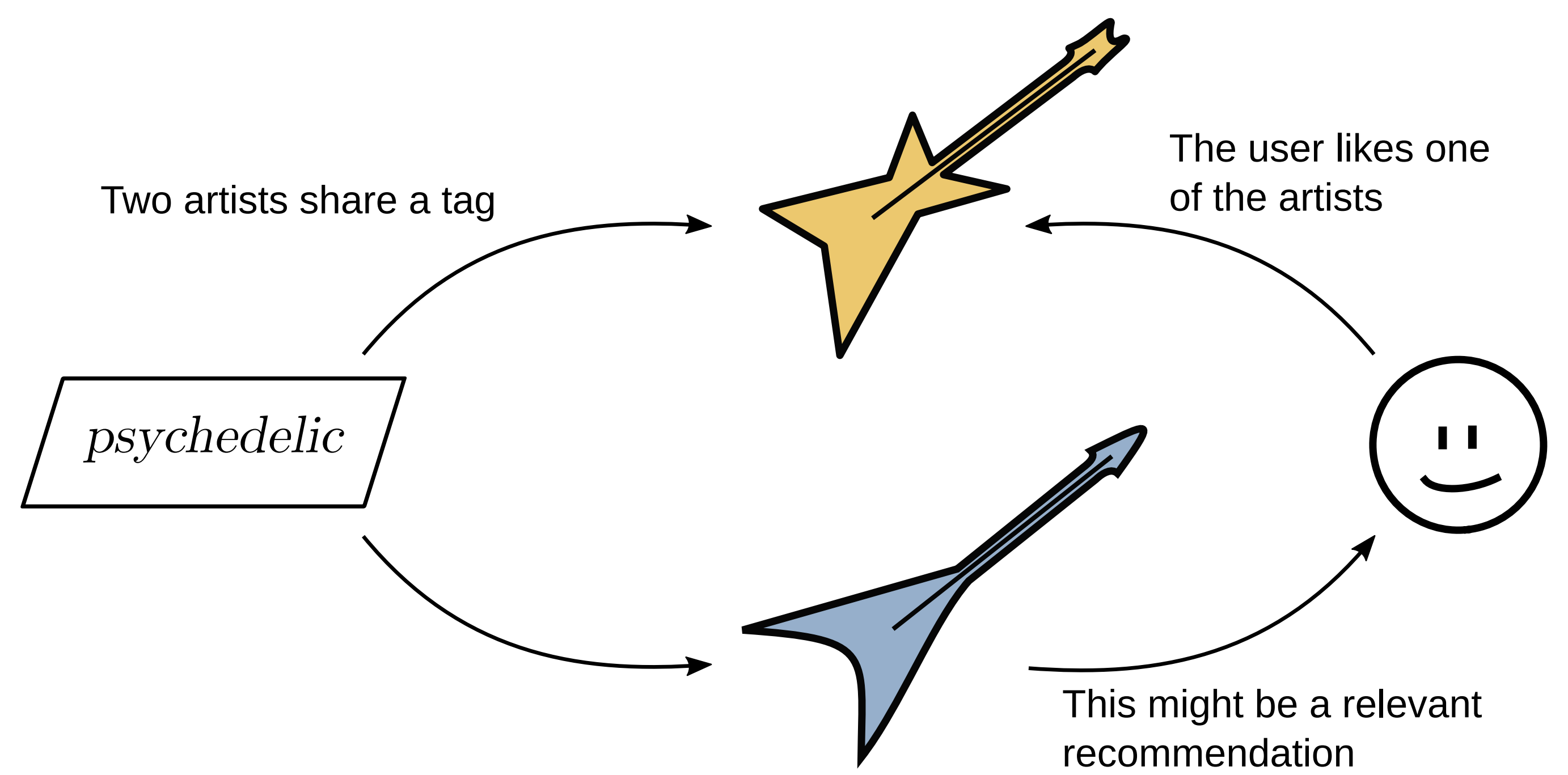


improving music recommendations with a weighted factorization of the tagging activity

