Fairness-Aware Tensor-Based Recommendation

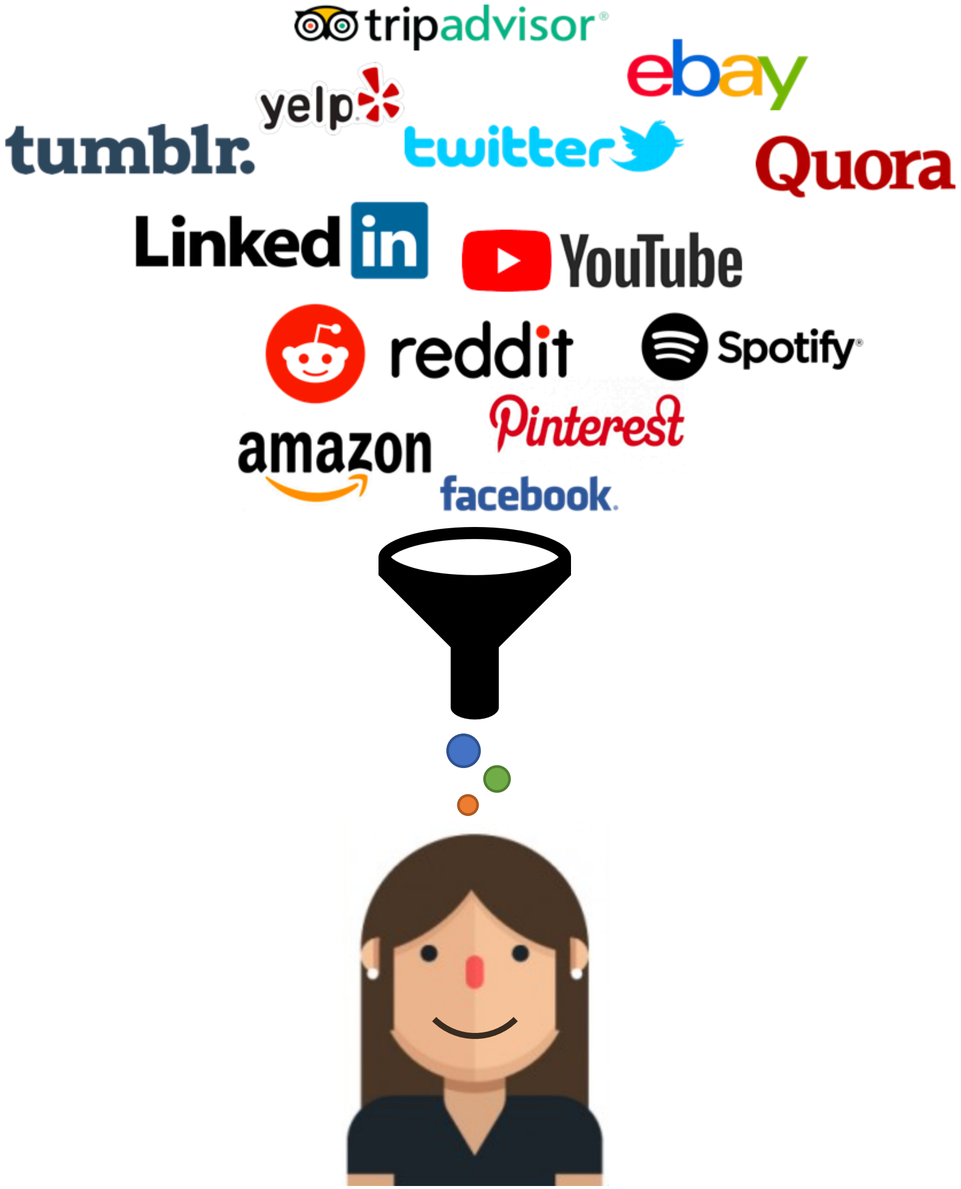
Ziwei Zhu, Xia Hu, and James Caverlee

Texas A&M University

TEXAS A&M UNIVERSITY

CIKM 2018

Recommenders – Essential Conduits

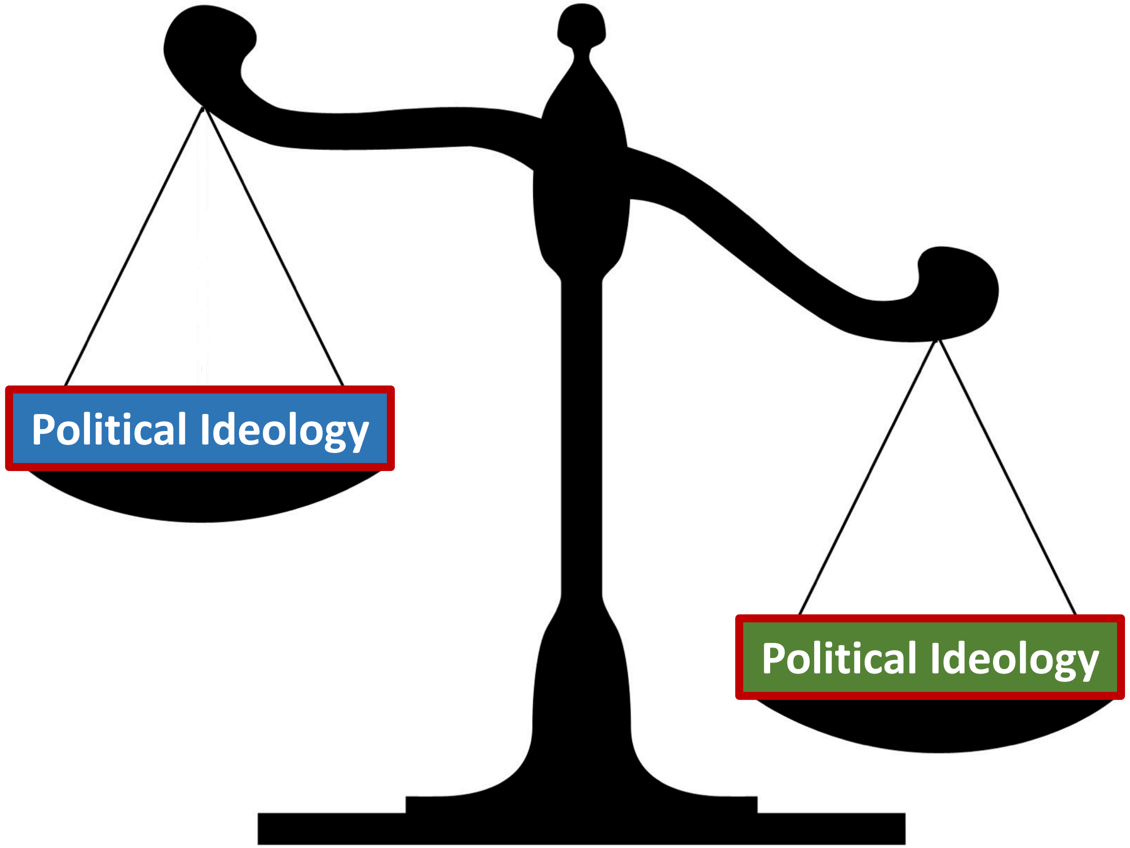


Algorithmic Bias in Recommenders

Unfair Recommenders

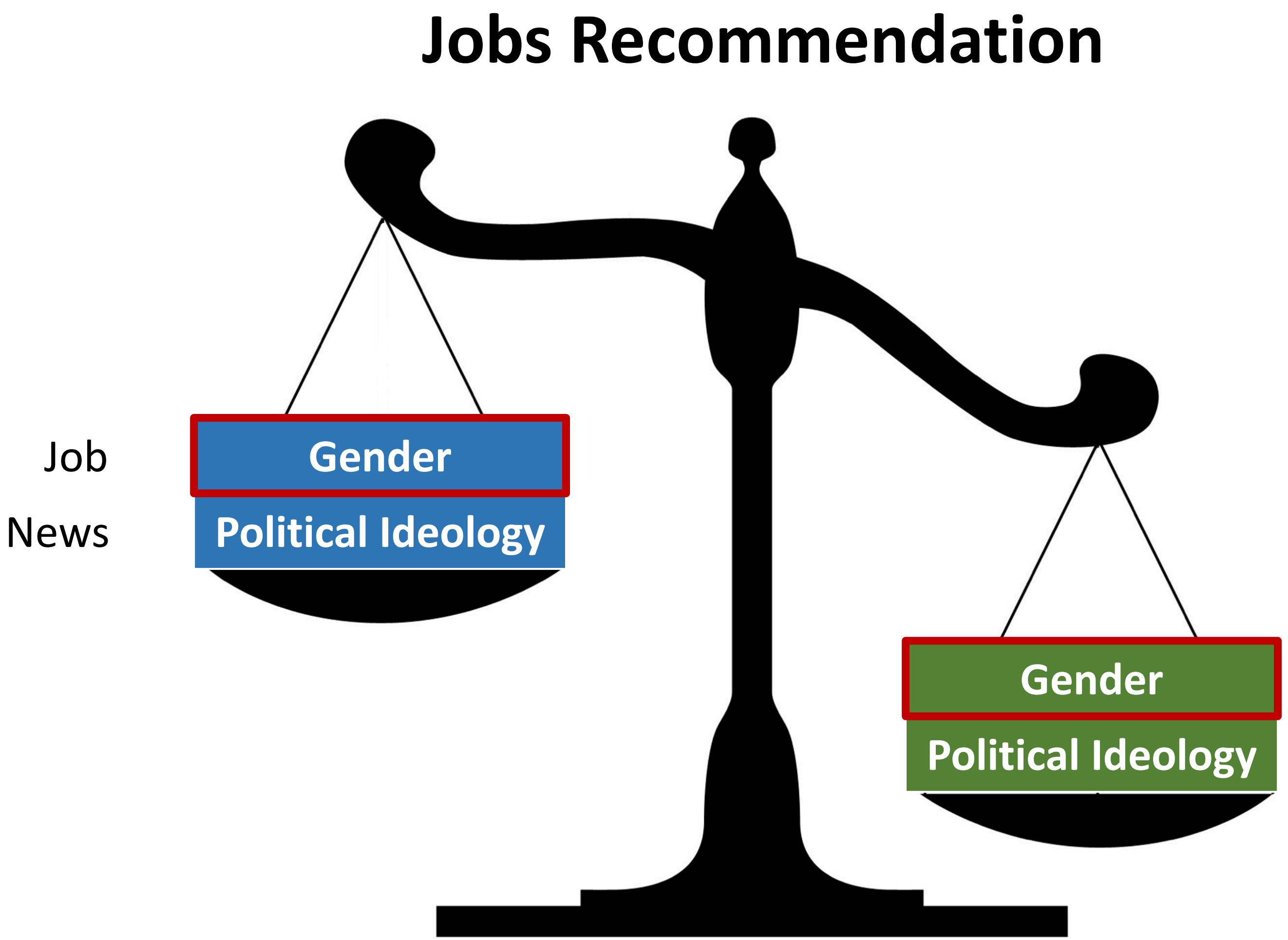
News Recommendation

News



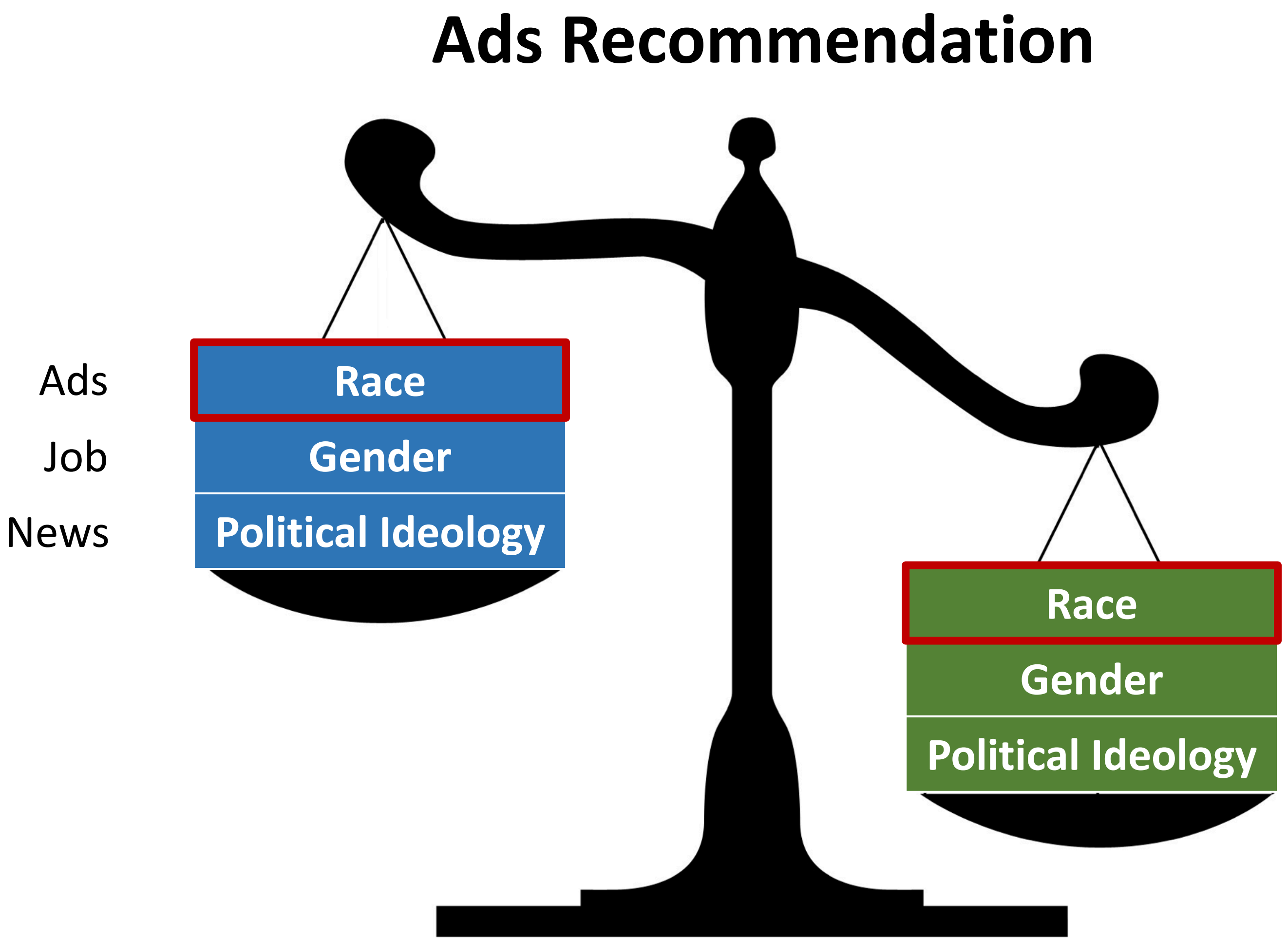
Refer to: Dan Bernhardt, Stefan Krasa, and Mattias Polborn. 2008. Political polarization and the electoral effects of media bias. Journal of Public Economics 92, 5-6 (2008), 1092–1104.

Unfair Recommenders



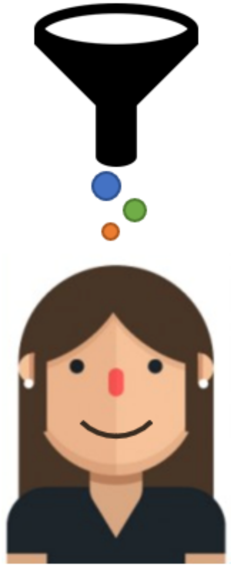
Refer to: L. Sweeney. 2013. Discrimination in online ad delivery. Queue 11, 3 (2013), 10.

Unfair Recommenders



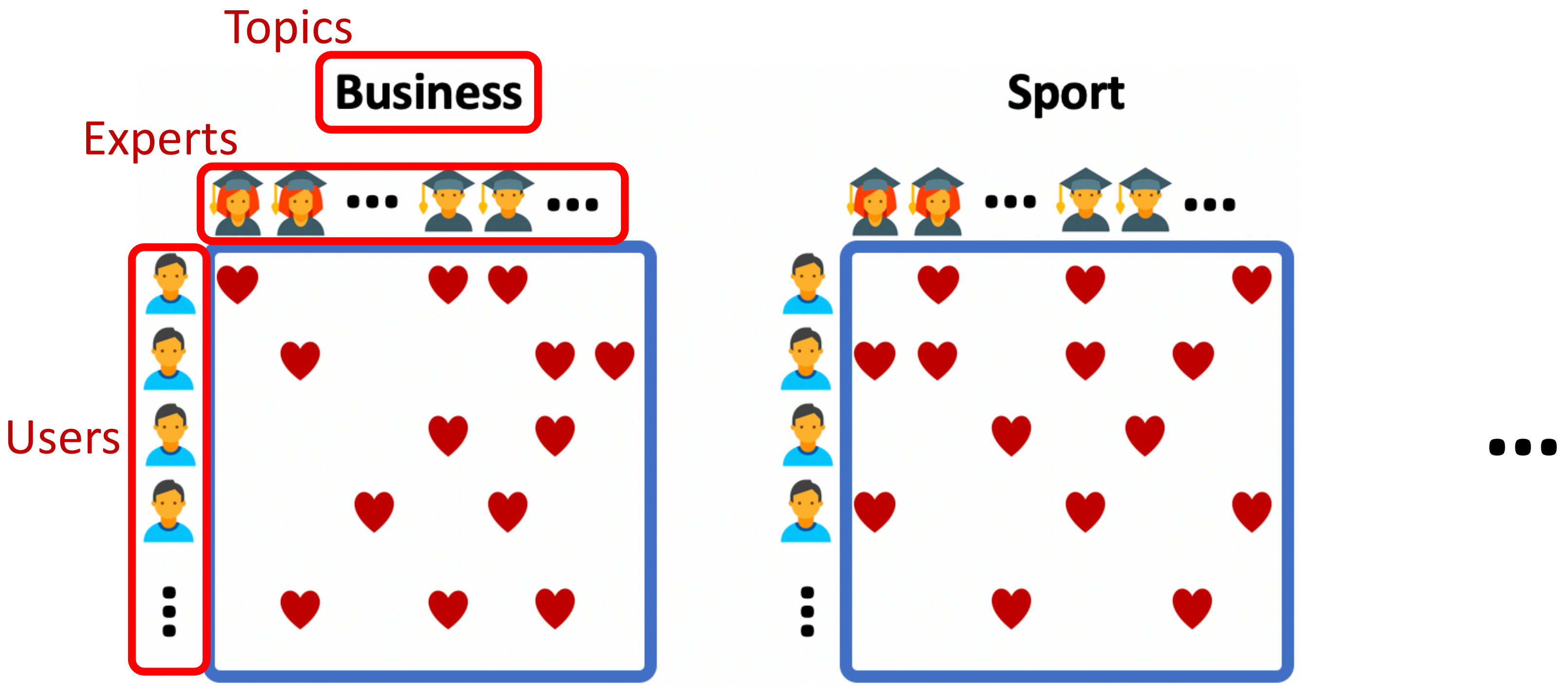
Refer to: Ayman Farahat and Michael C Bailey. 2012. How effective is targeted advertising?. In Proceedings of the 21st international conference on World Wide Web. ACM, 111– 120.

