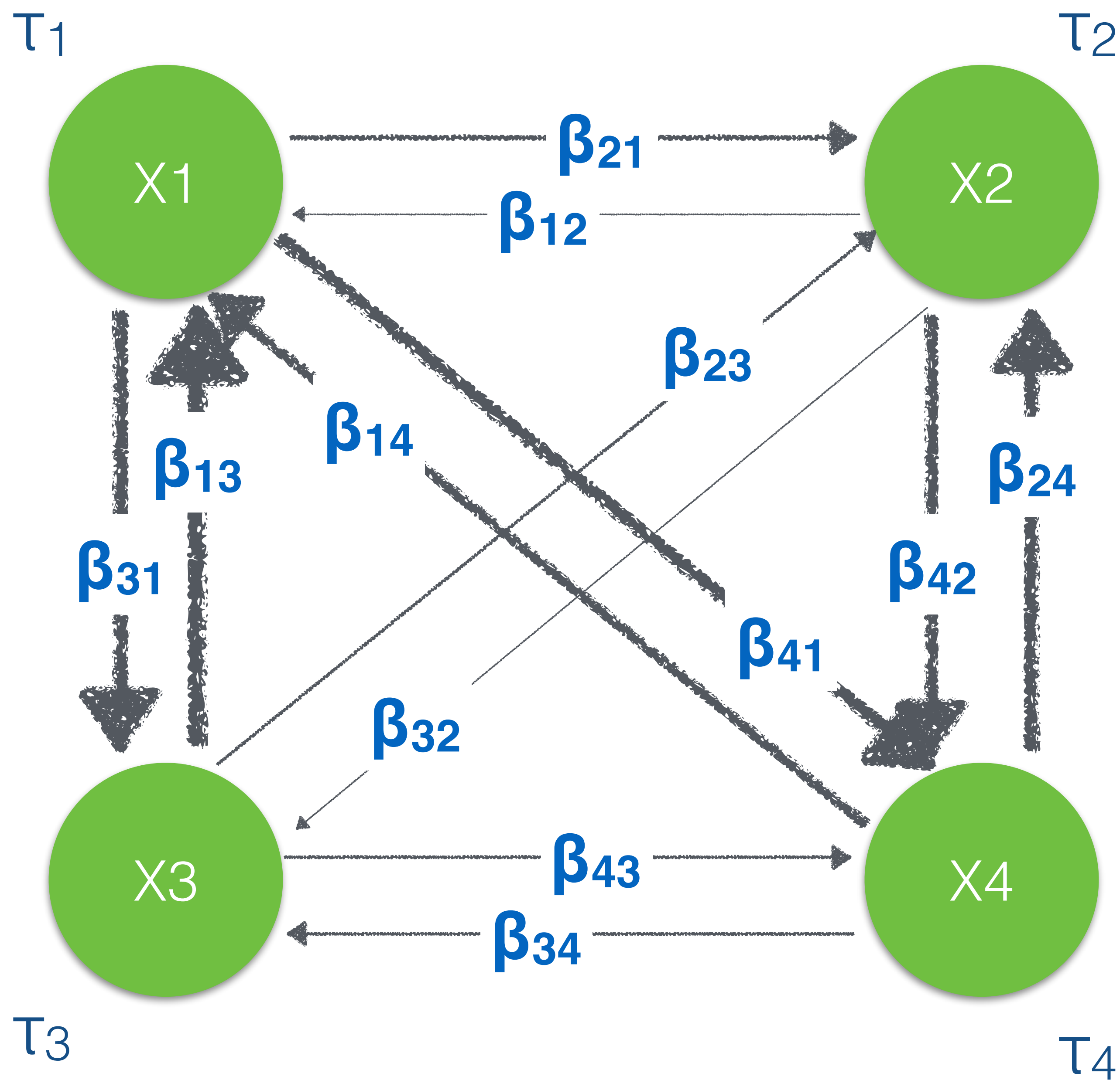


Repeat this for every variable

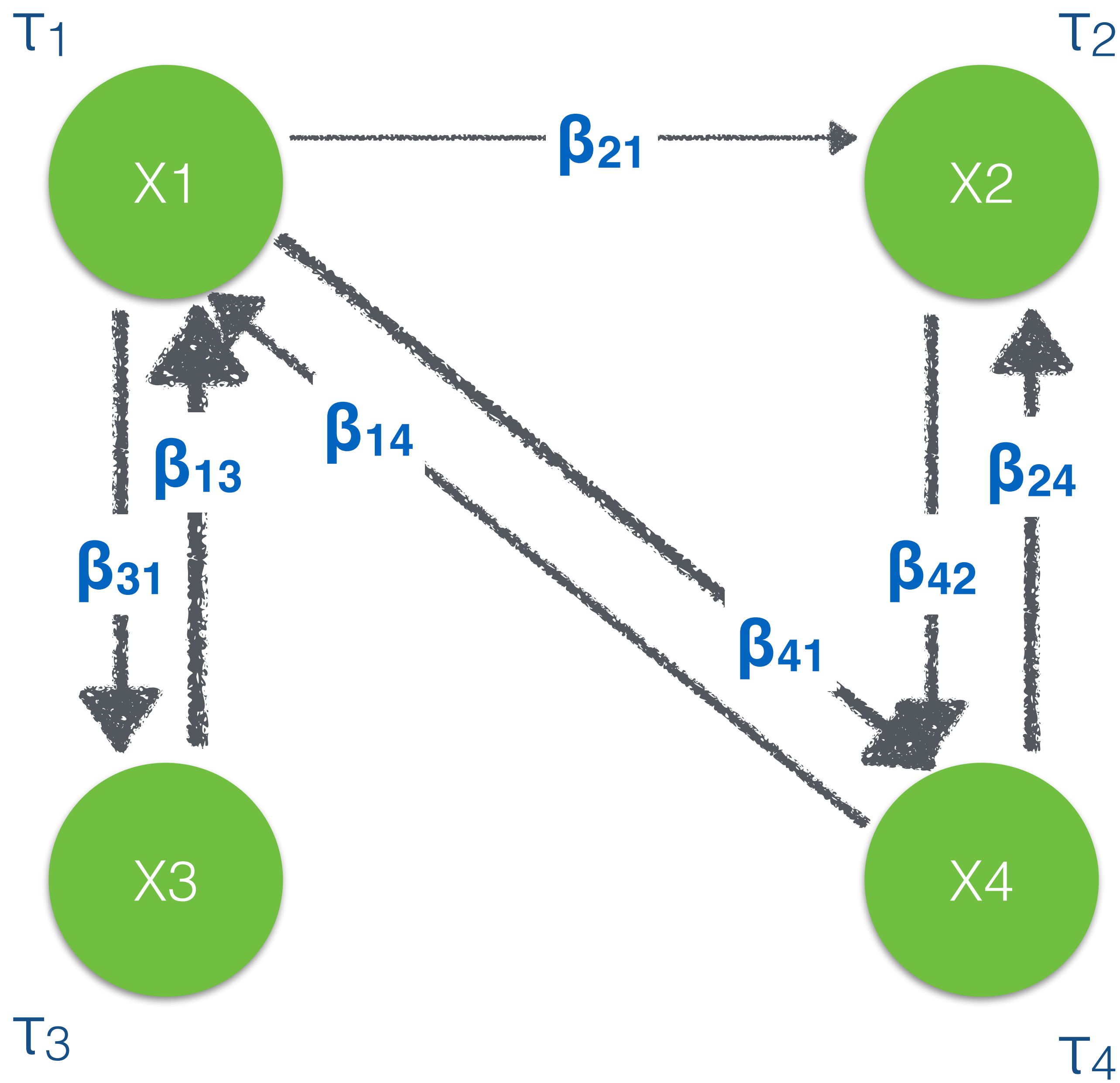


Repeat this for every variable


τ1
τ2
X1
X2
β21
β12
β23
β14
β13
β31
β24
β42
β41
β32
τ3
τ4
X3
X4
β43
β34

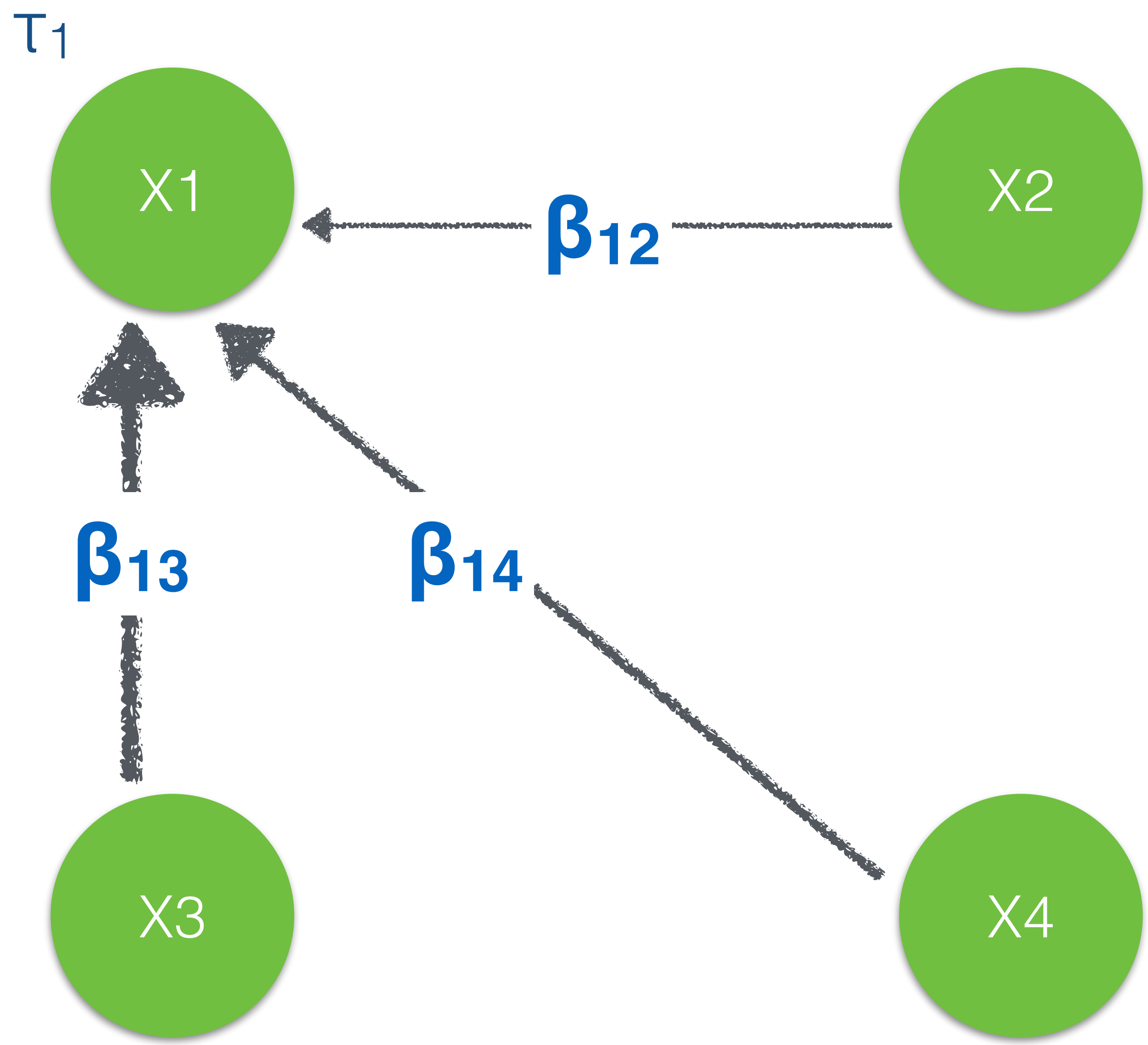


Control model complexity and prevent overfitting:
ℓ_1-**regularized** logistic regression



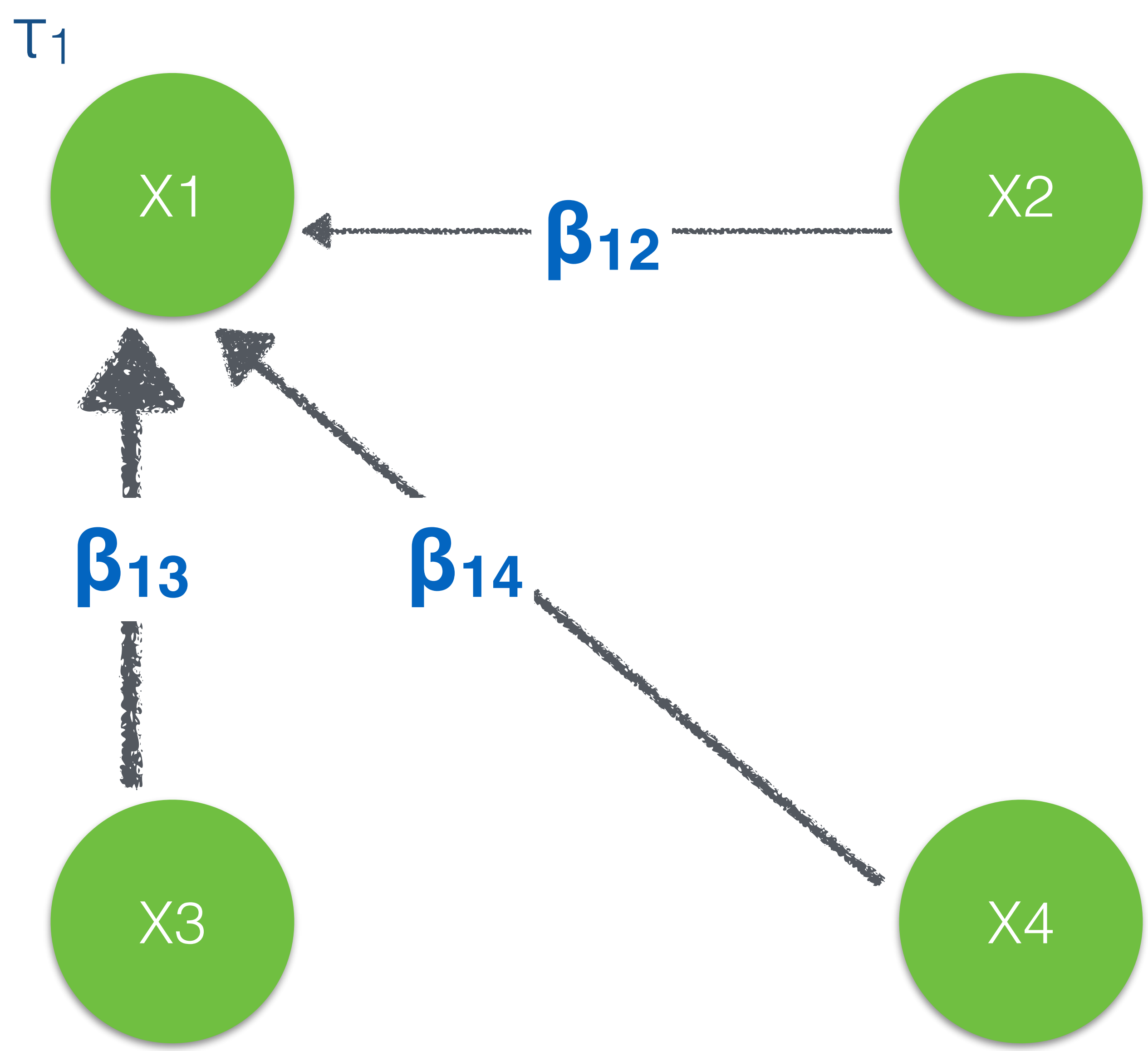
Control model complexity and prevent overfitting:
ℓ_1**-regularized** logistic regression

About regularization: shrinkage of coefficients



$$\hat{\Theta}_j^{\rho} = \arg\min_{\Theta_j} \left\{ -x_{ij} \cdot \left(\tau_j + \sum_{k \in V_{\setminus j}} \beta_{jk} x_{ik} \right) + \log\left(1 + \exp\left\{ \tau_j + \sum_{k \in V_{\setminus j}} x_{ik} \beta_{jk} \right\} \right) + \rho \sum_{k \in V_{\setminus j}} \left| \beta_{jk} \right| \right\}$$

