Exploration of the Robustness and Generalizability of the Additive Factors Model

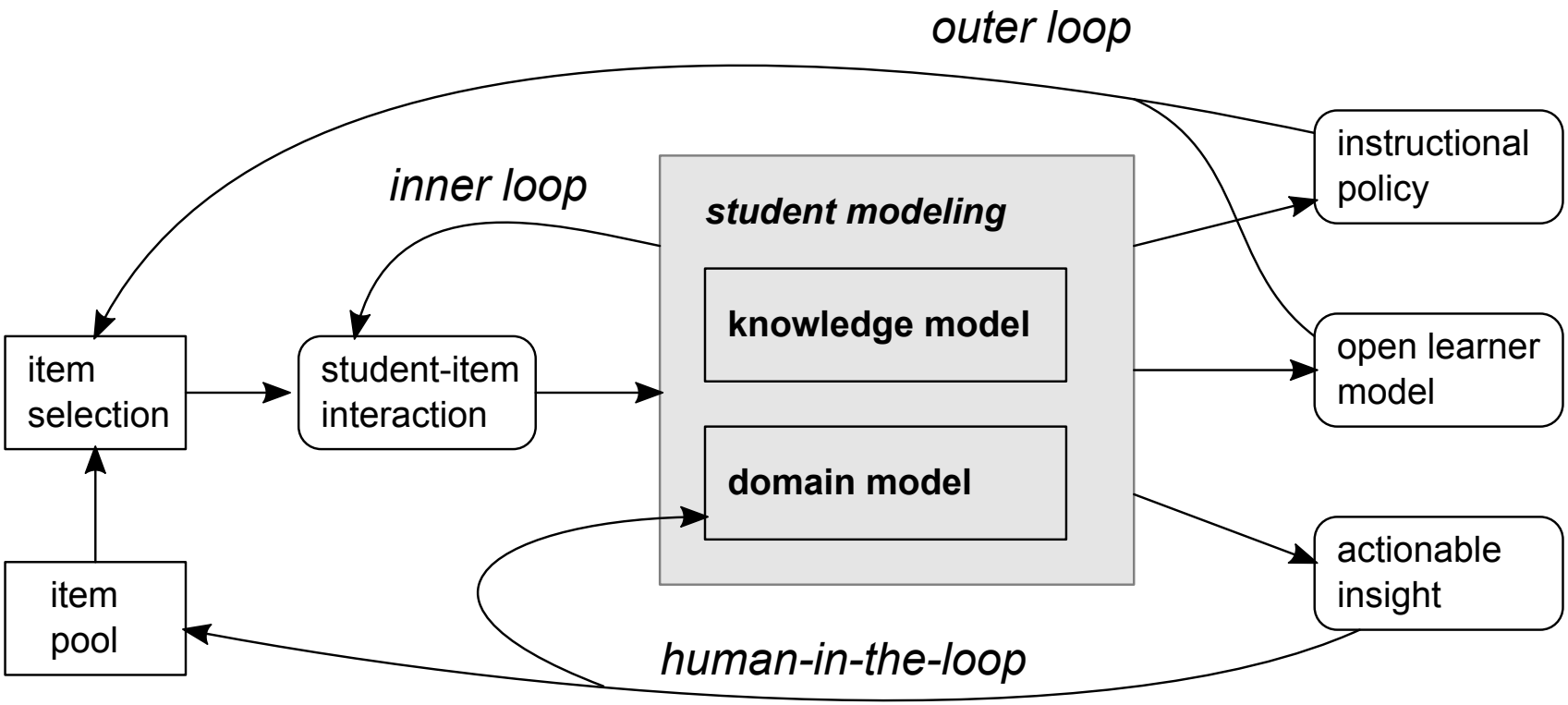
Tomáš Effenberger, **Radek Pelánek**, Jaroslav Čechák

