Network Autocorrelation Models

SOC 280: Analysis of Social Network Data

Carter T. Butts
Department of Sociology and
Institute for Mathematical Behavioral Sciences
University of California, Irvine

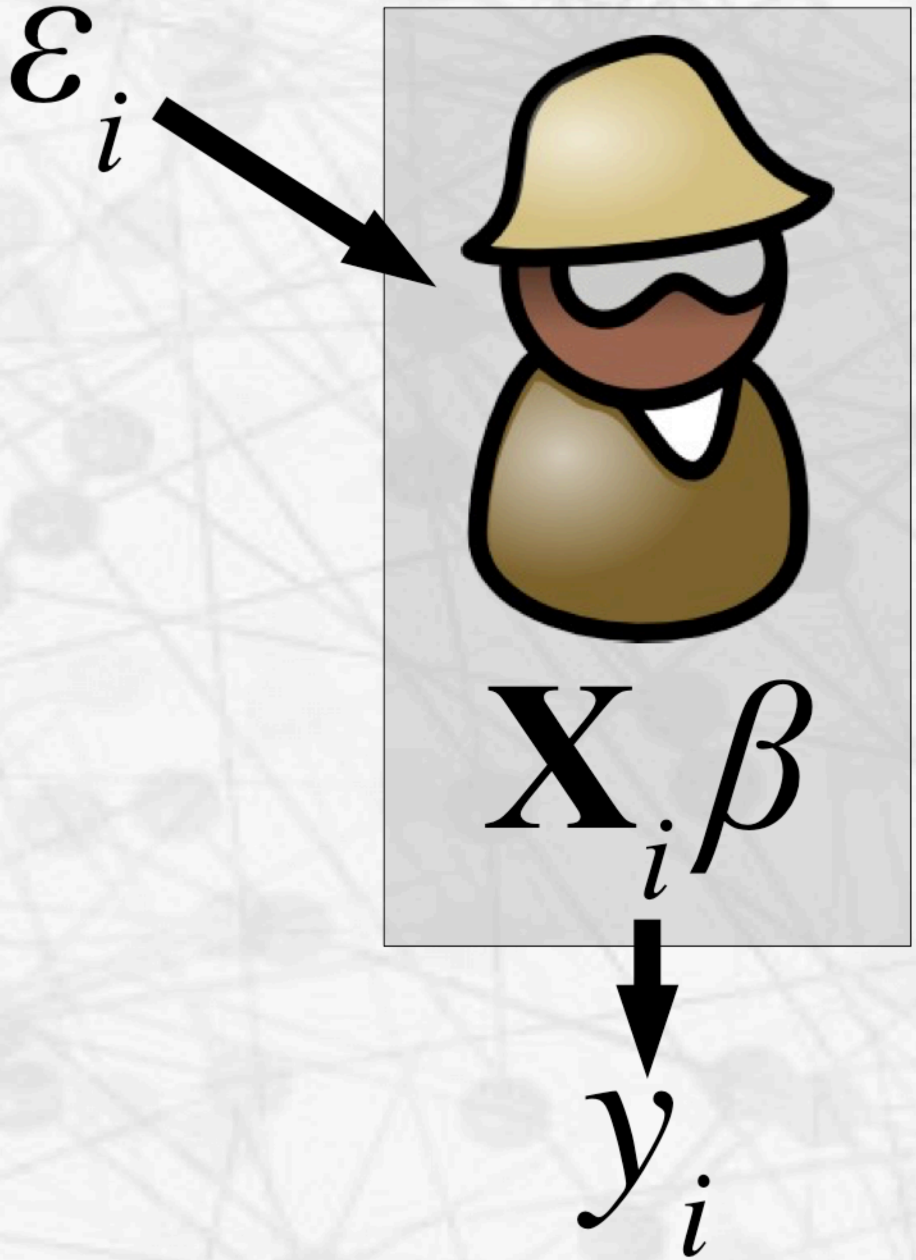
From Indices to Autocorrelation

- **Last time, we discussed how properties of positions could be related to covariates....**
  - Does being a gatekeeper help your career?
  - Do high-degree organizations make all the decisions?
- **Today: how covariates relate to each other through ties**
  - *Network autocorrelation*: (usually linear) dependence of individual attributes across positions
- **Applicable to many phenomena**
  - E.g., Social influence, organizational isomorphism, work-family spillover, neighborhood quality, economic productivity, etc.

