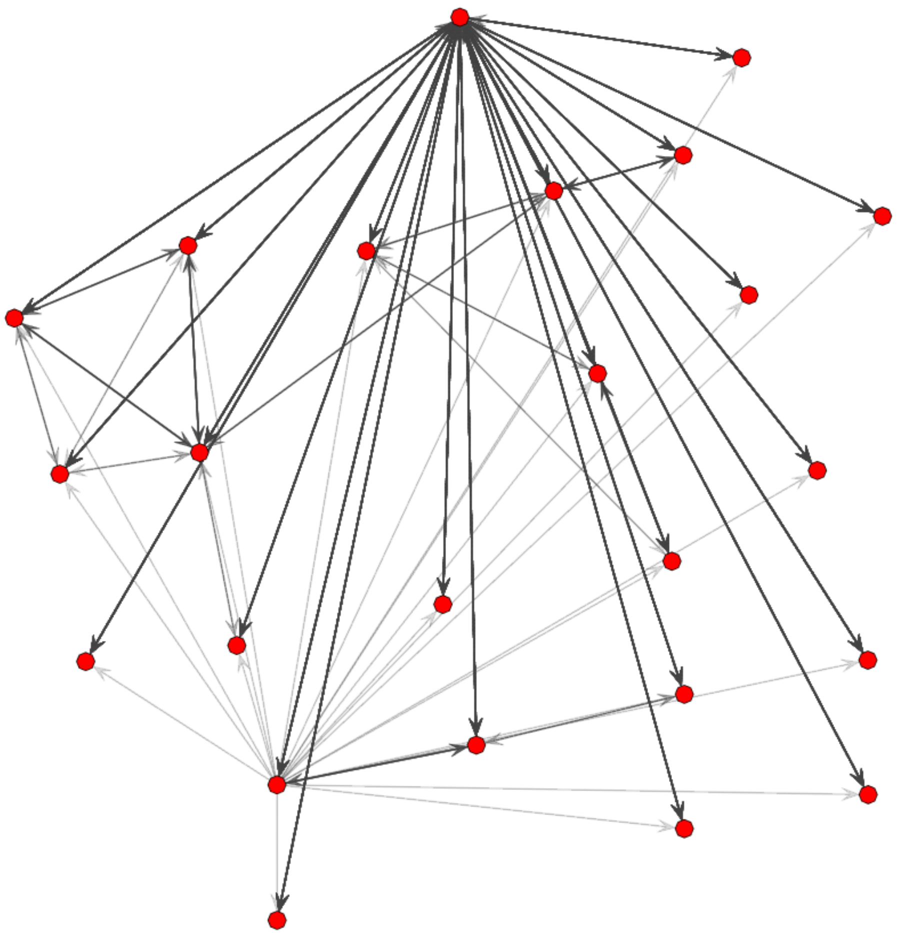
Multi-level Models for Classroom Dynamics

Christopher DuBois

Padhraic Smyth, UC Irvine
Carter Butts, UC Irvine
Nicole Pierski, UC Irvine
Dan McFarland, Stanford

The Data



High school interactions (McFarland 2001)

650 classroom sessions

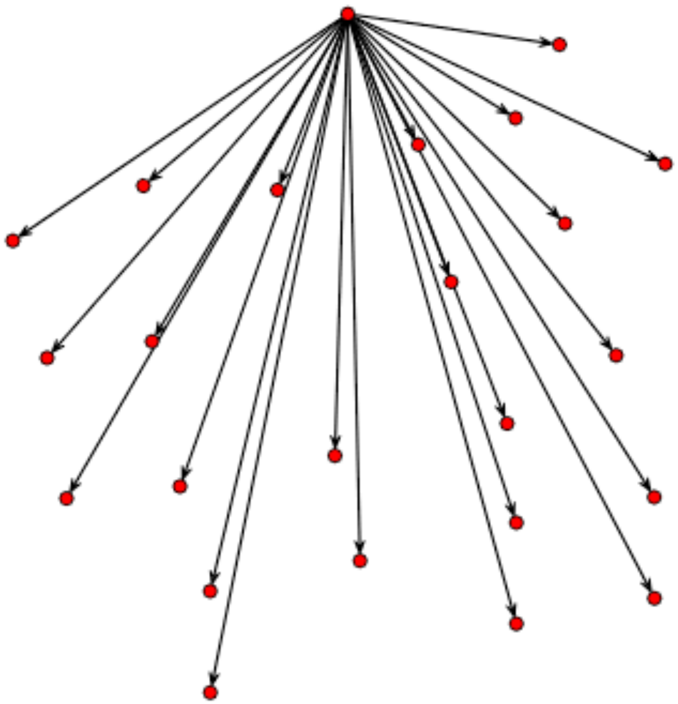
Covariates about class

- e.g. subject, teachers

Covariates about individuals

- e.g. race, extracurriculars

The Data



Nodes arranged (roughly) according to seating chart

Teacher interactions common

Local interactions common

Goals

Describe how the probability of each interaction varies with a set of covariates

Pull apart relative contribution of:

- actor covariates
- current context
- conversational dynamics

Make inferences about event sequences:

- within classroom sessions
- across classroom sessions

Long-term question:

- Given covariates about a classroom, can we predict aspects of the dynamics? (e.g. amount of reciprocity in interactions)

Notation

$\lambda_{ij}(t)$ Rate/hazard at time t of the interaction initiated by individual i and directed towards j

$x_{ij}(t)$ Covariates about interaction (i,j) at time t

Use a (positive) linear predictor to model the hazards:

$$\lambda_{ij}(t) = \exp\{\beta^T x_{ij}(t)\}$$

- Hazards depend on past history and covariates.
- Include rates that individuals "broadcast" to entire classroom.