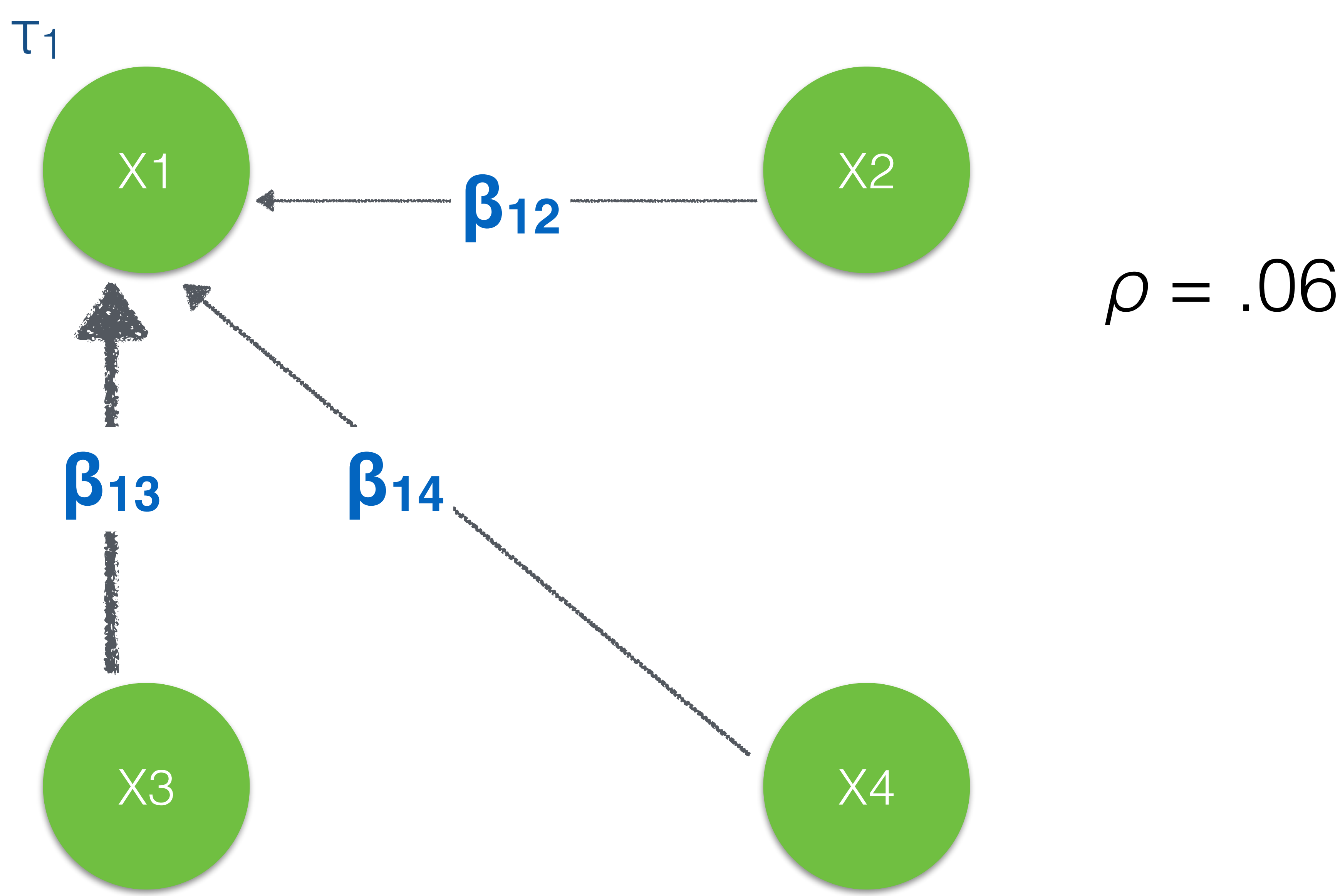
About regularization: shrinkage of coefficients



ρ = .008

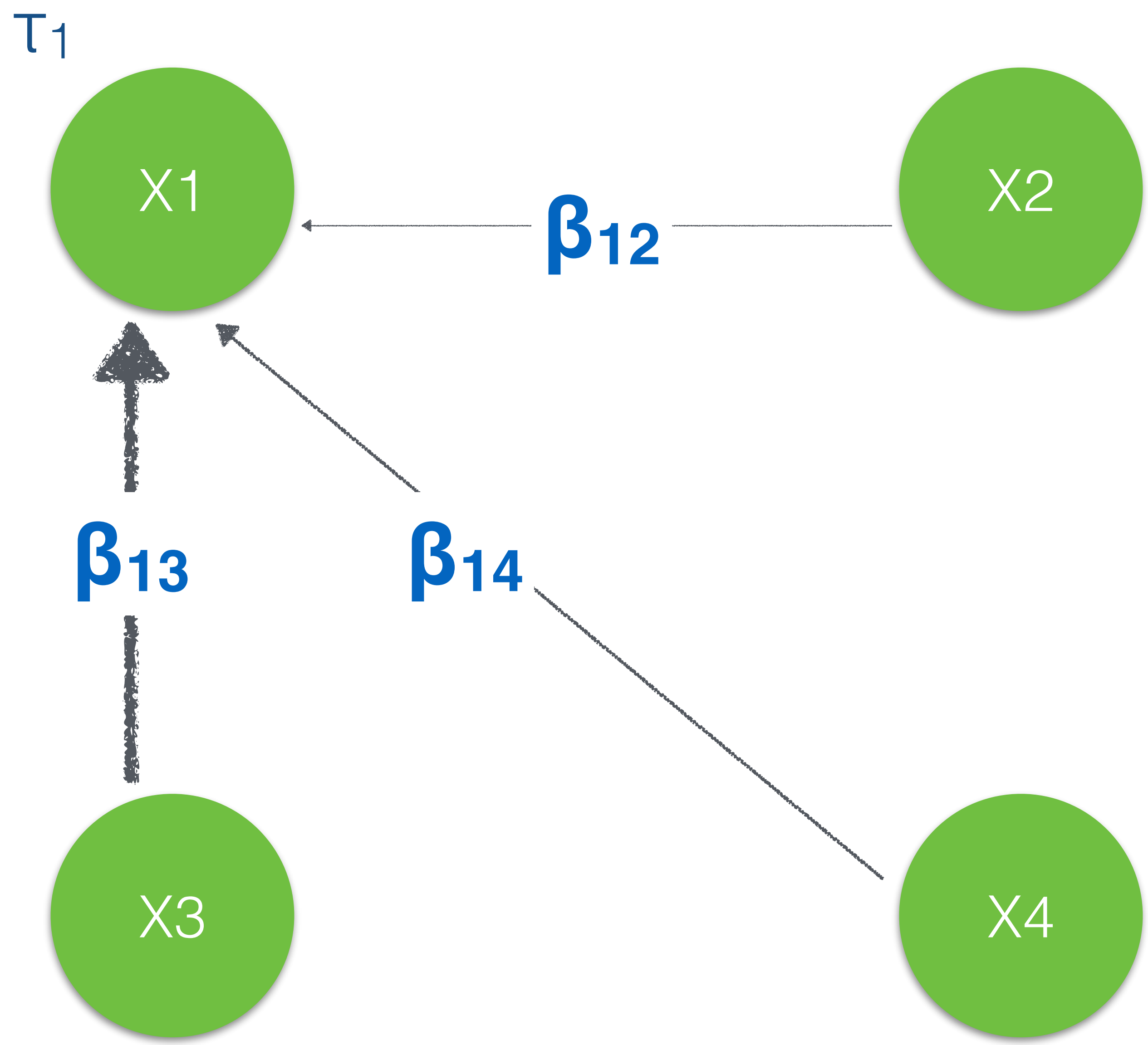
$$\hat{\Theta}_j^{\rho} = \arg\min_{\Theta_j} \left\{ -x_{ij} \cdot \left(\tau_j + \sum_{k \in V_{\setminus j}} \beta_{jk} x_{ik} \right) + \log\left(1 + \exp\left\{ \tau_j + \sum_{k \in V_{\setminus j}} x_{ik} \beta_{jk} \right\} \right) + \rho \sum_{k \in V_{\setminus j}} \left| \beta_{jk} \right| \right\}$$