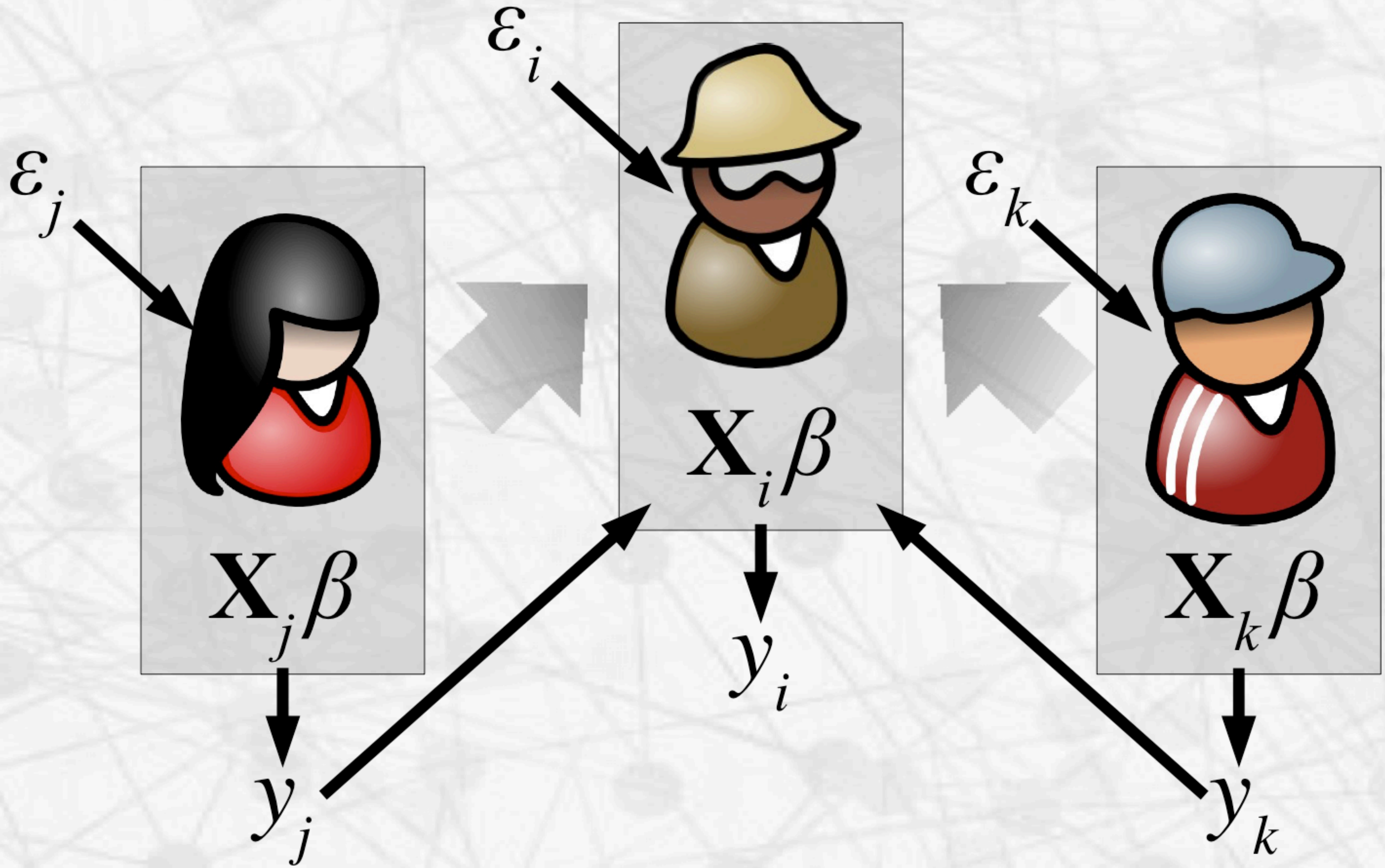
Network Autocorrelation Model

- **Simplest tool for modeling network autocorrelation: the (linear) network autocorrelation model**
  - Builds on standard (OLS) regression
  - Closely analogous to ARMA models in the time series literature
  - Identical to the spatial ARMA models of Cliff and Ord (1973), Anselin (1988), etc. (But we use a different name!)
- **Basic components**
  - Response variable (to be modeled)
  - “Weight” matrices (identical to or derived from original networks)
  - Predictive covariates
  - Parameters measuring nature/extent of dependence, covariate effects