LAK 2020

Overview

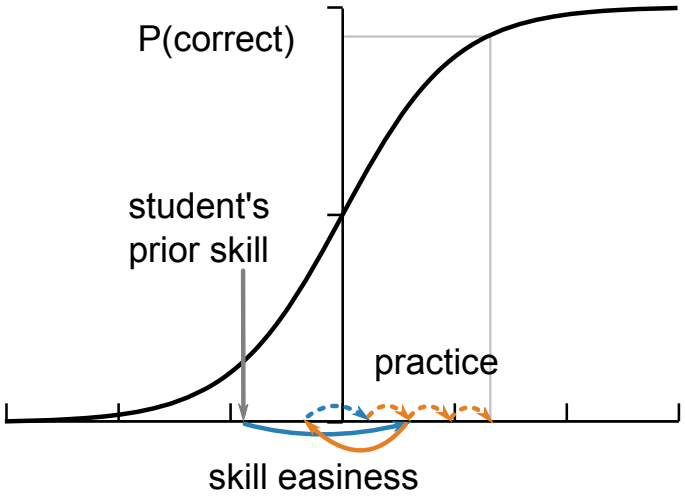
- AFM – widely used student model
- exploration of:
  - simplifying assumptions
  - robustness, generalizability (case study)
- message: more caution needed

Student Modeling



Bayesian Knowledge Tracing, Logistic Models, and Beyond: An Overview of Learner Modeling Techniques

Additive Factors Model



- family of "logistic models"
- multiple skills, Q-matrix
- used in many studies in last 10 years – see paper for overview

Q-matrix

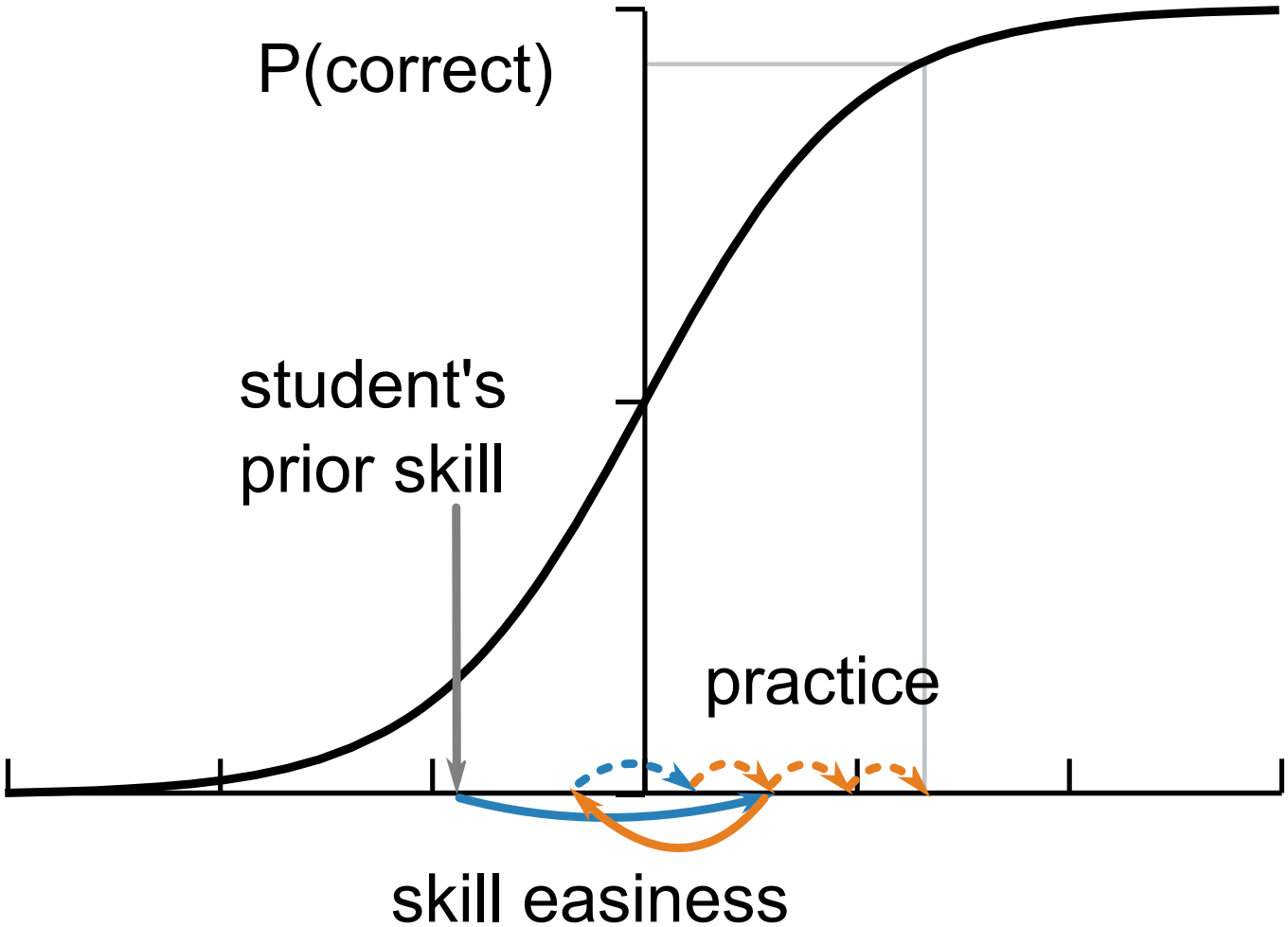
	$-$	$+$	$\times$	$()$
$10 + 3 \times 2$	0	1	1	0
$(7 - 4) \times 3$	1	0	1	1
$2 + (3 + 5)$	0	1	0	1
$8 - (6 + 2)$	1	1	0	1
$5 - 2 \times 6$	1	0	1	0