Goal



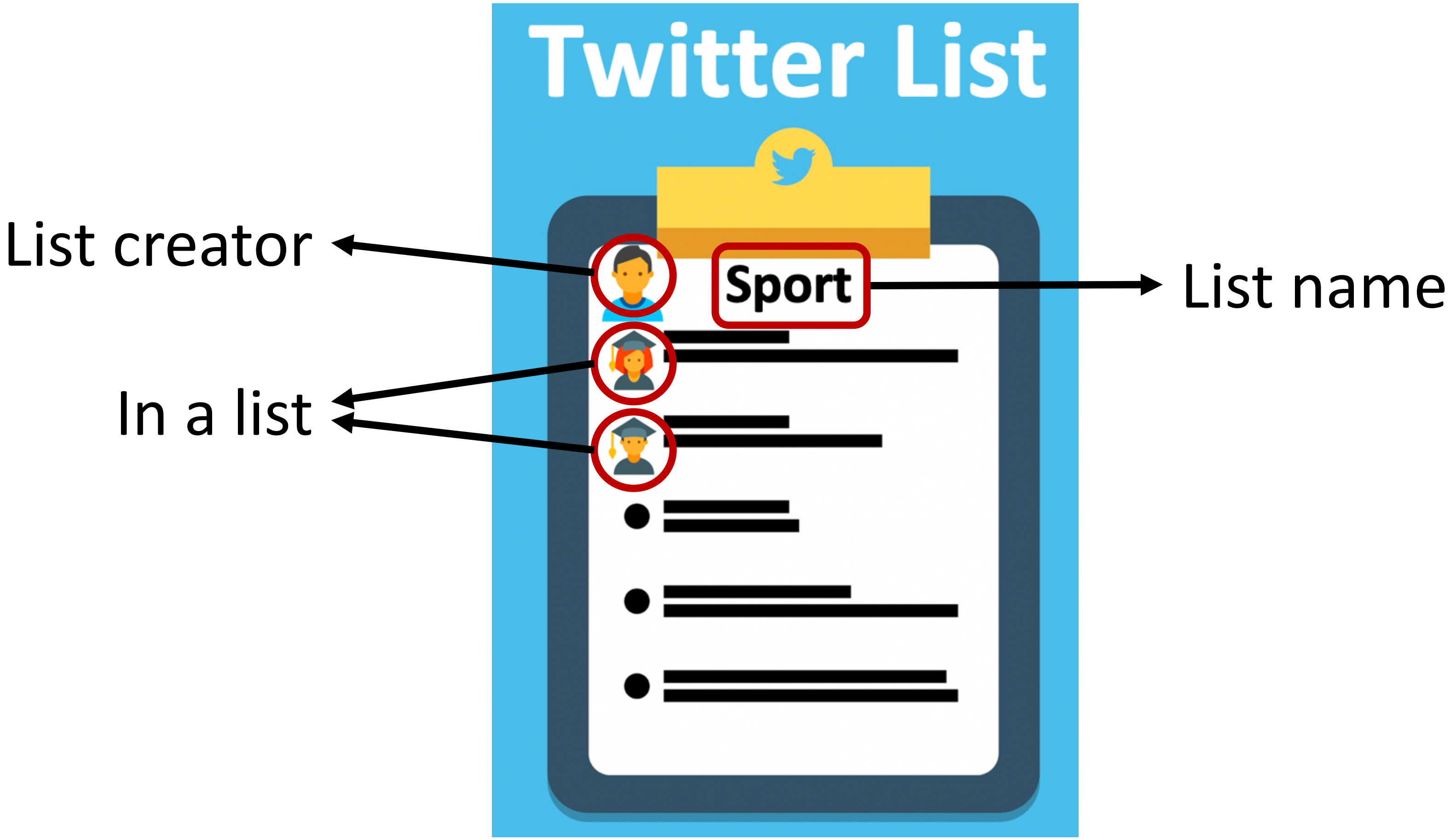
To enhance recommendation fairness while preserving recommendation quality.

Motivating Example: Expert Recommendation



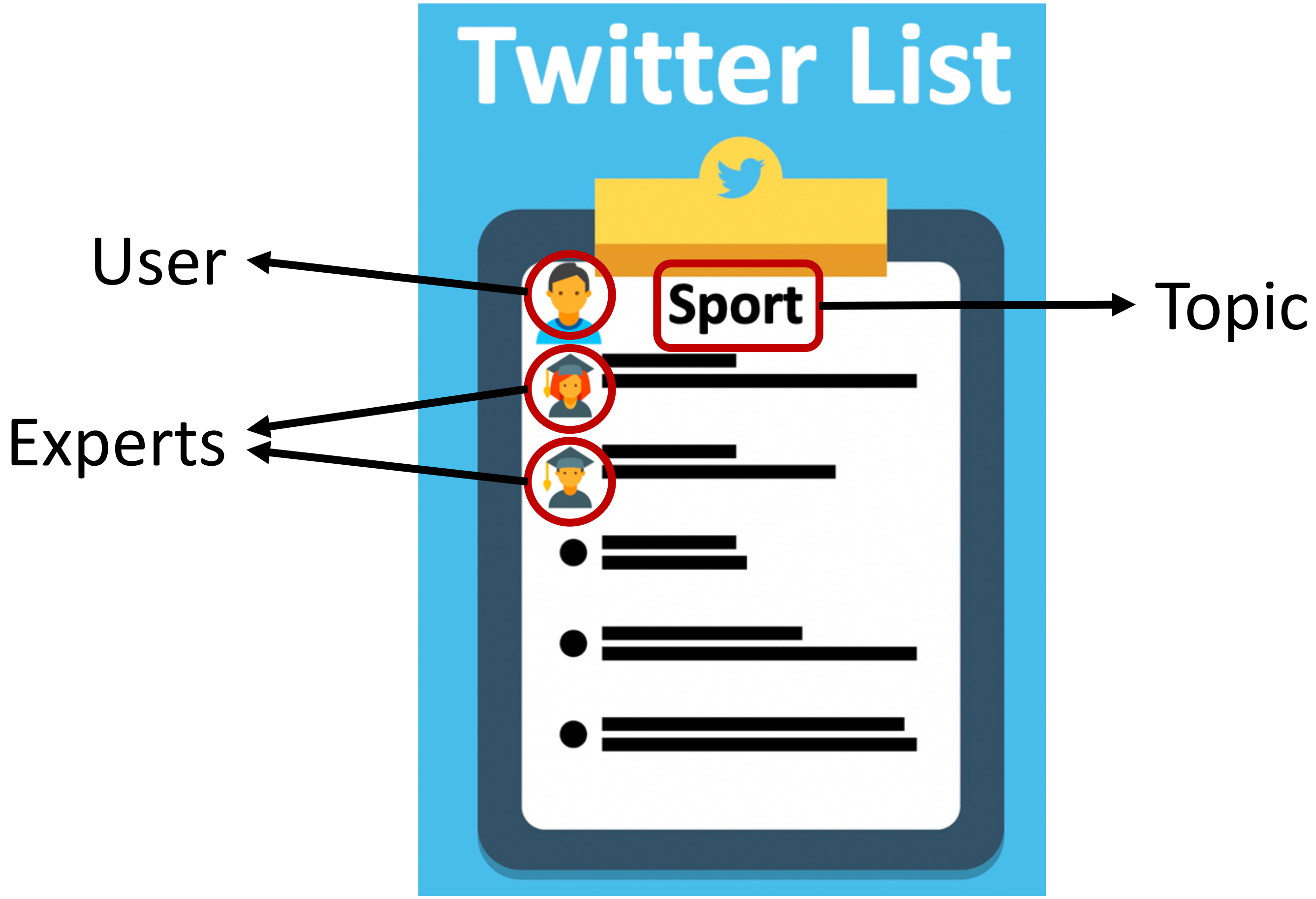
Recommend **experts** to **users** related to different **topics** based on historical user-expert interactions.

Motivating Example: Expert Recommendation



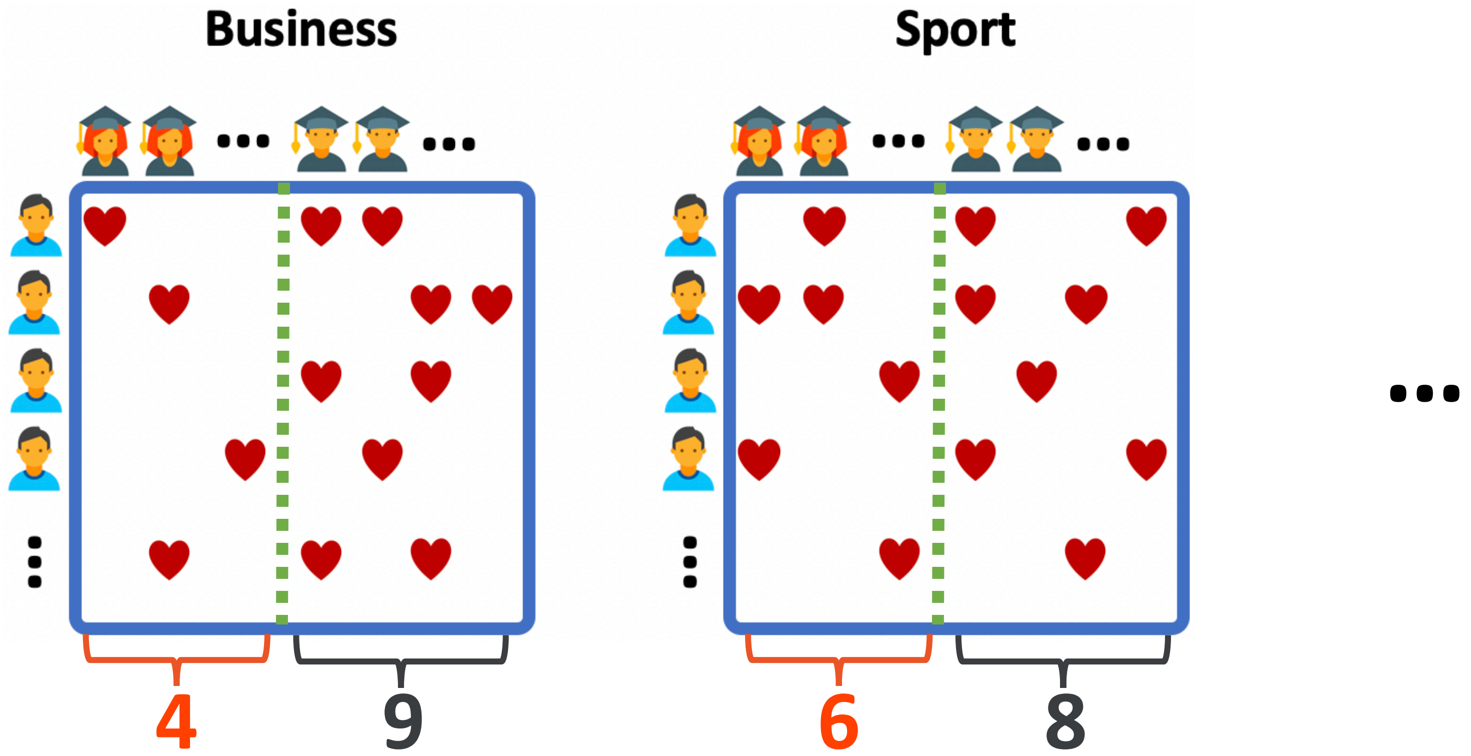
A twitter user is recognized as an **expert** related to a specific **topic** when he is added into a **Twitter List** with the topic name by **another user**.

Motivating Example: Expert Recommendation



A twitter user is recognized as an **expert** related to a specific **topic** when he is added into a **Twitter List** with the topic name by **another user**.

Motivating Example: Expert Recommendation



- We observed that distributions of liked experts with distinct **genders** and **races** are different.
- We care about recommendation **fairness** between different genders or races.

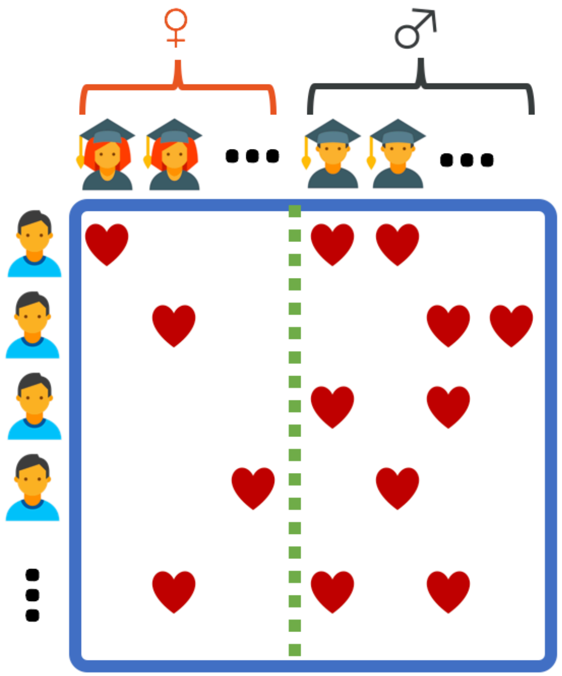
Question

How to define fairness for recommendation task?

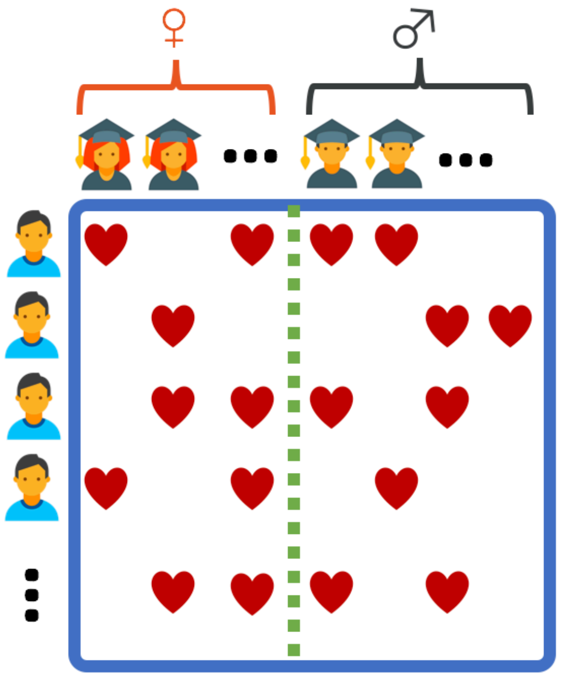
Statistical Parity

Statistical parity encourages a recommender to ensure **similar probability distributions** for both groups.

$$P[R|male] = P[R|female]$$



Unfair Example



Fair Example

Challenges

Existing approaches:

Challenges

Existing approaches:

i. focus on **two-dimensional** matrix completion; (we have user-expert-topic)

Challenges

Existing approaches:

i. focus on **two-dimensional** matrix completion; (we have user-expert-topic)

ii. assume there is only a **single binary sensitive feature** (gender: female vs. male); (we want to enhance sensitive feature of both gender and ethnicity, or even more)

Challenges

Existing approaches:

i. focus on **two-dimensional** matrix completion; (we have user-expert-topic)

ii. assume there is only a **single binary sensitive feature** (gender: female vs. male); (we want to enhance sensitive feature of both gender and ethnicity, or even more)

iii. trade-off considerable **recommendation quality** for improving fairness. (we want satisfactory recommendation utility)

Proposed Model – FATR

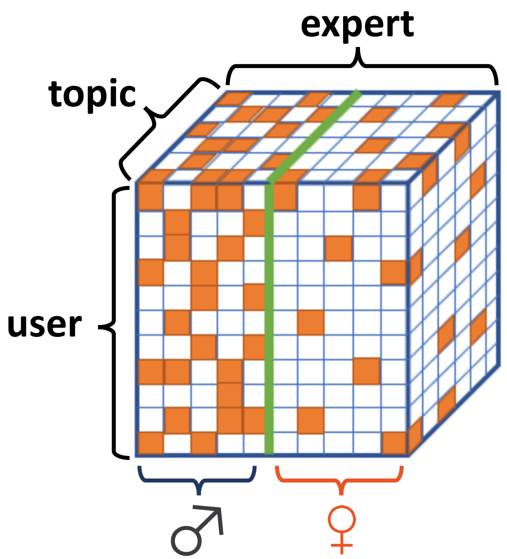
Fairness-**A**ware **T**ensor-based **R**ecommendation:

Proposed Model – FATR

Fairness-**A**ware **T**ensor-based **R**ecommendation:

i. leverages **tensor completion** as the foundation that models **multiple aspects** simultaneously;

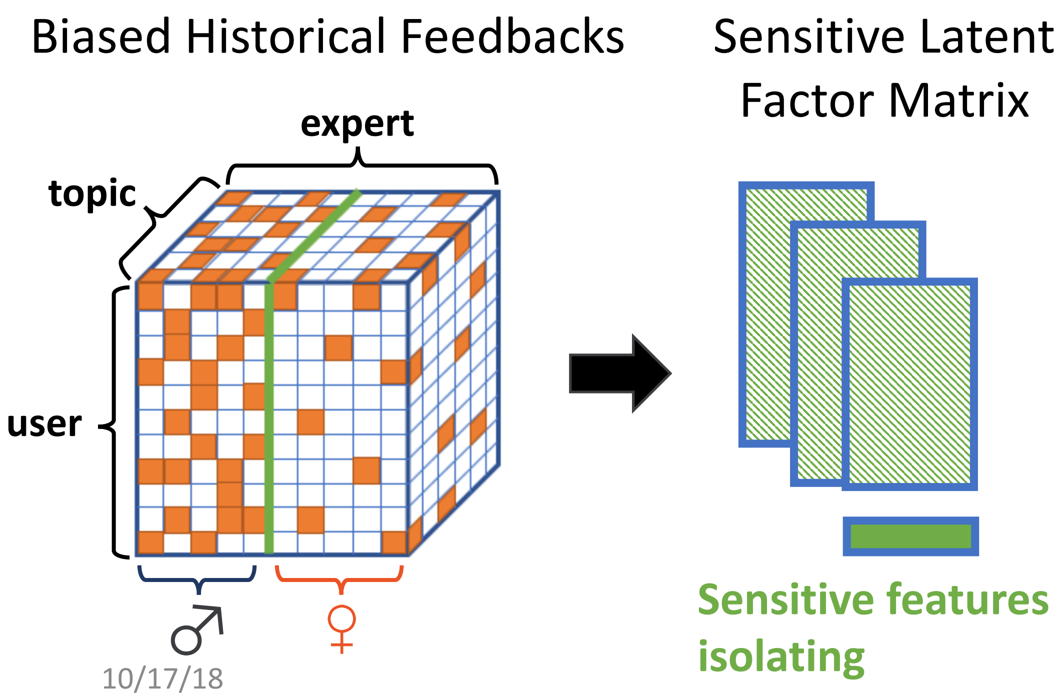
Biased Historical Feedbacks



Proposed Model – FATR

Fairness-**A**ware **T**ensor-based **R**ecommendation:

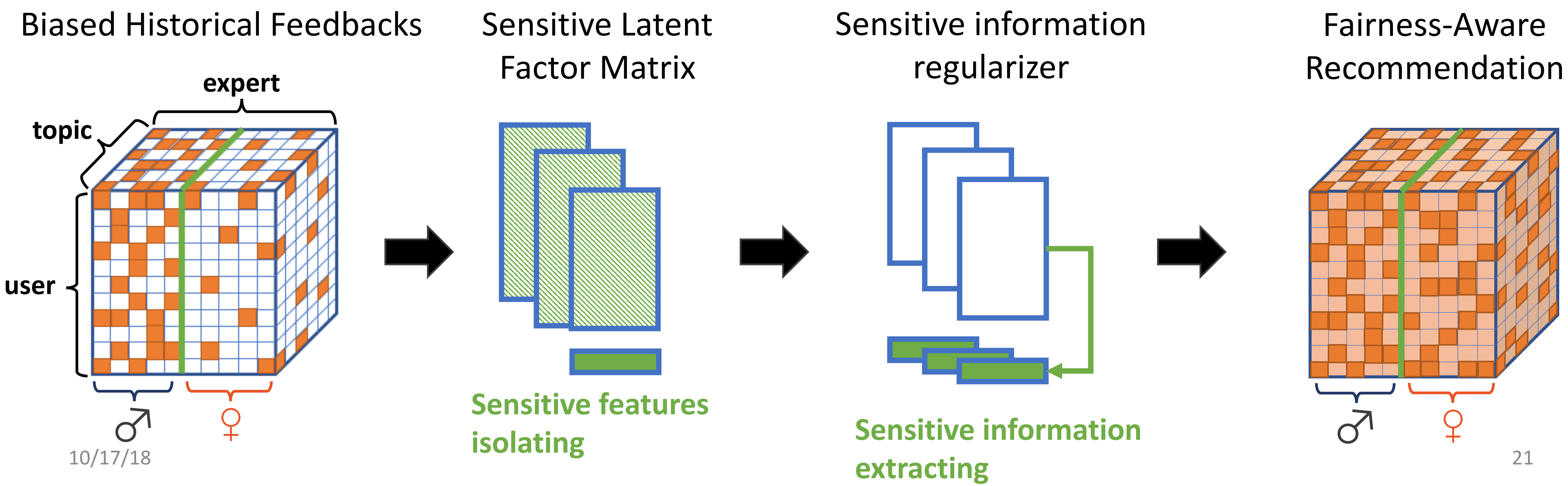
ii. uses a new **sensitive latent factor matrix** for isolating sensitive features that naturally adapt to **multi-feature** and **multi-category** cases;



Proposed Model – FATR

Fairness-**A**ware **T**ensor-based **R**ecommendation:

iii. utilizes a **sensitive information regularizer** for extracting sensitive information tainting other latent factors that promises **fairness enhancement** and **recommendation quality preserving**.