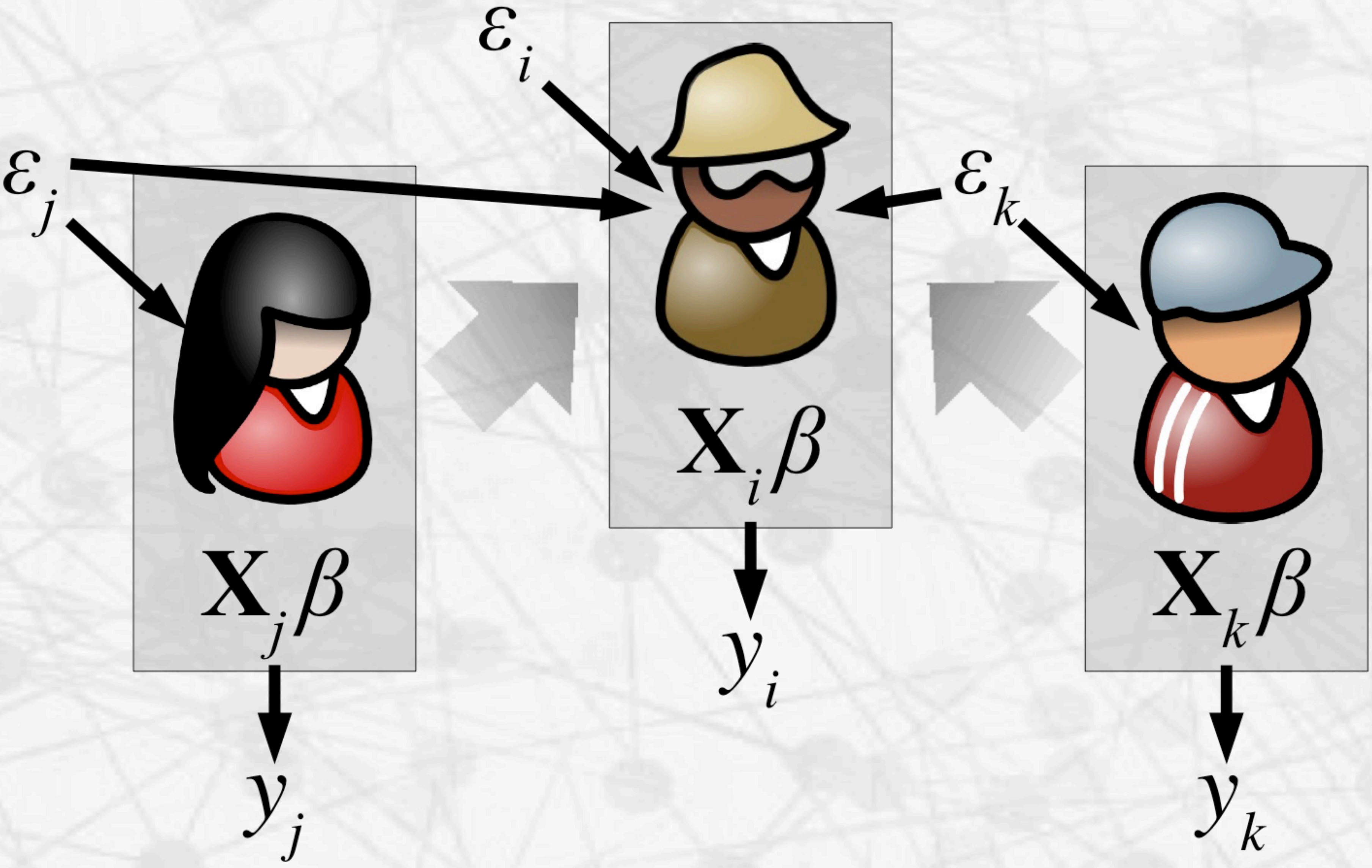
The Classical Regression Model



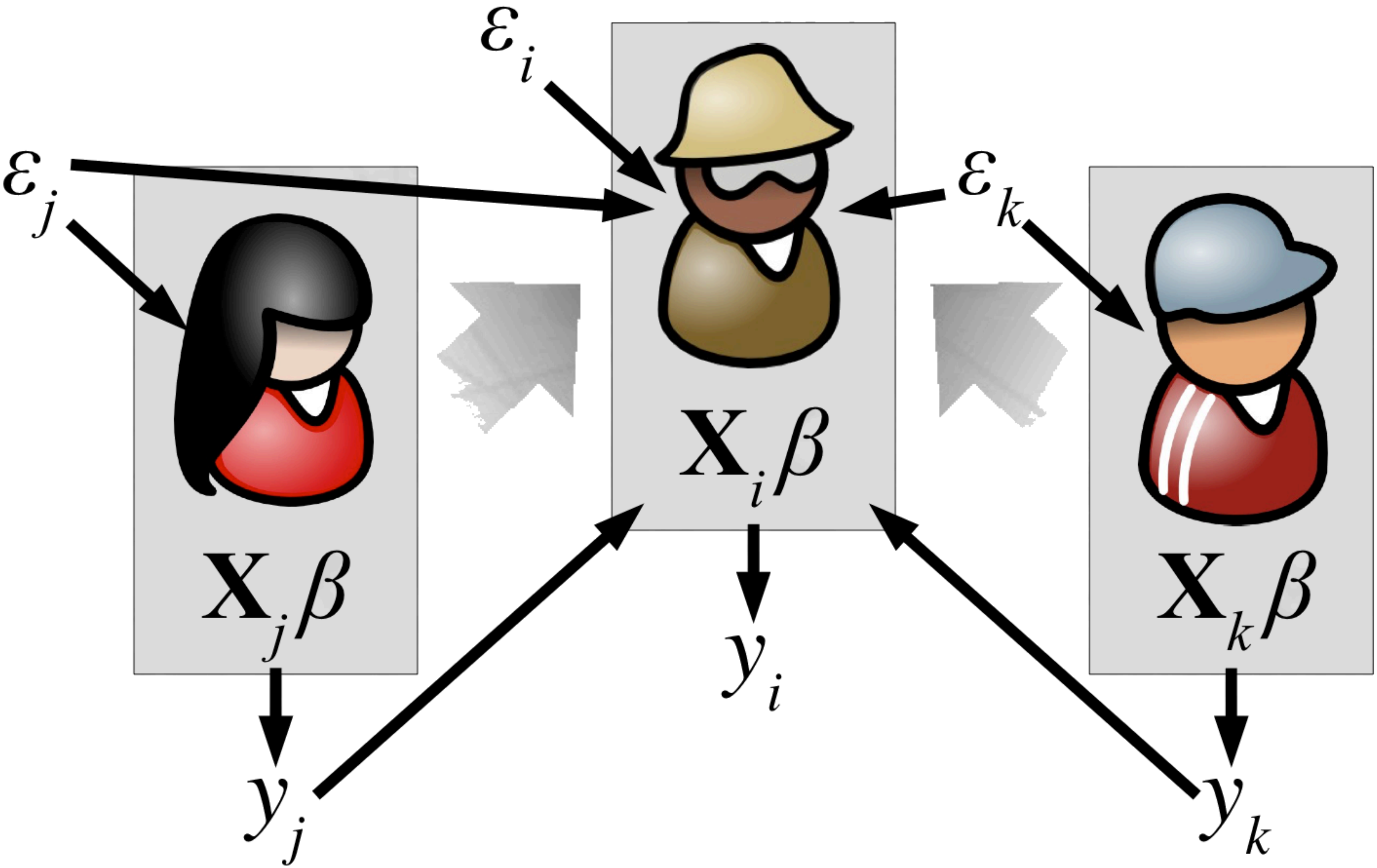
Adding Network AR Effects



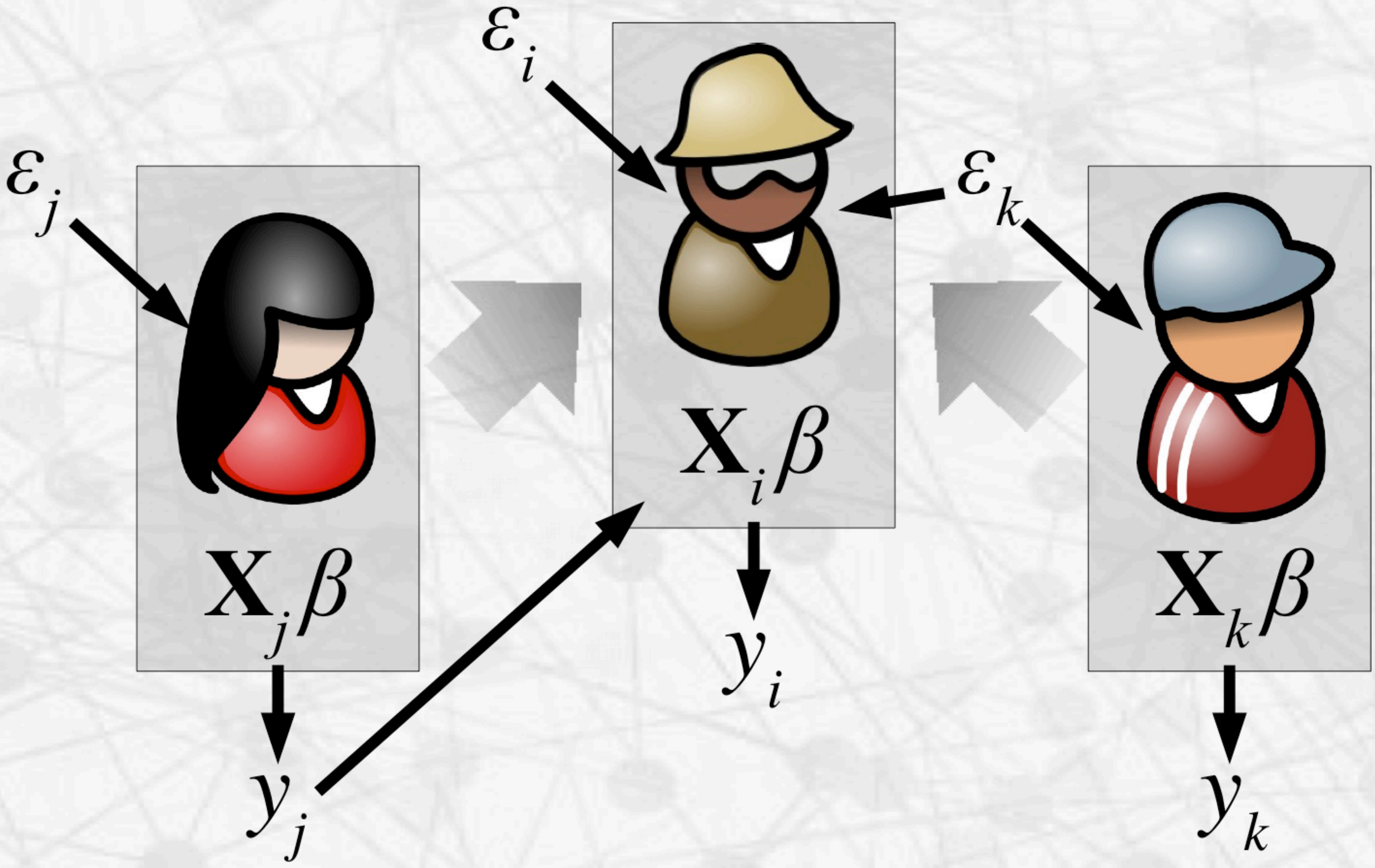
Adding Network MA Effects