Additive Factors Model

$$P(Y_{ij}|\alpha,\beta,\gamma) = \sigma\left(\alpha_i + \sum_{k=1}^{K}\beta_k q_{jk} + \sum_{k=1}^{K}\gamma_k q_{jk} t_{ik}\right)$$

- i is student index, j is item index,
- Y_{ij} is the binary response of a student i on a item j,
- $\sigma(x) = 1/(1+e^{-x})$ is the standard logistic function,
- K is the number of skills, J is the number of items,
- Q is the $J \times K$ binary matrix – q_{jk} is the indicator that item j uses skill k,
- α_i is the proficiency (prior skill) of a student i,
- β_k is the easiness of skill k,
- γ_k is the learning rate for skill k,
- t_{ik} is the number of times student i has practiced skill k (opportunity count).

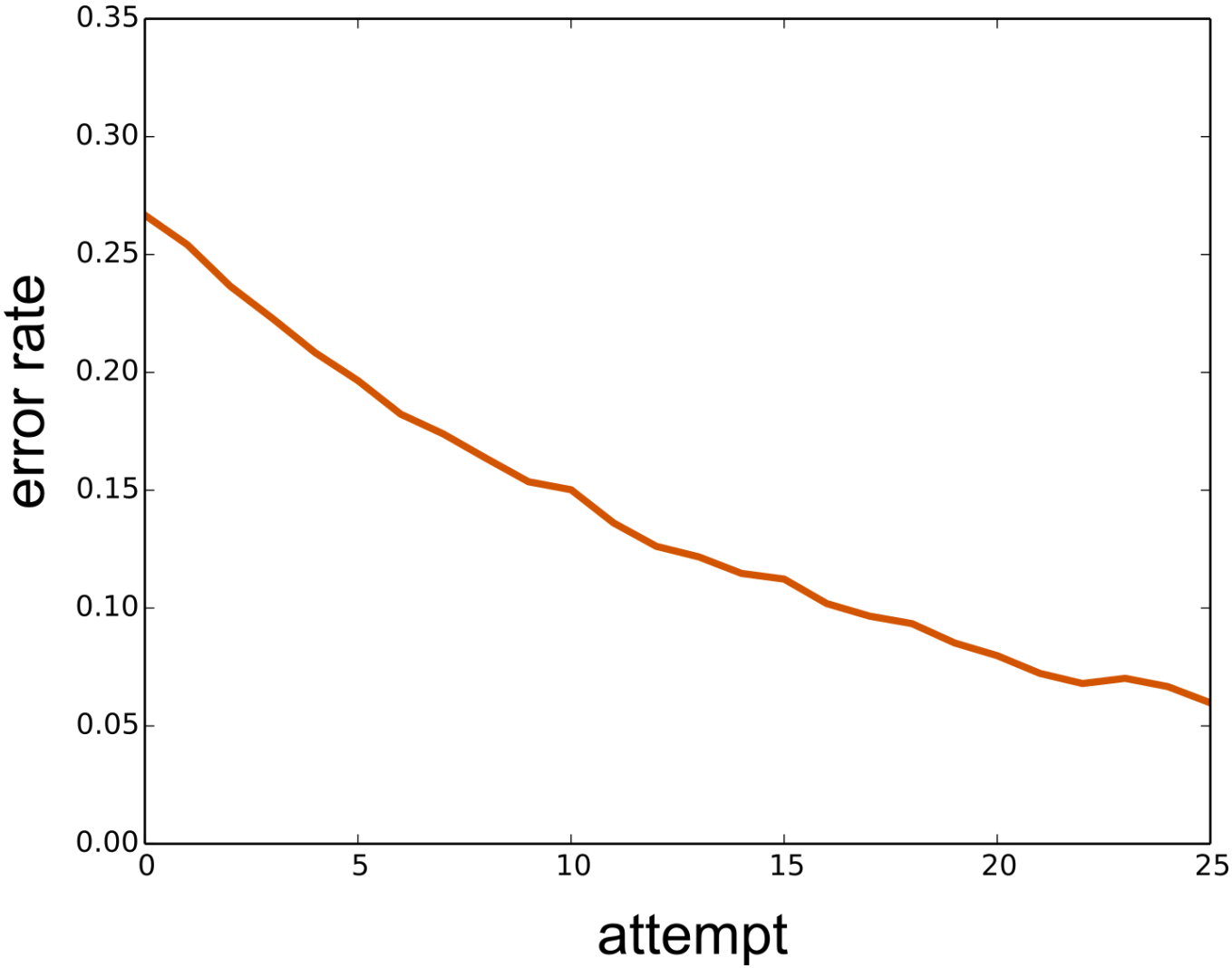
Additive Factors Model



AFM: Simplifying Assumptions

- learning is linear (on the logit scale)
- effect of practice not related to observed performance
- observed outcomes are binary (ignoring response time, common wrong answers)
- Q-matrix is binary
- compensatory model of skills
- ignores difficulty of items
- ignores biases in data (e.g., items solved in fixed order)

