A Probabilistic Model for Using Social Networks in Personalized Item Recommendation

Allison J.B. Chaney

Princeton University

Tina Eliassi-Rad

Rutgers University

David M. Blei

Columbia University

ajbc.io/spf

Personalized Item Recommendation

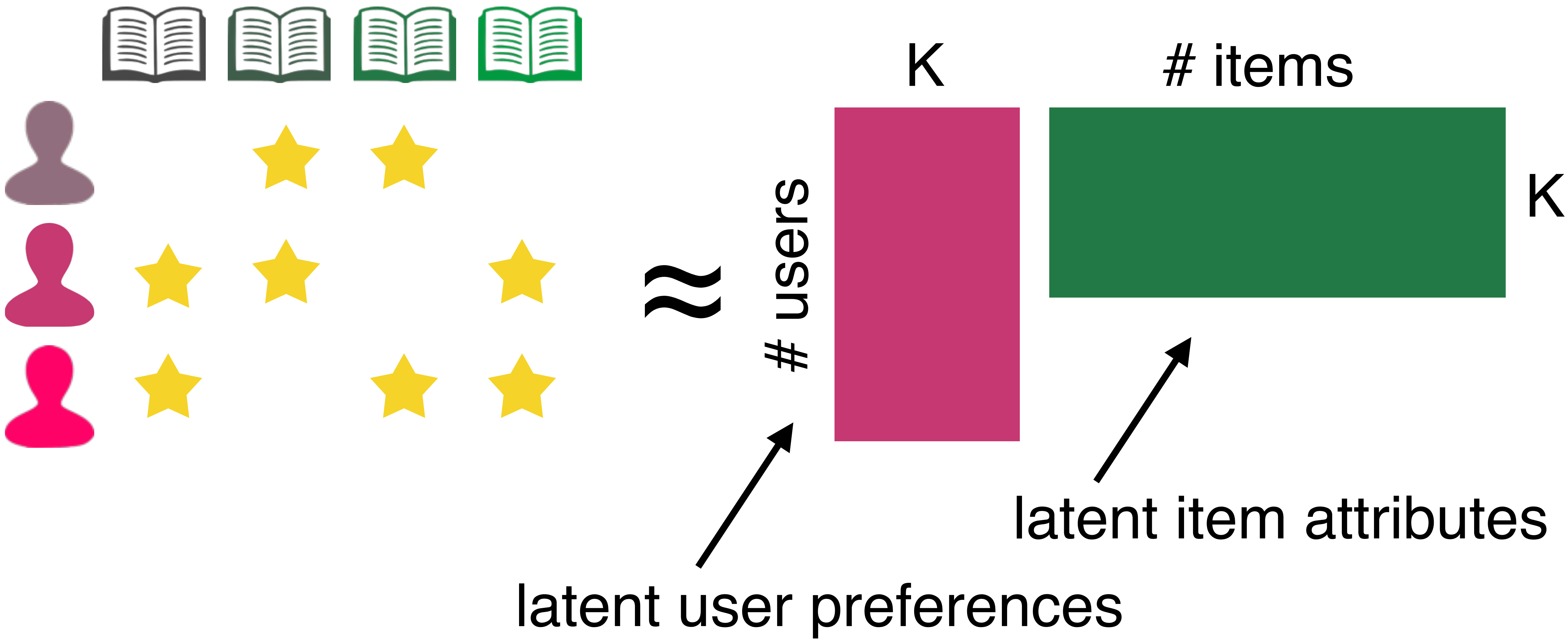
Anna Karenina

Winter’s Tale

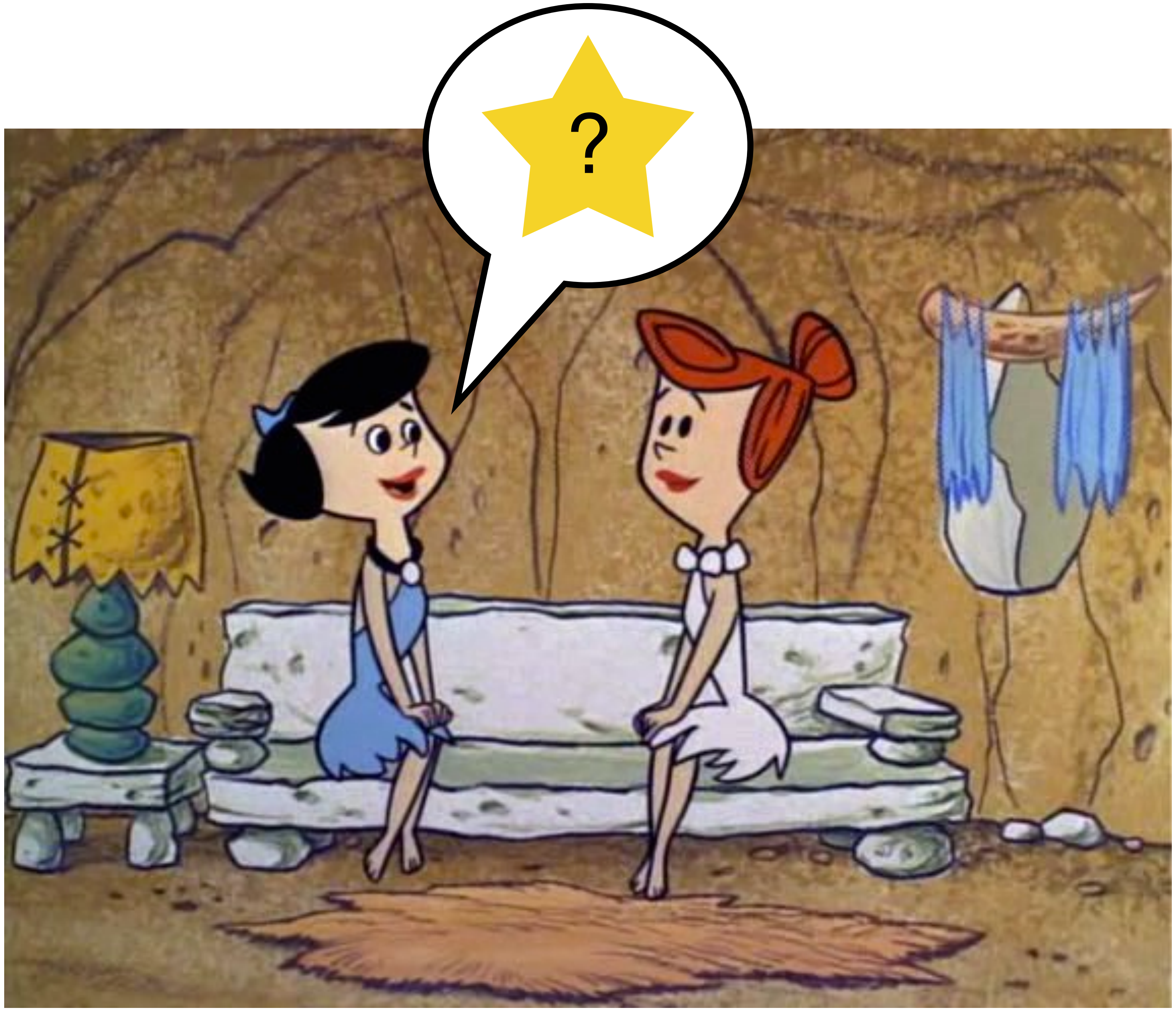
East of Eden

???

Matrix Factorization



Including Social Networks



Including Social Networks

- Matches our intuition

Including Social Networks

- Matches our intuition
- Introduces explainable serendipity

Including Social Networks

- Matches our intuition
- Introduces explainable serendipity
- Improves performance

Including Social Networks

- Matches our intuition
- Introduces explainable serendipity
- Improves performance
- Helps us learn about user behavior