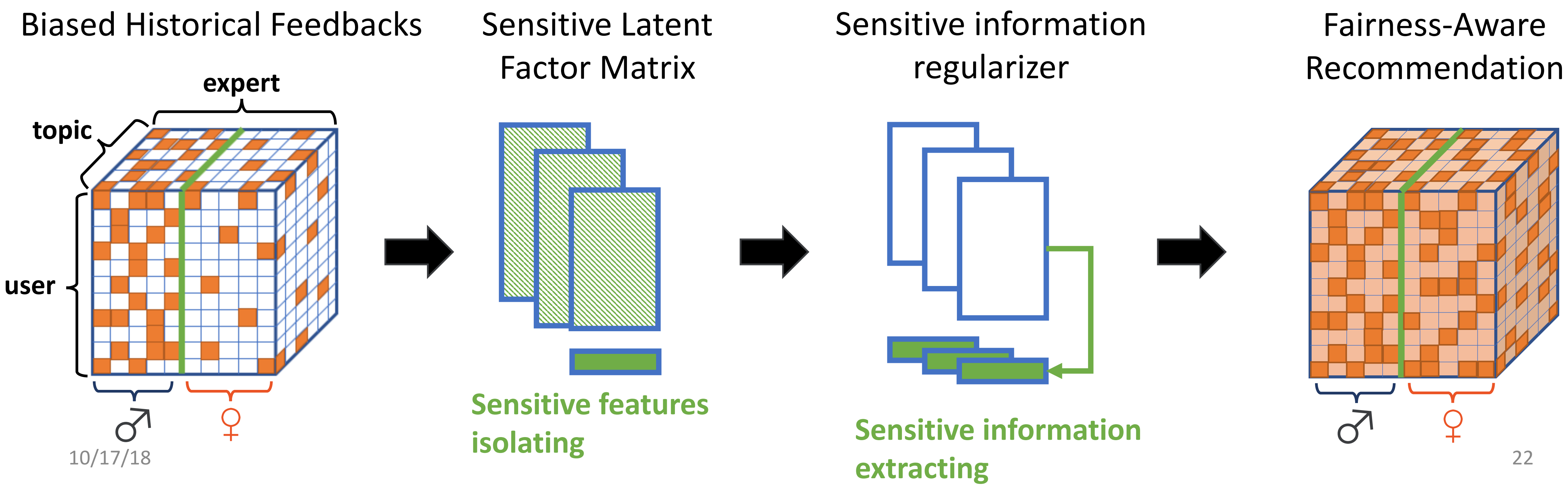
Proposed Model – FATR

Fairness-**A**ware **T**ensor-based **R**ecommendation:

i. tensor completion;

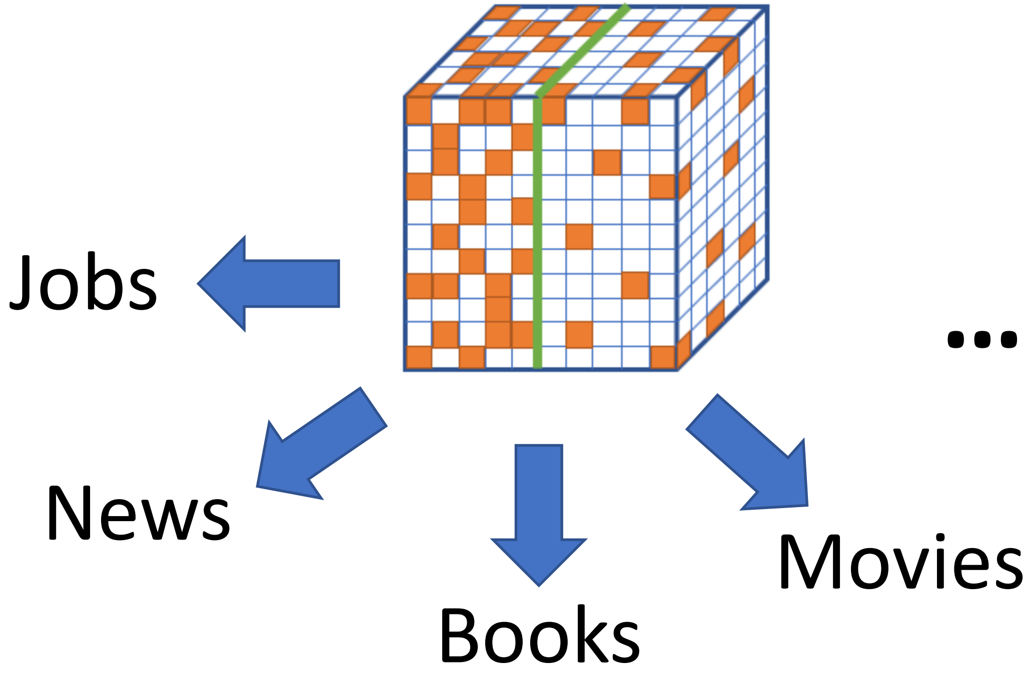
ii. sensitive latent factor matrix;

iii. sensitive information regularizer.



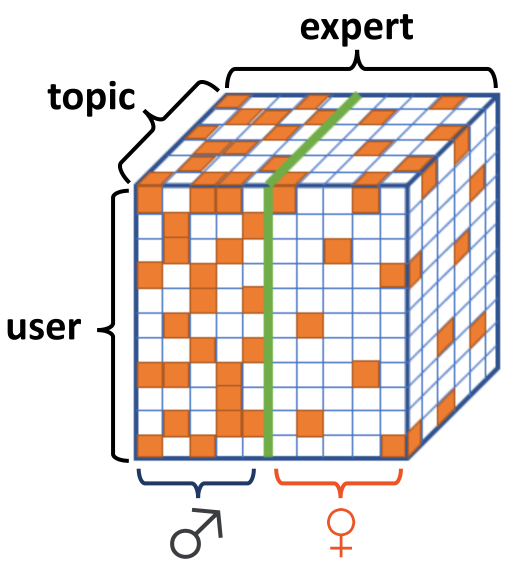
Proposed Model – FATR

FART can adapt to **other domains** with two, three or even **more dimensions**, not limited in the given expert recommendation task.

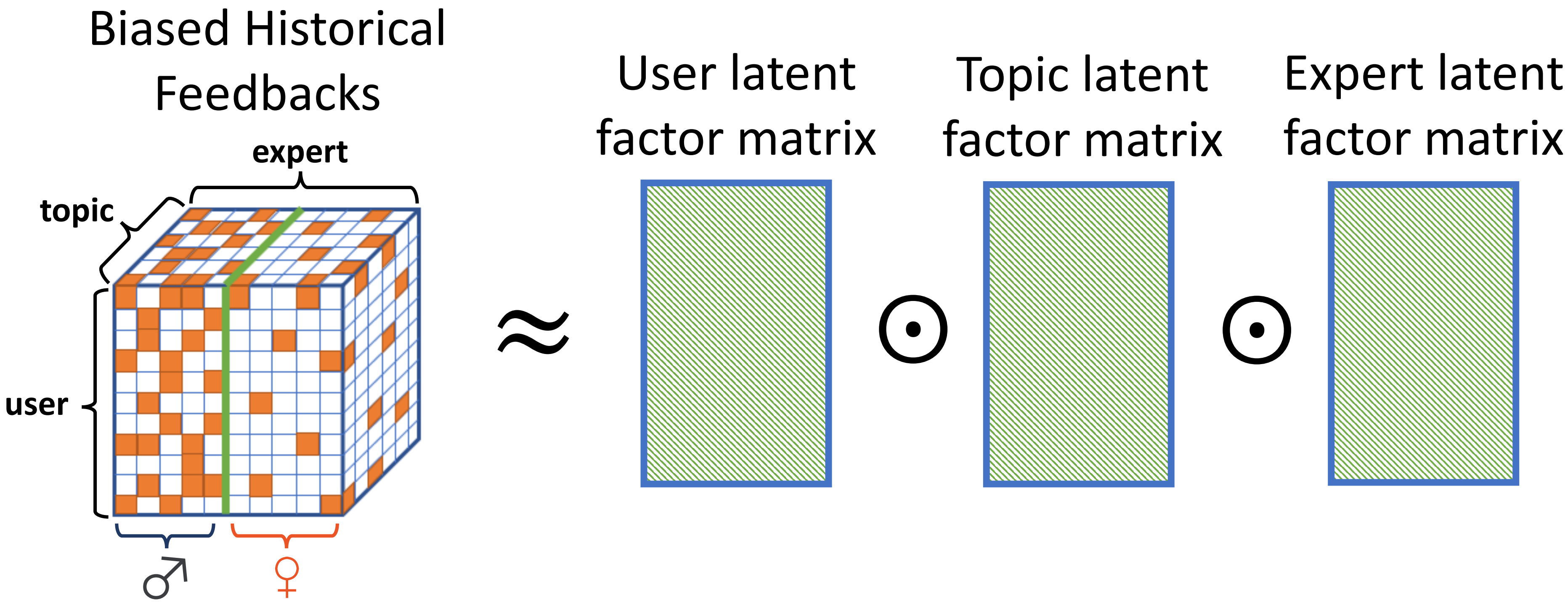


Conventional Tensor Completion

Biased Historical Feedbacks

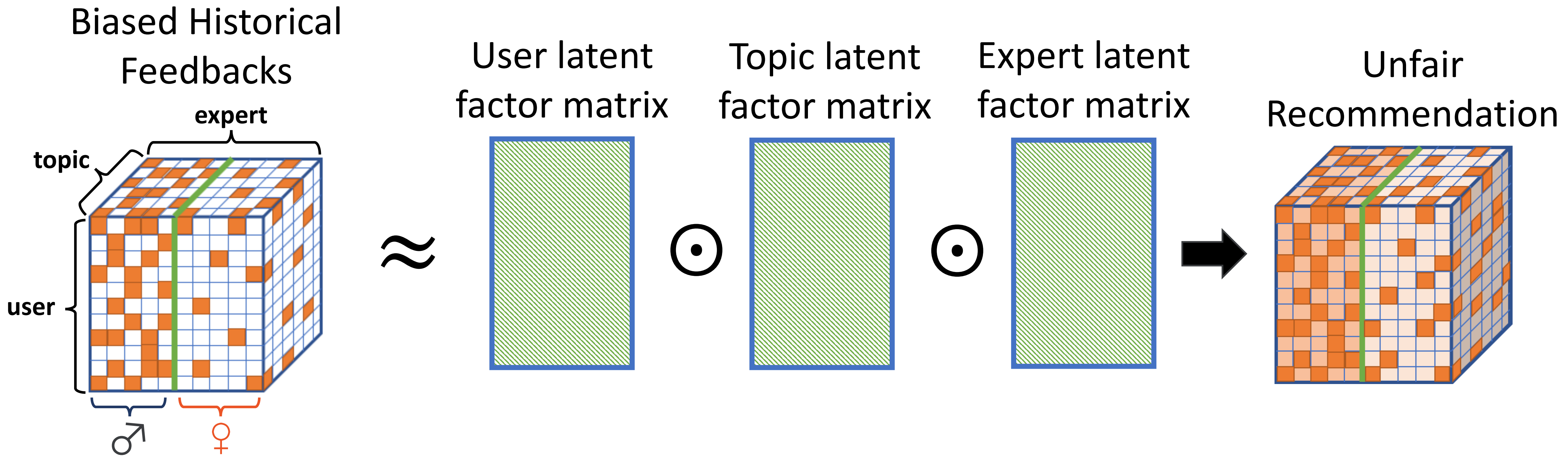


Conventional Tensor Completion



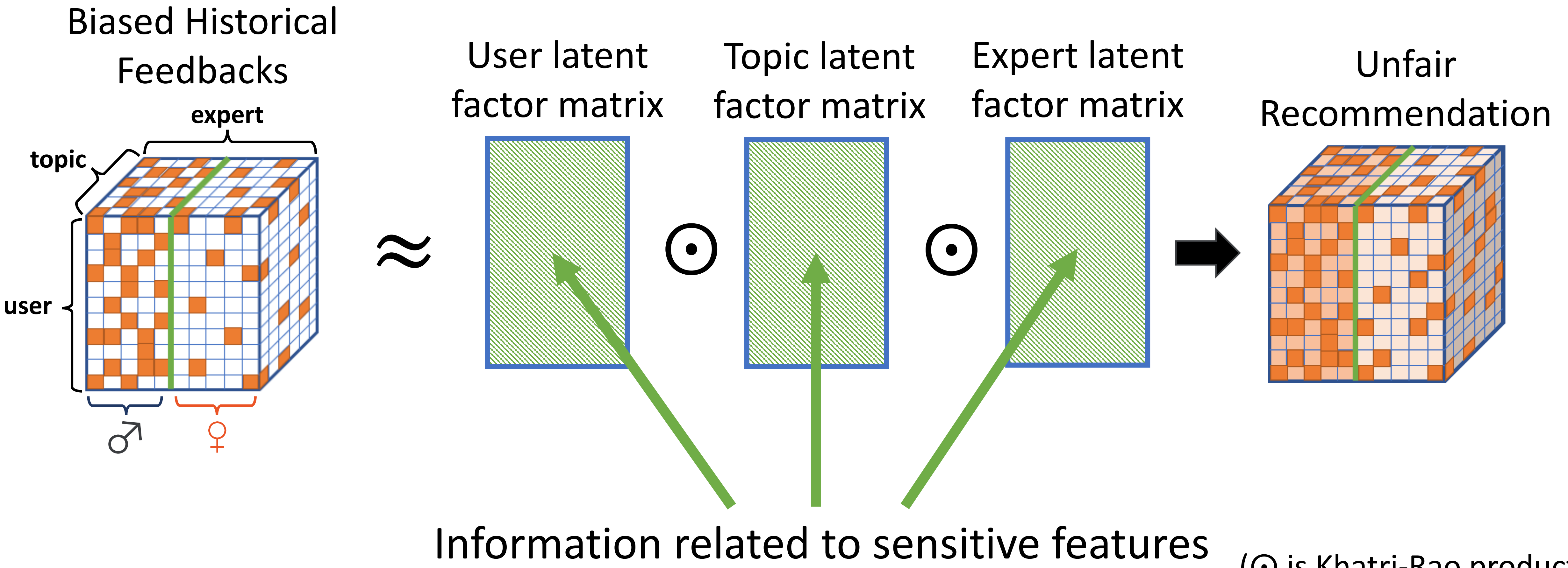
(⊙ is Khatri-Rao product)

Conventional Tensor Completion



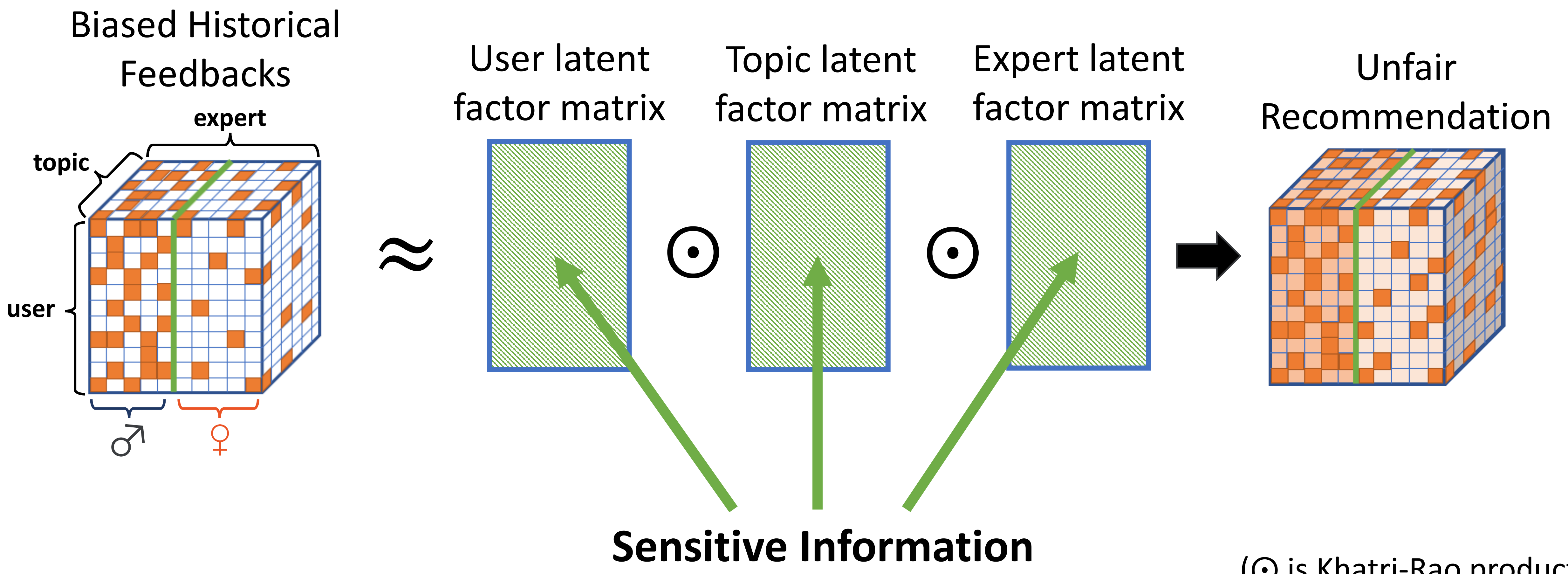
(⊙ is Khatri-Rao product)