Learning Curves



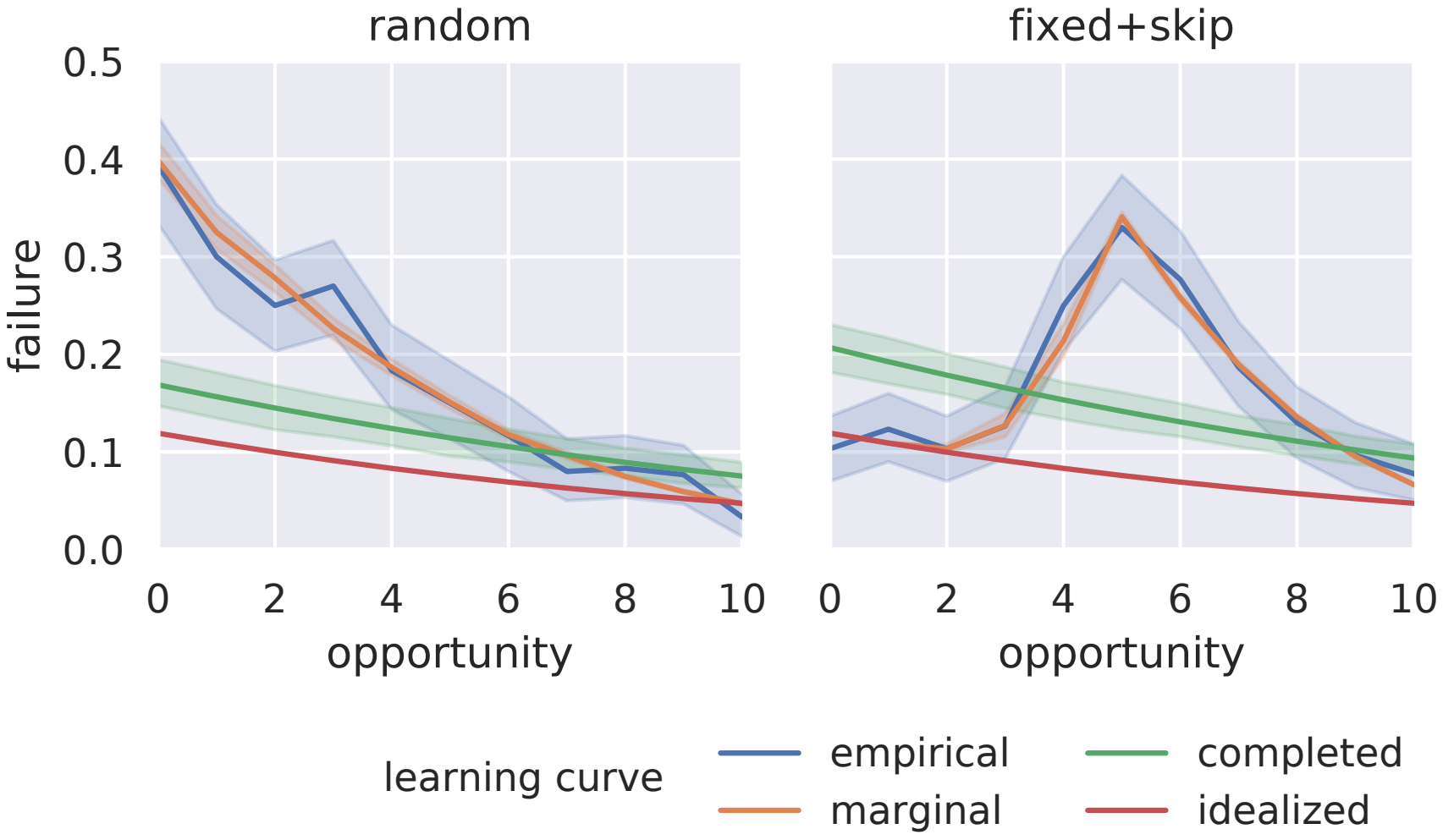
Types of Learning Curves

Type	Attempt	Opportunity	Success
empirical	observed	observed	observed
marginal	observed	observed	predicted
completed	observed	simulated	predicted
idealized	simulated	simulated	predicted

Simulation

- 2 skills
  - skill 1: easy
  - skill 2: difficult
- 15 items
- items 1-5: skill 1
- items 5-15: skill 1 & 2
- ordering: random, fixed+skip

Learning Curves



Typical AFM Application

- take the data
- fit AFM model with various Q-matricies
- take the Q-matrix with best fit
- analyze (one type of) learning curves

Case Study: Programming

- block-based programming
- 85000 attempts
- 5800 students
- 85 items

open source, data available:
`en.robomise.cz`
T. Effenberger. Blockly Programming Dataset. CSEDM 2019

Questions: Generalizability and Robustness

- Does the model (AFM) fit well data in this domain?
- Can we use the model to select good Q-matrix?
- Do the learning curves provide useful insight?
- How robust are the result with respect to methodological choices?