Model specification: $x_{ij}(t)$

Sender/recipient effects:

- race
- gender
- is_teacher

Event effects:

- teacher_student
- teacher_broadcast
- are_friends
- number_shared_activities

Participation shifts (Gibson 2003)

- Reciprocity (AB-BA)
- Turn taking (AB-BY)
- Others...

"Autocorrelation":

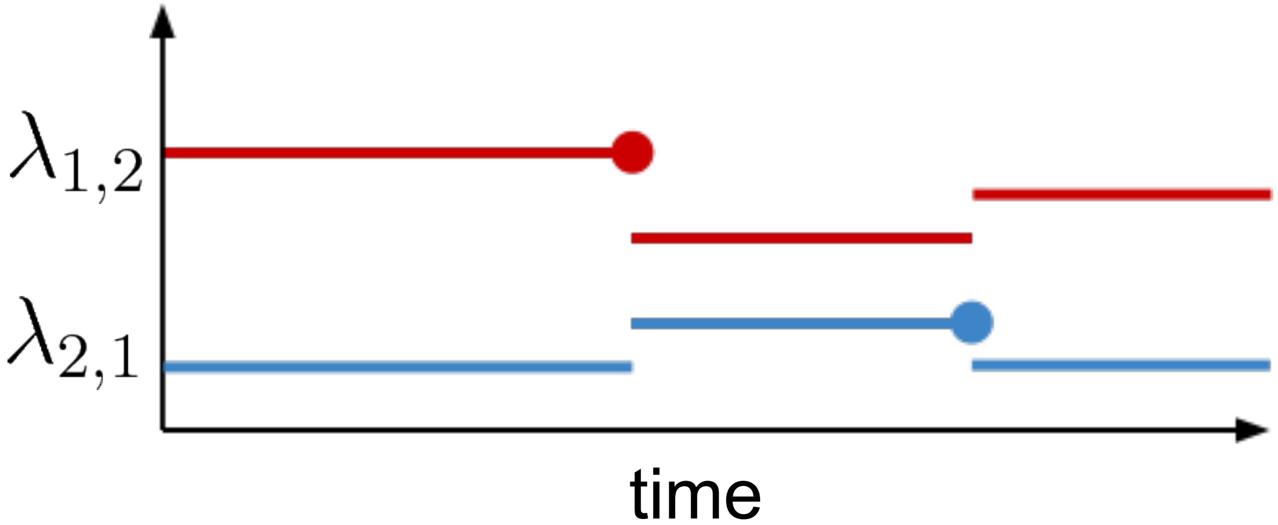
- recency (sender/receiver) (e.g. rank of individual in list of most recent)
- current event and previous event are both (teacher,broadcast)

"Context" of event:

- Lecture
- Silent time
- Groupwork

Model

Assume $\lambda_{ij}(t)$ is constant between events.



Model

Full likelihood for history of events:

$$p(\{(i_k, j_k, t_k)\}_{k=1,\ldots,M} | \beta, X) =$$

Model

Full likelihood for history of events:

$$p(\{(i_k, j_k, t_k)\}_{k=1,\ldots,M} | \beta, X) =$$

$$\prod_{k=1}^{M} \lambda_{i_k, j_k}(t_k) \prod_{ij} \exp\{-(t_k - t_{k-1})\lambda_{ij}(t_k)\}$$

Model

Full likelihood for history of events:

$$p(\{(i_k, j_k, t_k)\}_{k=1,\ldots,M} | \beta, X) =$$

$$\prod_{k=1}^{M} \boxed{\lambda_{i_k, j_k}(t_k)} \prod_{ij} \exp\{-(t_k - t_{k-1})\lambda_{ij}(t_k)\}$$

Hazard of k'th observed event

Model

Full likelihood for history of events:

$$p(\{(i_k, j_k, t_k)\}_{k=1,\ldots,M} | \beta, X) =$$

$$\prod_{k=1}^{M} \lambda_{i_k,j_k}(t_k) \prod_{ij} \exp\{-(t_k - t_{k-1})\lambda_{ij}(t_k)\}$$

Survival function for each event, representing the fact that no event occurred *between* event k-1 and event k

Model

Mapping to standard survival analysis methods:

- **Risk set**: all possible interactions among individuals
- Covariates are **time-varying** (and dependent on all previous events)
- Each event:
  - one observed **failure time**
  - times for other events are **censored**

Alternative perspectives:

- Continuous time process with N^2 states

Modeling Several Sequences

Parameter estimation:

- Can use standard techniques (e.g. Newton-Rapheson) to obtain maximum likelihood estimates

Problem:

- Some event sequences have few events
- Some effects may have few relevant events

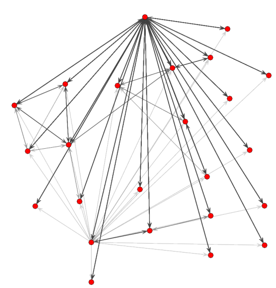
Today's approach:

- Share information across classroom sessions via a hierarchical model

Modeling Event Sequences

β — Event model parameters

↓

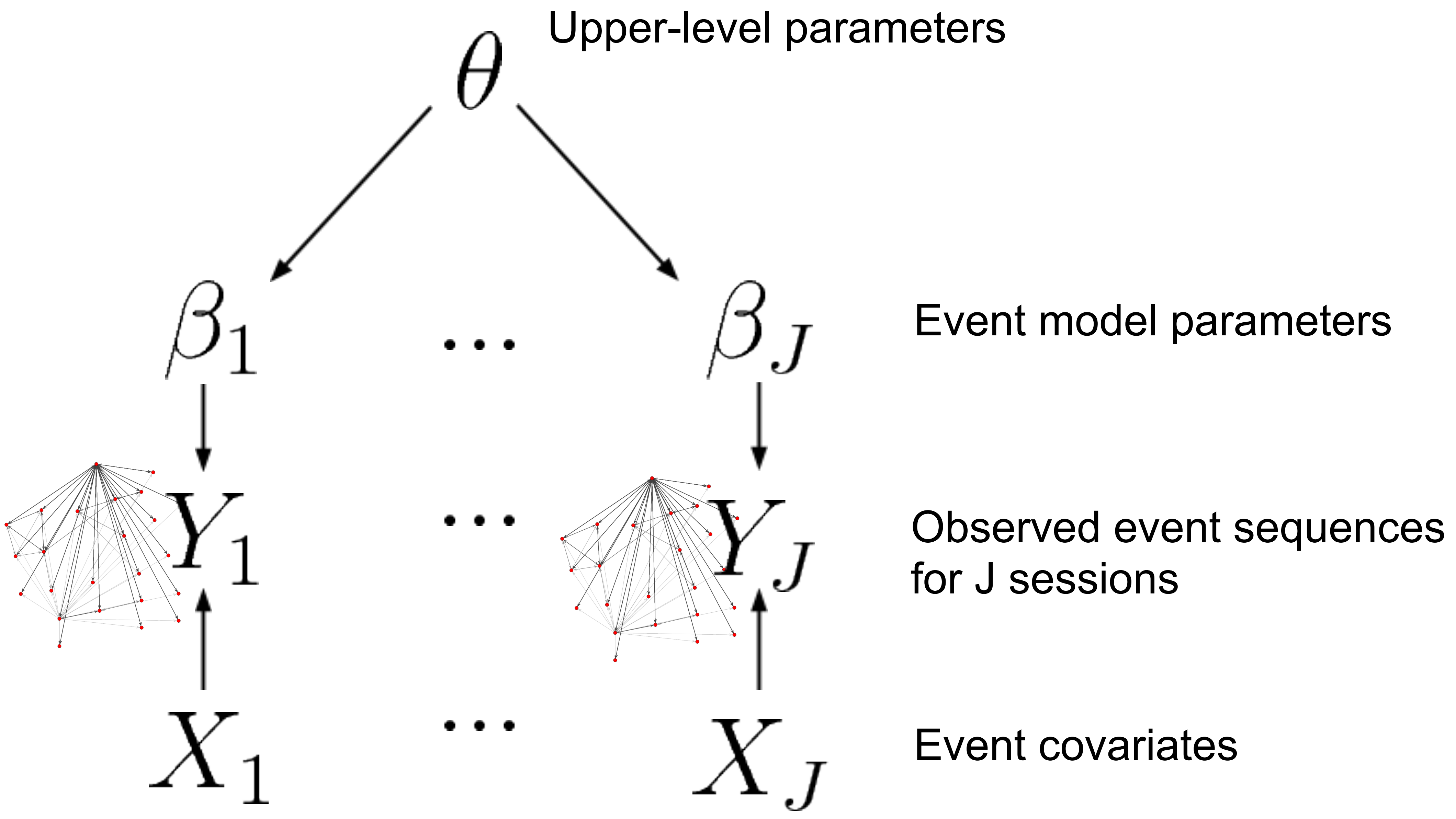


Y_j — Observed event sequence for session j

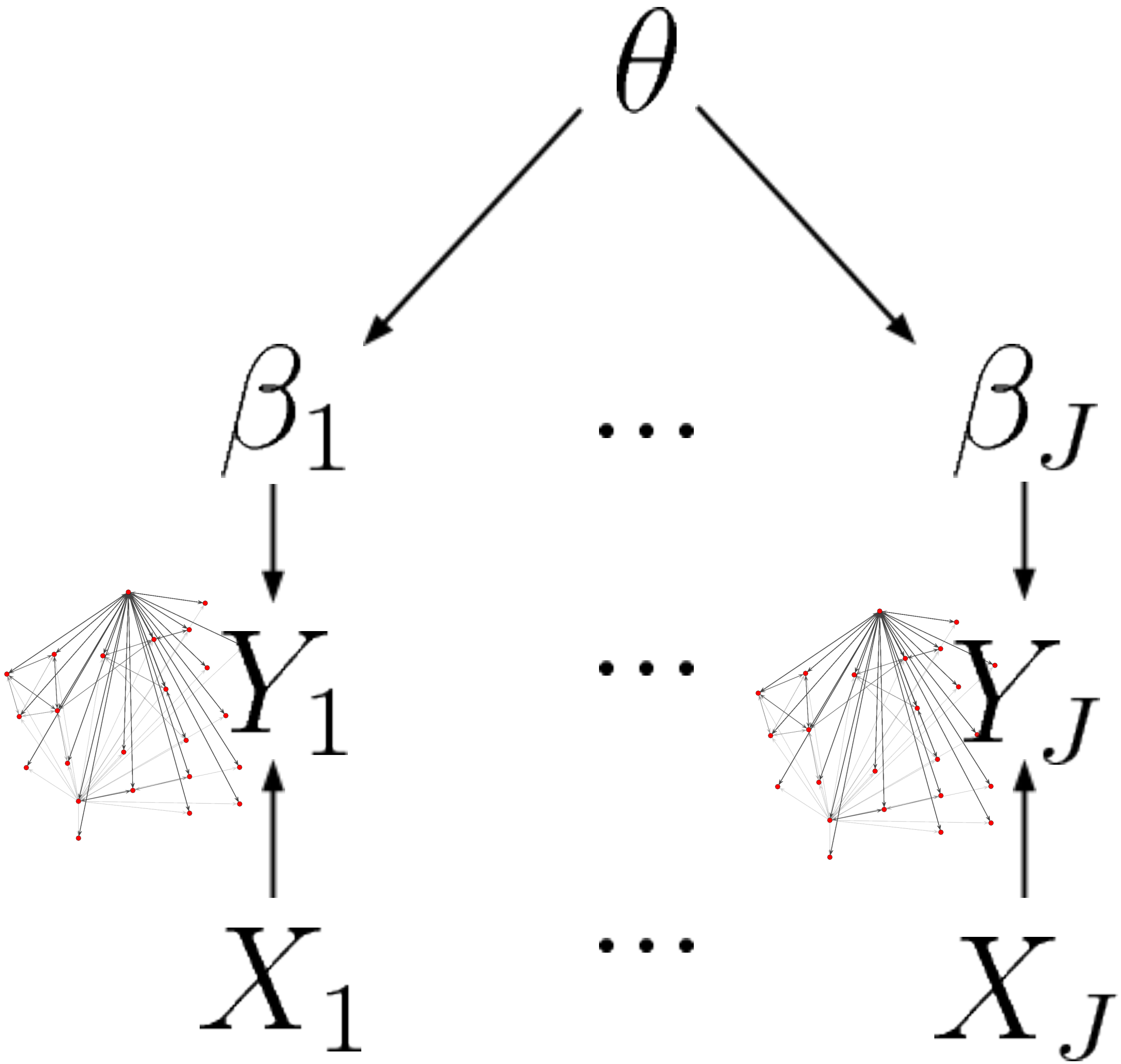
↑

X_j — Event covariates

Multilevel Relational Event Model