About regularization: shrinkage of coefficients



$$\hat{\Theta}_j^{\rho} = \arg\min_{\Theta_j} \left\{ -x_{ij} \cdot \left(\tau_j + \sum_{k \in V_{\setminus j}} \beta_{jk} x_{ik} \right) + \log \left(1 + \exp \left\{ \tau_j + \sum_{k \in V_{\setminus j}} x_{ik} \beta_{jk} \right\} \right) + \rho \sum_{k \in V_{\setminus j}} \left| \beta_{jk} \right| \right\}$$

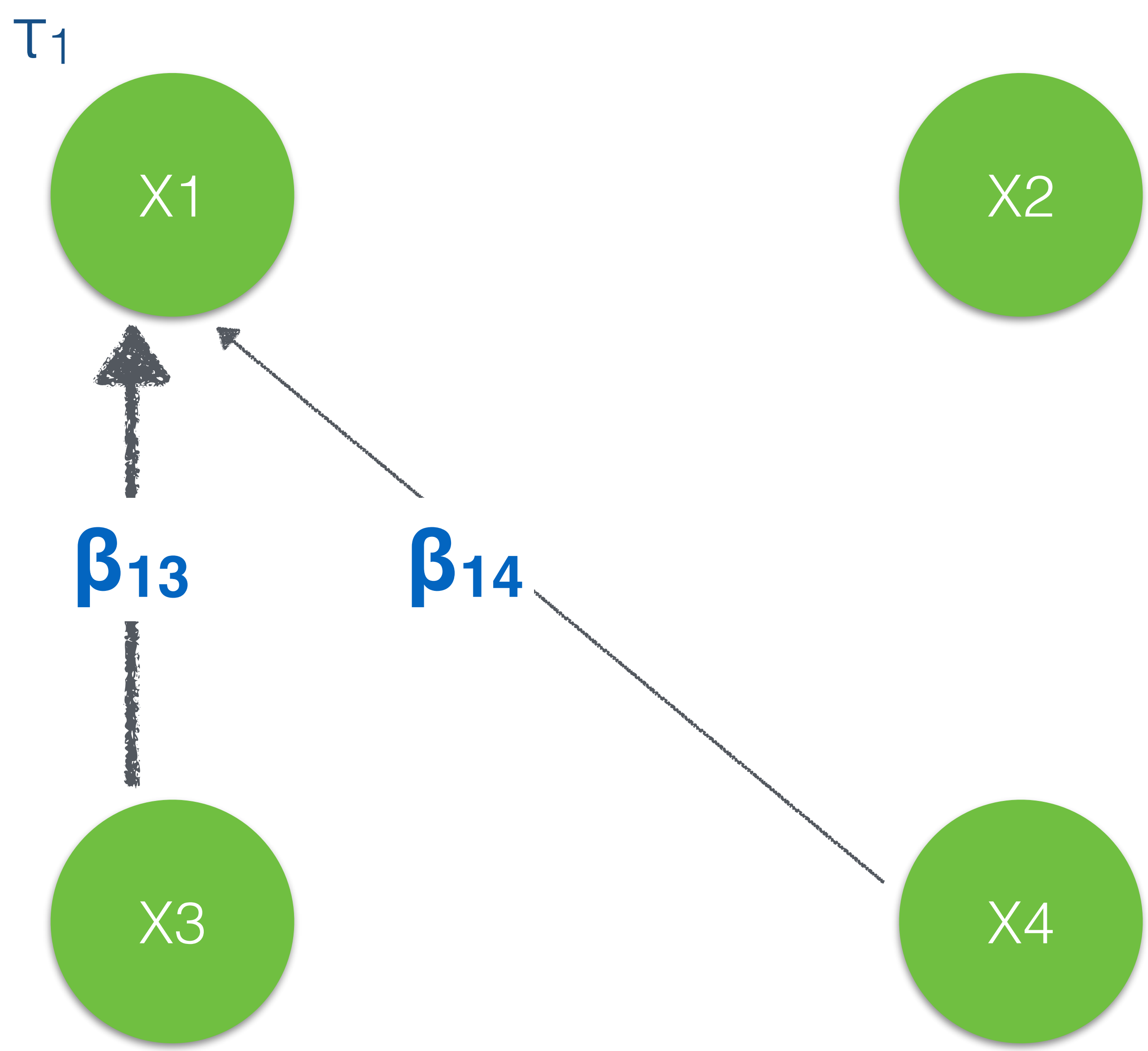
About regularization: shrinkage of coefficients



$\rho = .11$

$$\hat{\Theta}_j^{\rho} = \arg\min_{\Theta_j} \left\{ -x_{ij} \cdot \left(\tau_j + \sum_{k \in V_{\setminus j}} \beta_{jk} x_{ik} \right) + \log\left(1 + \exp\left\{ \tau_j + \sum_{k \in V_{\setminus j}} x_{ik} \beta_{jk} \right\} \right) + \rho \sum_{k \in V_{\setminus j}} \left| \beta_{jk} \right| \right\}$$

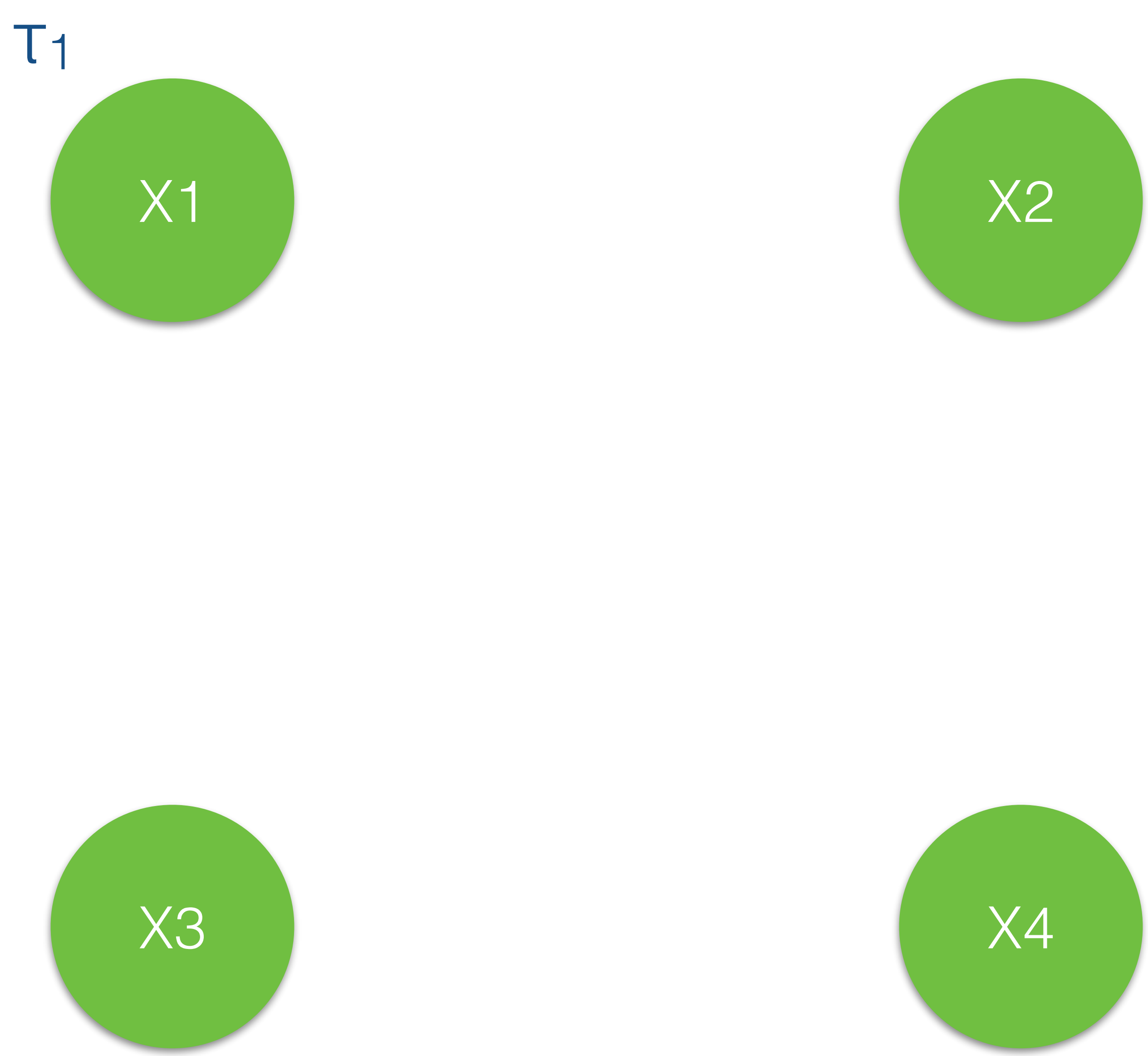
About regularization: shrinkage of coefficients