Conventional Tensor Completion



(⊙ is Khatri-Rao product)

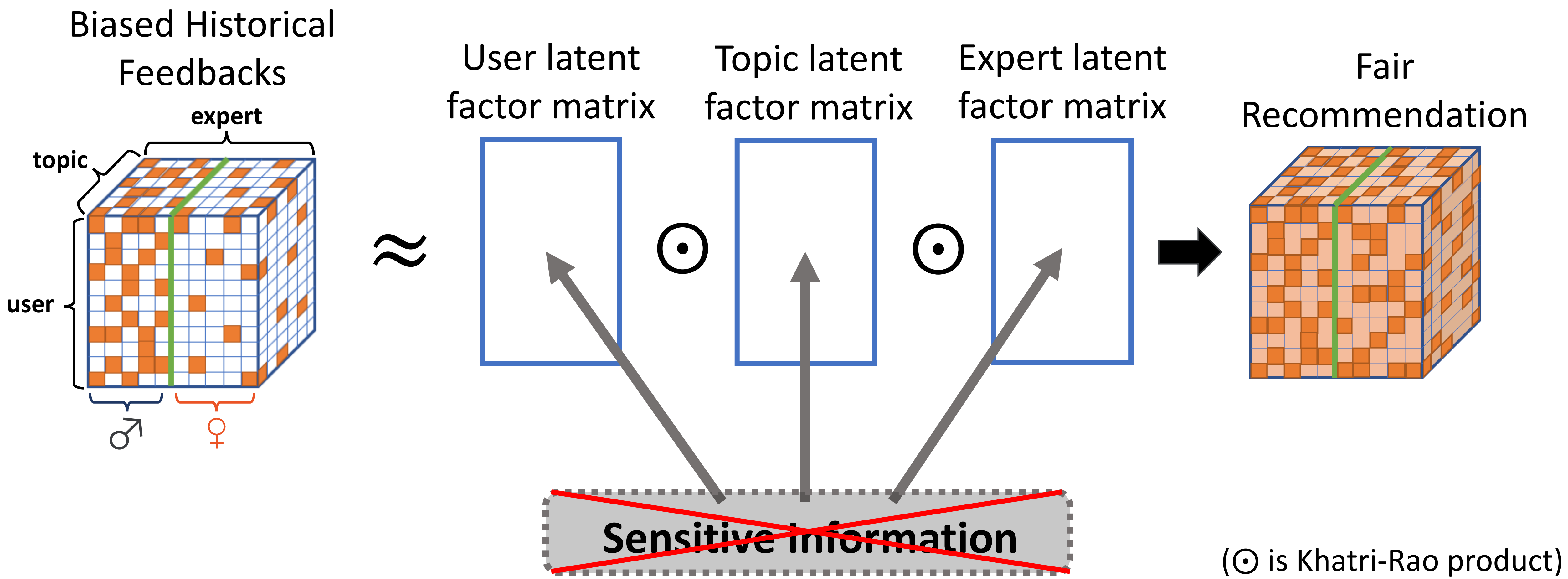
Conventional Tensor Completion



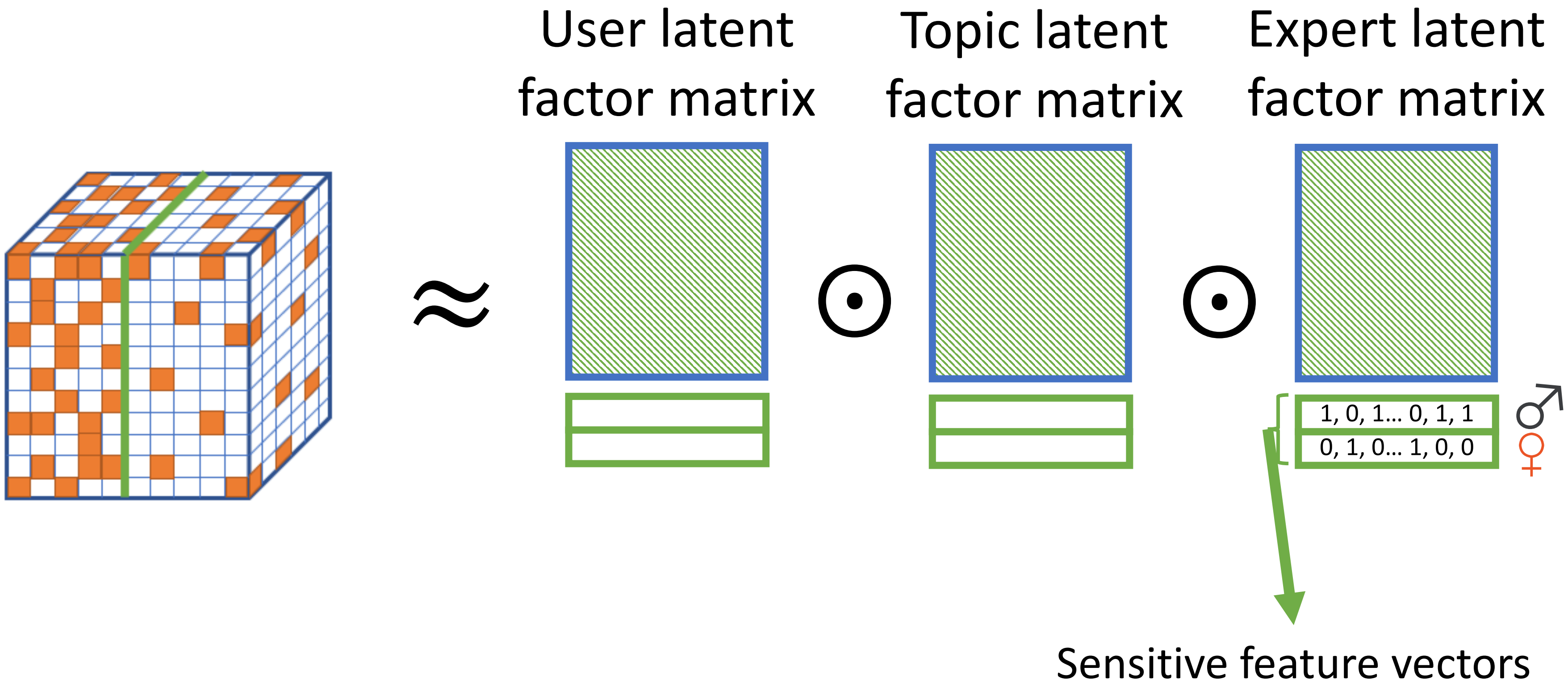
(⊙ is Khatri-Rao product)

Intuition

Remove Sensitive Information ➡ Enhance Fairness

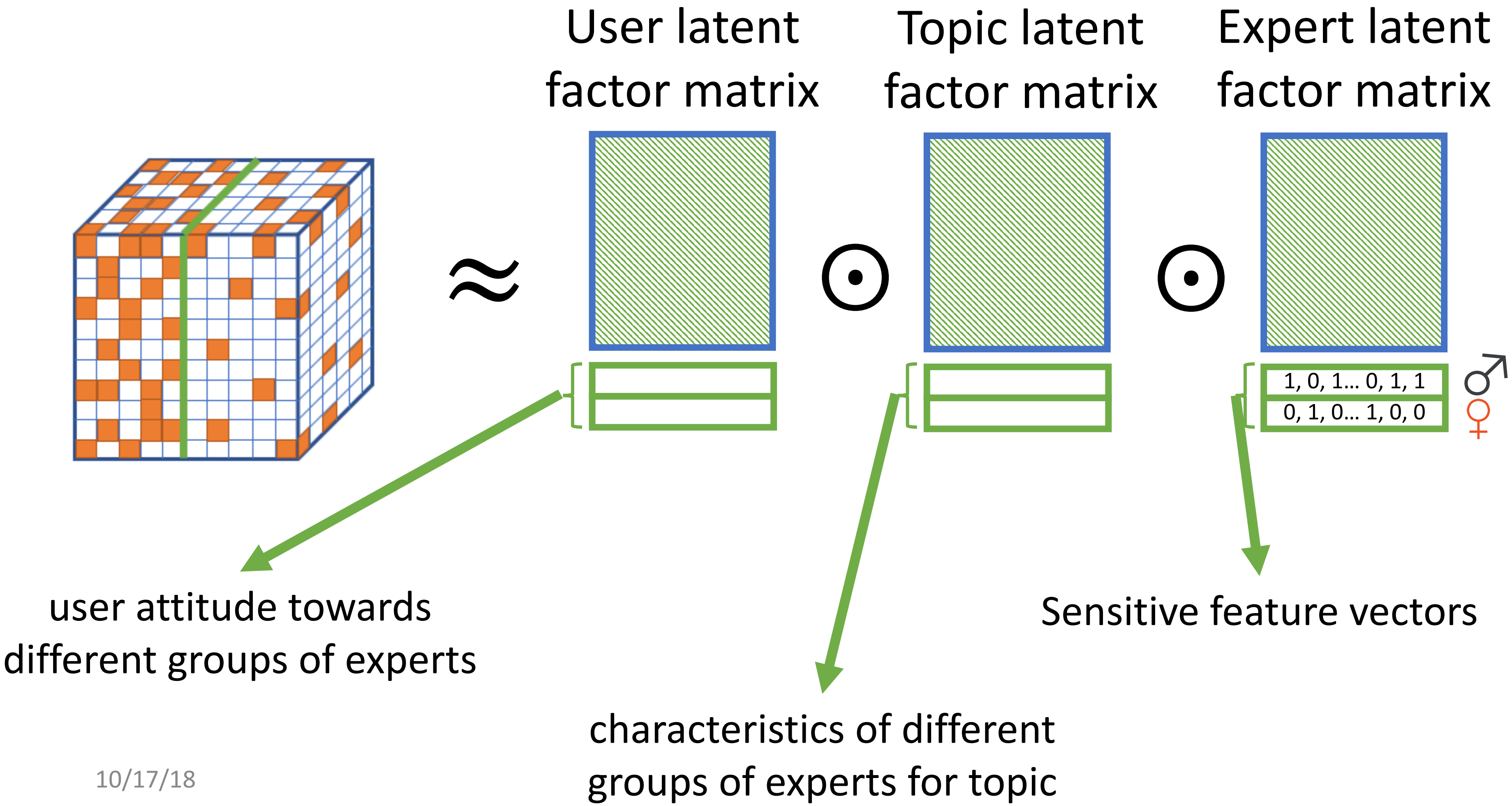


Sensitive Feature Isolation



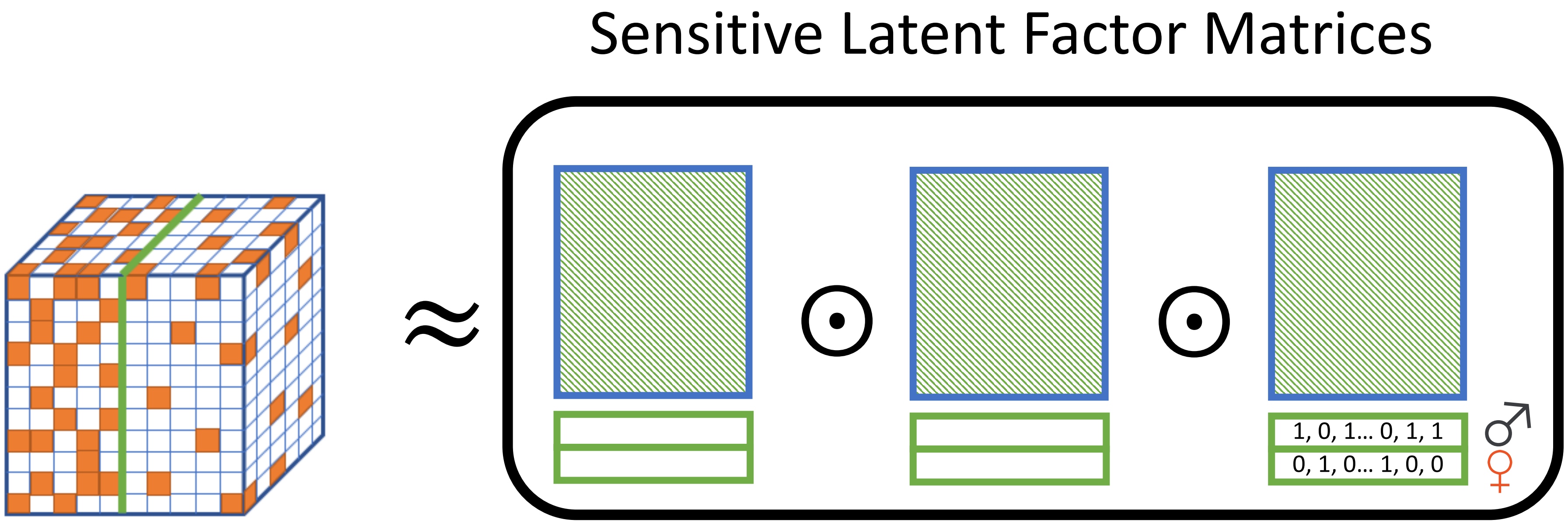
(⊙ is Khatri-Rao product)

Sensitive Feature Isolation



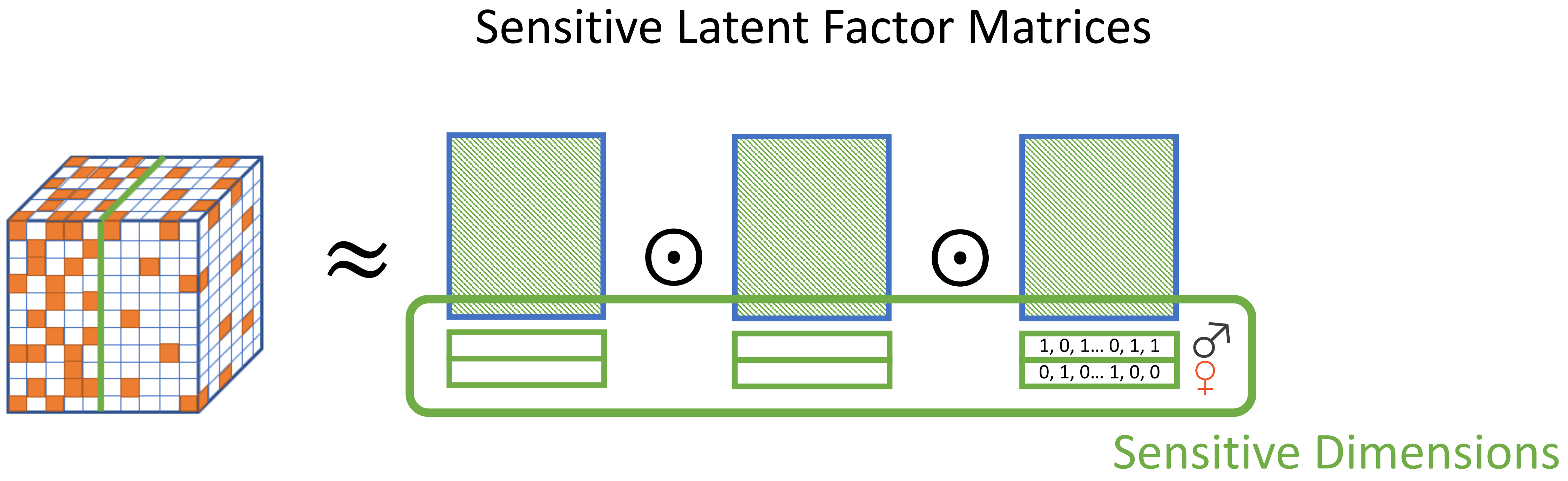
(⊙ is Khatri-Rao product)

Sensitive Feature Isolation



(⊙ is Khatri-Rao product)

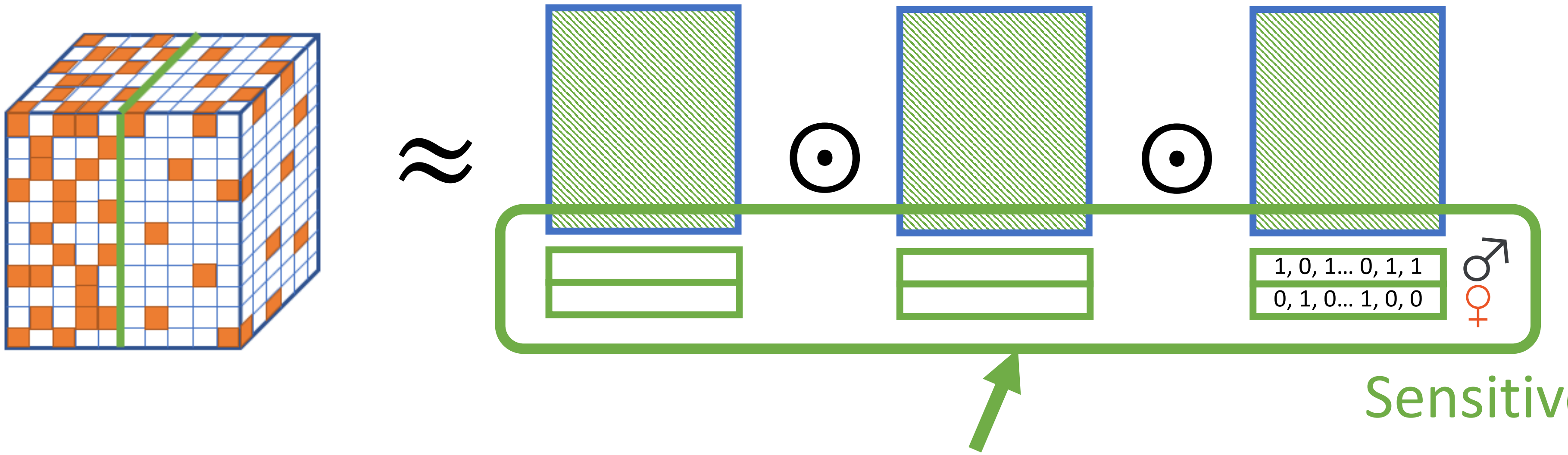
Sensitive Feature Isolation



(⊙ is Khatri-Rao product)