Multilevel Relational Event Model



$$\beta_{pi} \sim N(\theta_p, \sigma_p^2)$$

$$p(\theta_p, \sigma_p) \propto \frac{1}{\sigma_p}$$

Observed event sequences for J sessions

Event covariates

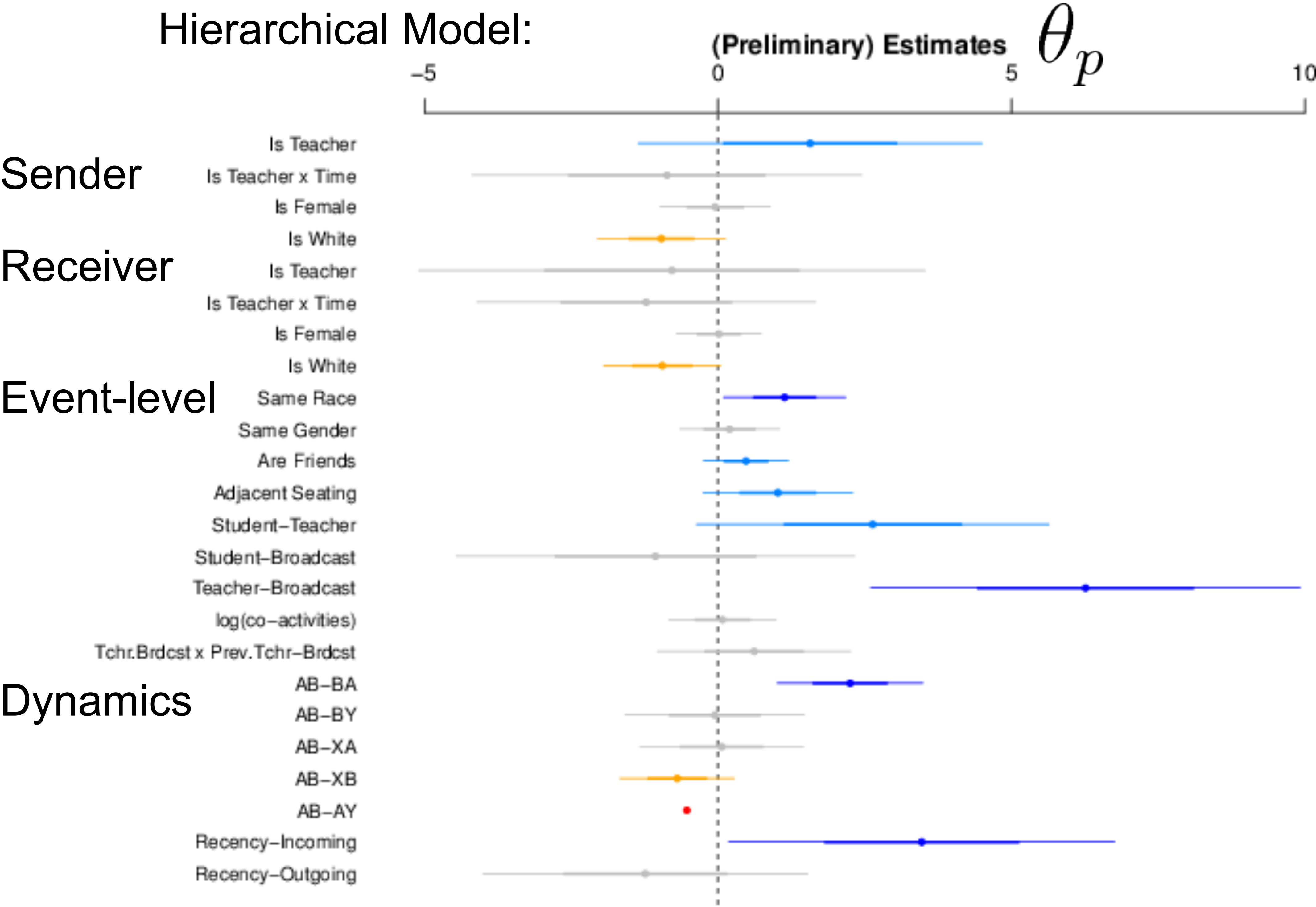
Inference

Iterated conditional modes (ICM):