An Example Etsy User

An Example Etsy User

An Example Etsy User

An Example Etsy User

An Example Etsy User

An Example Etsy User

An Example Etsy User

An Example Etsy User

An Example Etsy User

An Example Etsy User





An Example Etsy User







An Example Etsy User

An Example Etsy User









An Example Etsy User

An Example Etsy User

observed data

ratings

network

inference algorithm

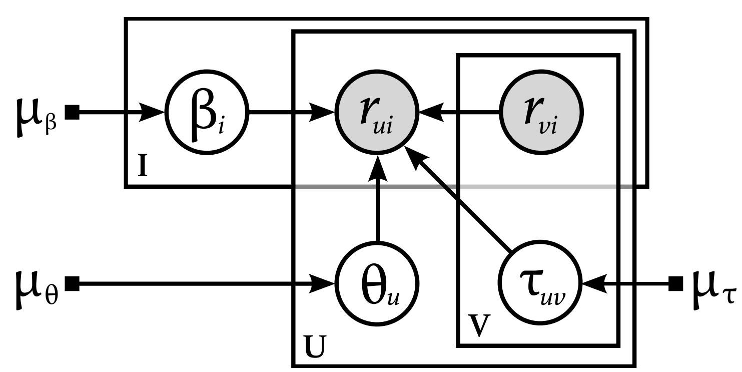
learned parameters

item attributes

user preferences

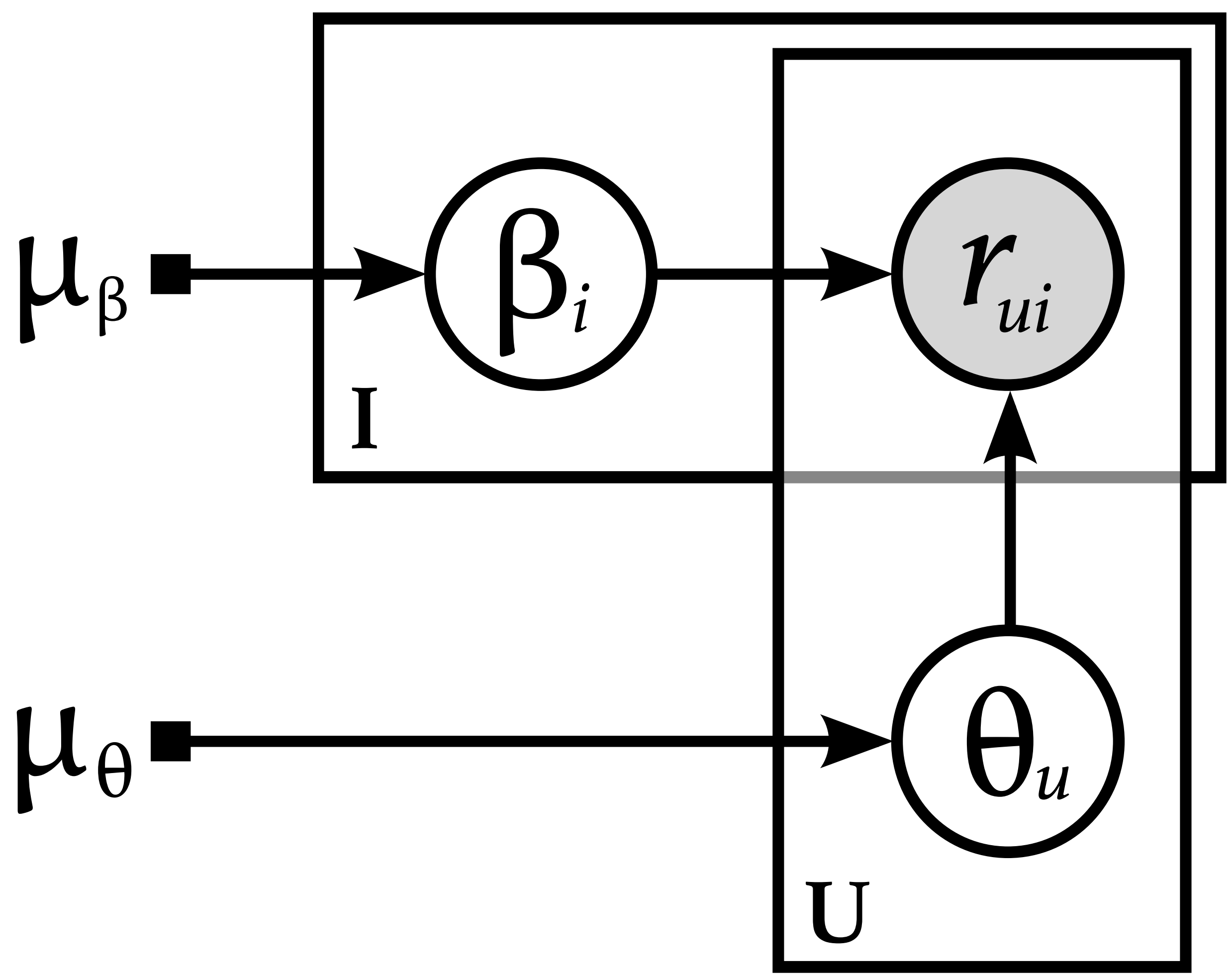
user influence

model assumptions

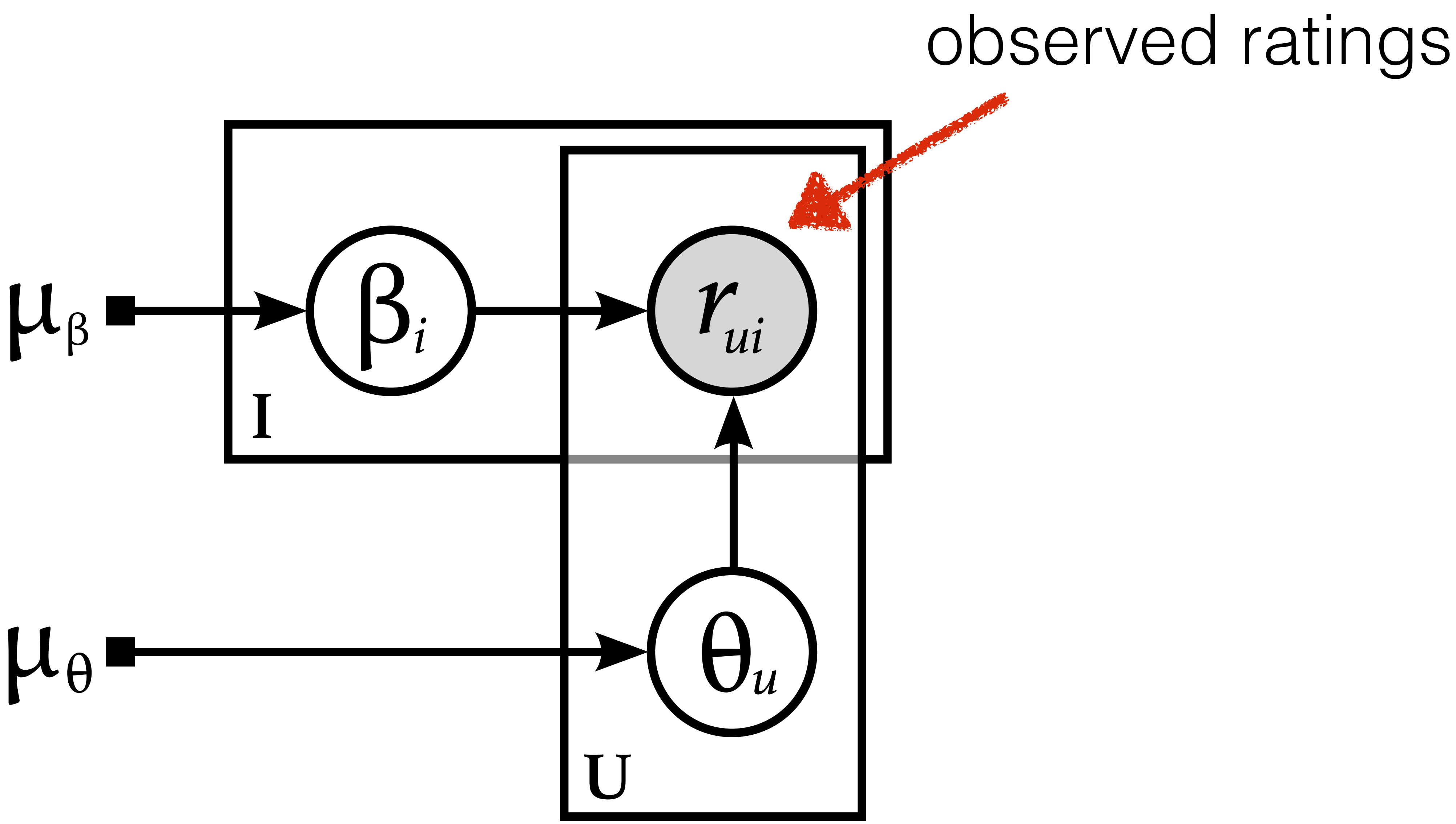