Sensitive Feature Isolation

Sensitive Latent Factor Matrices

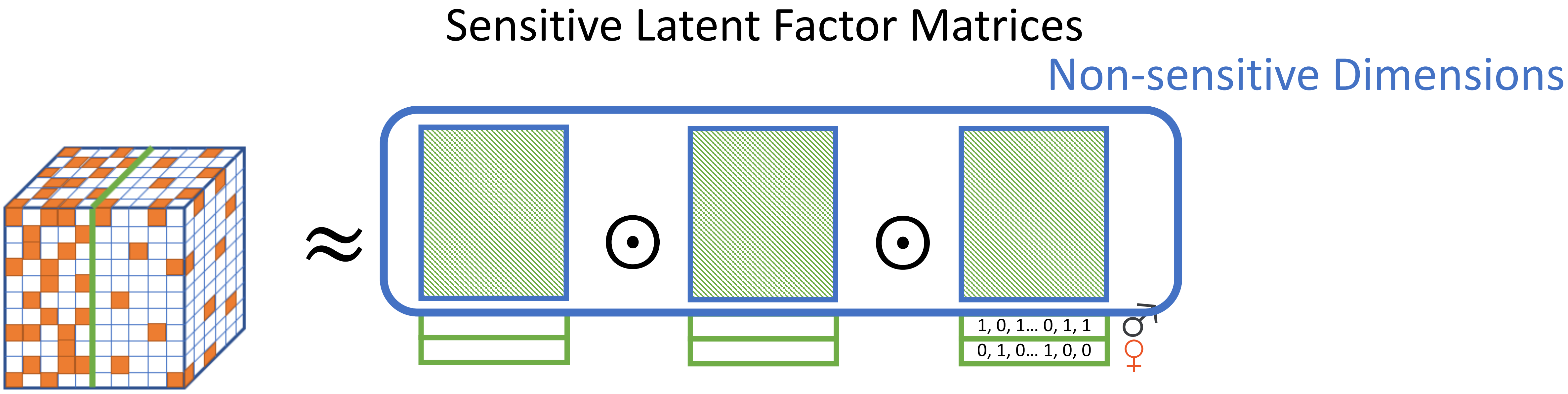


Sensitive Dimensions

Exclusively related to the **Sensitive Information**

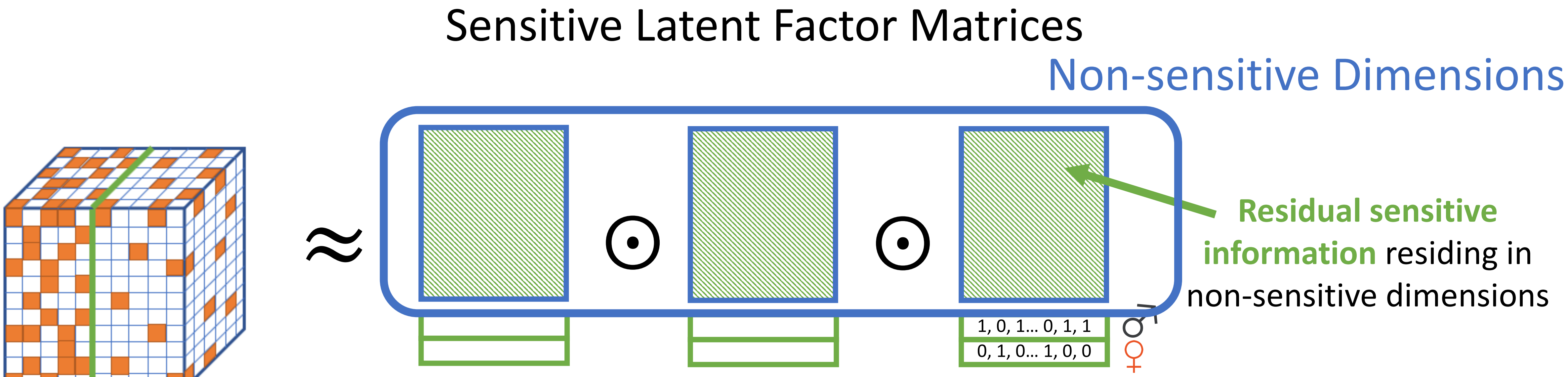
(⊙ is Khatri-Rao product)

Sensitive Feature Isolation



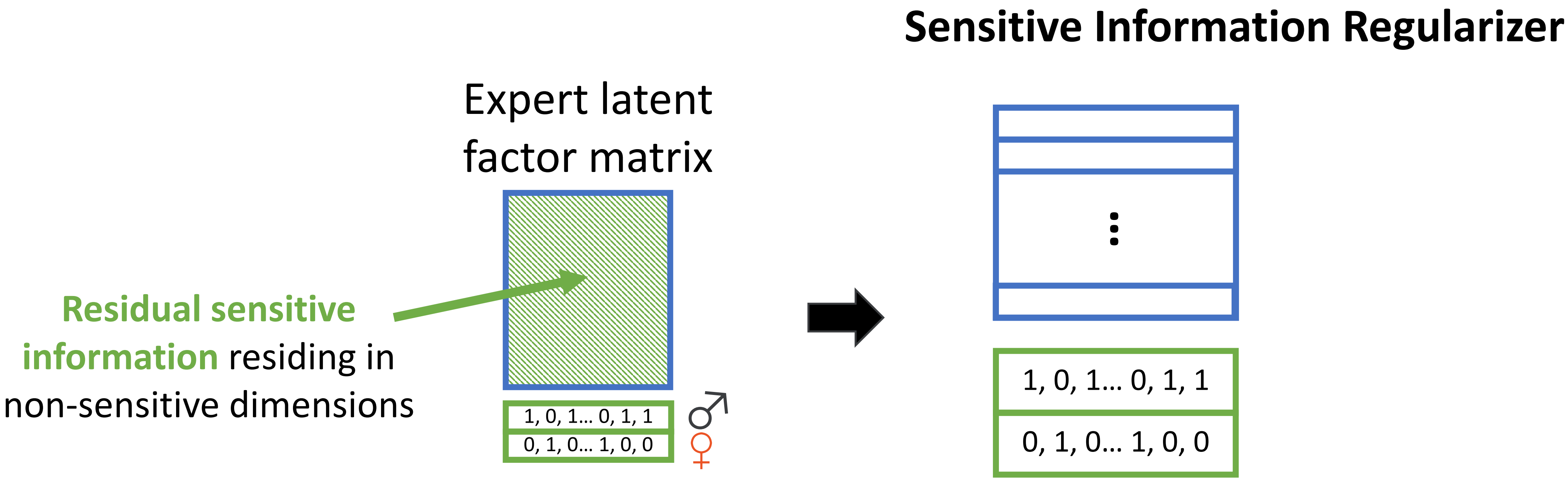
(⊙ is Khatri-Rao product)

Sensitive Feature Isolation

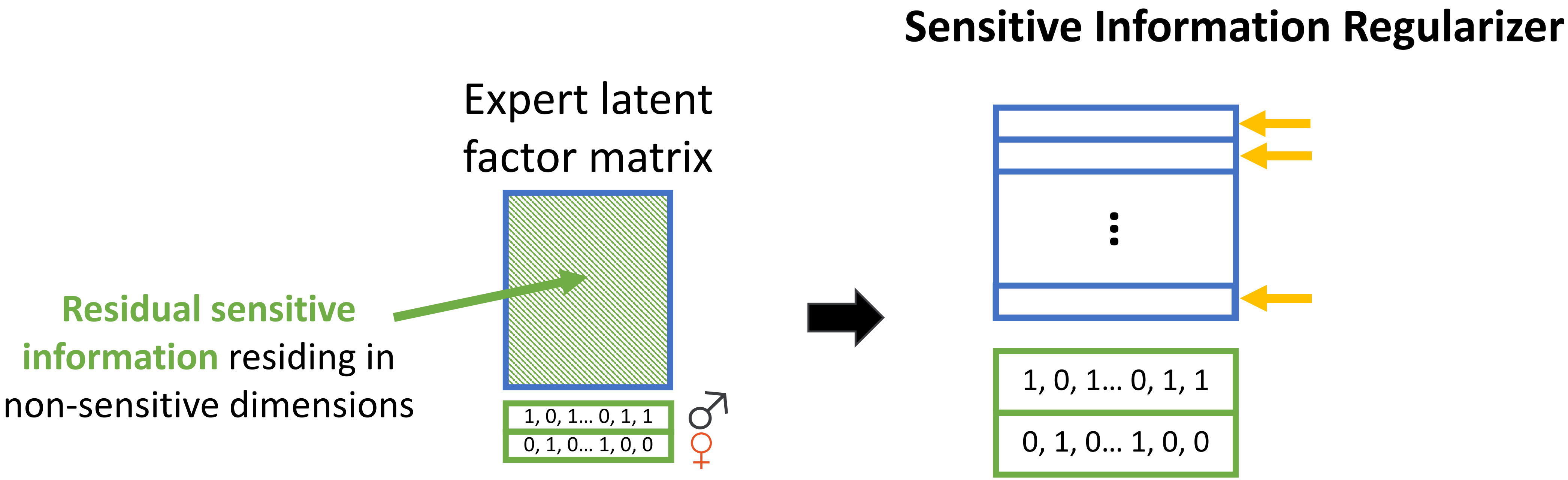


(⊙ is Khatri-Rao product)

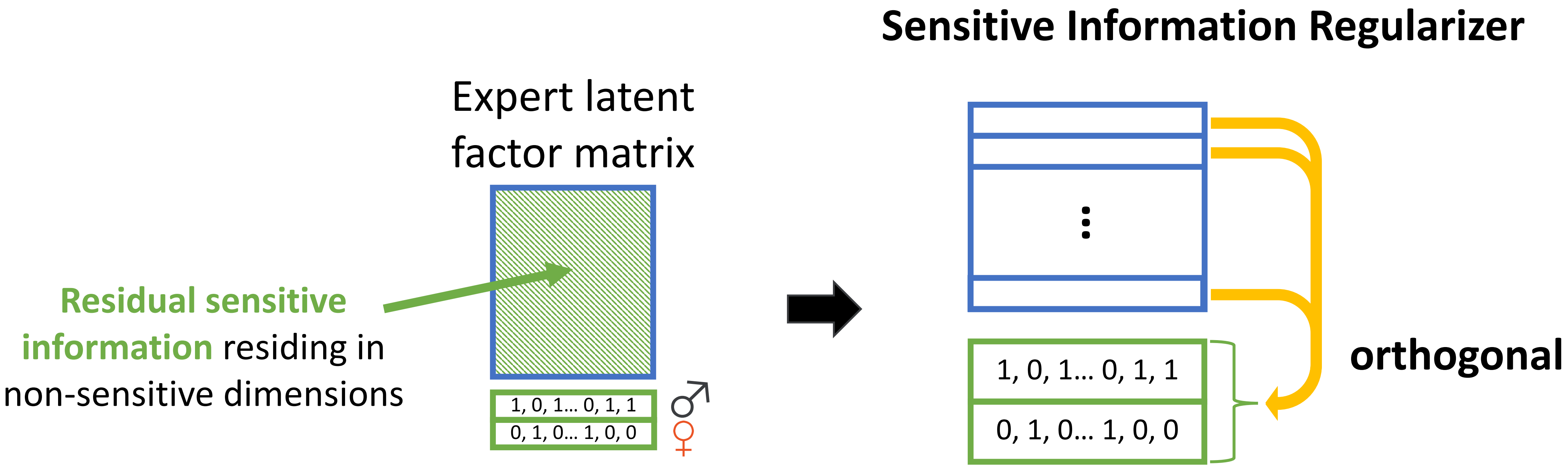
Sensitive Information Extraction



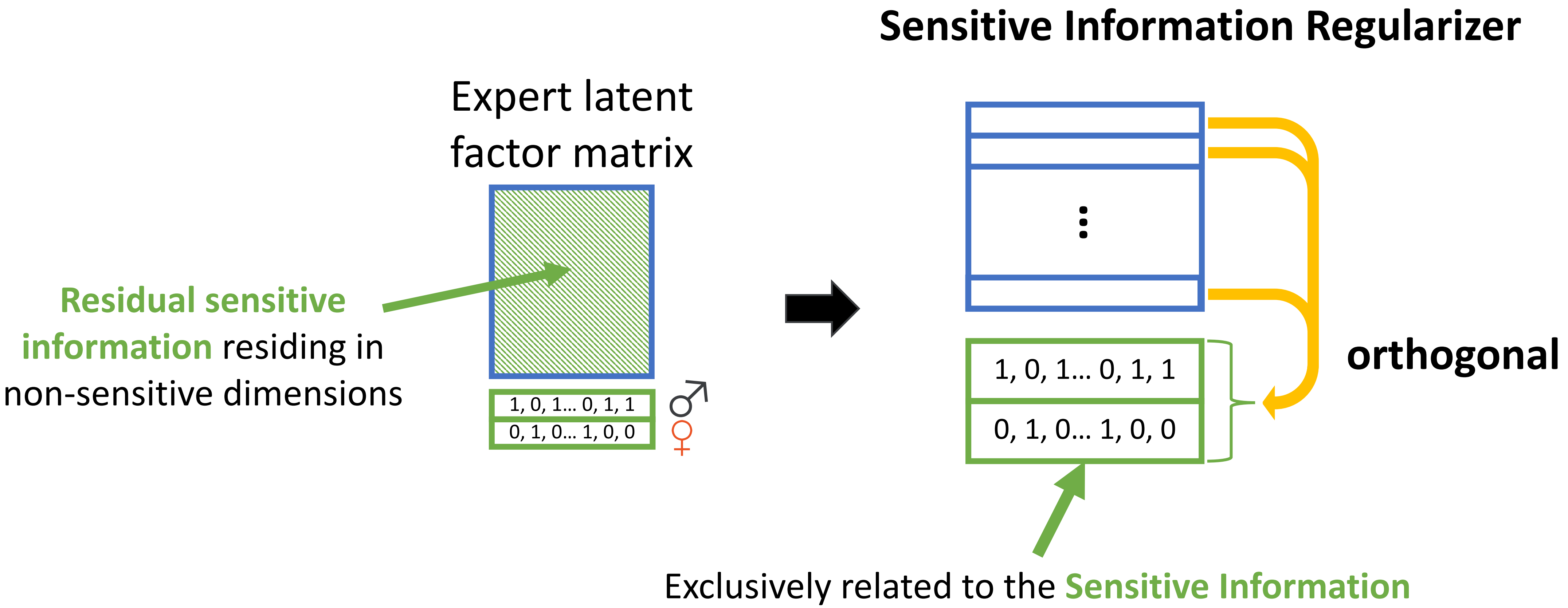
Sensitive Information Extraction



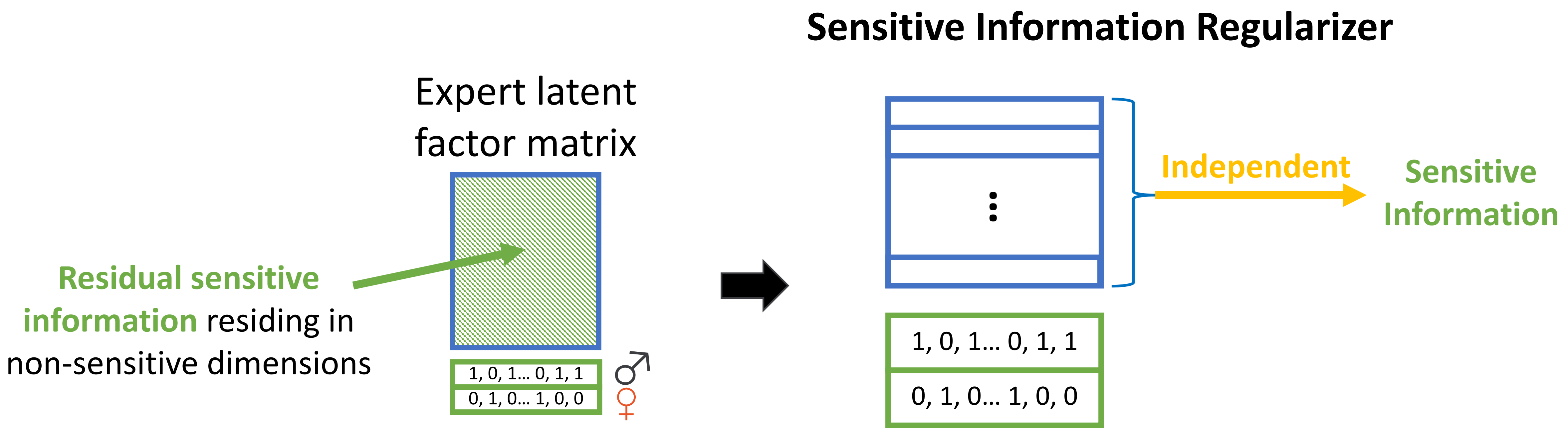
Sensitive Information Extraction



Sensitive Information Extraction

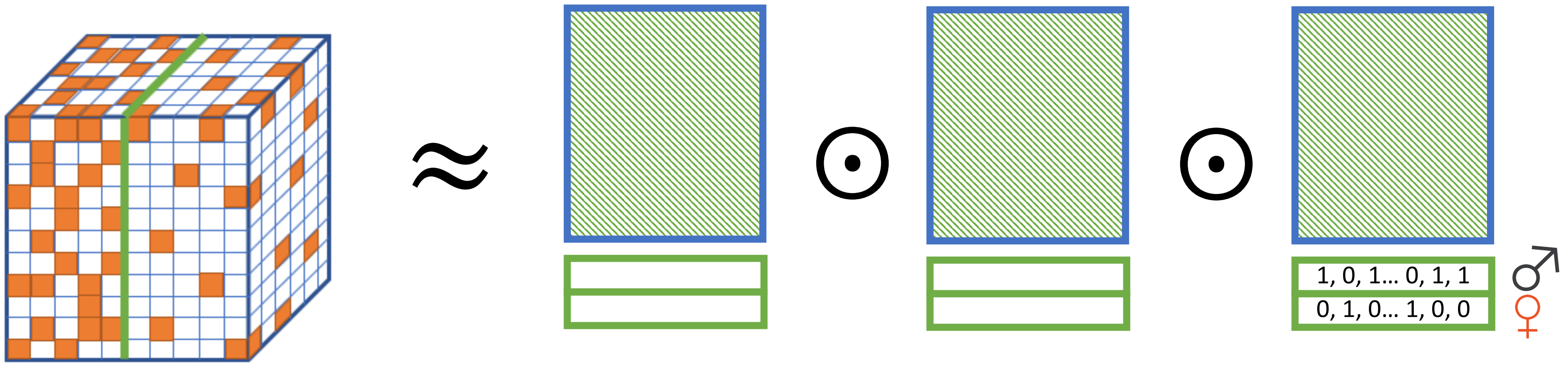


Sensitive Information Extraction



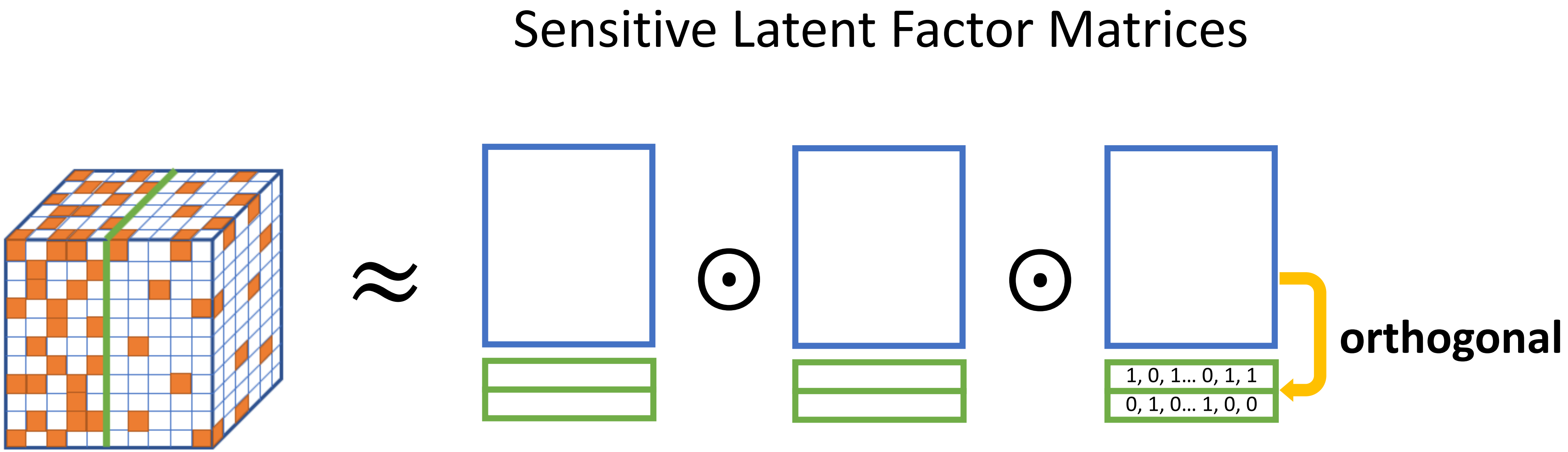
Sensitive Information Extraction

Sensitive Latent Factor Matrices



(⊙ is Khatri-Rao product)