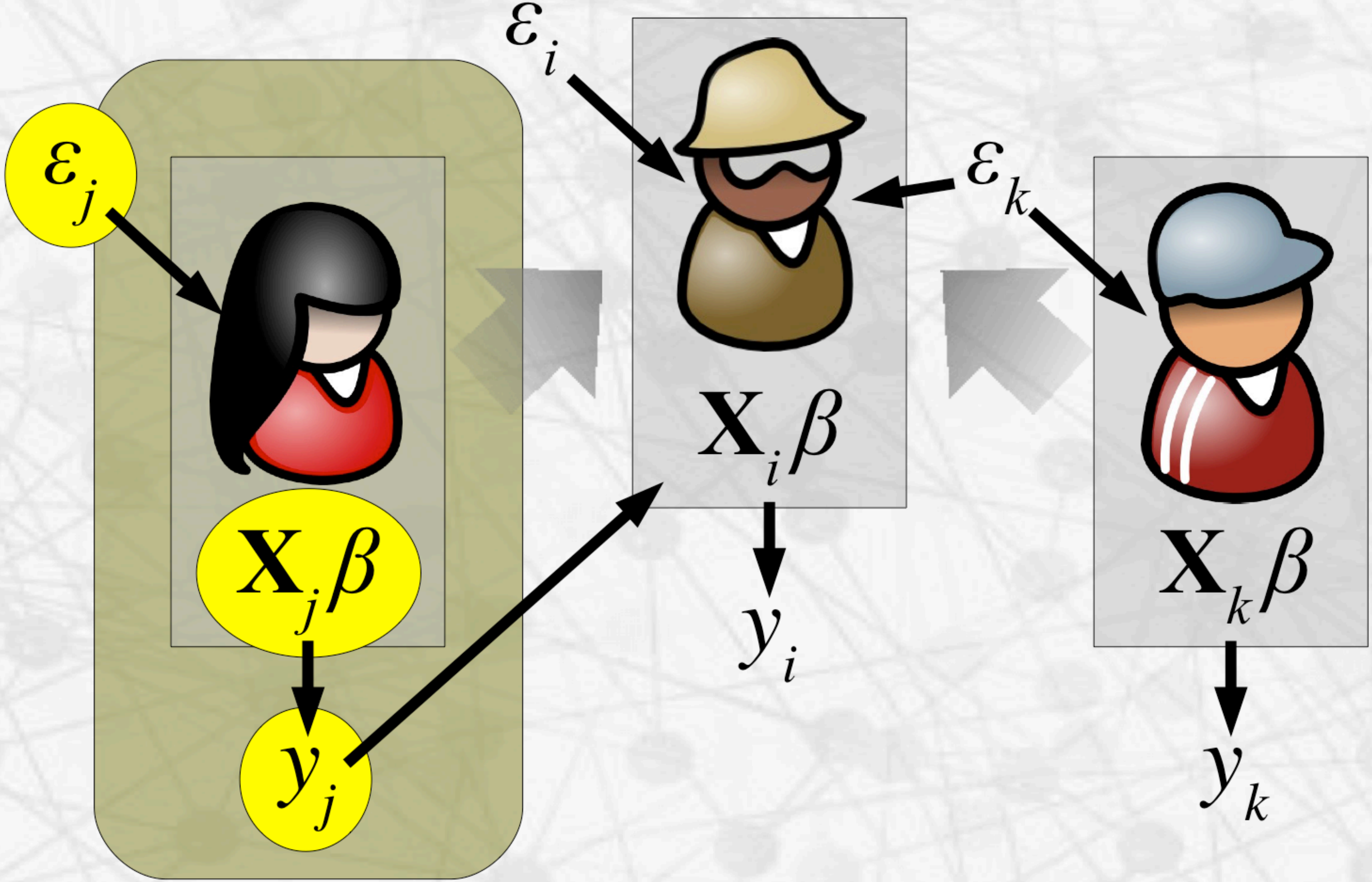
Network ARMA Model



AR Vs. MA

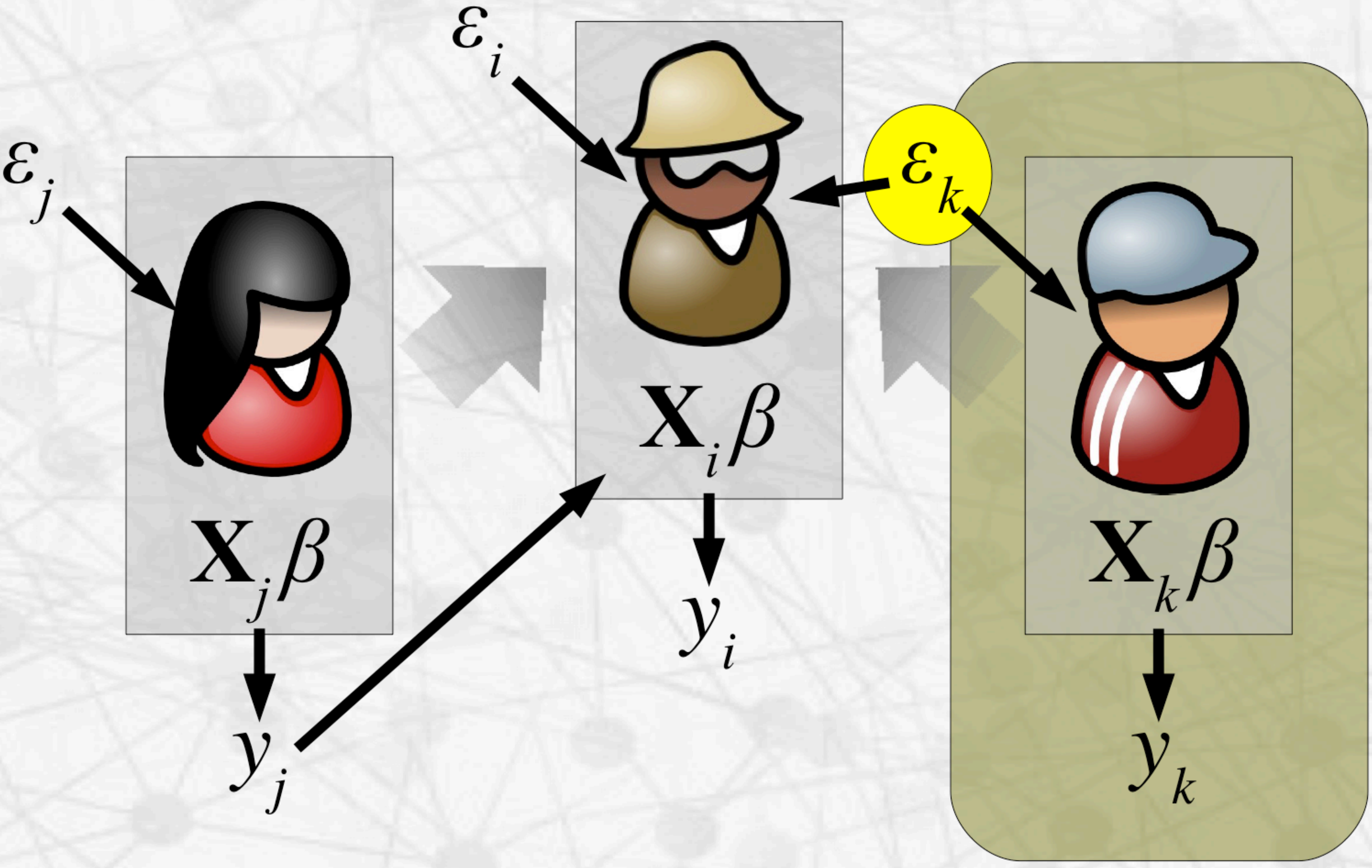

ε_i
ε_j
ε_k
$\mathbf{X}_j\beta$
$\mathbf{X}_i\beta$
$\mathbf{X}_k\beta$
y_j
y_i
y_k