recommendations

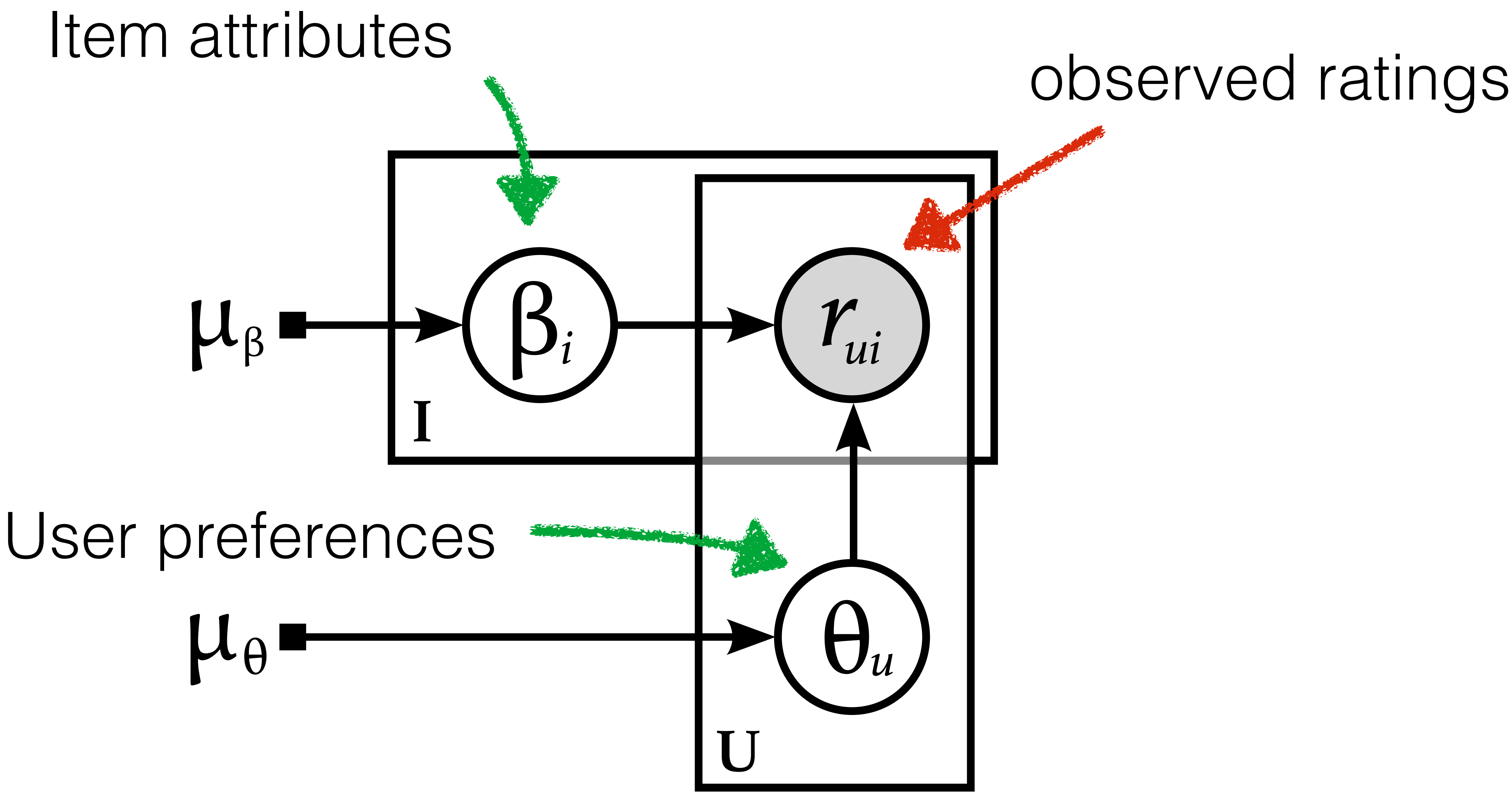
Matrix Factorization



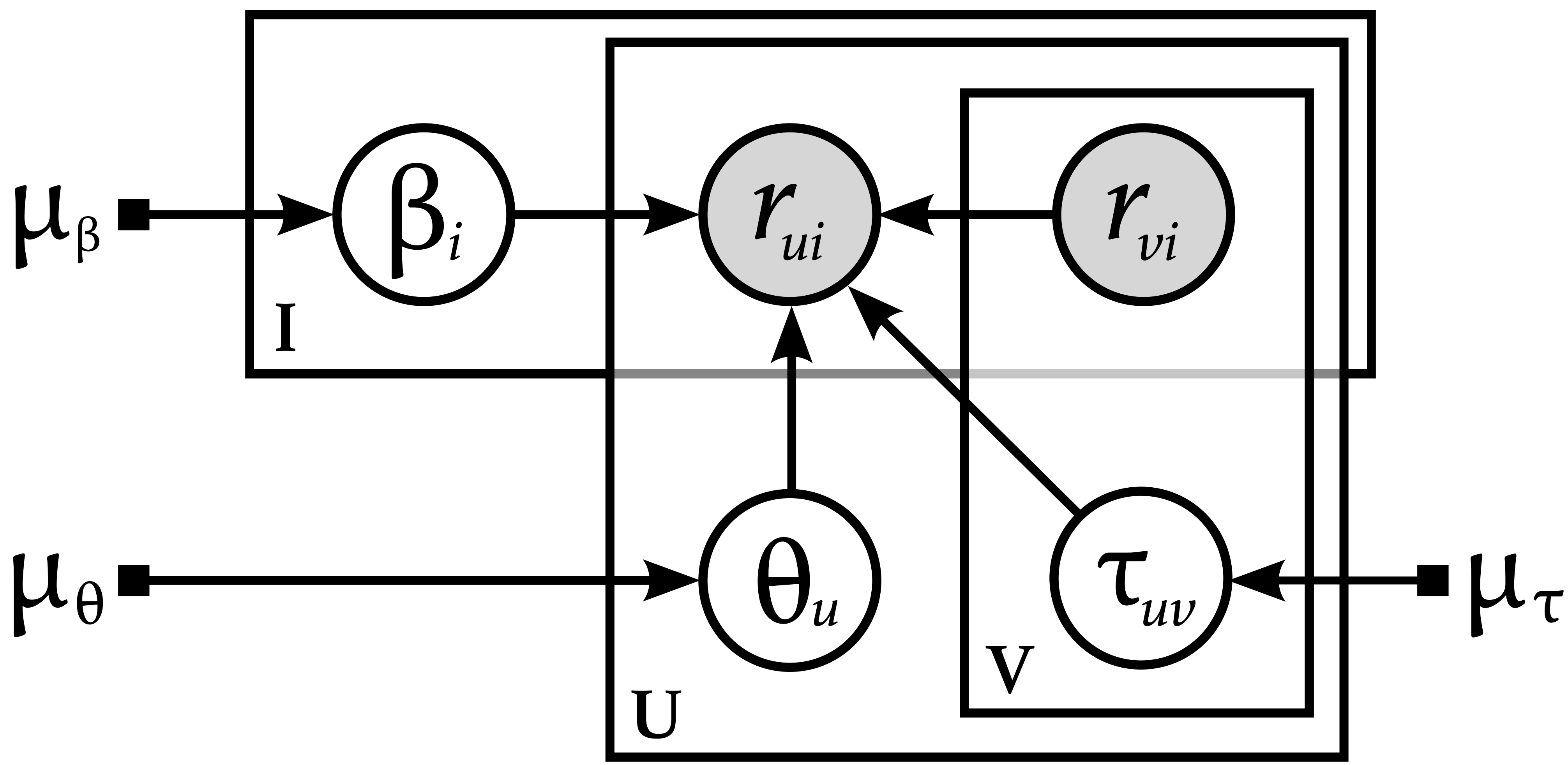
Matrix Factorization



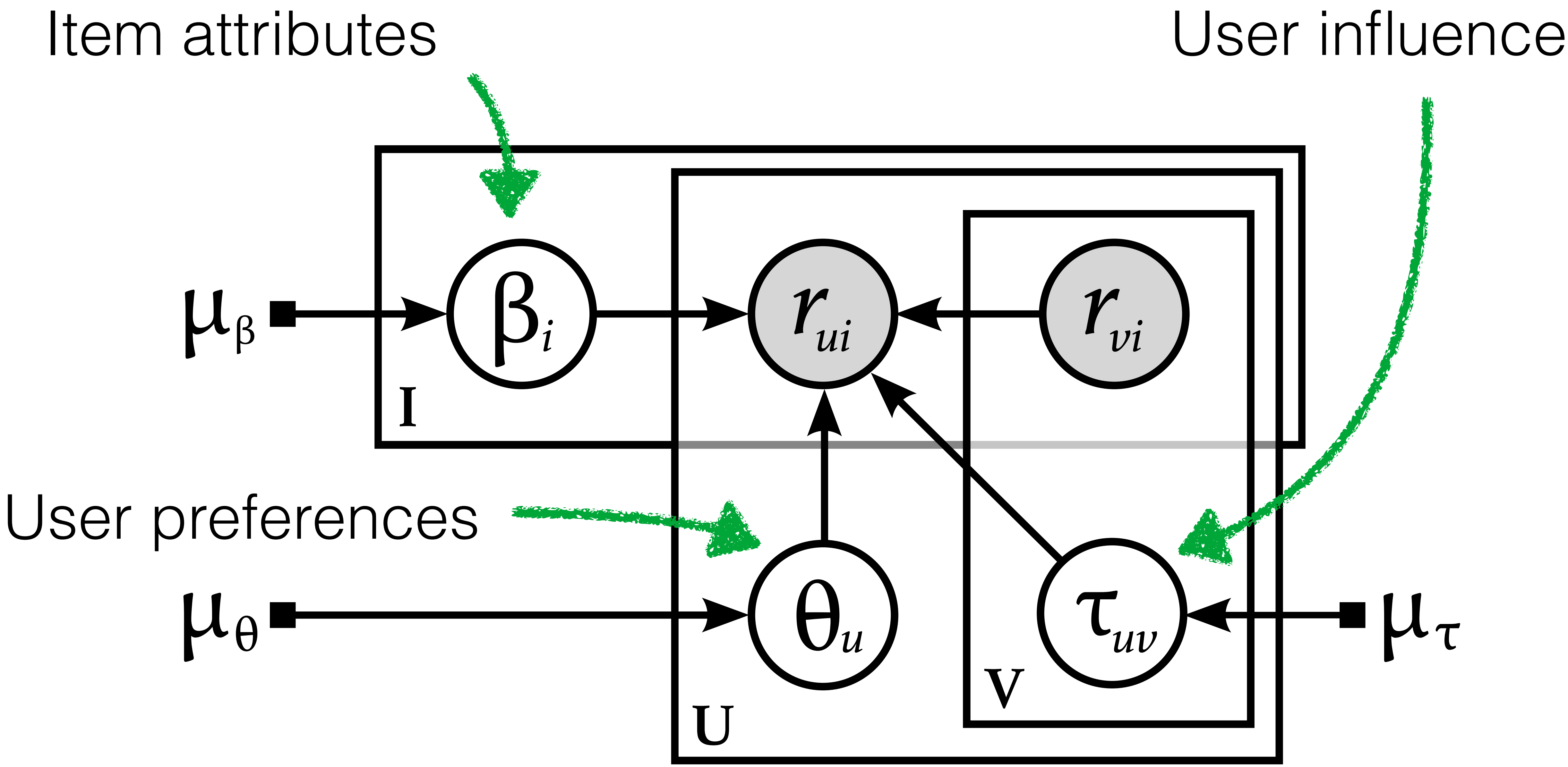
Matrix Factorization



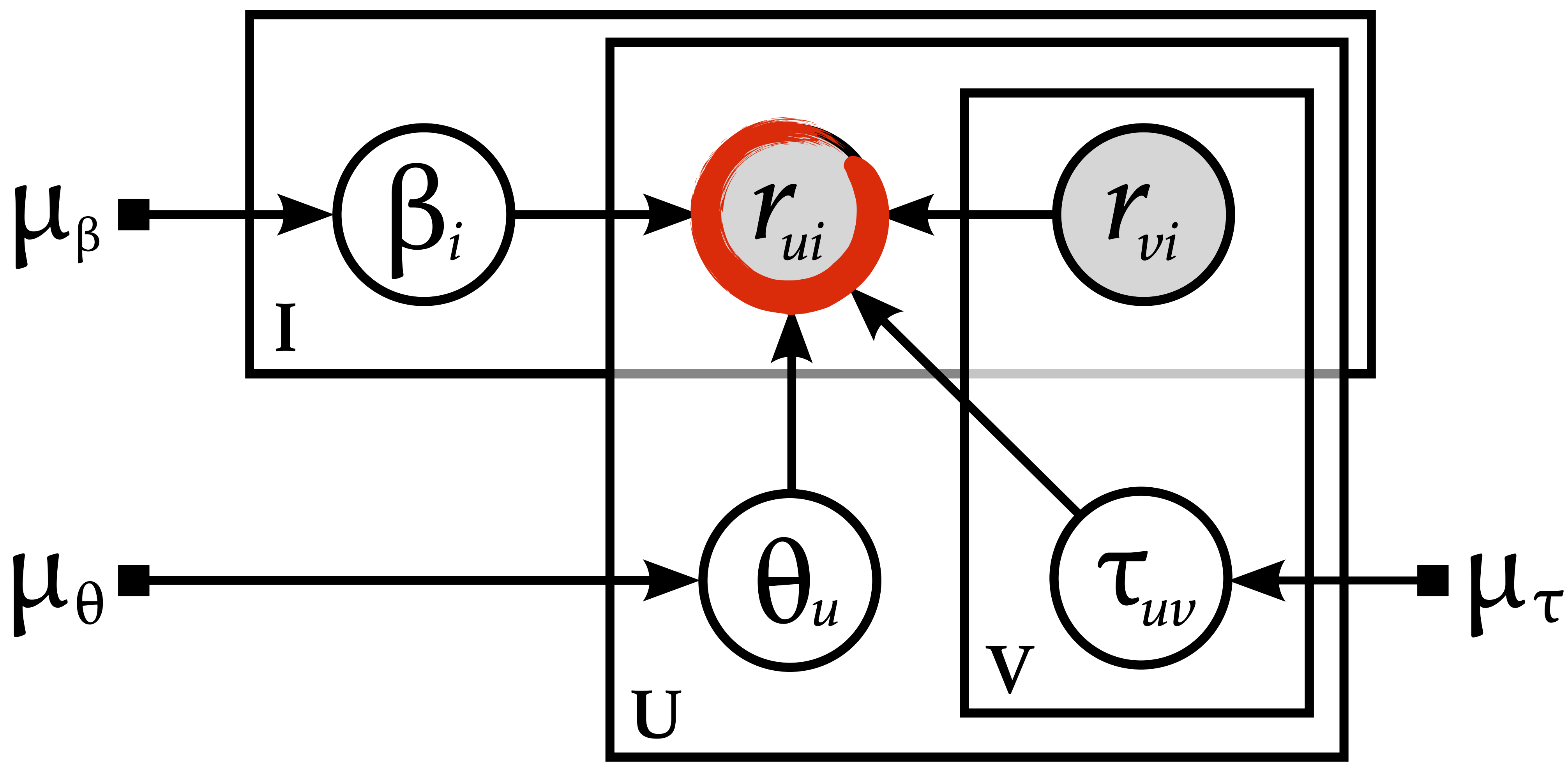
Social Poisson Factorization