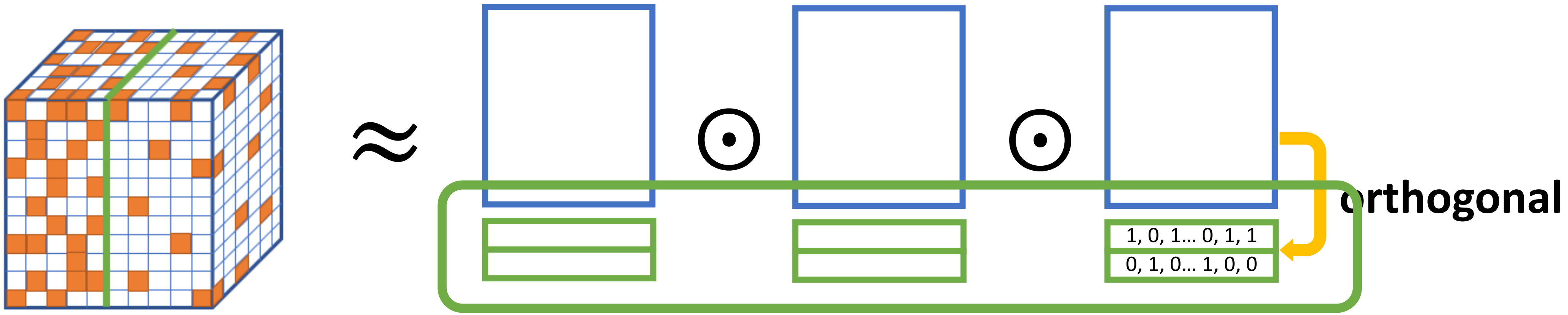
Sensitive Information Extraction



(⊙ is Khatri-Rao product)

Sensitive Information Extraction

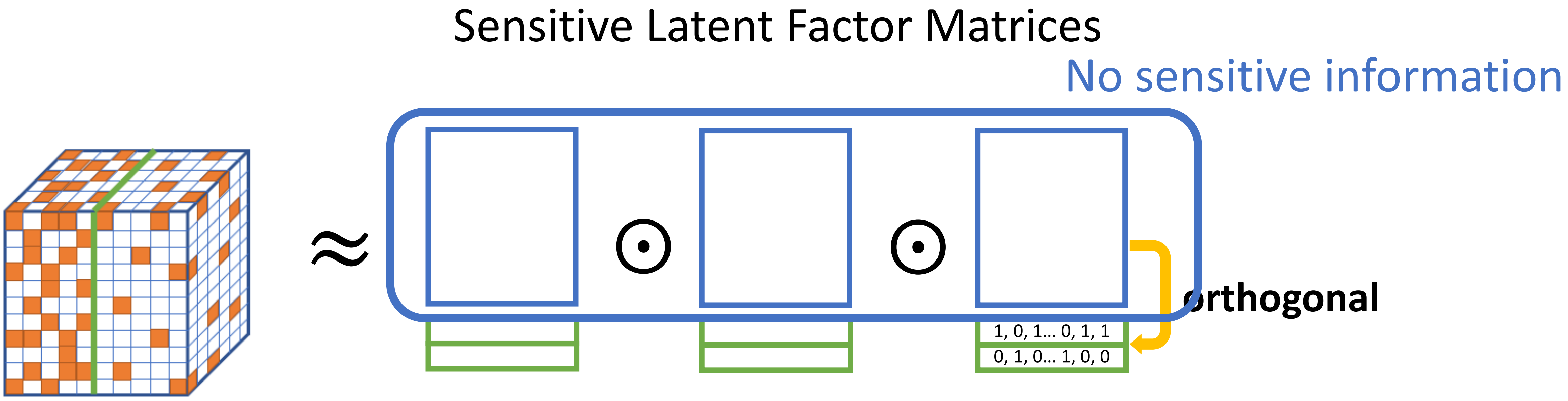
Sensitive Latent Factor Matrices



Sensitive information
extracted and isolated

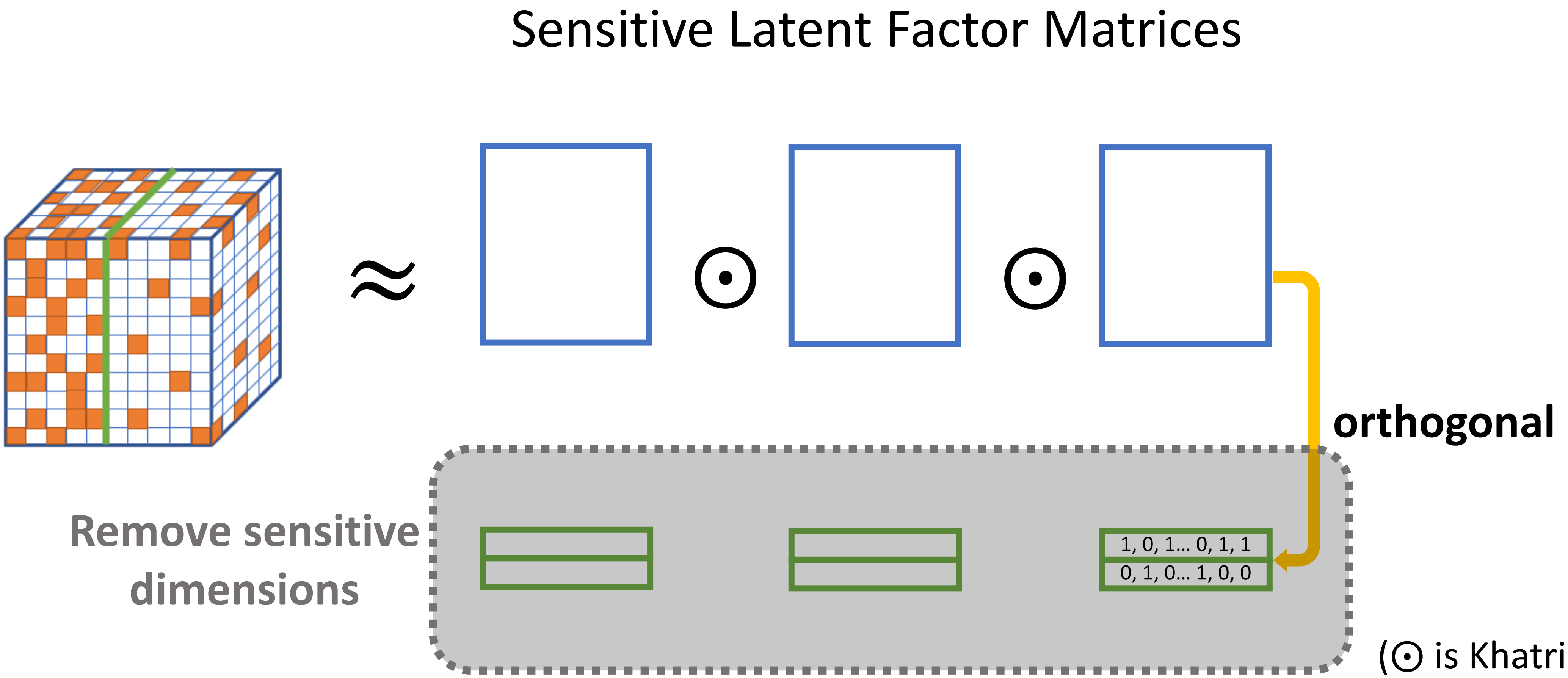
(⊙ is Khatri-Rao product)

Sensitive Information Extraction



(⊙ is Khatri-Rao product)

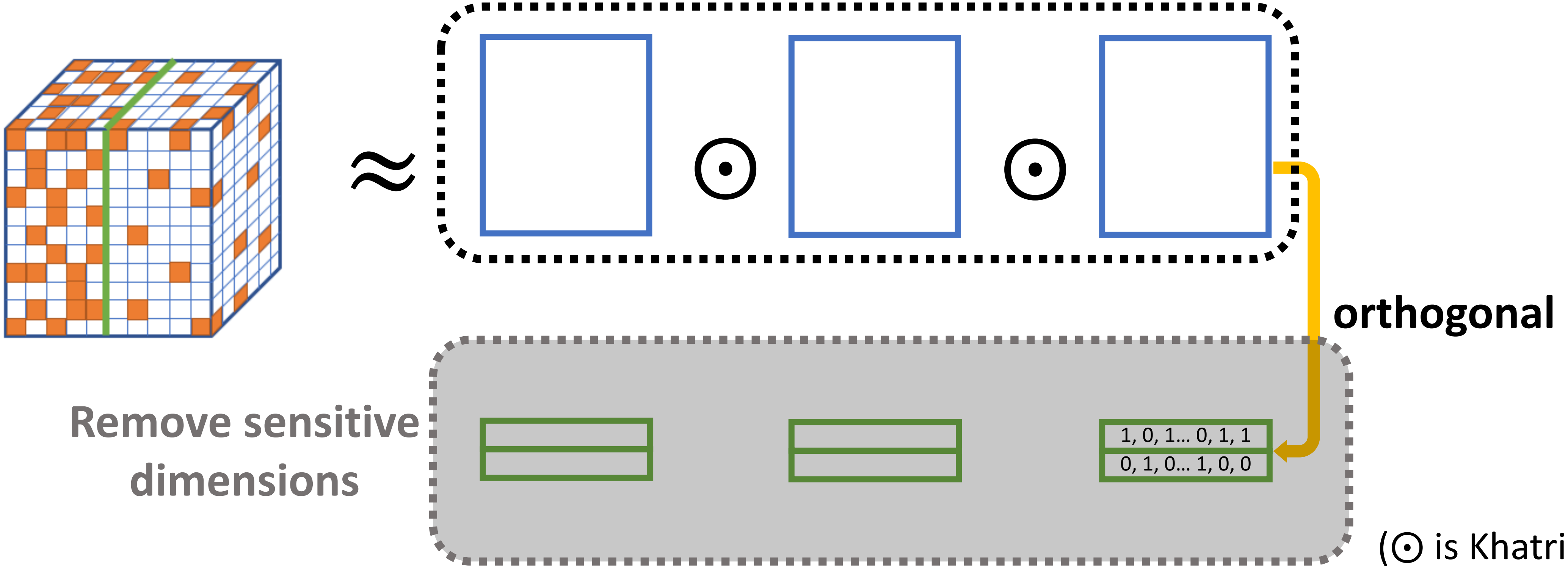
Fairness-aware Recommendation



(⊙ is Khatri-Rao product)

Fairness-aware Recommendation

Reconstruct the tensor by **non-sensitive** dimensions

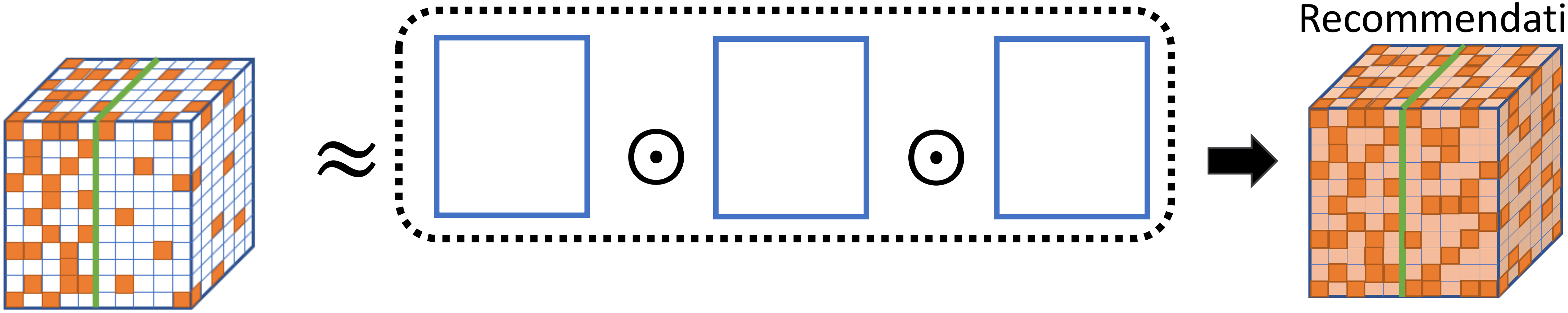


(⊙ is Khatri-Rao product)

Fairness-aware Recommendation

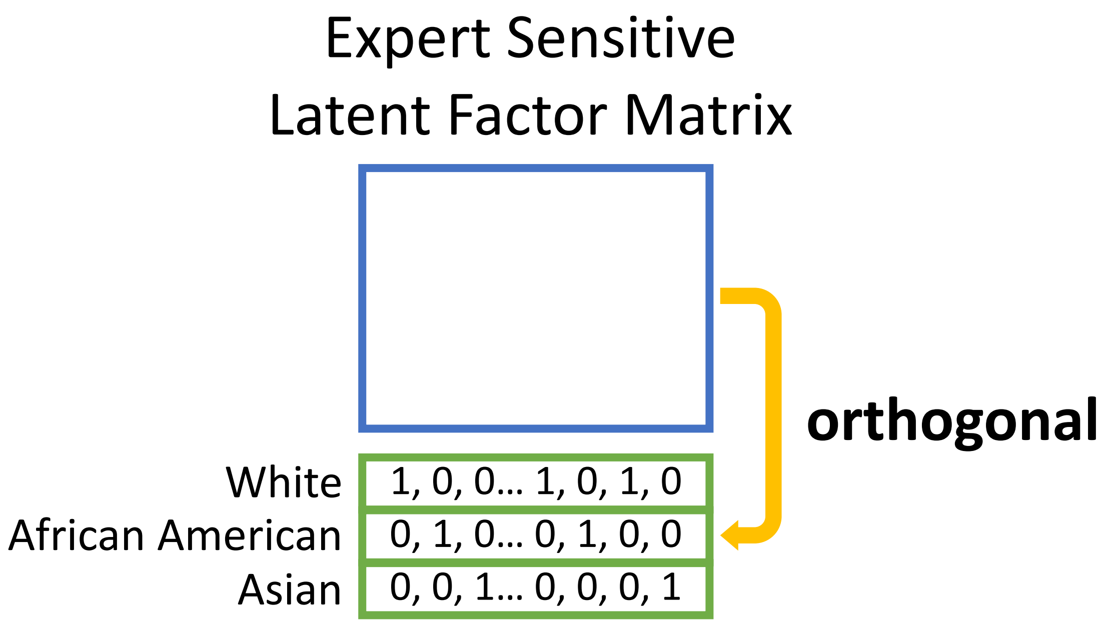
Reconstruct the tensor by **non-sensitive** dimensions

Fairness-Aware Recommendation

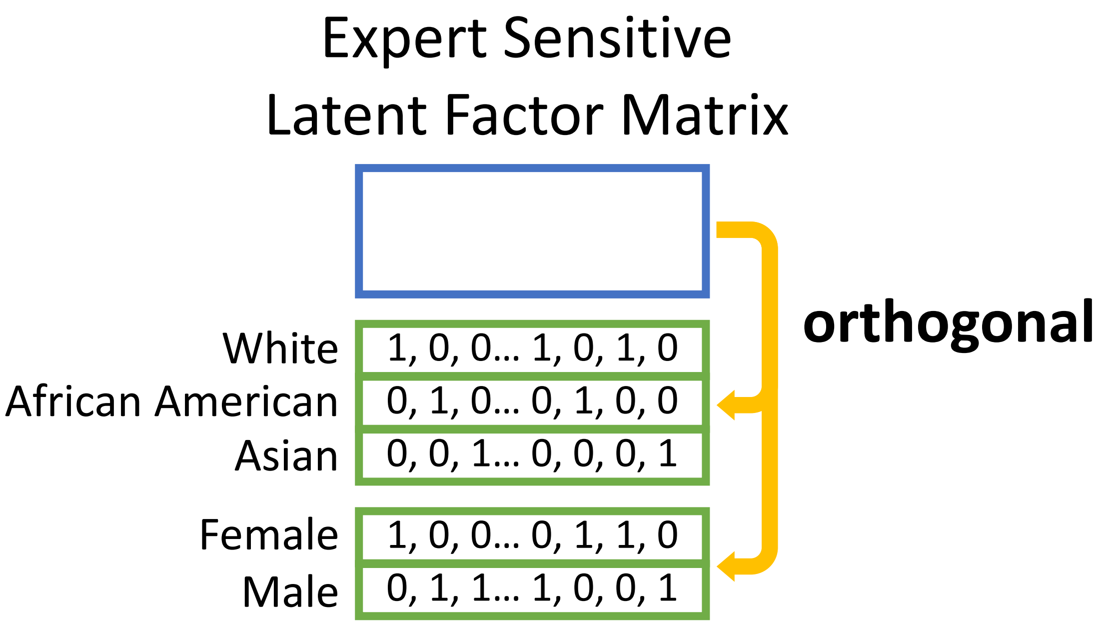


(⊙ is Khatri-Rao product)

Generalizing FATR – Multi-category



Generalizing FATR – Multi-feature



Experiment – Scenarios & Datasets

- 3D scenario: User-Expert-Topic Twitter – ethnicity of the expert as the sensitive feature (white vs. non-white);

Part of experiment is omitted. Refer to the paper for more details.