AR Vs. MA



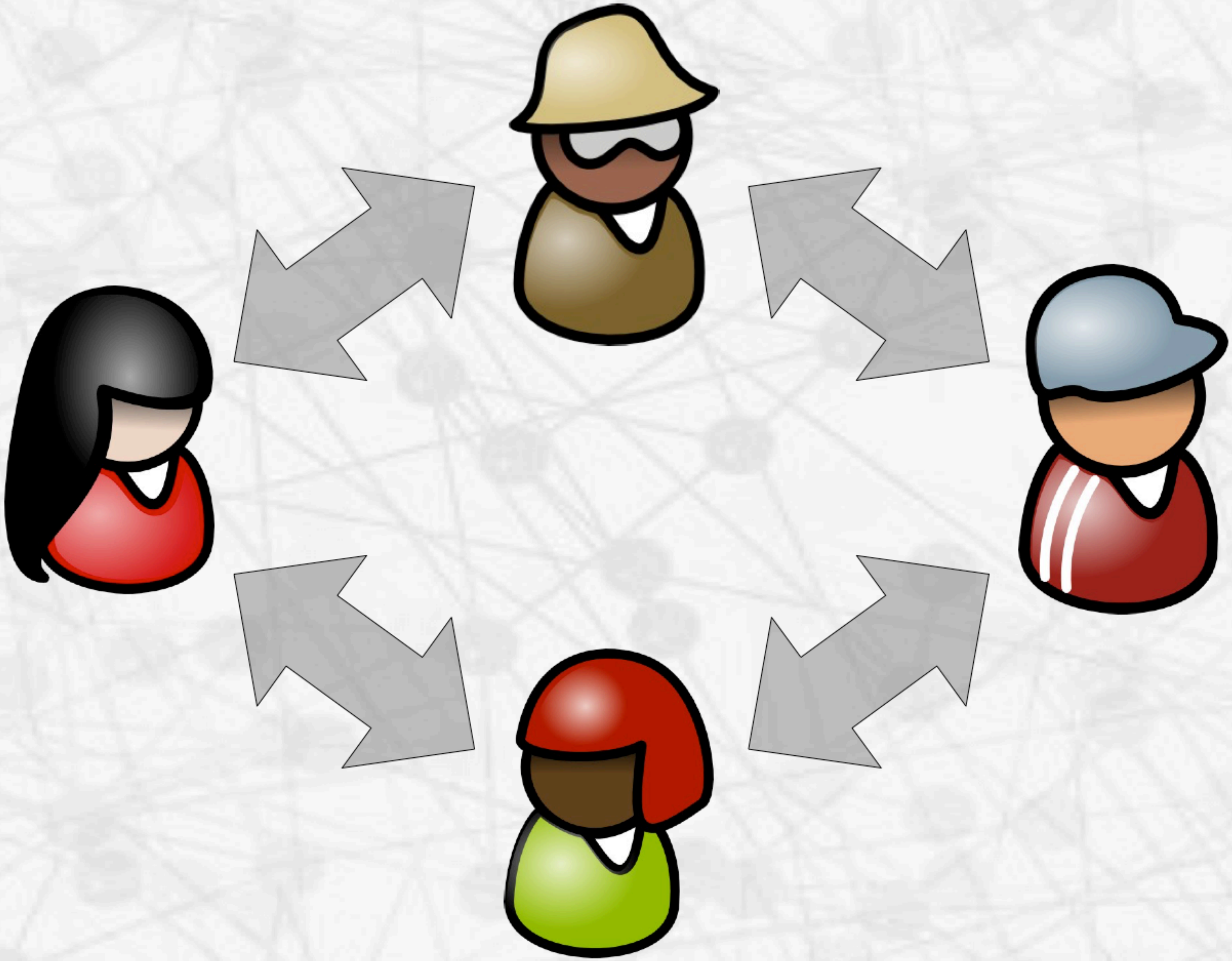
ε_i
ε_j
ε_k
$\mathbf{X}_j\beta$
$\mathbf{X}_i\beta$
$\mathbf{X}_k\beta$
y_j
y_i
y_k

AR Vs. MA

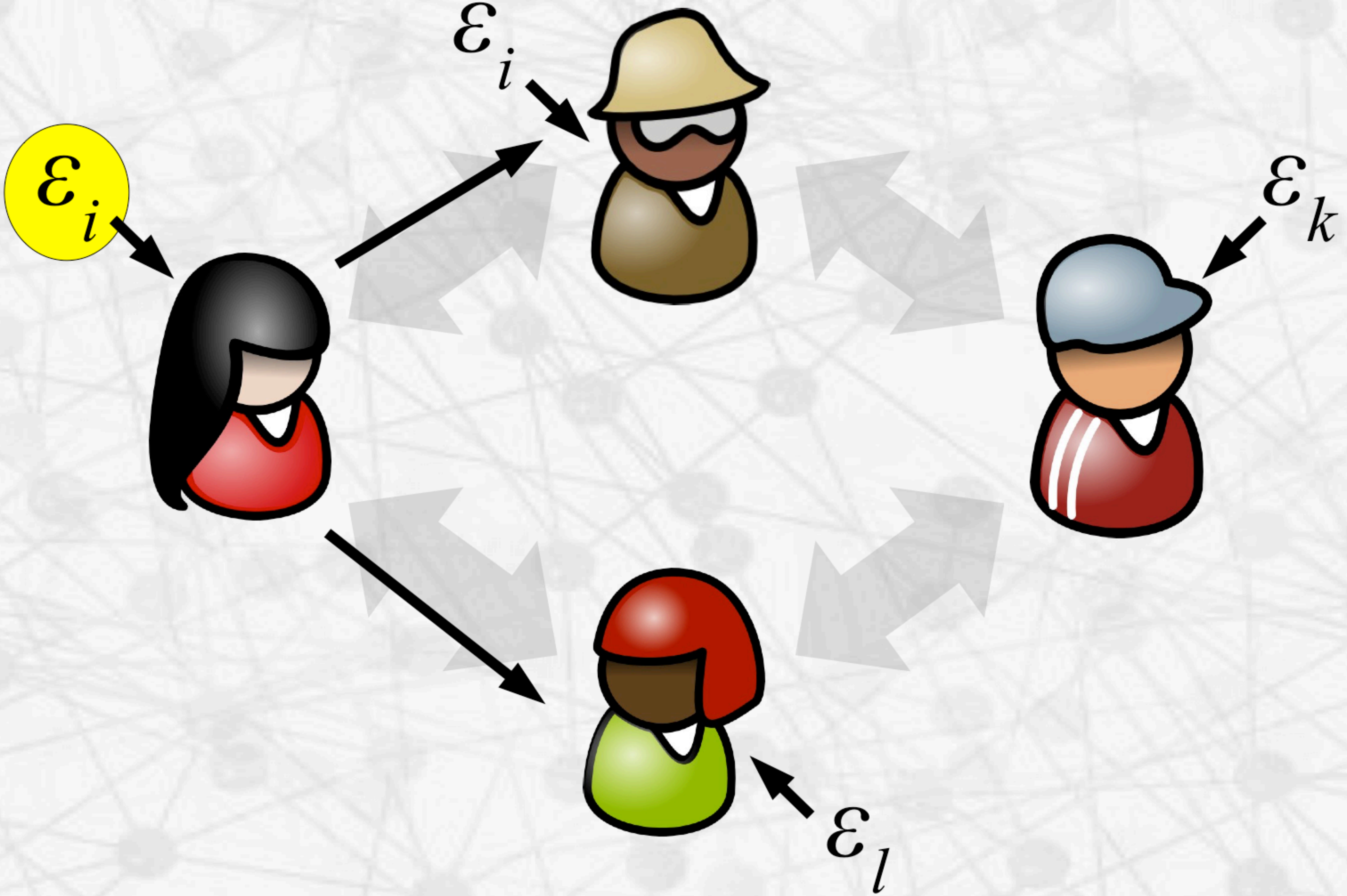


ε_i
ε_j
ε_k
$\mathbf{X}_j\beta$
$\mathbf{X}_i\beta$
$\mathbf{X}_k\beta$
y_j
y_i
y_k