$\rho = .19$

$$\hat{\Theta}_j^{\rho} = \arg\min_{\Theta_j} \left\{ -x_{ij} \cdot \left(\tau_j + \sum_{k \in V_{\setminus j}} \beta_{jk} x_{ik} \right) + \log\left(1 + \exp\left\{ \tau_j + \sum_{k \in V_{\setminus j}} x_{ik} \beta_{jk} \right\} \right) + \rho \sum_{k \in V_{\setminus j}} \left| \beta_{jk} \right| \right\}$$

About regularization: shrinkage of coefficients



$\rho = .24$

$$\hat{\Theta}_j^{\rho} = \arg\min_{\Theta_j} \left\{ -x_{ij} \cdot \left(\tau_j + \sum_{k \in V_{\setminus j}} \beta_{jk} x_{ik} \right) + \log\left(1 + \exp\left\{ \tau_j + \sum_{k \in V_{\setminus j}} x_{ik} \beta_{jk} \right\} \right) + \rho \sum_{k \in V_{\setminus j}} \left| \beta_{jk} \right| \right\}$$

l1-regularized logistic regression

- Optimizes neighbourhood selection by optimizing convex function

$$\hat{\Theta}_j^\rho = \arg\min_{\Theta_j} \left\{ -x_{ij} \cdot \left(\tau_j + \sum_{k \in V_{\setminus j}} \beta_{jk} x_{ik} \right) + \log\left(1 + \exp\left\{ \tau_j + \sum_{k \in V_{\setminus j}} x_{ik} \beta_{jk} \right\} \right) + \rho \sum_{k \in V_{\setminus j}} \left| \beta_{jk} \right| \right\}$$

- i: independent observations {1, 2, …, n}
- $\hat{\Theta}_j^\rho$: matrix with βjk and τj
- ρ: tuning parameter
- `R` package **`glmnet`**: 100 values of ρ

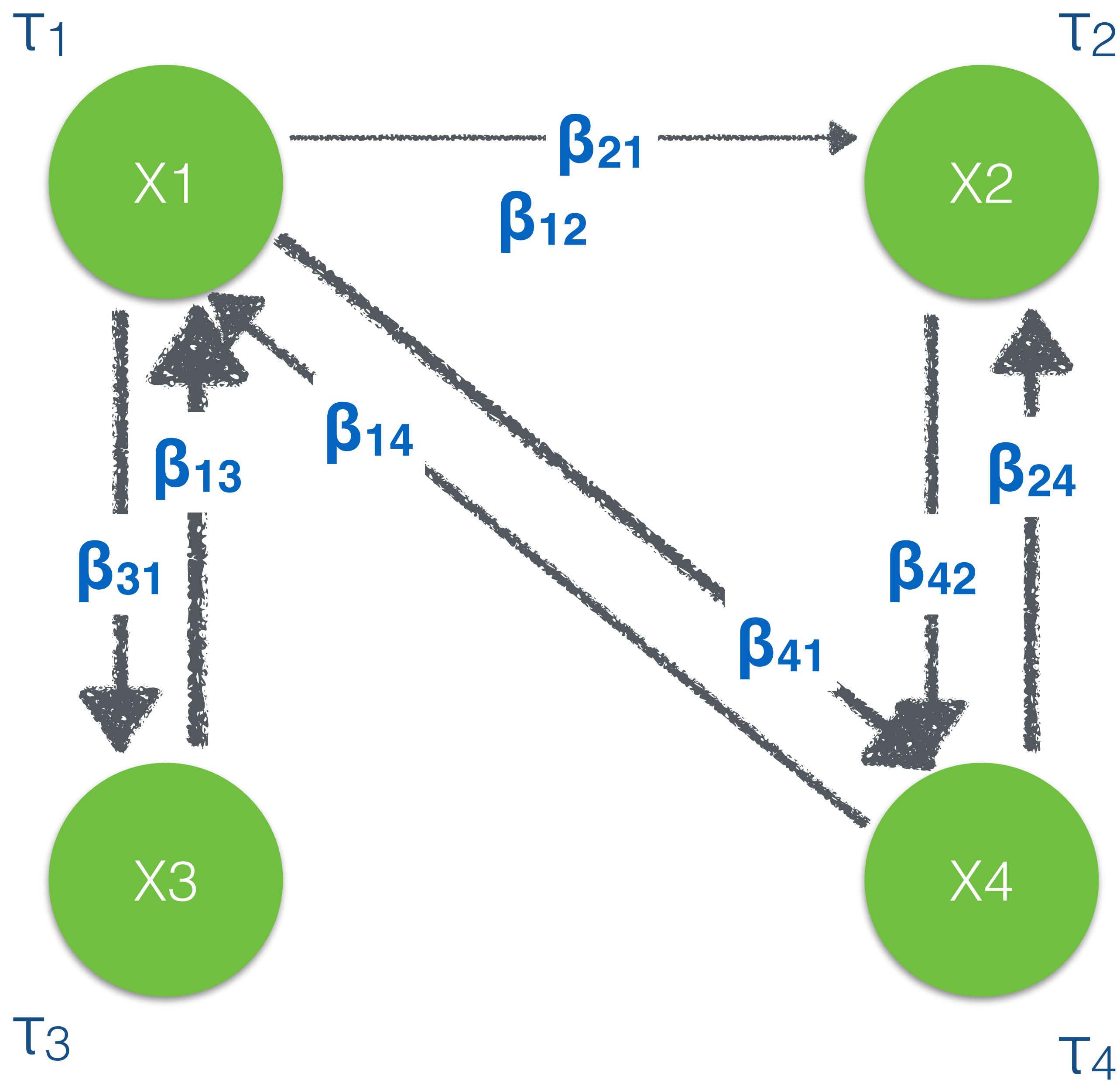
l1-regularized logistic regression

- Choose tuning parameter with *extended* Bayesian Information Criterion (EBIC)