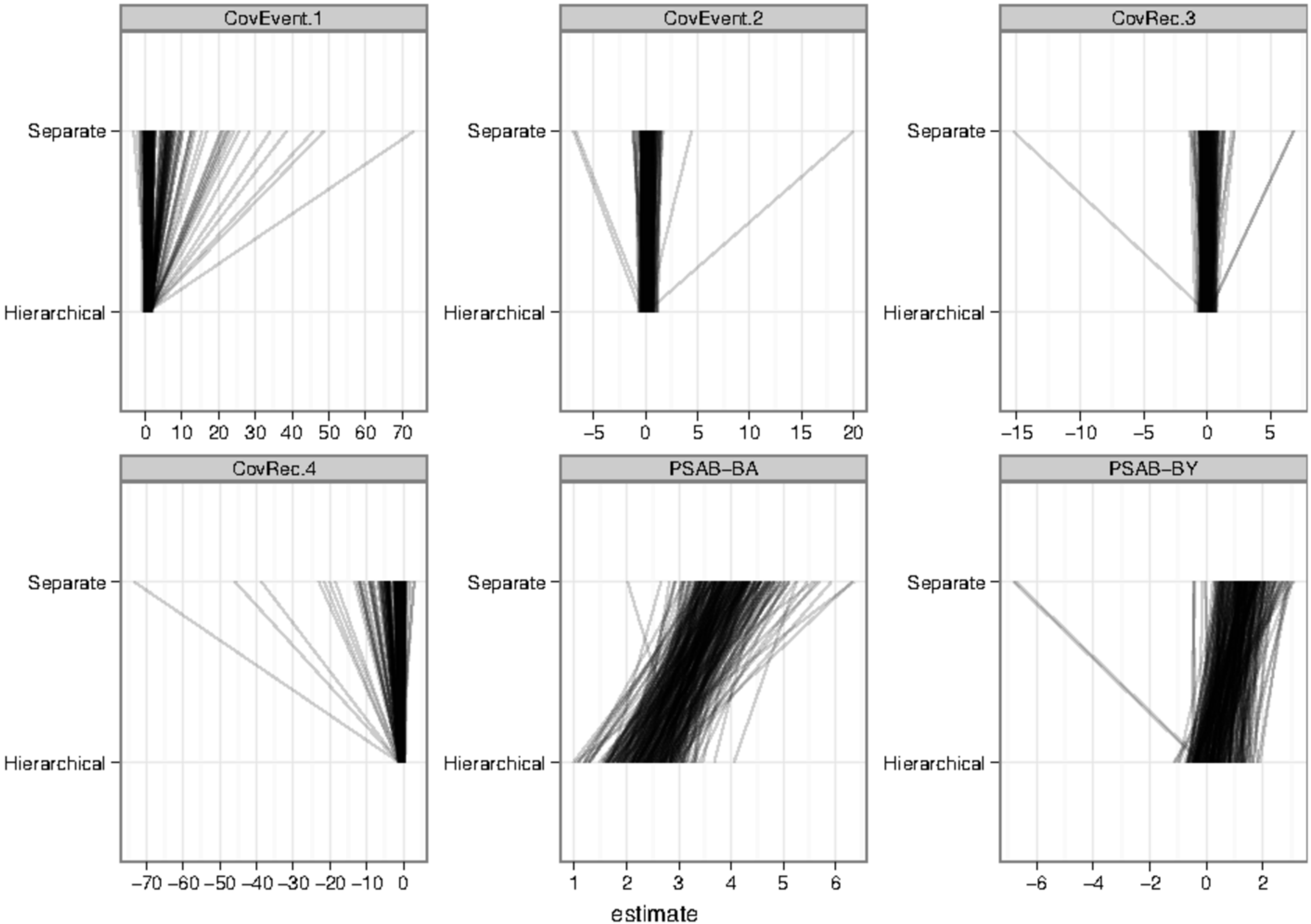
- Fit individual models to obtain beta for each session that maximizes the log posterior
- Obtain estimates for the upper-level model theta conditioned on the betas
- Iterate using theta as initial estimates for each beta.

Draw samples from posterior centered at mode via MH.