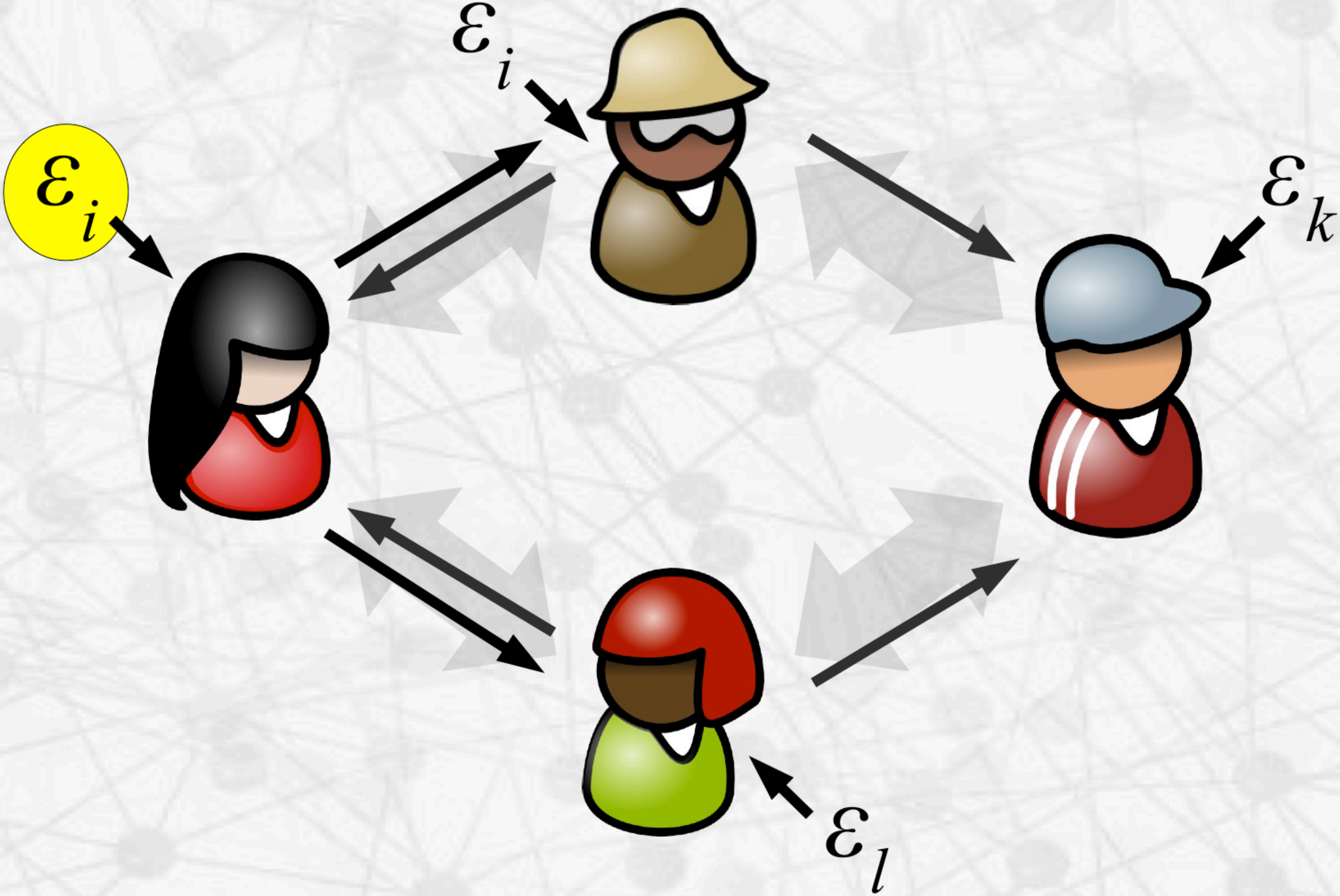
Network “Resonance”



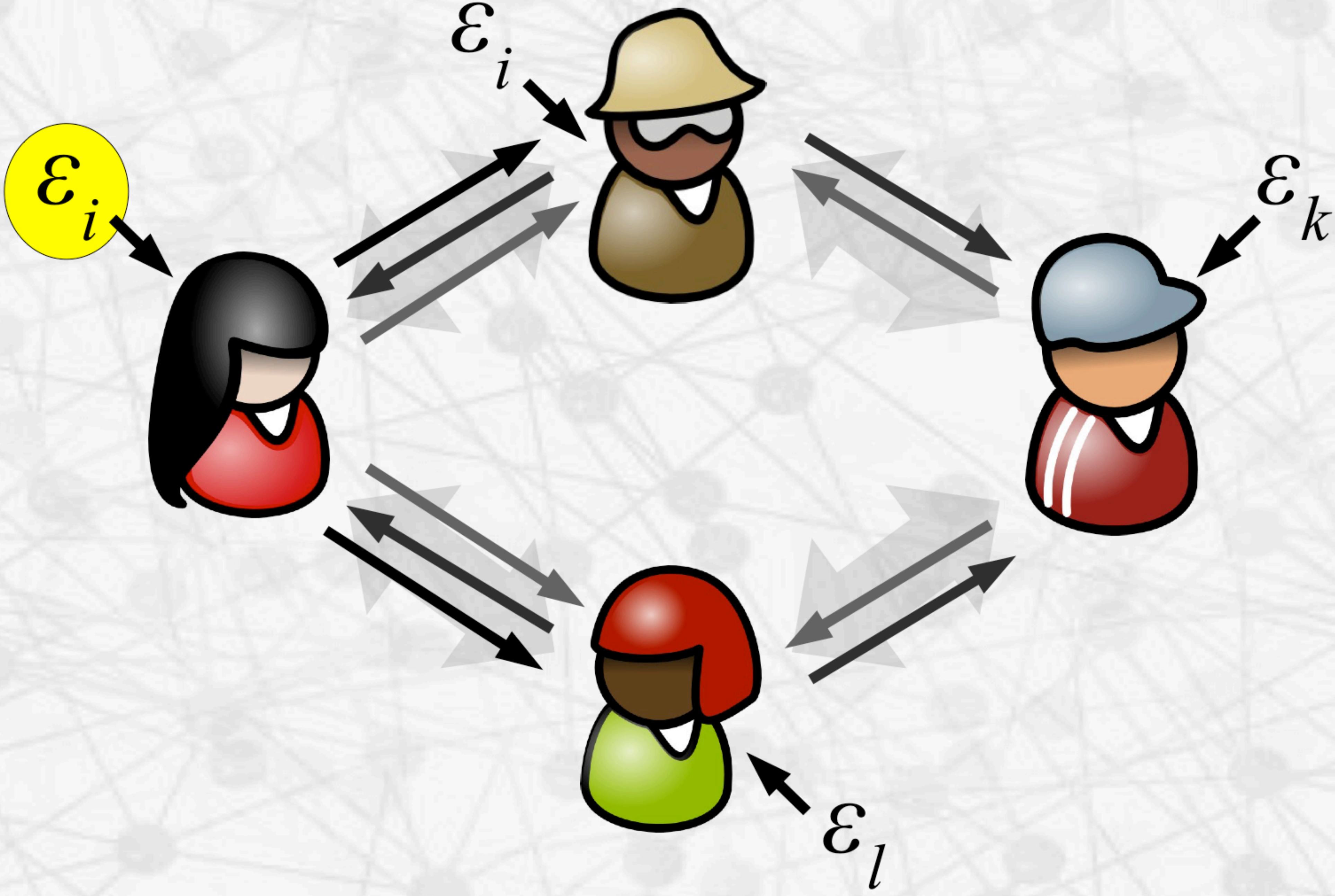
Network “Resonance”