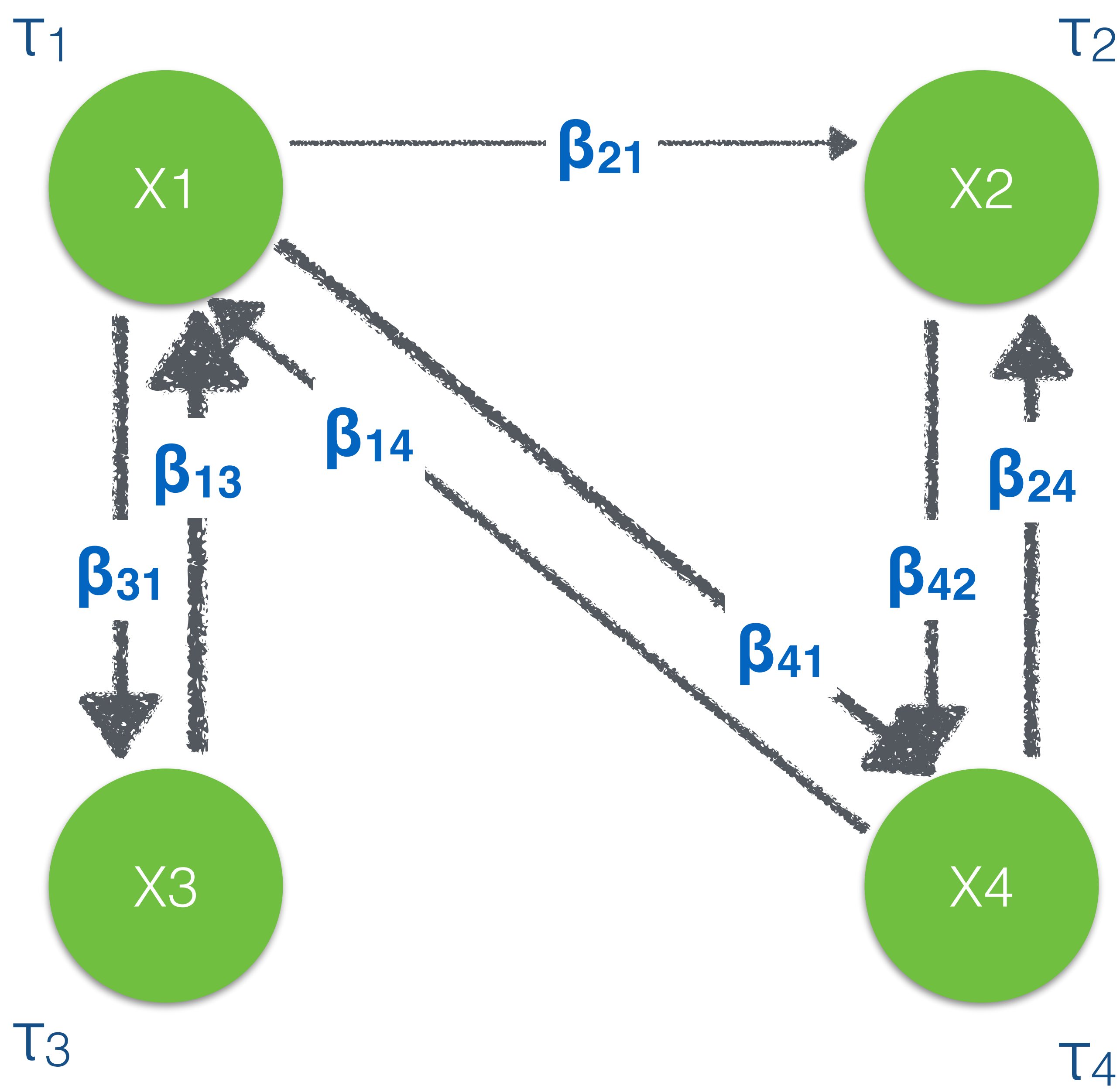
$$\mathrm{BIC}_\gamma(j) = -2\ell(\hat{\Theta}_J) + |J| \cdot \log(n) + 2\gamma|J| \cdot \log(p-1)$$

$$\ell(\hat{\Theta}_j) = \sum_{i=1}^{n}\left(\tau_j x_{ij} + \sum_{k \in V_{\setminus j}} \beta_{jk} x_{ij} x_{ik} - \log(1 + \exp\{\tau_j + \sum_{k \in V_{\setminus j}} x_{ik}\beta_{jk}\})\right)$$

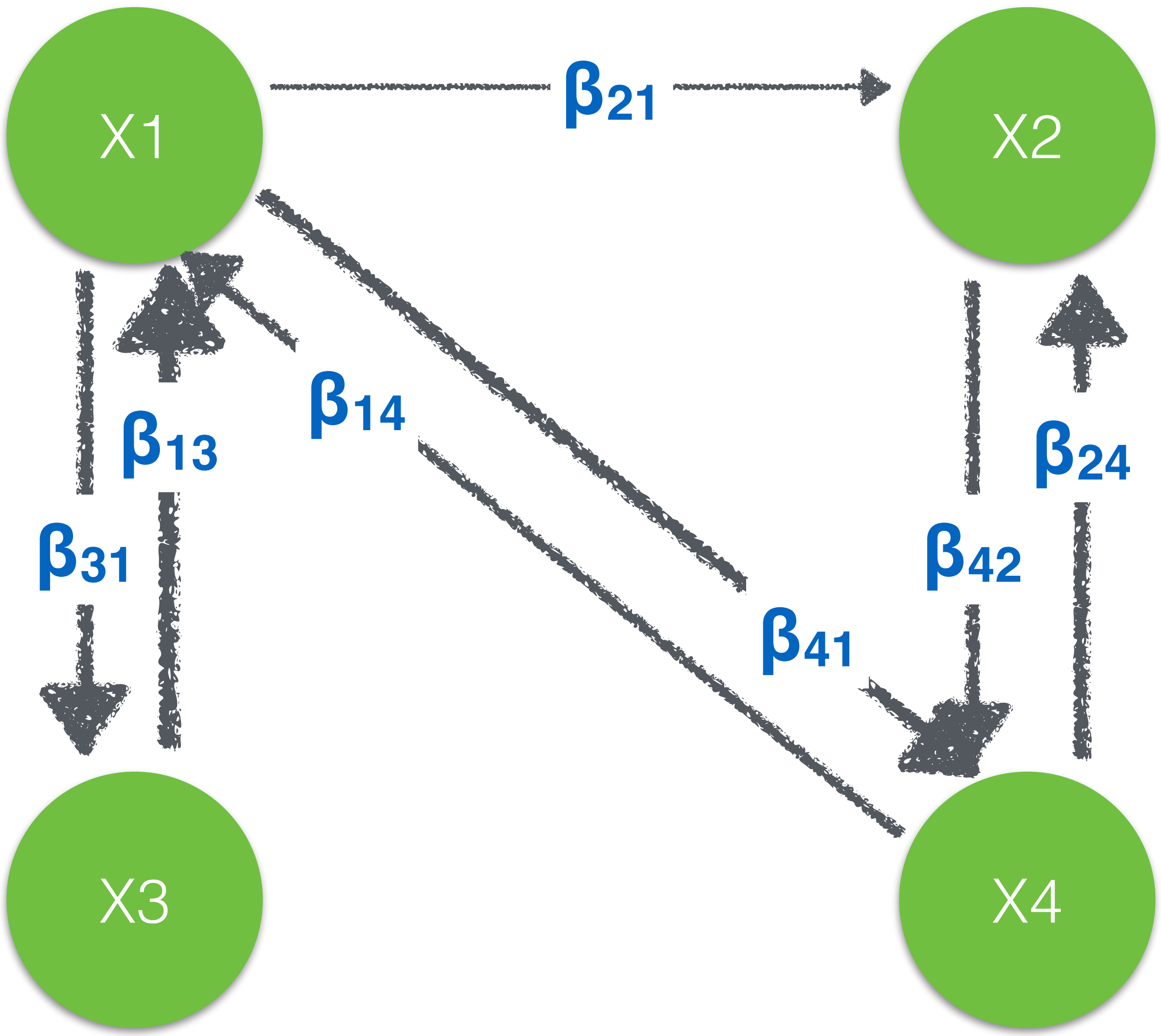
- |J|: number of neighbours
- n: number of observations
- p-1: number of covariates