Hierarchical Model:

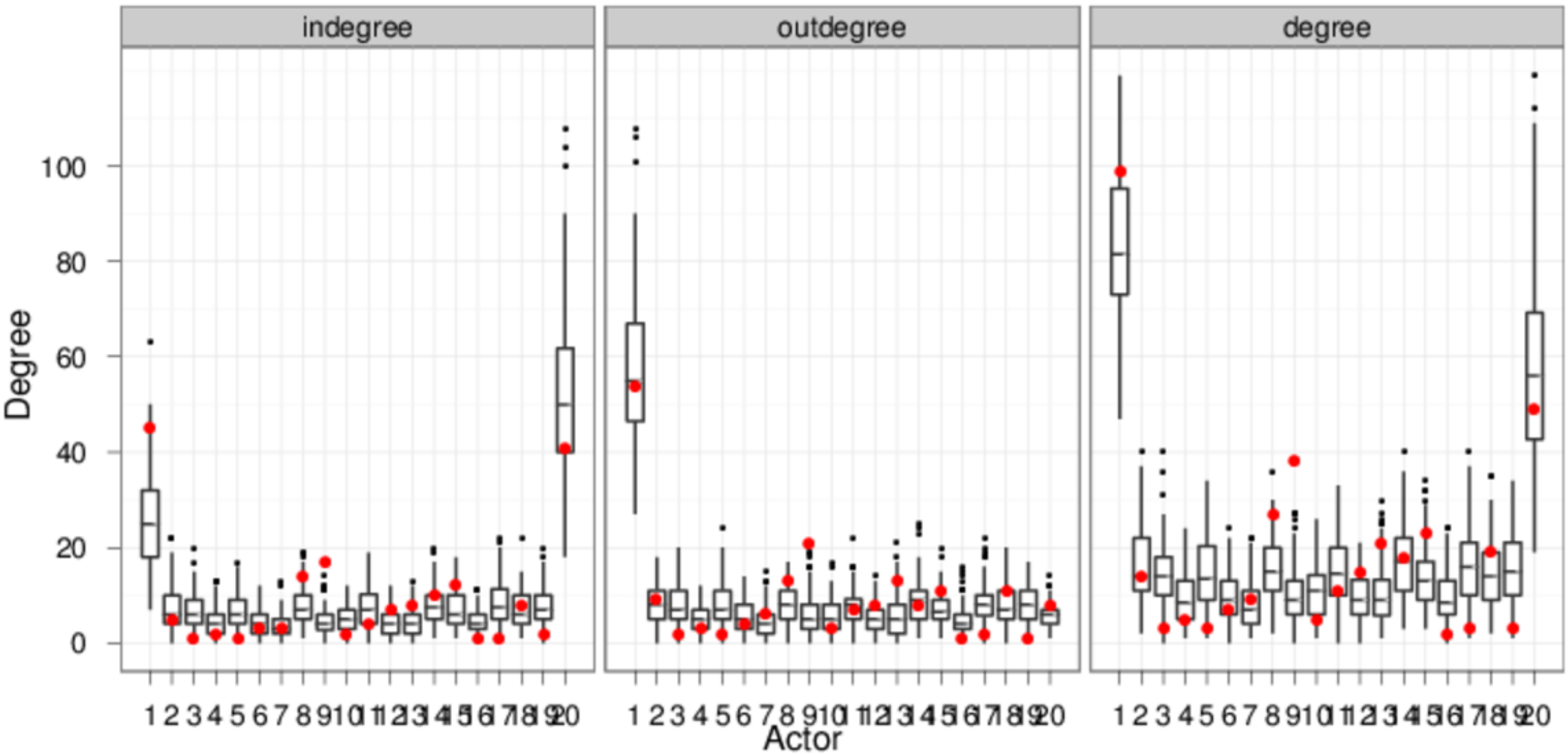


Shrinkage

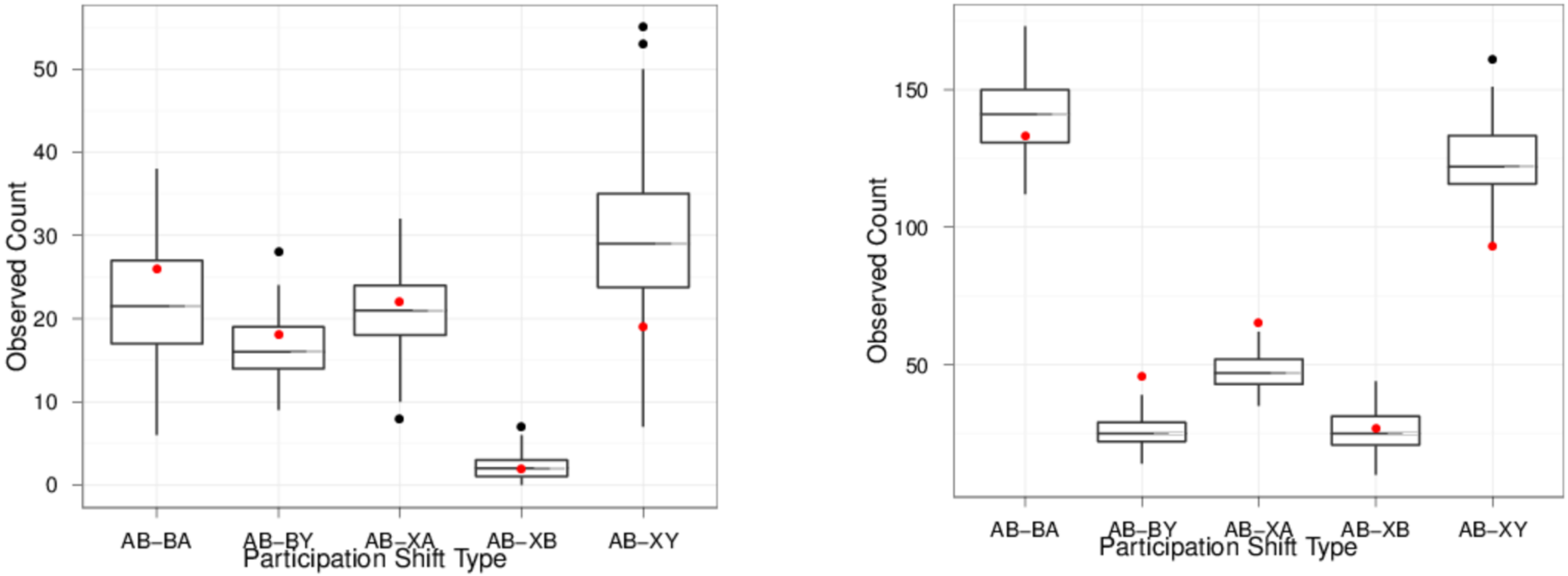


CovEvent.1
CovEvent.2
CovRec.3
CovRec.4
PSAB−BA
PSAB−BY
Separate
Hierarchical
0 10 20 30 40 50 60 70
−5 0 5 10 15 20
−15 −10 −5 0 5
−70 −60 −50 −40 −30 −20 −10 0
1 2 3 4 5 6
−6 −4 −2 0 2
estimate

Posterior-predictive checks: Degree



Posterior-predictive checks: "P-shifts"



Comparing p-shift statistics of observed data and data simulated using the parameter estimates for two classroom sessions.

Takeaways and future directions

Proof of concept:

- Can model event data using actor covariates and conversational dynamics
- Hierarchical modeling useful in this setting