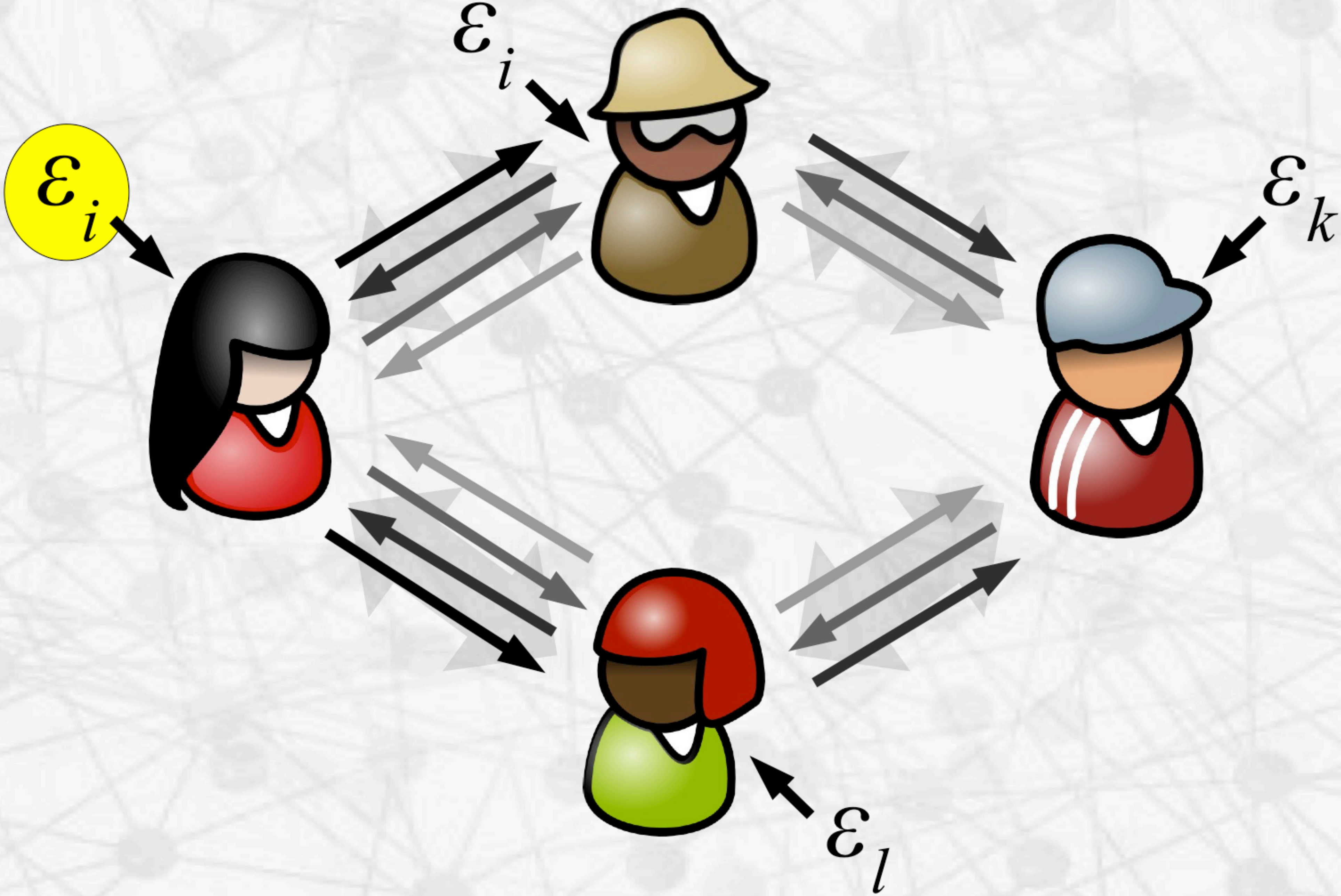
Network “Resonance”



Network "Resonance"



Network “Resonance”



The Formal Development

Individual Case Form:

$$y_i = \sum_{j=1}^{p} X_{ij} \beta_j + \epsilon_i$$

Group (Matrix/Vector) Form:

$$y = X \beta + \epsilon$$

The Formal Development

Individual Case Form:

Response

$$y_i = \sum_{j=1}^{p} X_{ij}\beta_j + \epsilon_i$$

Perturbation

Covariates

Group (Matrix/Vector) Form:

$$y = X\beta + \epsilon$$

The Formal Development

Individual Case Form:

$$y_i = \sum_{j=1}^{N} \theta \, W_{ij} \, y_j + \sum_{j=1}^{p} X_{ij} \beta_j + \epsilon_i$$

Group (Matrix/Vector) Form:

$$y = \theta \, W \, y + X \beta + \epsilon$$

The Formal Development

Individual Case Form:

AR Term

Covariates

$$y_i = \sum_{j=1}^{N} \theta\, W_{ij}\, y_j + \sum_{j=1}^{p} X_{ij}\, \beta_j + \epsilon_i$$

Response

Perturbation

Group (Matrix/Vector) Form:

$$y = \theta\, W\, y + X\, \beta + \epsilon$$

The Formal Development

Individual Case Form:

$$y_i = \sum_{j=1}^{N} \theta \ W_{ij} y_j + \sum_{j=1}^{p} X_{ij} \beta_j + \epsilon_i$$

$$\epsilon_i = \sum_{j=1}^{N} \psi \ Z_{ij} \epsilon_j + \nu_i$$

Group (Matrix/Vector) Form:

$$y = \theta \ W y + X \beta + \epsilon$$

$$\epsilon = \psi \ Z \epsilon + \nu$$

The Formal Development

Individual Case Form:

AR Term

Covariates

$$y_i = \sum_{j=1}^{N} \theta\, W_{ij}\, y_j + \sum_{j=1}^{p} X_{ij}\, \beta_j + \epsilon_i$$

