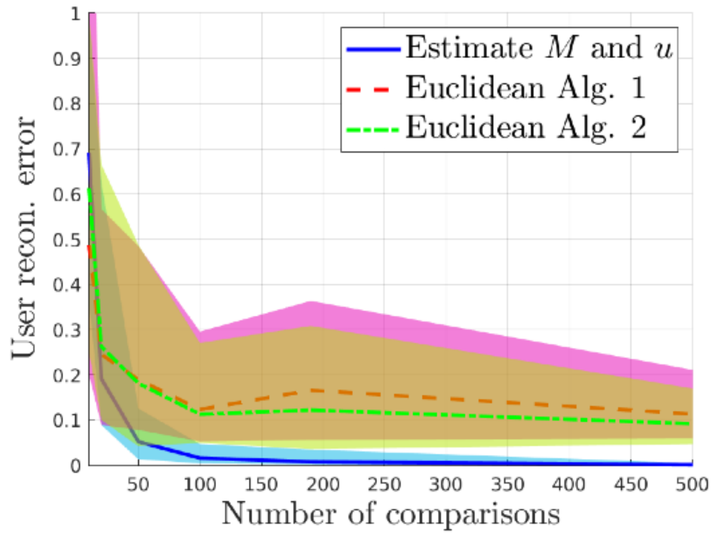
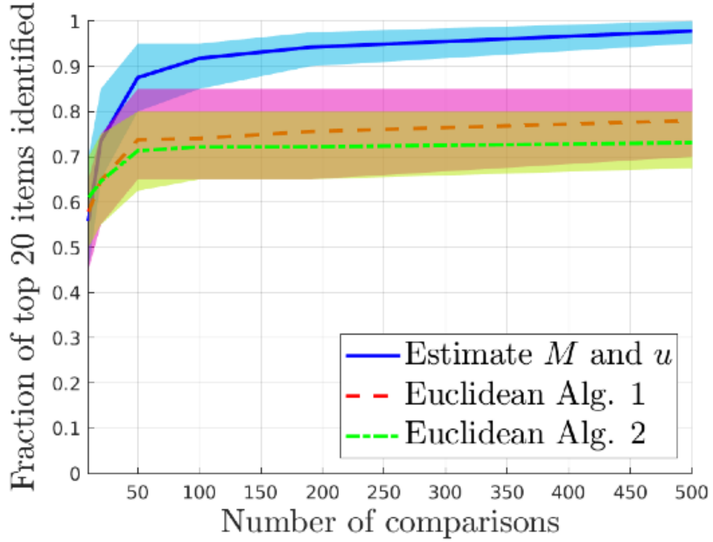
Synthetic experiments: baseline comparisons

- Comparison against *Davenport (2013)* and a variant of our estimator which assumes Euclidean distance
- Assume true metric is ***Mahalanobis***



"How well can we recover the ideal point?"



"How well can we recover the 20 most preferred items?"

Real-world example: graduate admissions

- 3,000+ candidates in 3 categories: Admitted with Fellowship, Admitted without Fellowship, and Denied
  - Form paired comparisons between categories
- 5 features: GPA, GRE verbal, GRE quantitative, GRE writing, letters of rec. score (LoR)

Feature interactions in $\widehat{M}$.	
$\lambda_1 = 1991$	0.909 GRE writing $- 0.392$ GPA
$\lambda_2 = 1971$	0.919 GPA $+ 0.393$ GRE writing
$\lambda_3 = 1178$	0.982 LoR
$\lambda_4 = 861$	0.942 GRE quant $- 0.310$ GRE verbal
$\lambda_5 = 286$	0.942 GRE verbal $+ 0.319$ GRE quant