- Can begin to ask questions at the network level:
  use models of observed networks to generalize to new networks

How do dynamics depend on the "context" of event?

- Lecture, Silent time, Groupwork

Multilevel modeling with session-level covariates:

- racial mixture
- survey results about the classroom session

Takeaways and future directions

Predictive evaluation:

- Predict out-of-sample events within a classroom
- Predict out-of-sample session information

"Big Data":

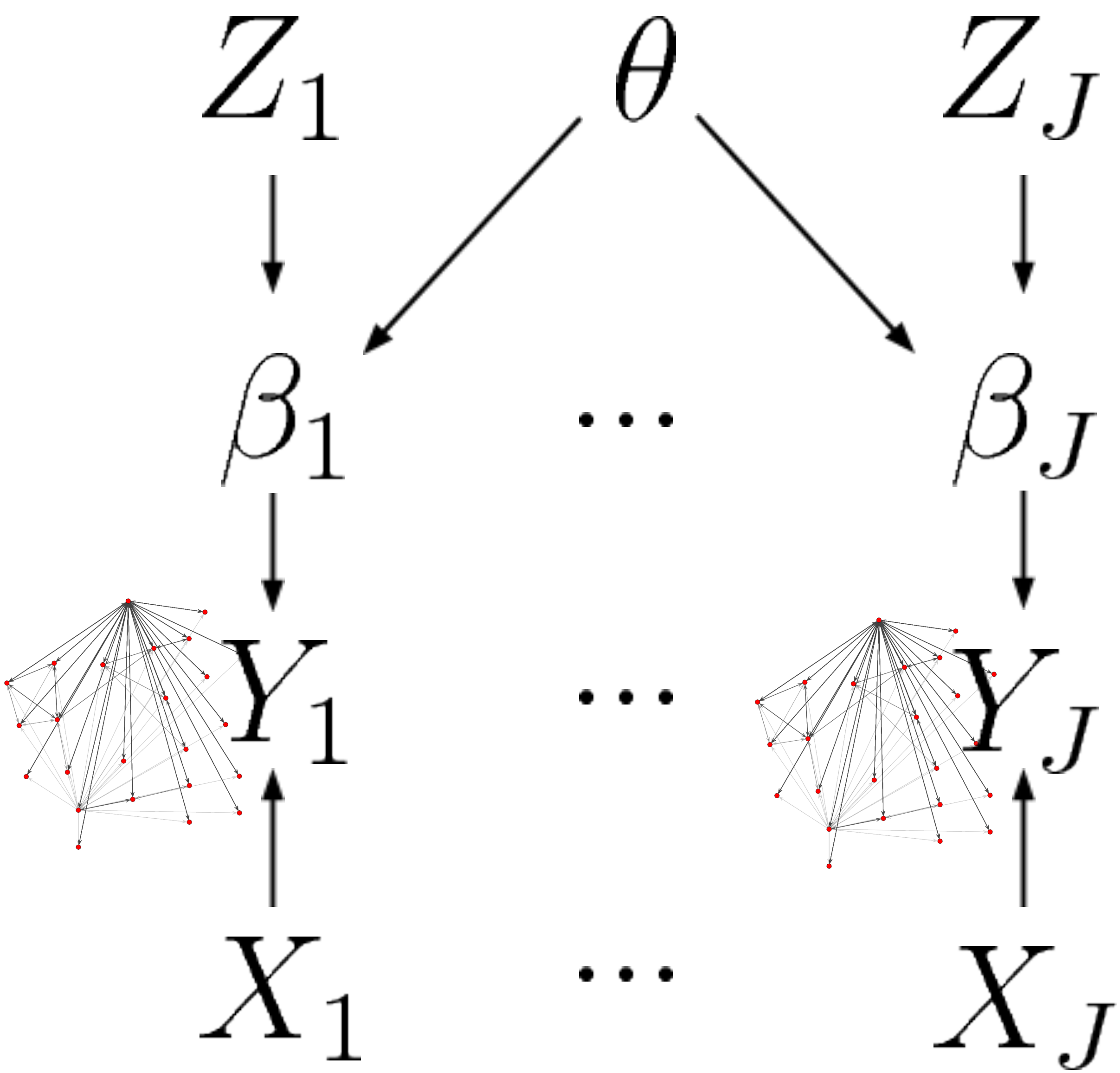
- Likelihood computations are intensive.
- Small group dynamics (~20 actors),
  but many networks (~280-600), many effects (~10-30)

What does the model predict?