- collect regularized parameters
- apply AND/OR rule



- collect regularized parameters
- apply AND/OR rule

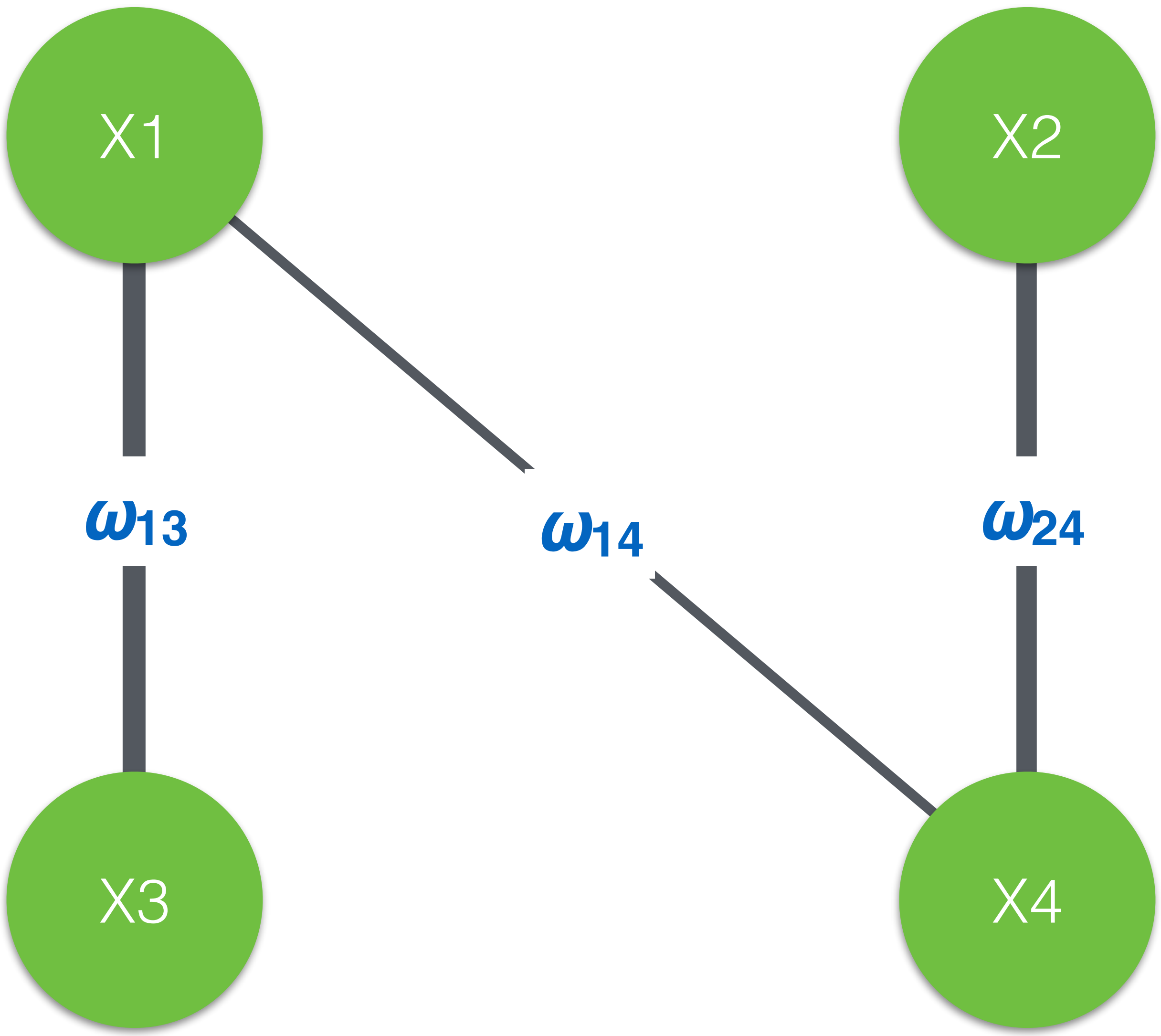


AND-rule:

if $\beta_{ij} \neq 0$ AND $\beta_{ji} \neq 0$

then $\omega_{ij} = (\beta_{ij} + \beta_{ji})/2$

else $\omega_{ij} = 0$



AND-rule:

if $\beta_{ij} \neq 0$ AND $\beta_{ji} \neq 0$

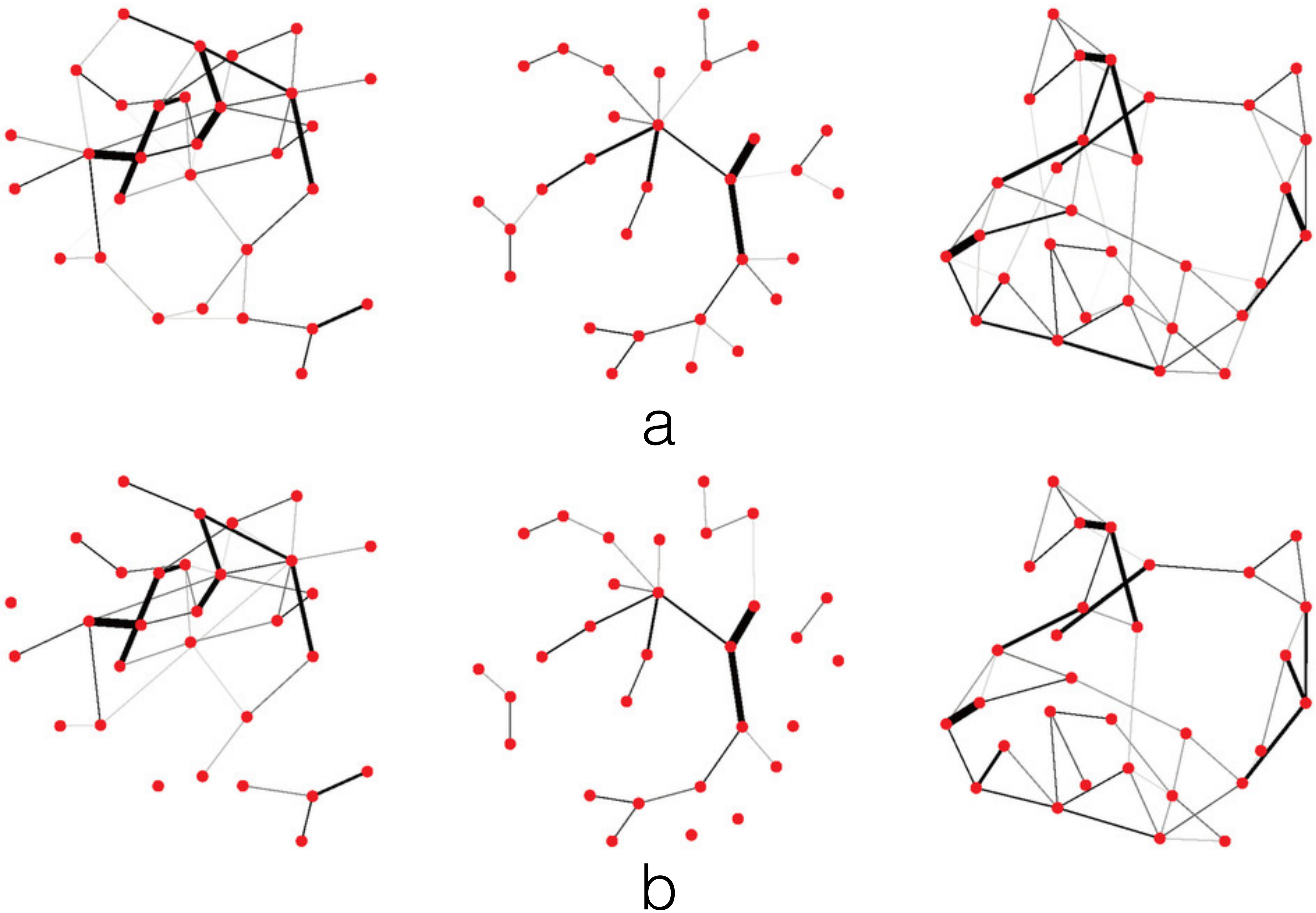
then $\omega_{ij} = (\beta_{ij} + \beta_{ji})/2$

else $\omega_{ij} = 0$

eLasso (`R` package `IsingFit`)