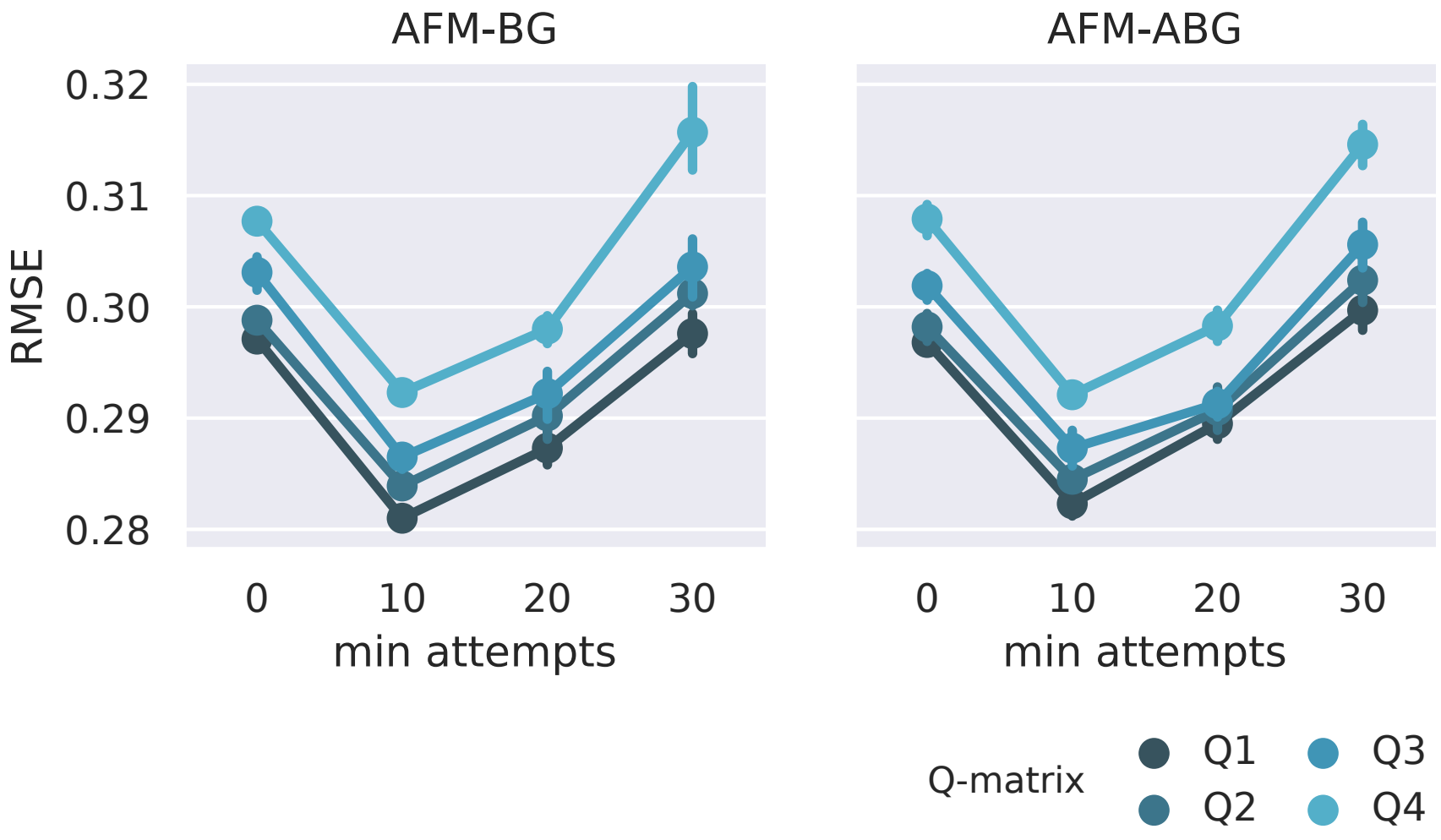
Q-matricies

Name	Skills
Q1	*teleports*, *collectables*, *obstacles*, *destructibles*, *program length limit*, sequences, while, repeat, loop, nested-loops, if, else, test, nested control structures, comparison
Q2	*teleports*, *collectables*, *obstacles*, *destructibles*, sequences, while, repeat, nested loops, if, else
Q3	sequences, while, repeat, nested loops, if, else
Q4	sequences, loop, nested loops, test