- Simulate ramifications (like in agent-based modeling)

Thank you

Multilevel Relational Event Model



Session-level covariates

Event model parameters

Observed event sequences for J sessions

Event covariates

Model

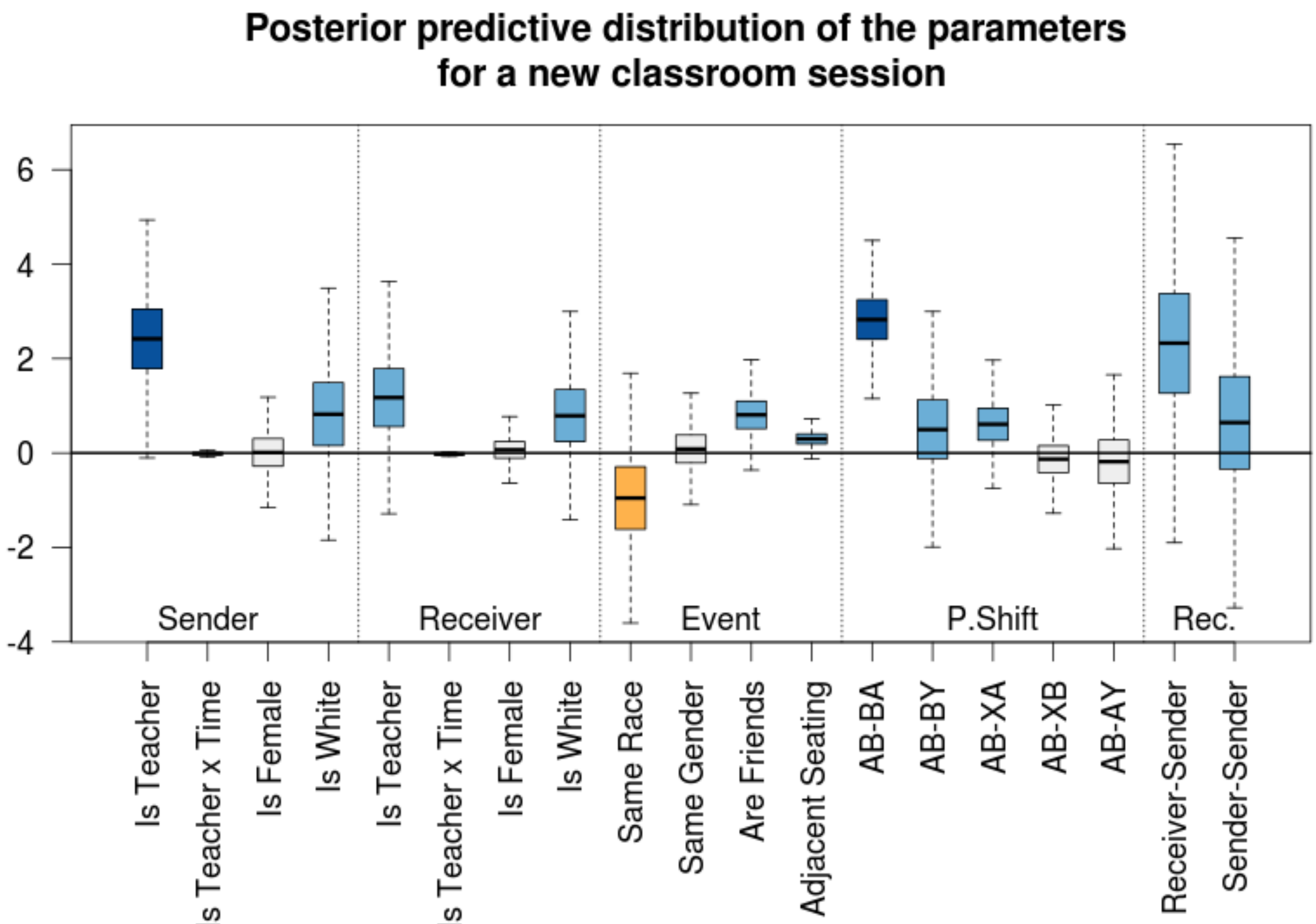
Partial likelihood for sequence of events A:

$$p(A|\beta) = \prod_{k=1}^{M} \frac{\exp\{\beta^T x_{a_k}(t_k)\}}{\sum_{a' \in R} \exp\{\beta^T x_{a'}(t_k)\}}$$

For each event, k: P(next event is a=(i,j) | some event occurs)

Alternatively, can consider a full likelihood where inter-arrival times have a parametric form (e.g. exponential).

Multilevel Relational Event Model



Outline

Data

Goals

Model
- Likelihood
- Specification
- Hierarchical extension
- Inference

Preliminary results

Future directions