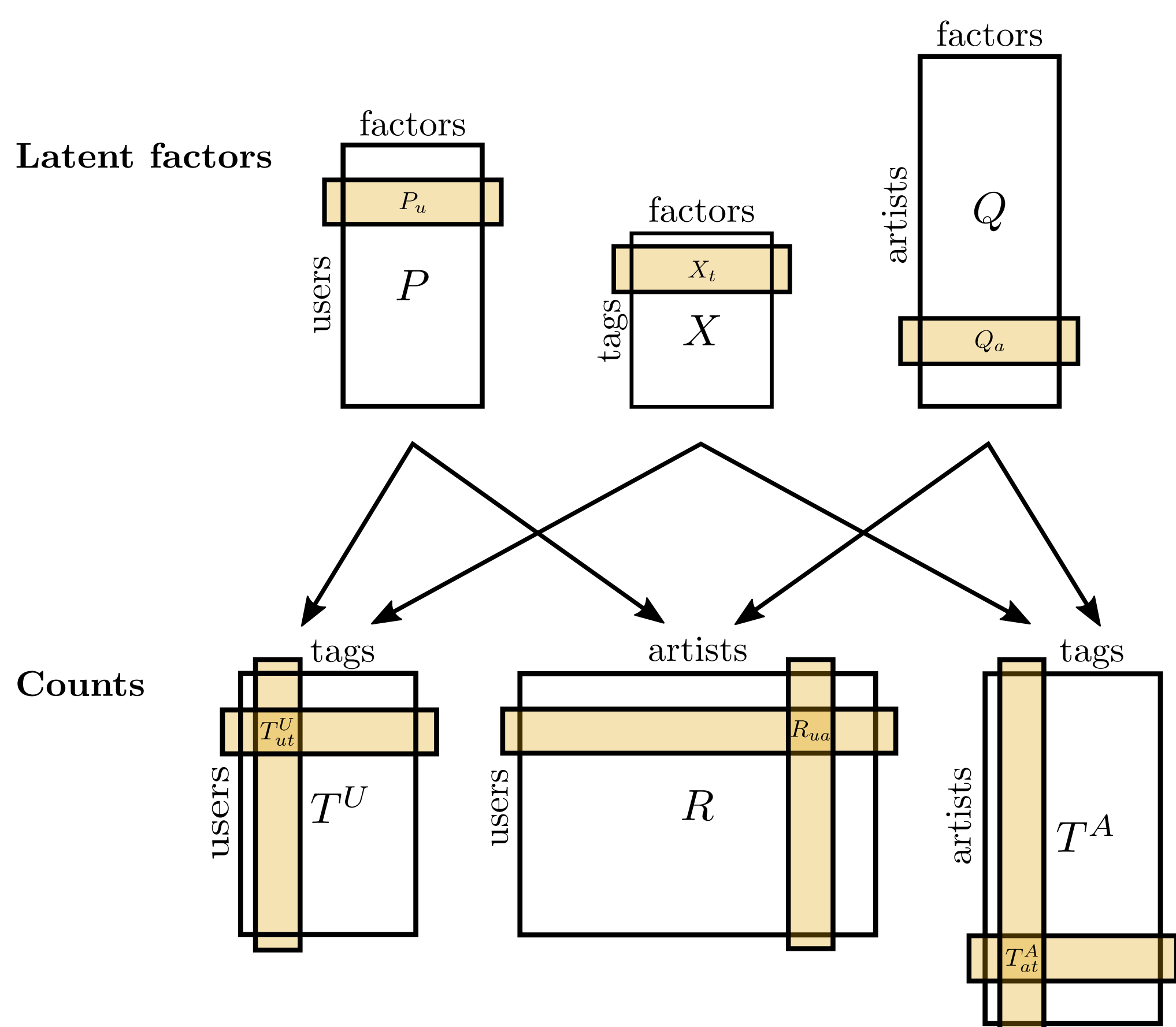
Andreu Vall, Marcin Skowron, Peter Knees, Markus Schedl
 Department of Computational Perception
 Johannes Kepler University, Linz, Austria
 andreu.vall@jku.at



Social tags provide information both about the users who apply them and about the artists that are labeled with them. Combining social tags with Collaborative Filtering (CF) into a hybrid model we are able to produce better artist recommendations. We identify relevant recommendations on the basis of tag connections, that could not be build on the basis of listening patterns only [1].

hybrid collaborative filtering with social tags

The proposed model combines user-artist playcounts with user-tag counts and artist-tag counts (i.e., how often a user has applied a tag and how often an artist has been labeled with a tag). Latent factors for users, artists and tags are estimated in a joint matrix factorization scheme to approximate binary versions of the data matrices.