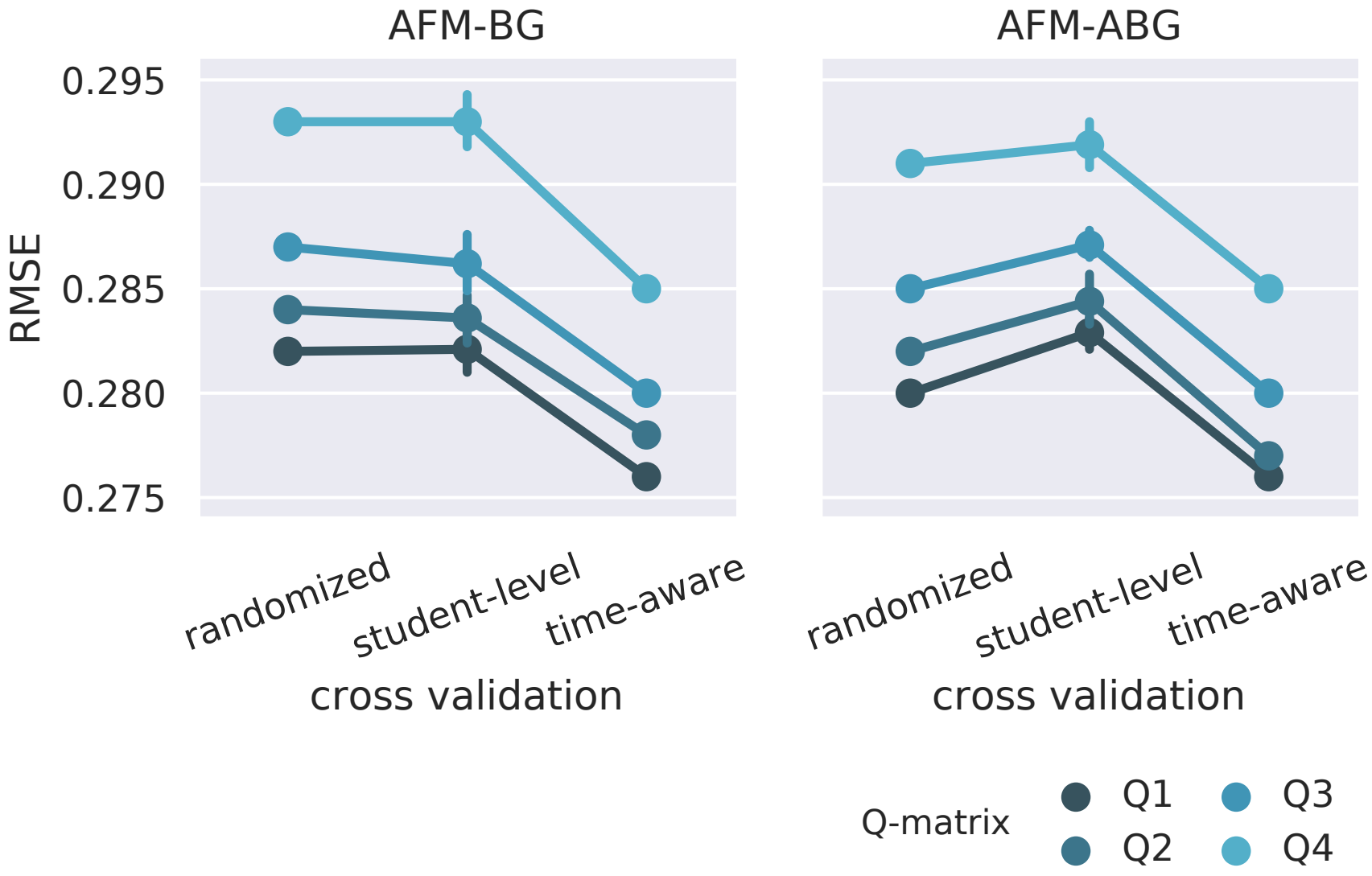
Models

- baselines: global average, item average
- AFM-BG: AFM without α parameter (student skill)
- AFM-ABG: full AFM
- AFM-BGT: AFM-BG model with item difficulties