Experiment – Scenarios & Datasets

- 3D scenario: User-Expert-Topic Twitter – ethnicity of the expert as the sensitive feature (white vs. non-white);

➢ Varying considerations: Twelve Synthetic Expert Datasets (four levels of bias & three levels of sparsity);

Part of experiment is omitted. Refer to the paper for more details.

Experiment – Scenarios & Datasets

- 3D scenario: User-Expert-Topic Twitter – ethnicity of the expert as the sensitive feature (white vs. non-white);
- Varying considerations: Twelve Synthetic Expert Datasets (four levels of bias & three levels of sparsity);

➢ Generalizing scenario: User-Expert-Topic Twitter Dataset – both ethnicity (three categories) and gender (two categories) as sensitive features.

Part of experiment is omitted. Refer to the paper for more details.

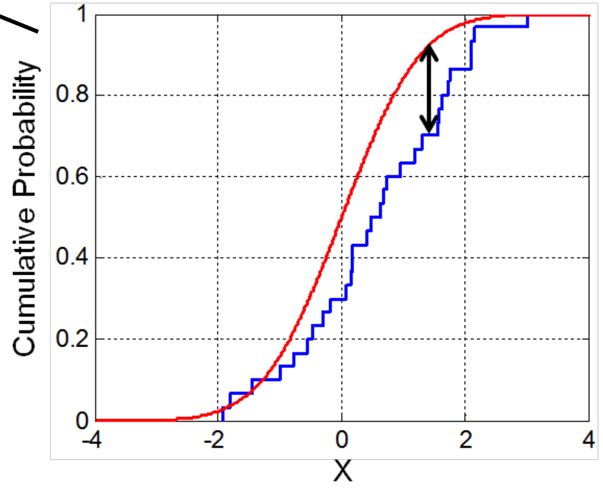
Experiment – Metrics

- Recommendation Quality: **Precision@k**, **Recall@k**, and **F1@k** (higher is better)

Experiment – Metrics

- Recommendation Quality: **Precision@k**, **Recall@k**, and **F1@k** (higher is better)
- Recommendation Fairness: **MAD** and **KS** (lower is better)

$$|Mean(R_1) - Mean(R_2)|$$



Experiment – Baselines

We consider two variations of FATR – **FT(G)** (using Gradient Descent) and **FT(N)** (using Newton's Method)

Experiment – Baselines

We consider two variations of FATR – **FT(G)** (using Gradient Descent) and **FT(N)** (using Newton's Method) – in comparison with

- Ordinary Tensor Completion (**OTC**) – Fairness-unaware
- Regularization-based Tensor Completion (**RTC**) – Fairness-aware

Experiment – Baselines

We consider two variations of FATR – **FT(G)** (using Gradient Descent) and **FT(N)** (using Newton's Method) – in comparison with

- Ordinary Tensor Completion (**OTC**) – Fairness-unaware
- Regularization-based Tensor Completion (**RTC**) – Fairness-aware
- Ordinary Matrix Completion (**OMC**) – Fairness-unaware
- Regularization-based Matrix Completion (**RMC**) – Fairness-aware
- Matrix-based variations of FATR – **FM(G)** and **FM(N)**

Experiment – Twitter

Research Questions:

- What is the different between **Matrix-based** vs. **Tensor-based**?
- How does **FATR** perform in comparison with **baselines**?

Experiment – Twitter (Matrix vs. Tensor)

- What is the different between **Matrix-based** vs. **Tensor-based**?

Recommendation Quality: Tensor-based **better** than Matrix-based

(higher is better)

Methods	R@15	P@15
OMC	0.3467	0.0842
OTC	0.4384	0.0958
RMC	0.1609	0.0702
RTC	0.3003	0.0515
FM(G)	0.4045	0.0891
FT(G)	0.4180	0.0870
FM(N)	0.3298	0.0687
FT(N)	0.3975	0.0786

Experiment – Twitter (Matrix vs. Tensor)

- What is the different between **Matrix-based** vs. **Tensor-based**?

Recommendation Fairness: Tensor-based **worse** than Matrix-based for baselines
(lower is better)

Methods	KS	MAD
OMC	0.1660	0.0122
OTC	0.3662	0.0333
RMC	0.1521	0.0086
RTC	0.2003	0.0171
FM(G)	0.0523	0.0037
FT(G)	0.0195	0.0024
FM(N)	0.0245	0.0044
FT(N)	0.0173	0.0029

Experiment – Twitter (Matrix vs. Tensor)

- What is the different between **Matrix-based** vs. **Tensor-based**?

Recommendation Fairness: Tensor-based **better** than Matrix-based for FATR
(lower is better)

Methods	KS	MAD
OMC	0.1660	0.0122
OTC	0.3662	0.0333
RMC	0.1521	0.0086
RTC	0.2003	0.0171
FM(G)	0.0523	0.0037
FT(G)	0.0195	0.0024
FM(N)	0.0245	0.0044
FT(N)	0.0173	0.0029

Experiment – Twitter (FATR vs. Baselines)

- How does **FATR** perform in comparison with **baselines**?

Recommendation Quality: FATR **better** than RTC
(higher is better)

Methods	R@15	P@15
OMC	0.3467	0.0842
OTC	0.4384	0.0958
RMC	0.1609	0.0702
RTC	0.3003	0.0515
FM(G)	0.4045	0.0891
FT(G)	0.4180	0.0870
FM(N)	0.3298	0.0687
FT(N)	0.3975	0.0786

Experiment – Twitter (FATR vs. Baselines)

- How does **FATR** perform in comparison with **baselines**?

Recommendation Quality: FATR **slightly worse** than OTC
(higher is better)

Methods	R@15	P@15
OMC	0.3467	0.0842
OTC	0.4384	0.0958
RMC	0.1609	0.0702
RTC	0.3003	0.0515
FM(G)	0.4045	0.0891
FT(G)	0.4180	0.0870
FM(N)	0.3298	0.0687
FT(N)	0.3975	0.0786

Experiment – Twitter (FATR vs. Baselines)

- How does **FATR** perform in comparison with **baselines**?

Recommendation Fairness: FATR **better** than RTC and OTC
(lower is better)

Methods	KS	MAD
OMC	0.1660	0.0122
OTC	0.3662	0.0333
RMC	0.1521	0.0086
RTC	0.2003	0.0171
FM(G)	0.0523	0.0037
FT(G)	0.0195	0.0024
FM(N)	0.0245	0.0044
FT(N)	0.0173	0.0029

Experiment – Synthetic

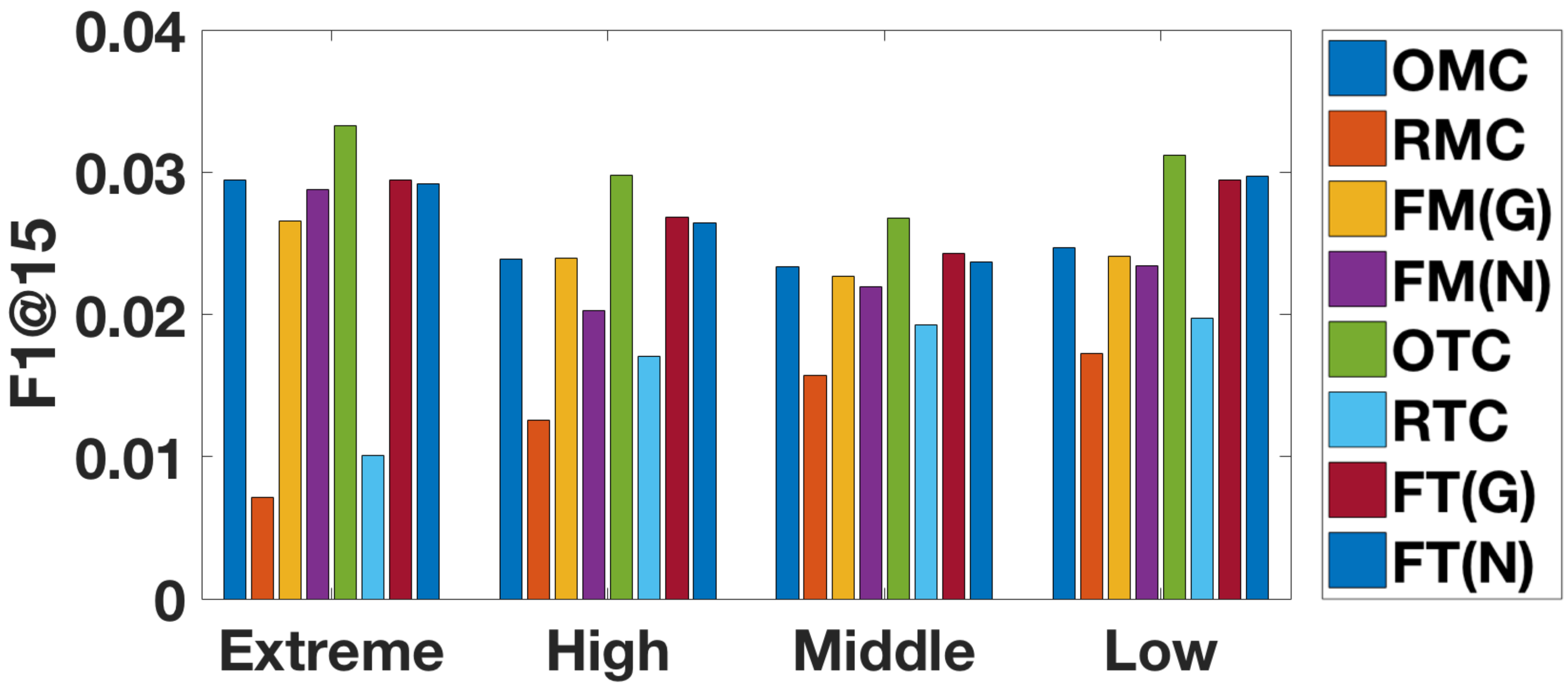
Research Questions:

- Is FATR **robust** to the impact of data **bias** (four levels: low, medium, high, extreme)?
- Is FATR **robust** to the impact of data **sparsity** (three levels: low, medium, high)?

Experiment – Synthetic (Impact of Bias)

- Is FATR **robust** to the impact of data **bias**?

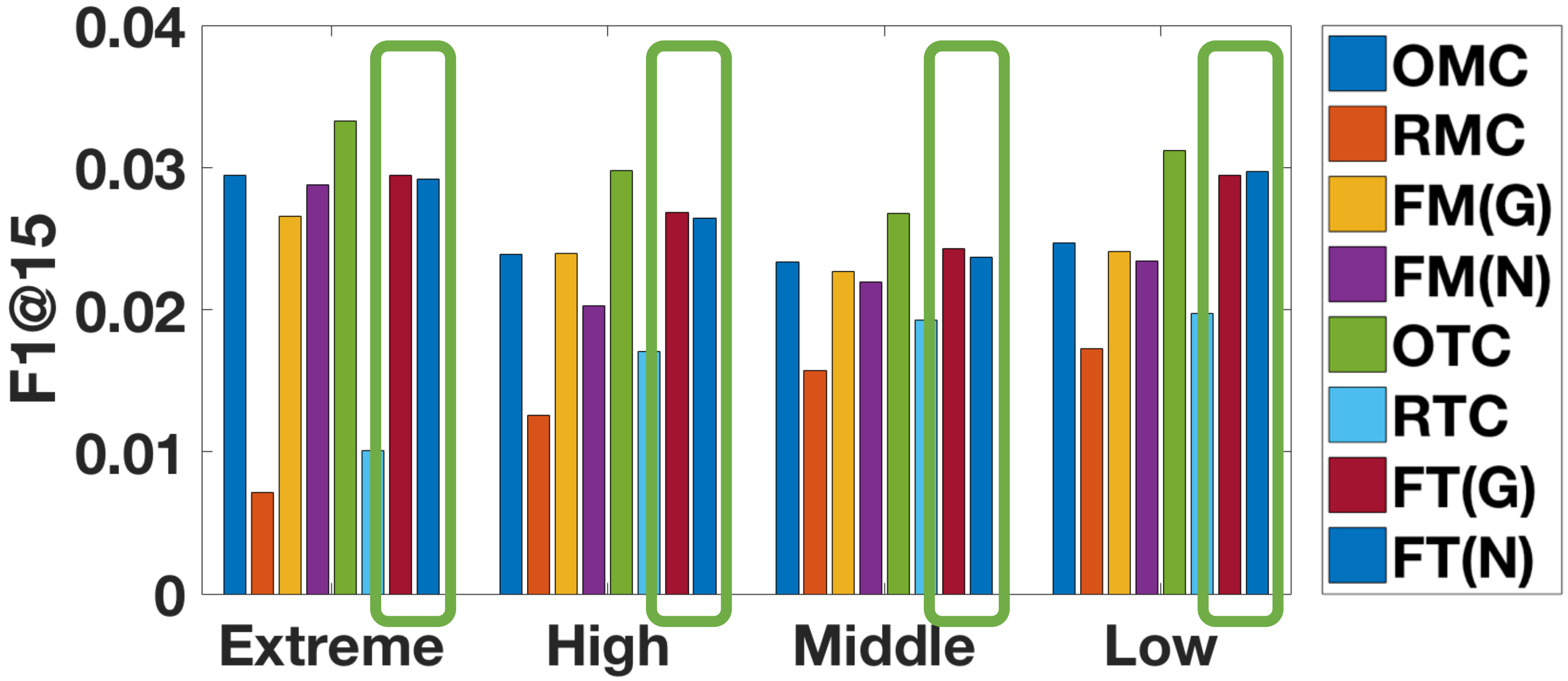
Recommendation Quality (higher is better)



Experiment – Synthetic (Impact of Bias)

- Is FATR **robust** to the impact of data **bias**?

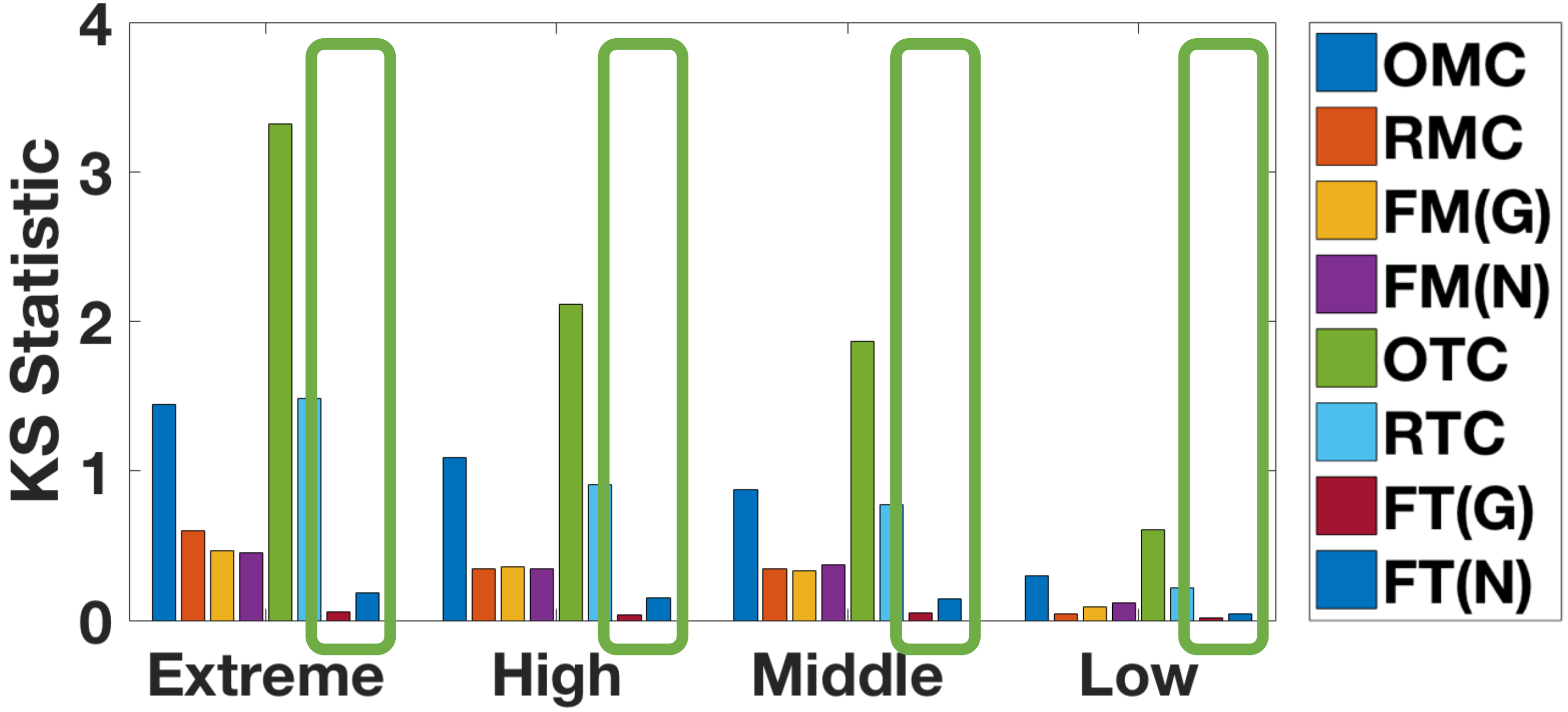
FATR provides **relatively high** recommendation quality under different bias situations



Experiment – Synthetic (Impact of Bias)

- Is FATR **robust** to the impact of data **bias**?

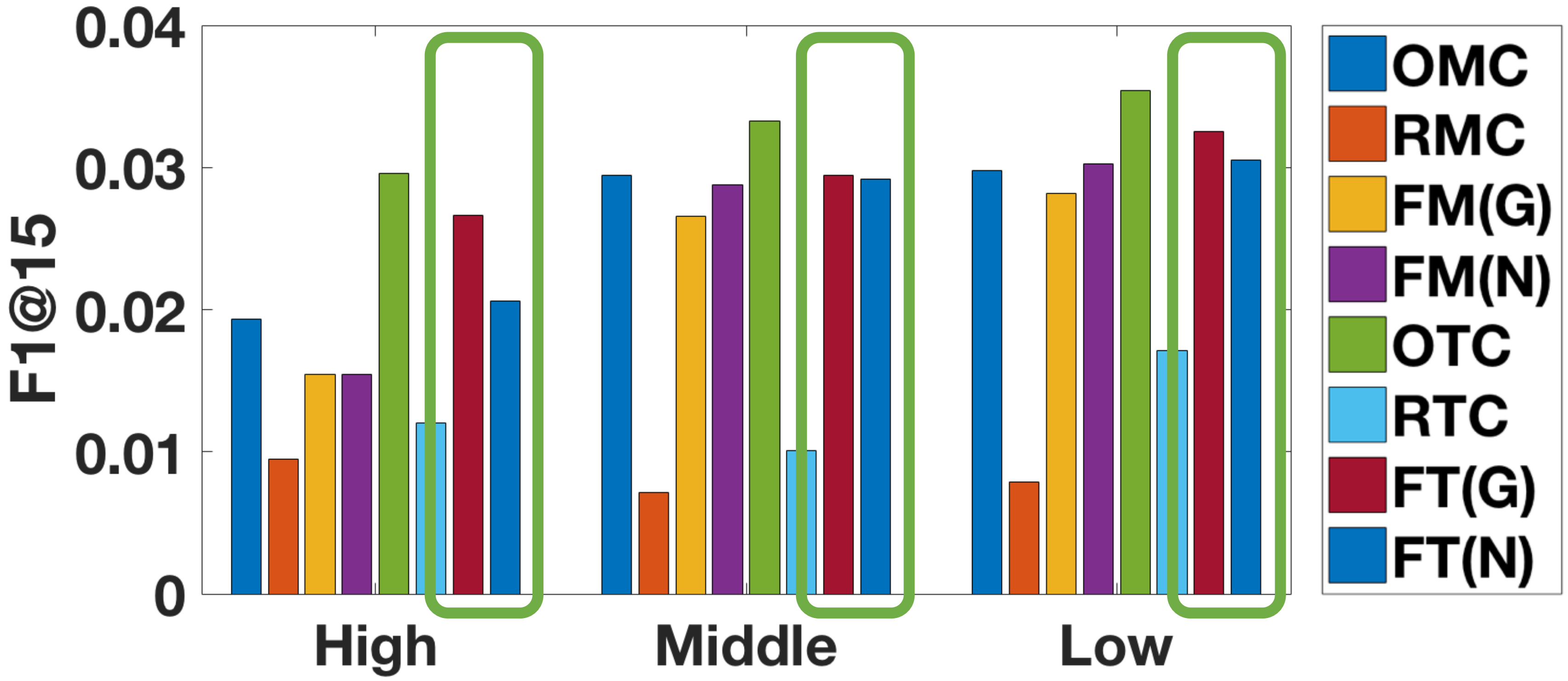
FATR **enhances fairness** to a great extent under different bias situations



Experiment – Synthetic (Impact of Sparsity)

- Is FATR **robust** to the impact of data **sparsity**?