Social Poisson Factorization



$$r_{ui} \mid r_{-u,i} \sim \text{Poisson}\left(\theta_u^\top \beta_i + \sum_{v \in N(u)} \tau_{uv} r_{vi}\right)$$



Posterior Inference:
How do we go from a generative model to finding the values of the variables that best fit our data?

Posterior Distribution

latent model parameters

easy to compute

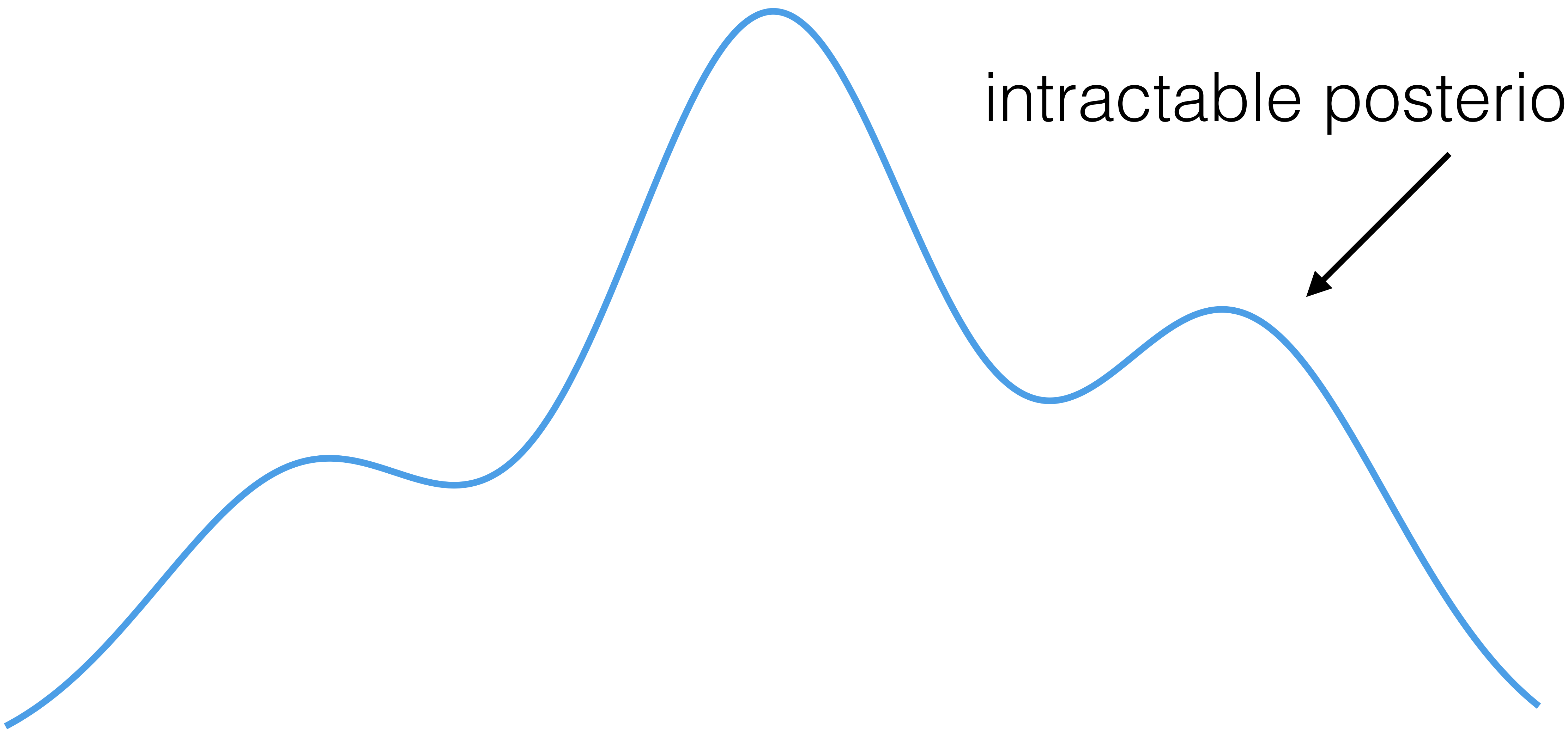
$$p(\beta, \theta, \tau \mid \mathbf{R}, \mathbf{N}, \mu) = \frac{p(\beta, \theta, \tau, \mathbf{R}, \mathbf{N} \mid \mu)}{\int_\beta \int_\theta \int_\tau p(\beta, \theta, \tau, \mathbf{R}, \mathbf{N} \mid \mu)}$$