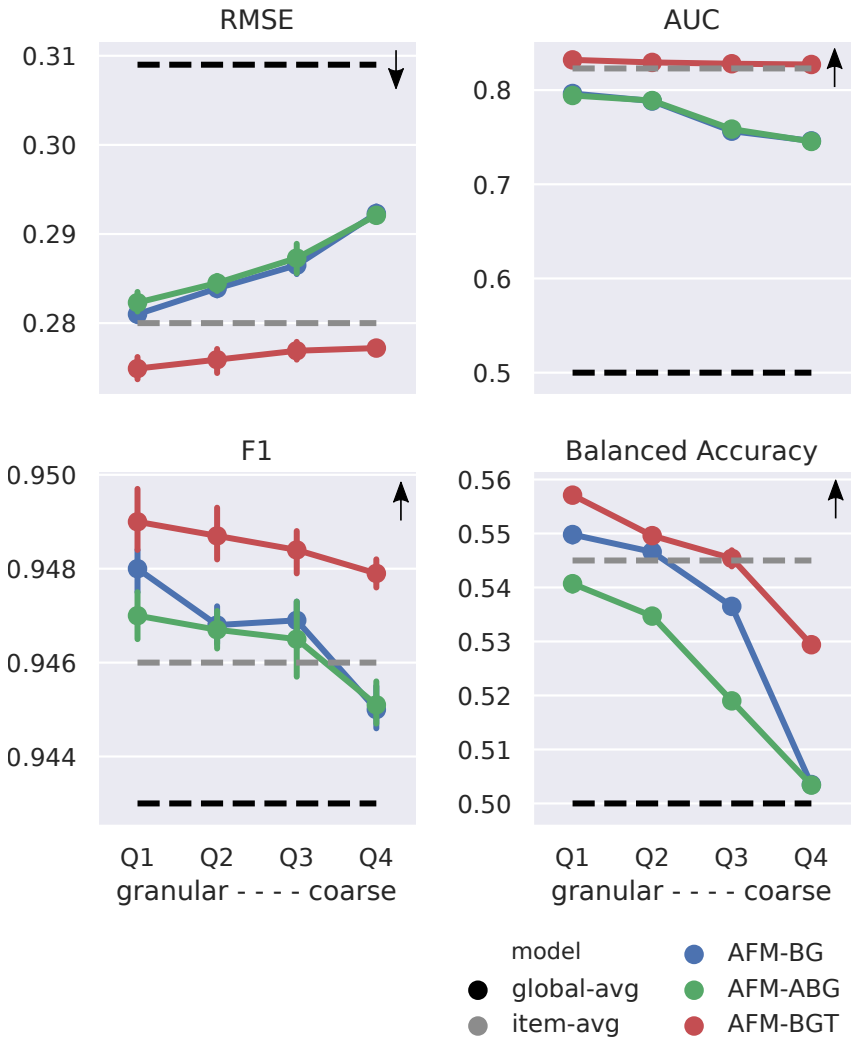
Results: Effect of Data Filtering



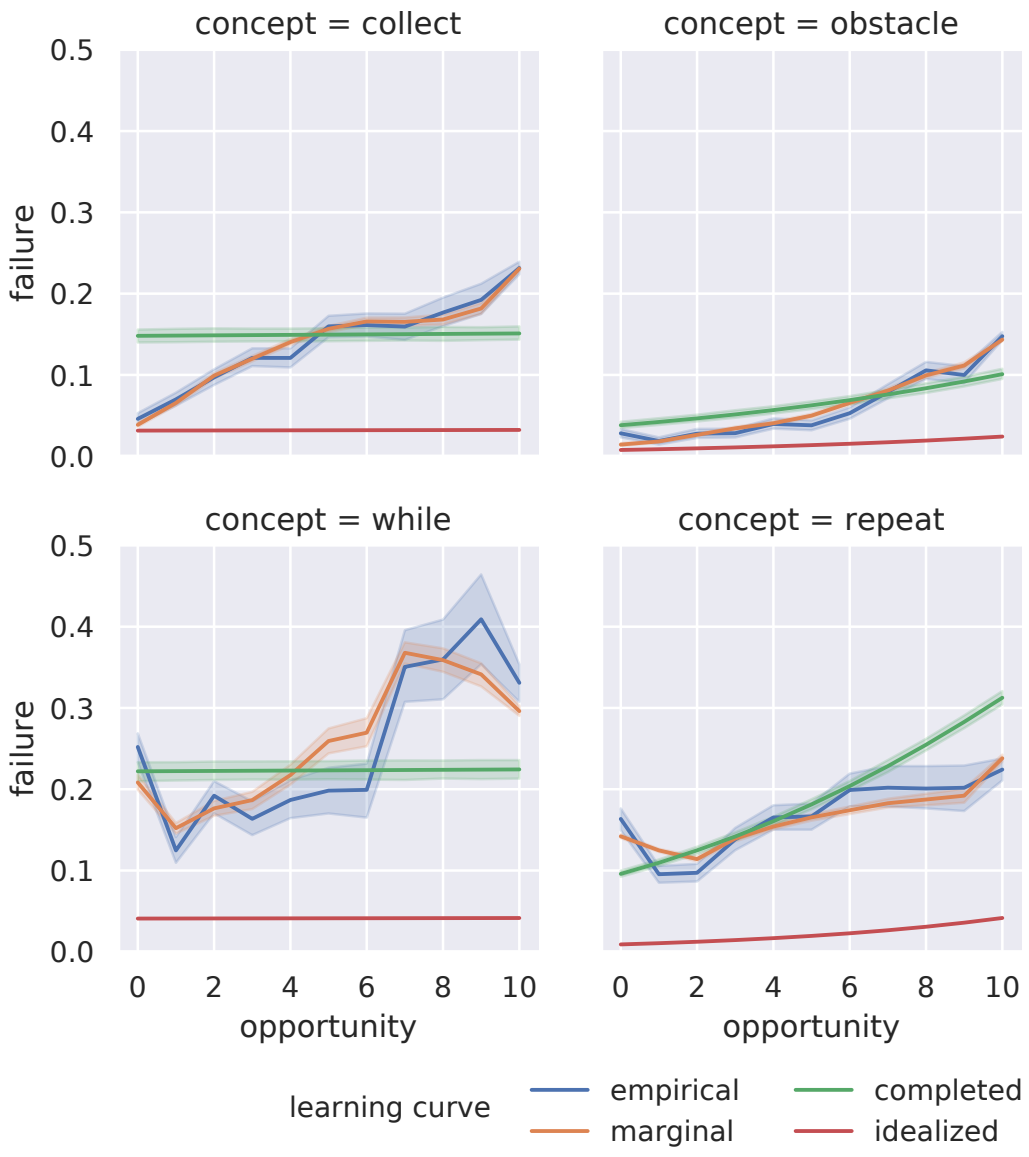
Results: Cross-validation Setting



Results: Model Comparison



Results: Learning Curves



Conclusions

- studies using AFM: more caution necessary
- AFM has many simplifying assumptions, not satisfied in practice
- out-of-the-box application $\Rightarrow$ possibly misleading conclusions
- basic precaution: comparison with "item average" predictor

Questions

Contact:
Radek Pelánek
xpelanek@fi.muni.cz
https://www.fi.muni.cz/adaptivelearning/