Response

MA Term

$$\epsilon_i = \sum_{j=1}^{N} \psi\, Z_{ij}\, \epsilon_j + \nu_i$$

Total Perturbation

Individual Perturbation

Group (Matrix/Vector) Form:

$$y = \theta\, W\, y + X\, \beta + \epsilon$$

$$\epsilon = \psi\, Z\, \epsilon + \nu$$

Resolving the Resonance: Total Perturbations

$$\epsilon = \psi Z \epsilon + \nu$$

Resolving the Resonance: Total Perturbations

$$\epsilon = \psi \, Z \epsilon + \nu$$

$$\epsilon - \psi \, Z \epsilon = \nu$$

Resolving the Resonance: Total Perturbations

$$\epsilon = \psi Z \epsilon + \nu$$

$$\epsilon - \psi Z \epsilon = \nu$$

$$\left(\mathrm{I} - \psi Z\right) \epsilon = \nu$$

Resolving the Resonance: Total Perturbations

$$\epsilon = \psi Z \epsilon + \nu$$

$$\epsilon - \psi Z \epsilon = \nu$$

$$(\mathrm{I} - \psi Z)\epsilon = \nu$$

$$\epsilon = (\mathrm{I} - \psi Z)^{-1} \nu$$

Resolving the Resonance: Total Perturbations

$$\epsilon = \psi\ Z \epsilon + \nu$$

$$\epsilon - \psi\ Z \epsilon = \nu$$

$$\left(I - \psi\ Z\right) \epsilon = \nu$$

$$\epsilon = \left(I - \psi\ Z\right)^{-1} \nu$$

Total Perturbations

Matrix of "Equilibrium MA Weights"

Individual Perturbations

Resolving the Resonance, II: Responses

$$y = \theta W y + X \beta + \epsilon$$

Resolving the Resonance, II: Responses

$$y = \theta W y + X\beta + \epsilon$$

$$y - \theta W y = X\beta + \epsilon$$

Resolving the Resonance, II: Responses

$$\mathbf{y} = \theta\,\mathbf{W}\,\mathbf{y} + \mathbf{X}\,\beta + \epsilon$$

$$\mathbf{y} - \theta\,\mathbf{W}\,\mathbf{y} = \mathbf{X}\,\beta + \epsilon$$

$$\left(\mathbf{I} - \theta\,\mathbf{W}\right)\mathbf{y} = \mathbf{X}\,\beta + \epsilon$$

Resolving the Resonance, II: Responses

$$y = \theta W y + X\beta + \epsilon$$

$$y - \theta W y = X\beta + \epsilon$$

$$(I - \theta W) y = X\beta + \epsilon$$

$$y = (I - \theta W)^{-1} (X\beta + \epsilon)$$

Resolving the Resonance, II: Responses

$$y = \theta W y + X\beta + \epsilon$$

$$y - \theta W y = X\beta + \epsilon$$

$$(I - \theta W) y = X\beta + \epsilon$$

$$y = (I - \theta W)^{-1} (X\beta + \epsilon)$$

Responses

Matrix of "Equilibrium AR Weights"

Covariates

Total Perturbations

Resolving the Resonance, II: Responses

$$y = \theta W y + X\beta + \epsilon$$

$$y - \theta W y = X\beta + \epsilon$$

$$(I - \theta W) y = X\beta + \epsilon$$

$$y = (I - \theta W)^{-1} (X\beta + \epsilon)$$

$$\boxed{y = (I - \theta W)^{-1} \left(X\beta + (I - \psi Z)^{-1} \nu \right)}$$

Resolving the Resonance, II: Responses

Matrix of "Equilibrium MA Weights"

Matrix of "Equilibrium AR Weights"

Covariates

Individual Perturbations

Responses

$$\mathrm{y} = \left(\mathrm{I} - \theta\ \mathrm{W}\right)^{-1} \left(\mathrm{X}\beta + \left(\mathrm{I} - \psi\ \mathrm{Z}\right)^{-1} \nu\right)$$

Using Multiple Weight Matrices

$$\theta\ \mathrm{W} \rightarrow \sum_{i=1}^{w} \theta_i \mathrm{W}_{\mathrm{i}}$$

$$\psi\ \mathrm{Z} \rightarrow \sum_{i=1}^{z} \psi_i \mathrm{Z}_{\mathrm{i}}$$

Using Multiple Weight Matrices

$$\theta\,\mathrm{W} \rightarrow \sum_{i=1}^{w} \theta_i \mathrm{W}_\mathrm{i} \qquad\qquad \psi\,\mathrm{Z} \rightarrow \sum_{i=1}^{z} \psi_i \mathrm{Z}_\mathrm{i}$$

$$\mathrm{y} = \left(\mathrm{I} - \sum_{i=1}^{w} \theta_i \mathrm{W}_\mathrm{i}\right)^{-1} \left(\mathrm{X}\beta + \left(\mathrm{I} - \sum_{i=1}^{z} \psi_i \mathrm{Z}_\mathrm{i}\right)^{-1} \nu\right)$$

Inference with the Network Autocorrelation Model

- **In practice, rarely given parameters**
  - Usually observe $\mathbf{y}$, $\mathbf{X}$, and $\mathbf{W}$ and/or $\mathbf{Z}$, want to infer β, θ, ψ
  - Need each $\mathbf{I}$-$\mathbf{W}$, $\mathbf{I}$-$\mathbf{Z}$ invertible for solution to exist
- **Typical approach: maximum likelihood**
  - Generally take ν as iid, $\nu_i \sim N(0,\sigma^2)$
  - Standard errors based on the inverse information matrix at the MLE
  - Compare models in the usual way (e.g., AIC, BIC)
  - (Could also do this in a Bayesian way, could substitute other forms for ν, etc.)

Choosing the Weight Matrix

- **An important modeling issue: choosing the form for $\mathbf{W}$ and/or $\mathbf{Z}$**
  - Must satisfy invertibility condition
  - i,j cell must express the dependence of i on j
- **Many suggestions given by Leenders (2002)**
- **Ex: Friedkin-Johnson criteria**
  - Based on small-group social influence experiments
  - Induce quasi-convexity in equilibrium attitudes
  - $\mathbf{W}_{ij} \geq 0$, $\mathbf{W}_{ij} \leq 1$
  - $\sum_j \mathbf{W}_{ij} \leq 1$
  - Diagonal can be >0

Important Errors to Avoid

- **Using NAMs when you need NLIs (or vice versa)**
  - Am I affected by my *position*, or by the *properties of my neighbors*?
  - Nearly all influence theories imply some sort of autocorrelation; ditto for some theories of competition
- **Confusing correlation with influence**
  - If influence occurs through direct ties, structurally equivalent positions will correlate; NAM will show an SE effect
- **Confusing AR and MA terms**
  - AR terms imply diffusion of response values, MA terms imply diffusion of perturbations
  - Do my neighbors' covariates affect me? If so, this is an AR process! (If not, it's an MA process.)
- **Forgetting selection effects**
  - NAM assumes that weight matrix is fixed
  - If selection is present, NAM results will be misleading