observed data

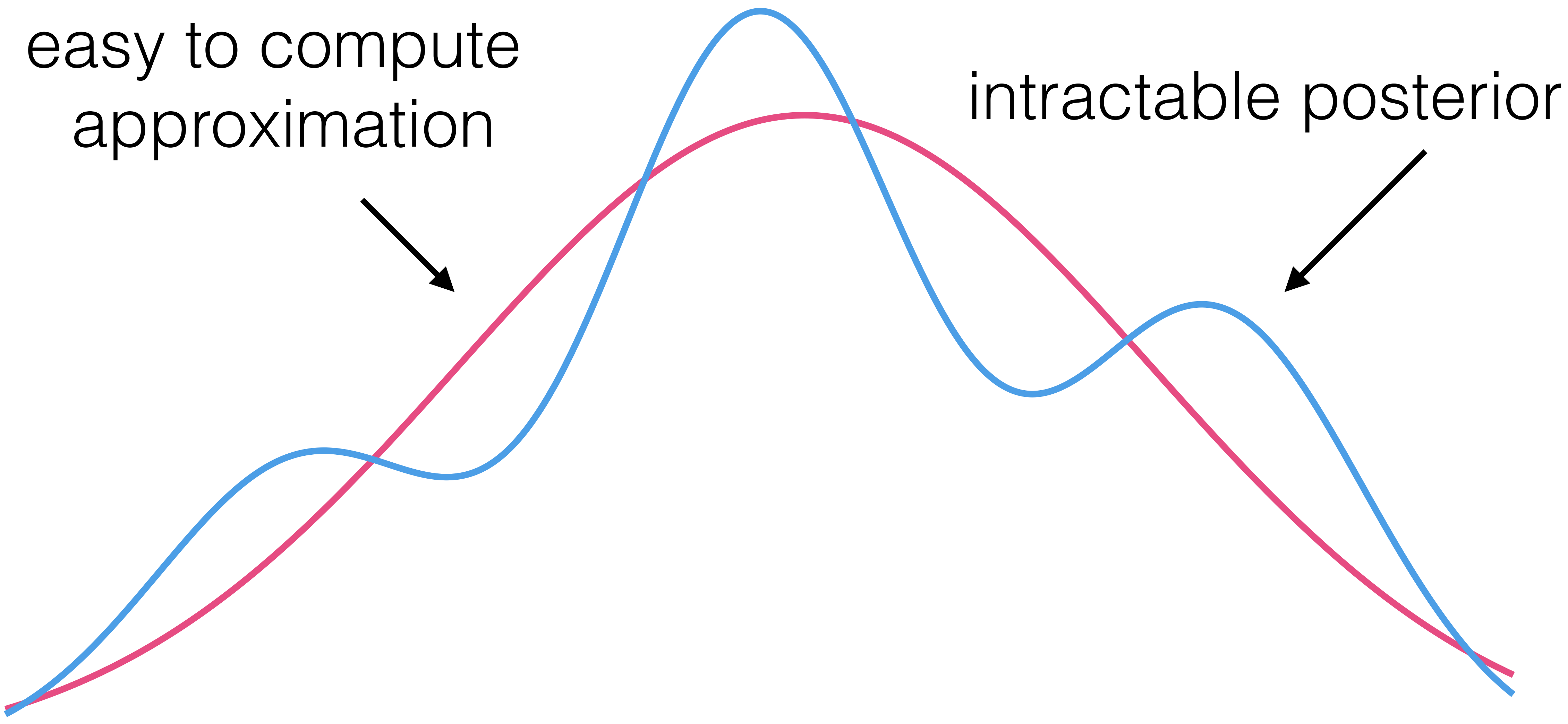
intractable

model hyperparameters

Mean Field Variational Inference



Mean Field Variational Inference



Recommendation

$$\mathbf{E}[r_{ui}] = \mathbf{E}[\theta_u]^\top \mathbf{E}[\beta_i] + \sum_{v \in N(u)} \mathbf{E}[\tau_{uv}] r_{vi}$$

Data

source	# users	# items	% ratings	% edges
Ciao	7,000	98,000	0.038%	0.103%
Epinions	39,000	131,000	0.012%	0.011%
Flixster	132,000	42,000	0.122%	0.006%
Douban	129,000	57,000	0.221%	0.016%
Social Reader	122,000	6,000	0.065%	0.001%
Etsy	40,000	5,202,000	0.009%	0.300%

etsy.com and librec.net/datasets.html

Existing Methods for Including Social Networks

SoRec	Ma et al., SoRec: Social Recommendation Using Probabilistic Matrix Factorization, SIGIR 2008.
RSTE	Ma et al., Learning to Recommend with Social Trust Ensemble, SIGIR 2009.
SocialMF	Jamali and Ester, A Matrix Factorization Technique with Trust Propagation for Recommendation in Social Networks, RecSys 2010.
TrustMF	Yang et al., Social Collaborative Filtering by Trust, IJCAI 2013.
TrustSVD	Guo et al., TrustSVD: Collaborative Filtering with Both the Explicit and Implicit Influence of User Trust and of Item Ratings, AAAI 2015.