Cost function

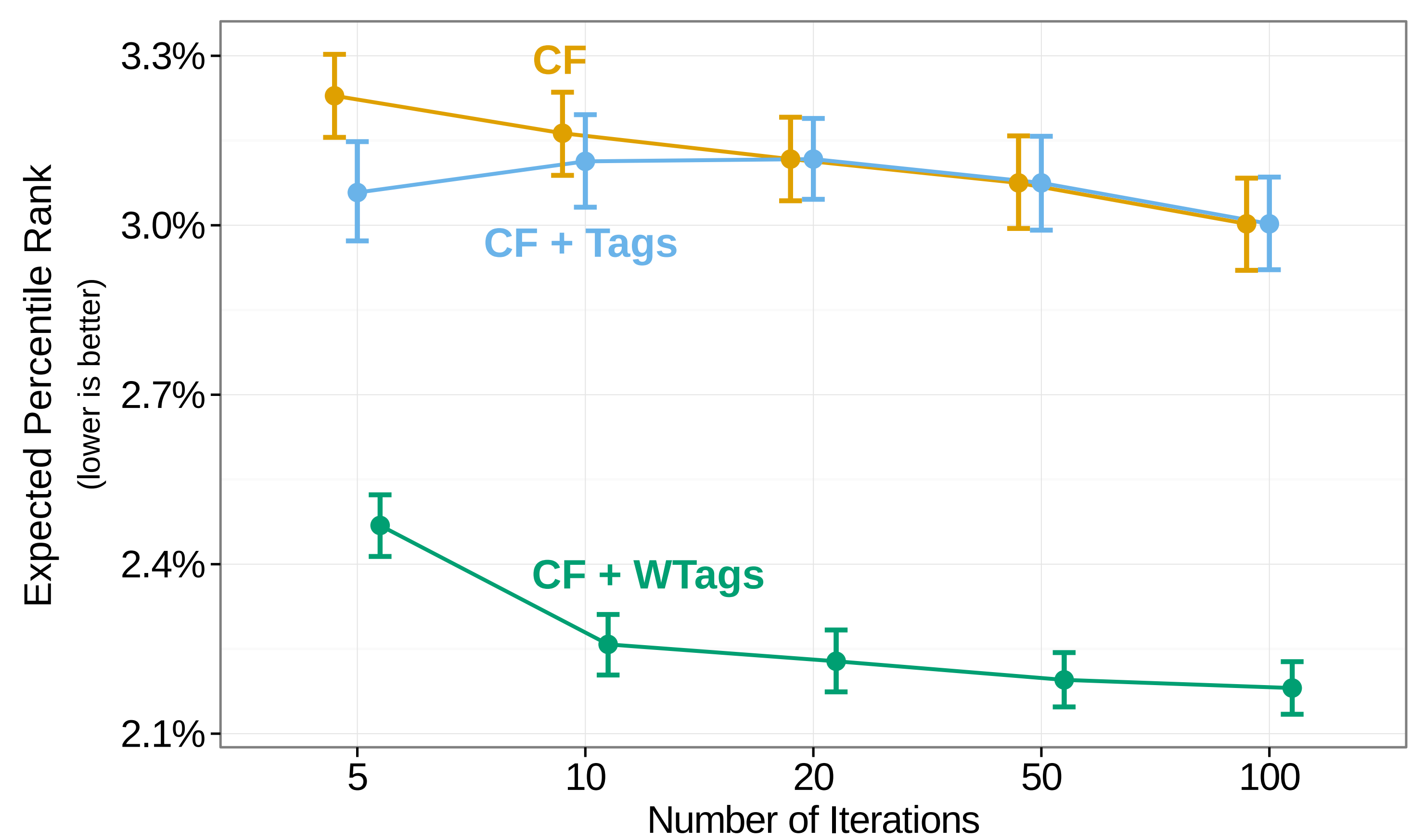
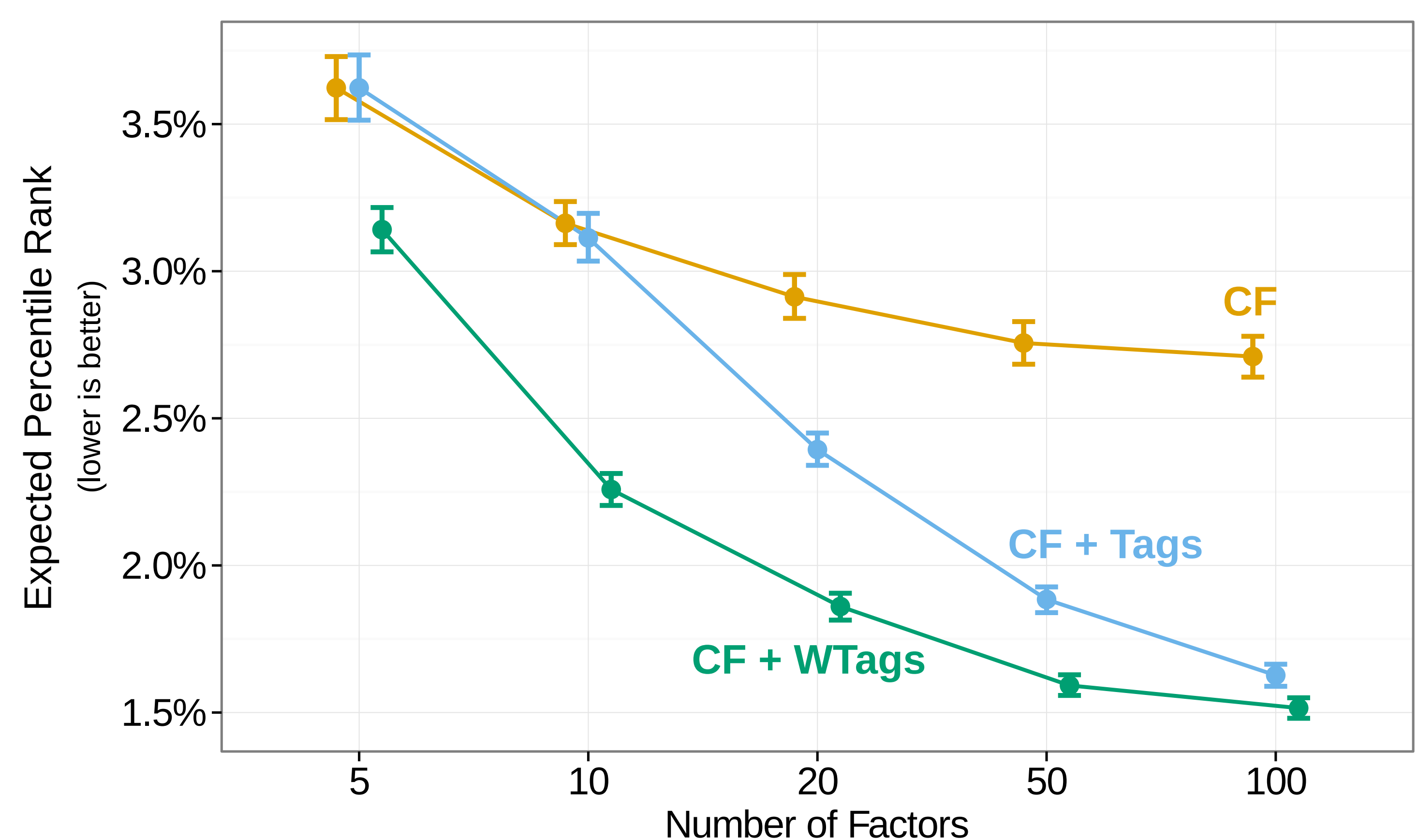
$$J(P, Q, X) = \sum_{ua \in R} w(\alpha, R_{ua}) (\tilde{R}_{ua} - P_u Q_a^T)^2 + \mu_1 \sum_{ut \in T^U} w(\beta, T_{ut}^U) (\tilde{T}_{ut}^U - P_u X_t^T)^2 + \mu_2 \sum_{at \in T^A} w(\gamma, T_{at}^A) (\tilde{T}_{at}^A - Q_a X_t^T)^2 + \lambda (\|P\|_F^2 + \|Q\|_F^2 + \|X\|_F^2)$$