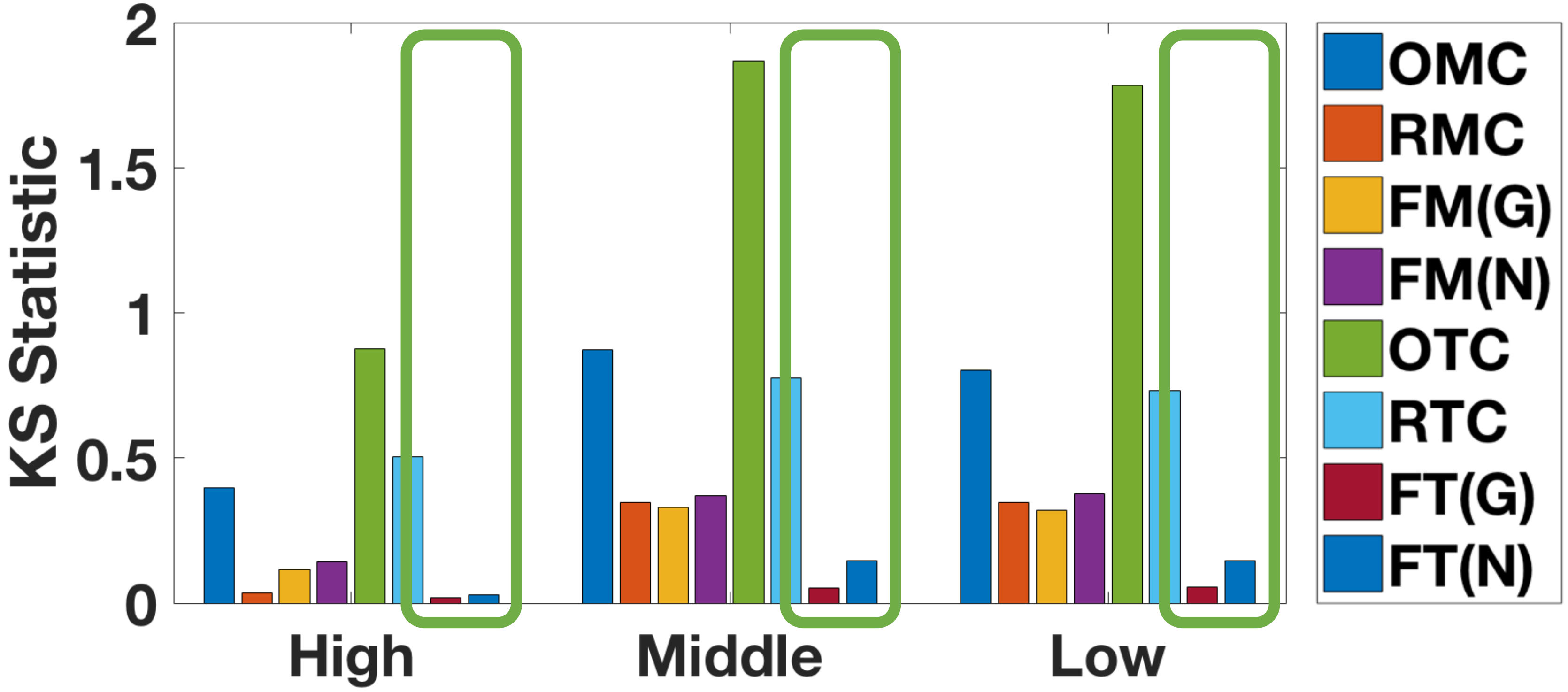
FATR provides **relatively high** recommendation quality under different sparsity situations



Experiment – Synthetic (Impact of Sparsity)

- Is FATR **robust** to the impact of data **sparsity**?

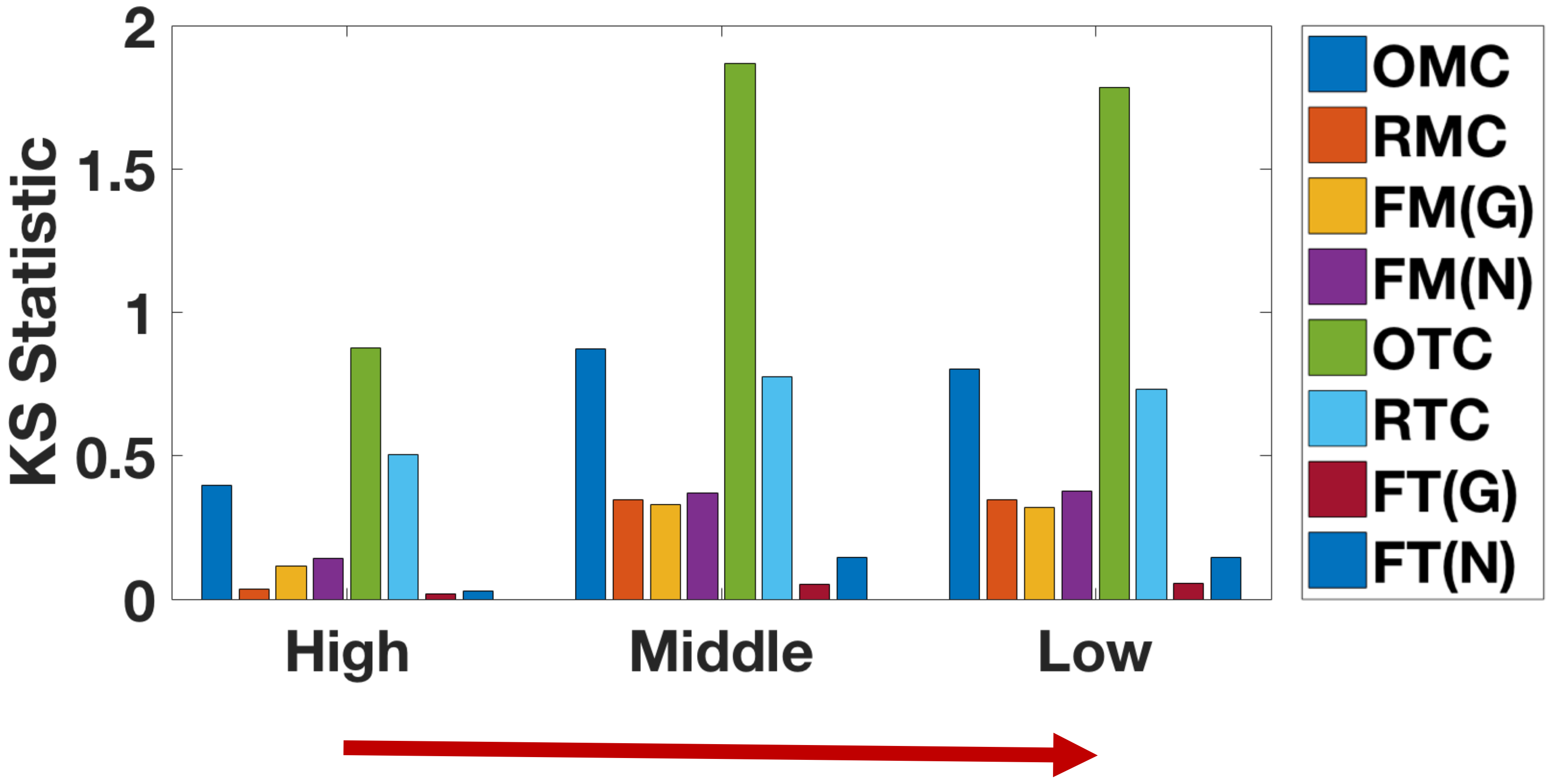
FATR **enhances fairness** to a great extent under different sparsity situations



Experiment – Synthetic (Impact of Sparsity)

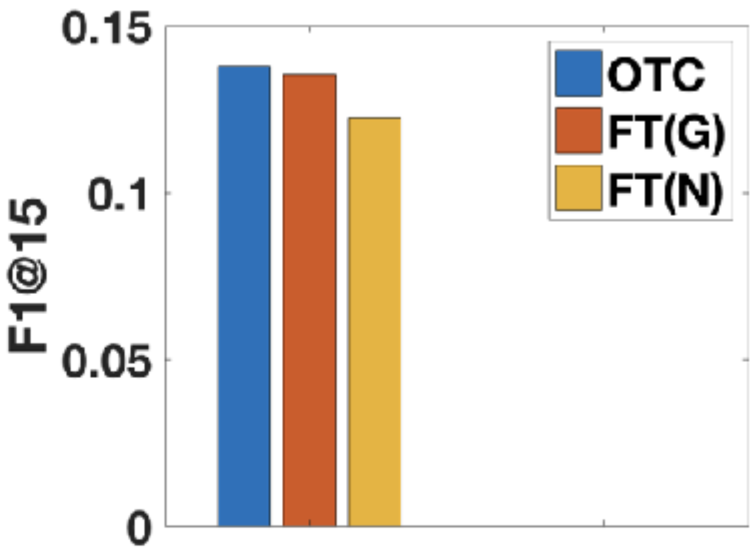
Besides …

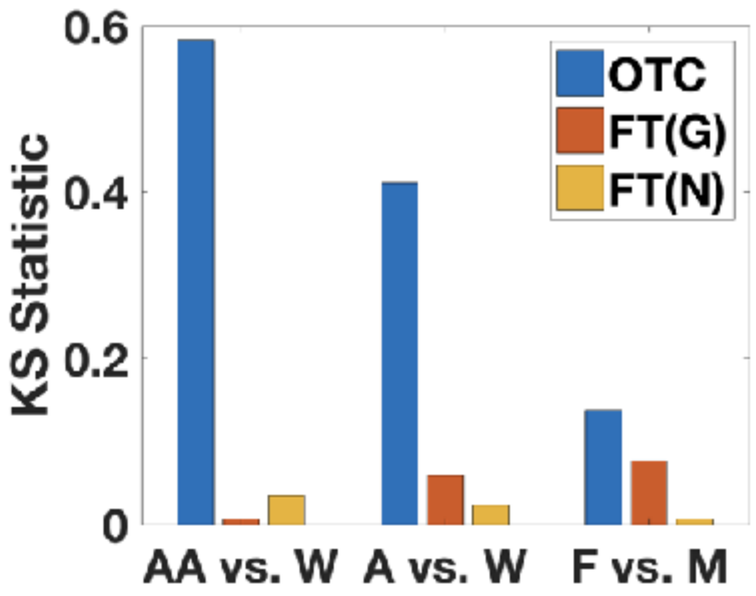
With data denser, recommendation **unfairness** goes more severe.



Experiment – Generalizing

- How does FATR perform for **multi-feature** and **multi-category** case?



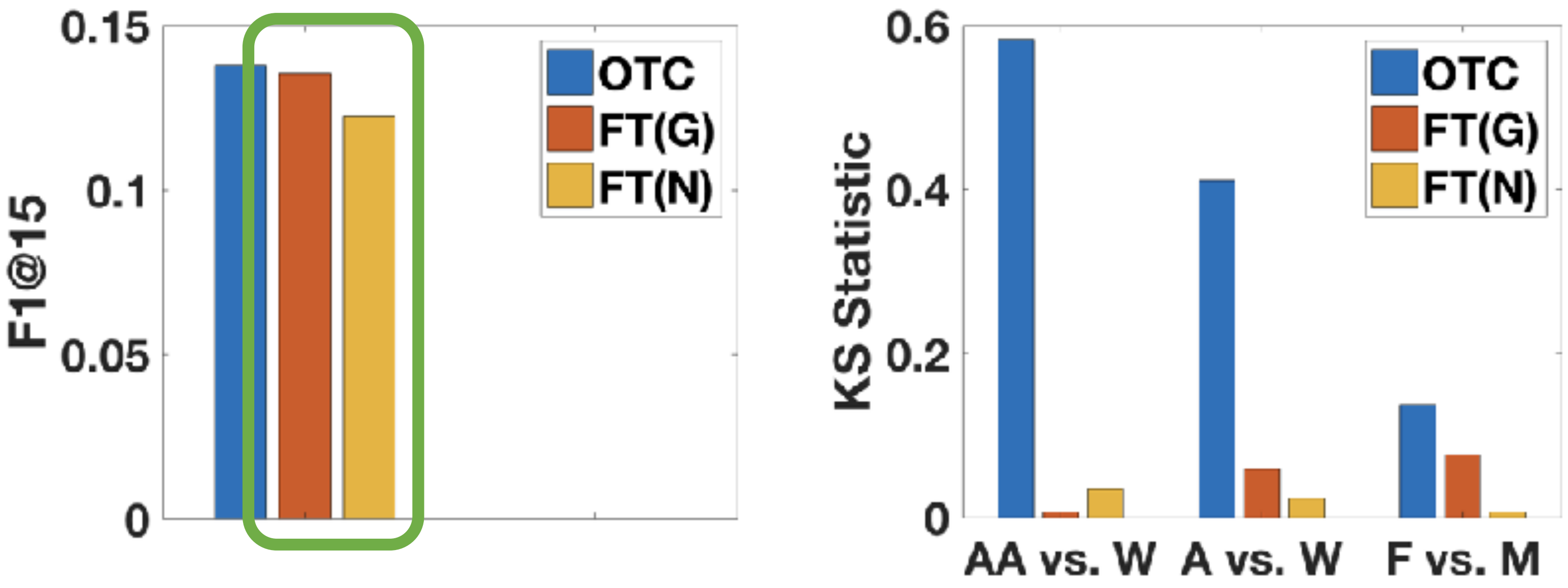


F: Female; **M**: Male
AA: American African; **W**: White; **A**: Asian

Experiment – Generalizing

- How does FATR perform for **multi-feature** and **multi-category** case?

FATR preserves **high** recommendation quality.

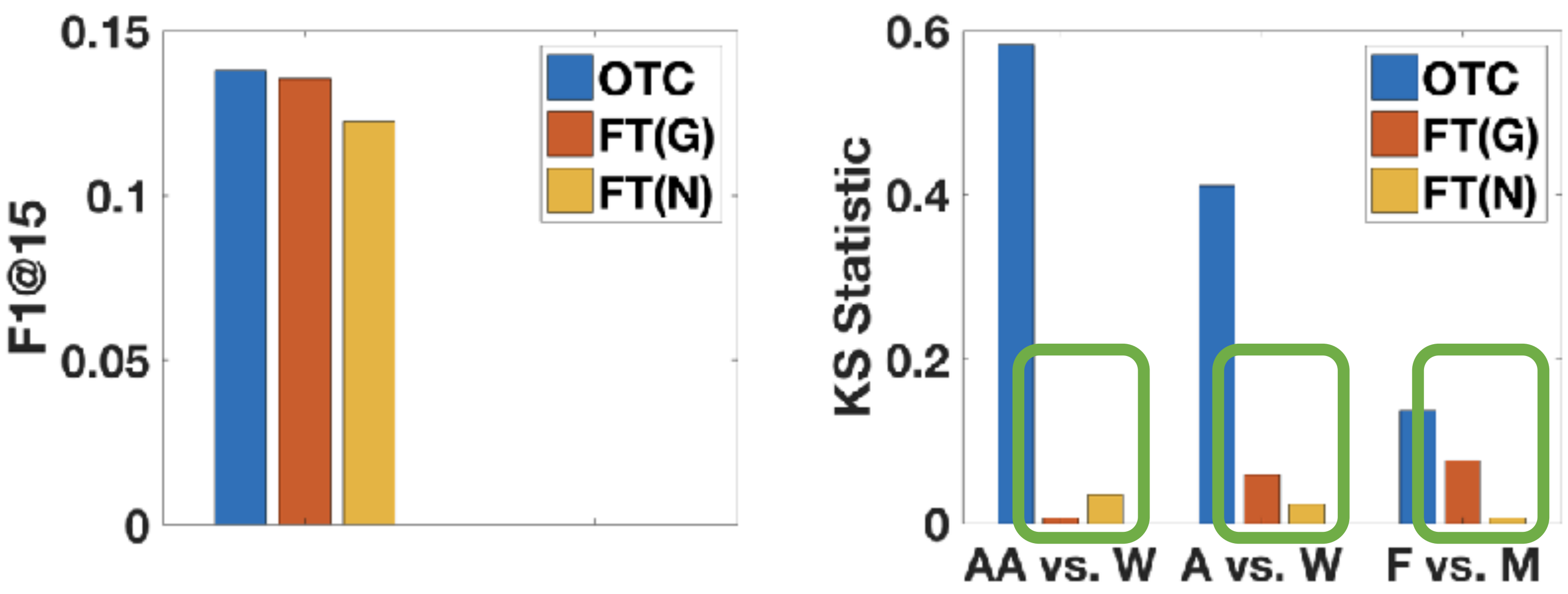


F: Female; **M**: Male
AA: American African; **W**: White; **A**: Asian

Experiment – Generalizing

- How does FATR perform for **multi-feature** and **multi-category** case?

FATR **enhances** fairness well for both ethnicity and gender.



F: Female; **M**: Male
AA: American African; **W**: White; **A**: Asian

Conclusion and Future Work

Conclusions:

- Propose a novel **tensor**-based framework – FATR – to enhance **fairness** while maintaining **recommendation quality**;
- FATR can handle **multi-feature** and **multi-category** scenarios;
- Extensive experiments show the **effectiveness** of FATR and **robustness** to data **bias** and **sparsity**.

Future Work:

- Extend to **alternative notions of fairness** beyond statistical parity;
- Extend to **rating prediction** tasks for recommenders with explicit rating data.

Thank You!

Ziwei Zhu, Xia Hu, and James Caverlee

Texas A&M University