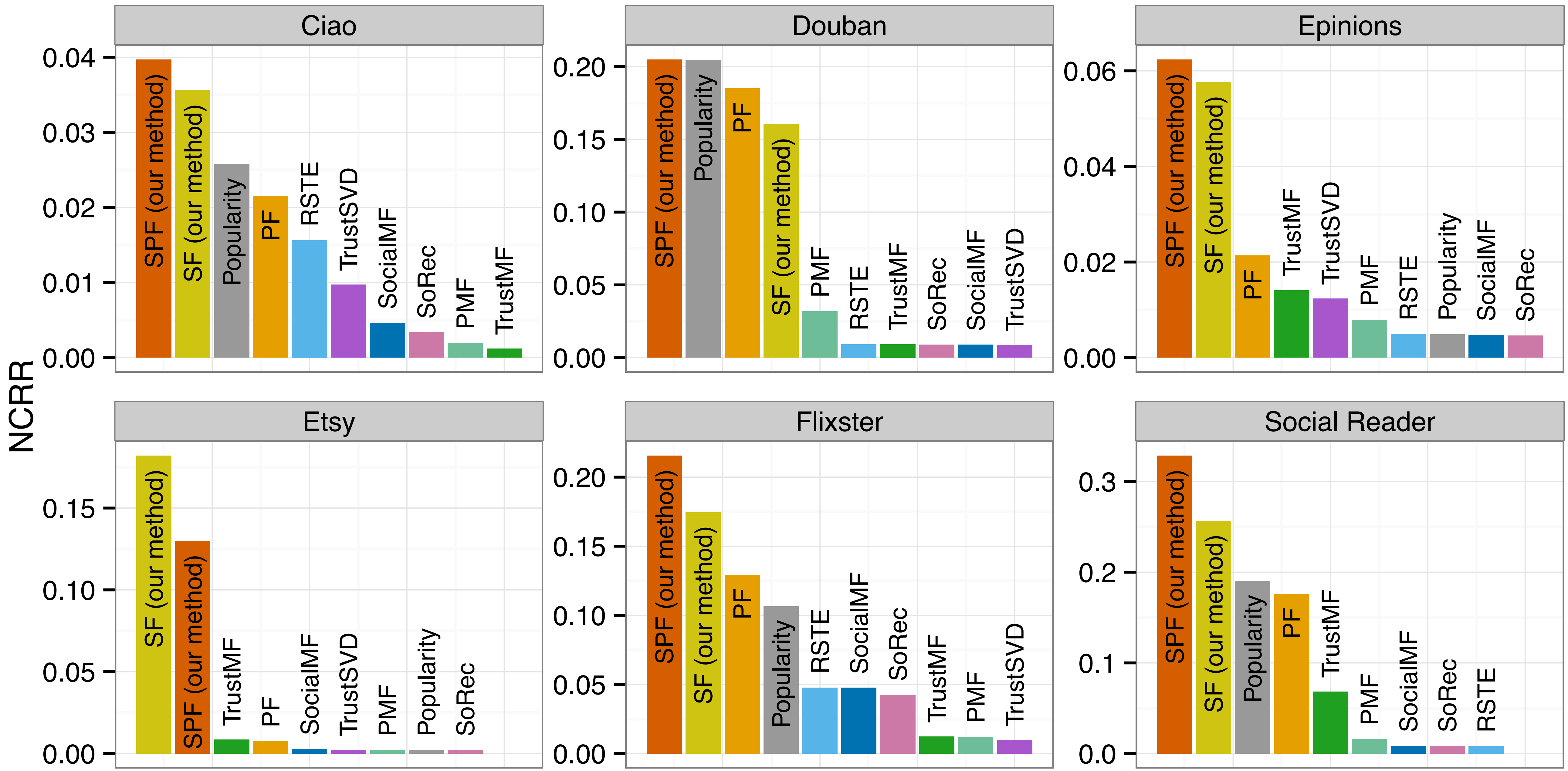
librec.net

Evaluation on held-out data

$$CRR(user) = \sum_{n=1}^{N} \frac{\mathbf{1}[rec_n \in \mathcal{H}]}{n} = \sum_{i \in \mathcal{H}} \frac{1}{rank(i)}$$

$$NCRR(user) = \frac{CRR(user)}{\text{ideal } CRR(user)}$$

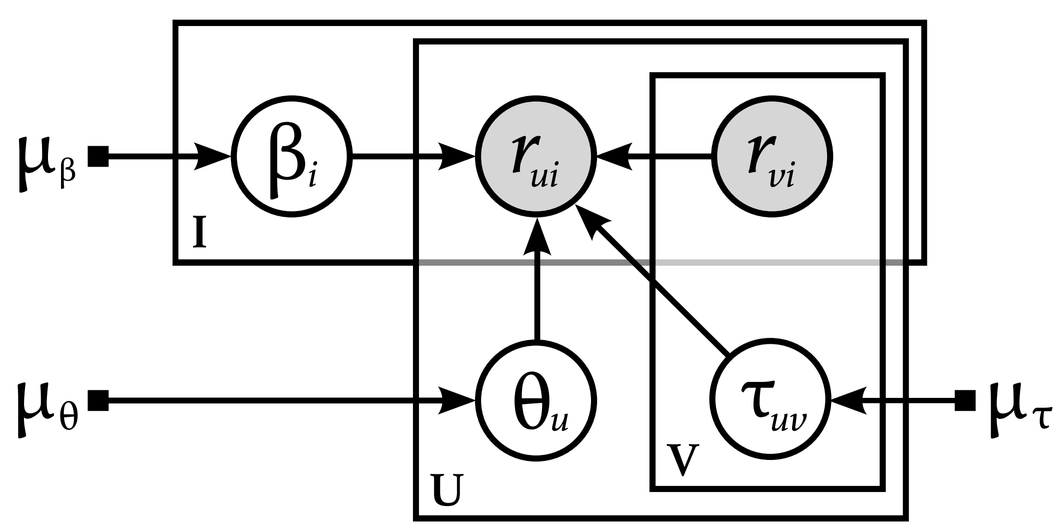
Results



NCRR
Ciao
SPF (our method)
SF (our method)
Popularity
PF
RSTE
TrustSVD
SocialMF
SoRec
PMF
TrustMF
0.00
0.01
0.02
0.03
0.04
Douban
SPF (our method)
Popularity
PF
SF (our method)
PMF
RSTE
TrustMF
SoRec
SocialMF
TrustSVD
0.00
0.05
0.10
0.15
0.20
Epinions
SPF (our method)
SF (our method)
PF
TrustMF
TrustSVD
PMF
RSTE
Popularity
SocialMF
SoRec
0.00
0.02
0.04
0.06
Etsy
SF (our method)
SPF (our method)
TrustMF
PF
SocialMF
TrustSVD
PMF
Popularity
SoRec
0.00
0.05
0.10
0.15
Flixster
SPF (our method)
SF (our method)
PF
Popularity
RSTE
SocialMF
SoRec
TrustMF
PMF
TrustSVD
0.00
0.05
0.10
0.15
0.20
Social Reader
SPF (our method)
SF (our method)
Popularity
PF
TrustMF
PMF
SocialMF
SoRec
RSTE
0.0
0.1
0.2
0.3

Summary

- SPF performs better than comparison models
- SPF is interpretable and has explainable serendipity
- SPF scales well to large data
- Source code available at ajbc.io/spf



Thank you!

Questions and suggestions welcome.

Thank you to Blei Lab colleagues
and Guibing Guo (LibRec creator)

ajbc.io/spf