Predicted preference

$$Z = PQ^T$$

numerical study

- 2,902 users
- 71,223 artists
- 687,833 user-artist pairs
- 630 unique user tags
- 12,902 unique artist tags
- Data collected through the Last.fm API

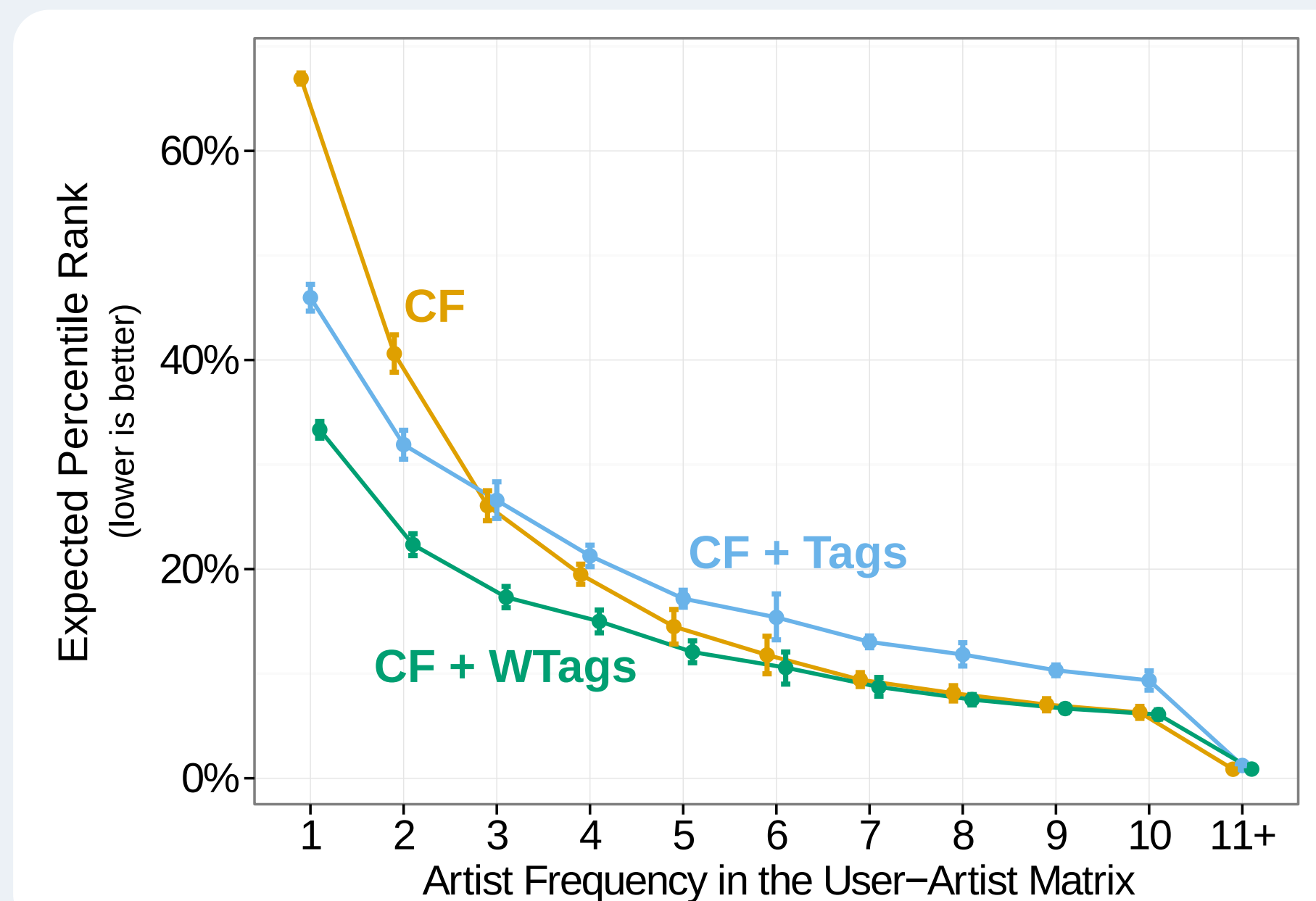


qualitative example

We trained *CF*, *CF+Tags* and *CF+WTags* using 80% of the data. The table shows the predicted expected preference Z_{ua} of a selected user u for four different artists she listened to, but belong to the 20% of data withheld for test. Only *CF+WTags* consistently identifies that the artists are relevant for the user.

Artist Name	Z_{ua}^{CF}	$Z_{ua}^{CF+Tags}$	$Z_{ua}^{CF+WTags}$
Feliu Ventura	0.00	0.09	0.61
Joan Colomo	0.35	0.58	0.75
Manos de Topo	0.23	0.44	0.64
Mazoni	0.00	0.22	0.69

We evaluated the performance of the models as a function of the artist frequency in the user-artist playcounts matrix. The plot shows that the hybrid model *CF+WTags* performs significantly better when there is little listening information [2].



[1] A. Vall, M. Skowron, P. Knees, and M. Schedl. Improving music recommendations with a weighted factorization of the tagging activity. In Proc. ISMIR, 2015.
 [2] A. Vall. Listener-inspired automated music playlist generation. In Proc. RecSys, 2015.