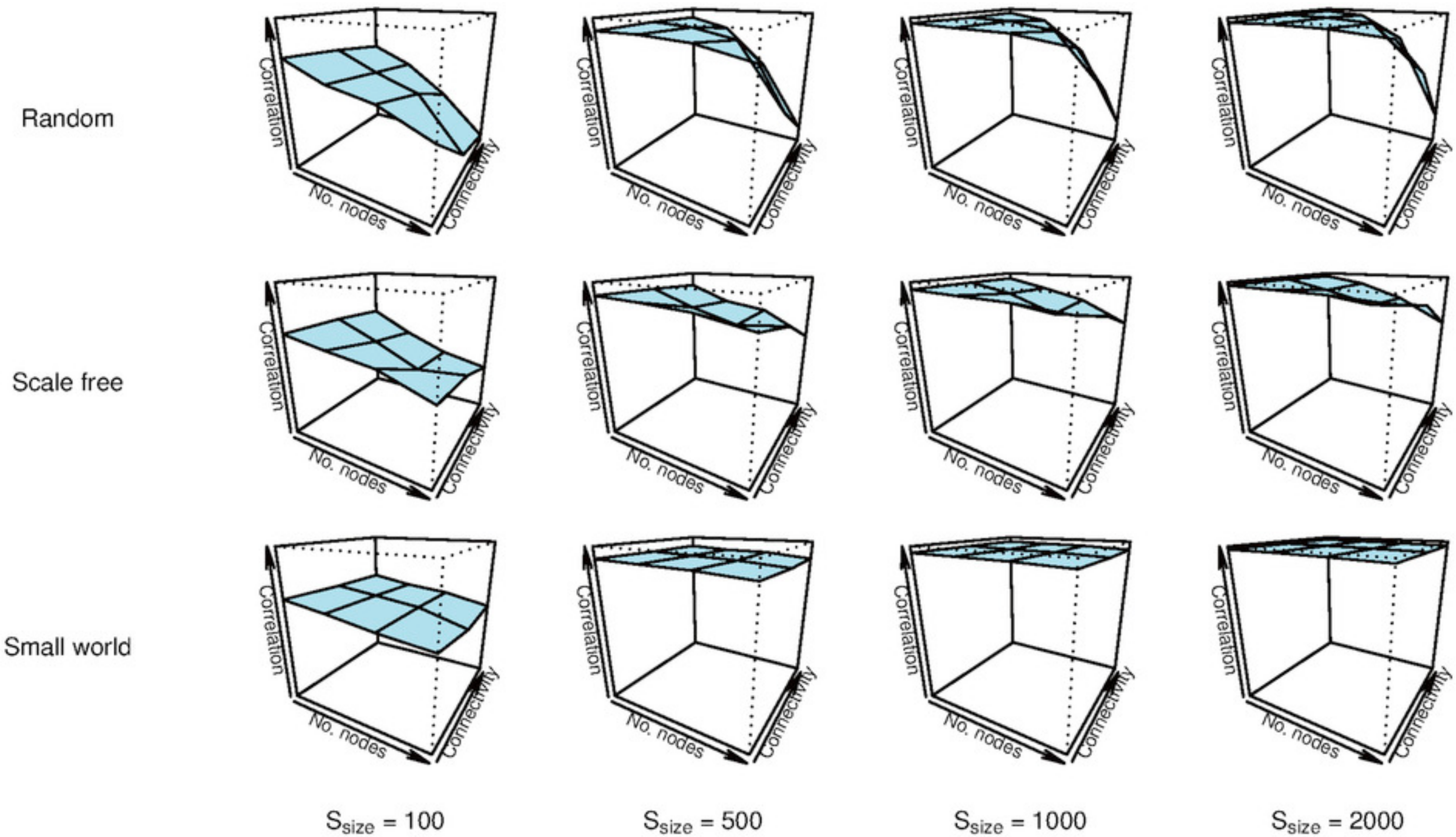
True

Estimated



Borkulo, C. D. van, Borsboom, D., Epskamp, S., Blanken, T. F., Boschloo, L., Schoevers, R. A., & Waldorp, L. J. (2014). A new method for constructing networks from binary data. Scientific Reports, 4(5918), 1–10. doi:10.1038/srep05918

Validation

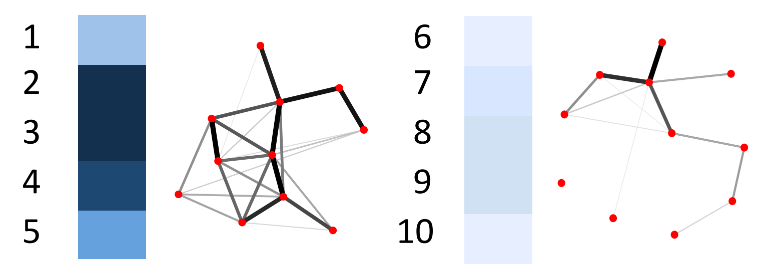


Borkulo, C. D. van, Borsboom, D., Epskamp, S., Blanken, T. F., Boschloo, L., Schoevers, R. A., & Waldorp, L. J. (2014). A new method for constructing networks from binary data. Scientific Reports, 4(5918), 1–10. doi:10.1038/srep05918

Network Comparison Test

Observed data and networks

Persisters Remitters



Measure of difference

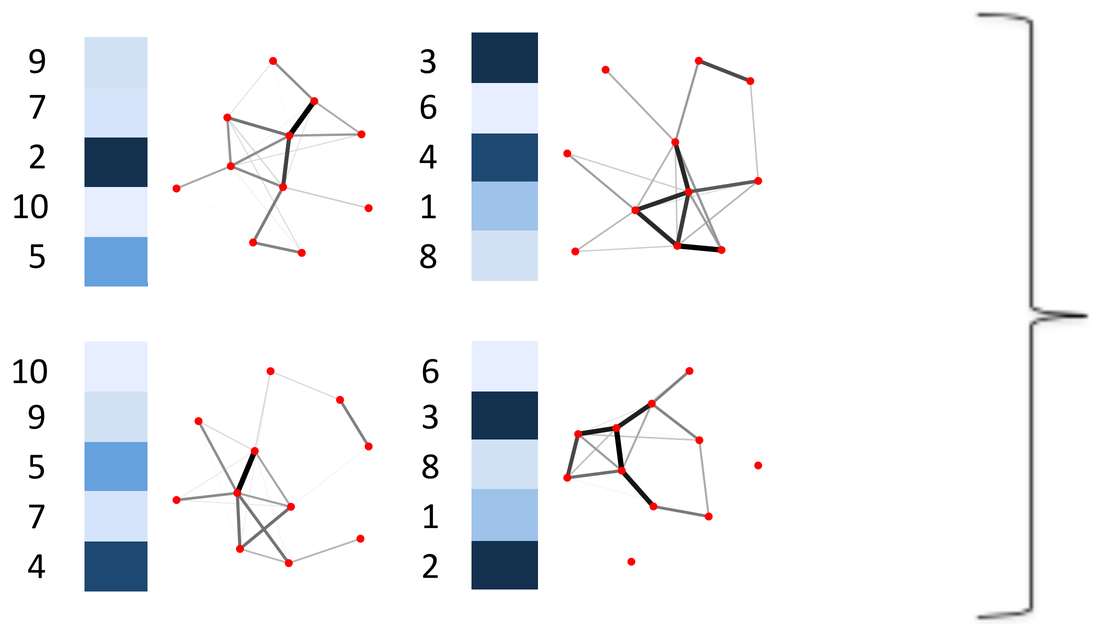
difference in global

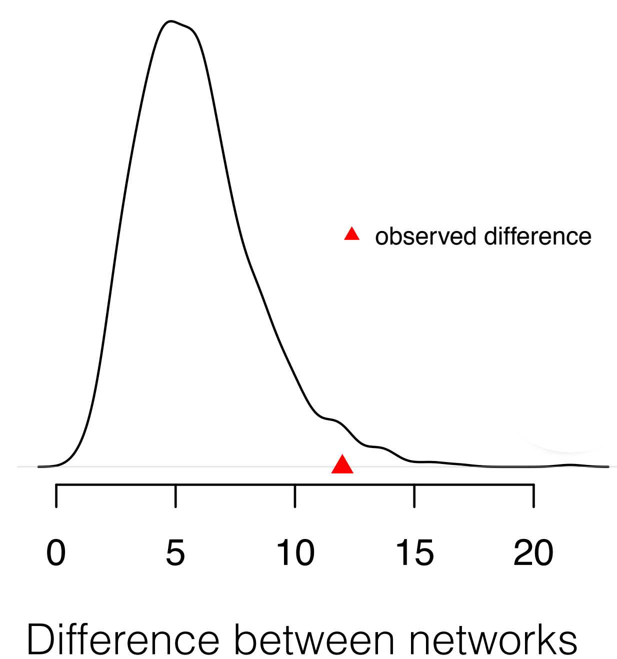
strength of networks

$|(\sum|\beta_{jk}^{pers}| - \sum|\beta_{jk}^{rem}|)|$

Permuted data and networks

'Persisters' 'Remitters'





Van Borkulo CD, Waldorp LJ, Boschloo L, Kossakowski J, Tio P, Schoevers RA Borsboom D. (2015). Distinguishing networks: a permutation test for comparing network structures. Manuscript in preparation.

Thanks for your attention!

谢谢

Borkulo, C. D. van, Borsboom, D., Epskamp, S., Blanken, T. F., Boschloo, L., Schoevers, R. A., & Waldorp, L. J. (2014). A new method for constructing networks from binary data. Scientific Reports, 4(5918), 1–10. doi:10.1038/srep05918

Van Borkulo CD, Waldorp LJ, Boschloo L, Kossakowski J, Tio P, Schoevers RA Borsboom D. (2015). Distinguishing networks: a permutation test for comparing network structures. Manuscript in preparation.

Claudia van Borkulo
http://cvborkulo.com
cvborkulo@gmail.com