

Delta distancing: A lifting approach to localizing items from user comparisons

Andrew D. McRae^g, Austin Xu^g, Jihui Jin^g, Namrata Nadagouda^g, Nauman Ahad^g, Peimeng Guan^g, Santhosh Karnik^m, Mark A. Davenport^g

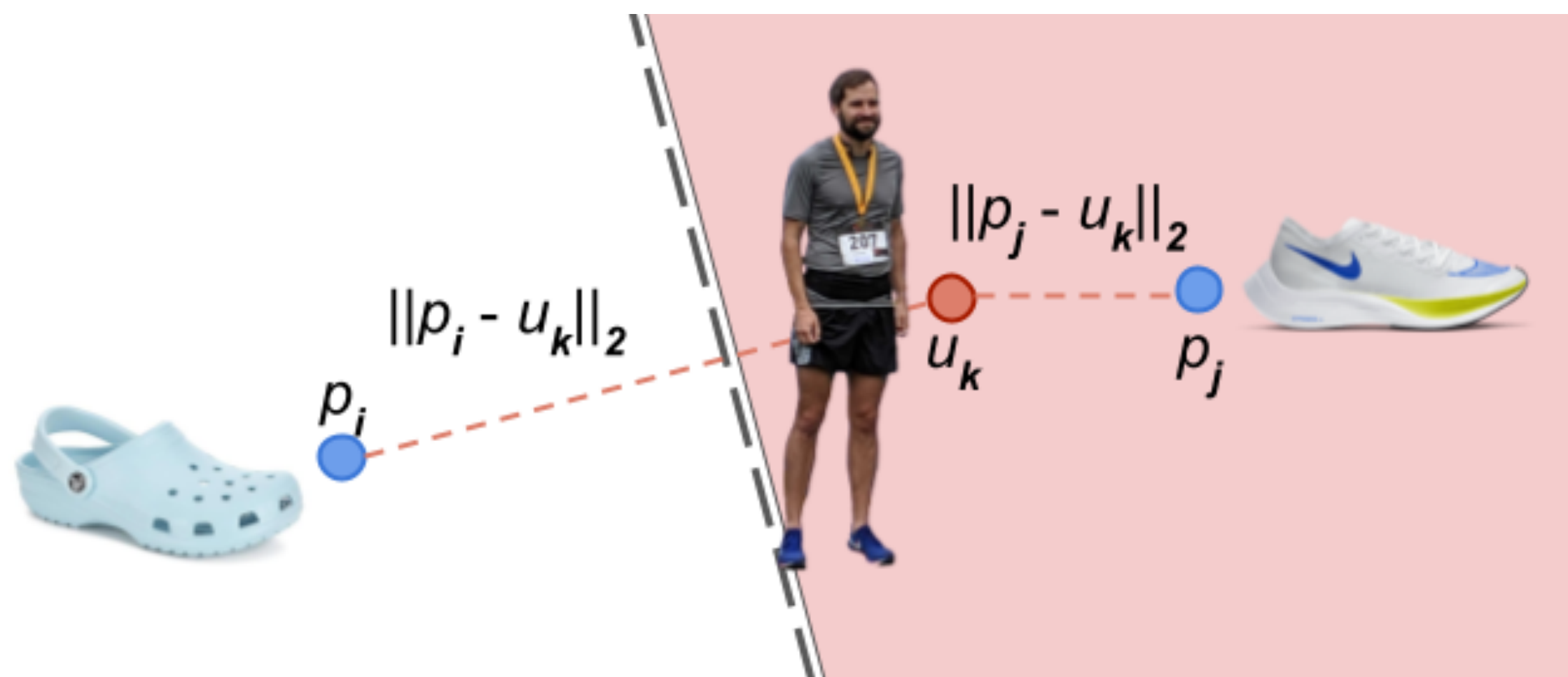
Georgia Institute of Technology^g, Michigan State University^m

Overview



Ideal point model:
The closer an item is to a user's *ideal point*, the more preferred

Item localization from paired comparisons:
Each paired comparison response indicates which item is closer to the user.



Task: Estimate a set of points $p_1, \dots, p_n \in \mathbb{R}^d$ from paired comparisons from a set of user ideal points $u_1, \dots, u_\ell$ of the form

Bradley-Terry Model: We assume paired comparison outcomes $y_k \in \{-1, +1\}$ are Bernoulli distributed with mean

$$\mu_k = \frac{1}{1 + \exp(-\delta_{i(k)j(k)k})}$$

where $\delta_{i(k)j(k)k} = \|p_j - u_k\|_2^2 - \|p_i - u_k\|_2^2$

Delta Distancing

“Lifting” the problem:

In order to estimate $p_1, \dots, p_n$, we lift the problem to the space of PSD matrices by

$$\begin{aligned} \|p_i - u_k\|_2^2 &= \|p_i\|_2^2 - 2\langle p_i, u_k \rangle + \|u_k\|_2^2 \\ &= \left\langle \begin{bmatrix} p_i p_i^T & p_i \\ p_i^T & 1 \end{bmatrix}, \begin{bmatrix} I_d & -u_k \\ -u_k^T & \|u_k\|_2^2 \end{bmatrix} \right\rangle \\ &=: \langle P_i, U_k \rangle \end{aligned}$$

Now we wish to estimate n symmetric positive semidefinite matrices in $\mathbb{R}^{d+1 \times d+1}$. Our problem is now linear in matrices P_i

A maximum-likelihood motivated optimization program:

To solve for $P_1, \dots, P_n$, we solve

$$\begin{aligned} \min_{P_1, \dots, P_n} \sum_{k=1}^m \log \left(1 + \exp \left(y_k \langle P_{i(k)} - P_{j(k)}, U_k \rangle \right) \right) \\ + \lambda \sum_{i=1}^n \text{tr}(P_i) \\ \text{subject to } P_i \succeq 0 \\ (P_i)_{d+1, d+1} = 1, i = 1, \dots, n \end{aligned}$$

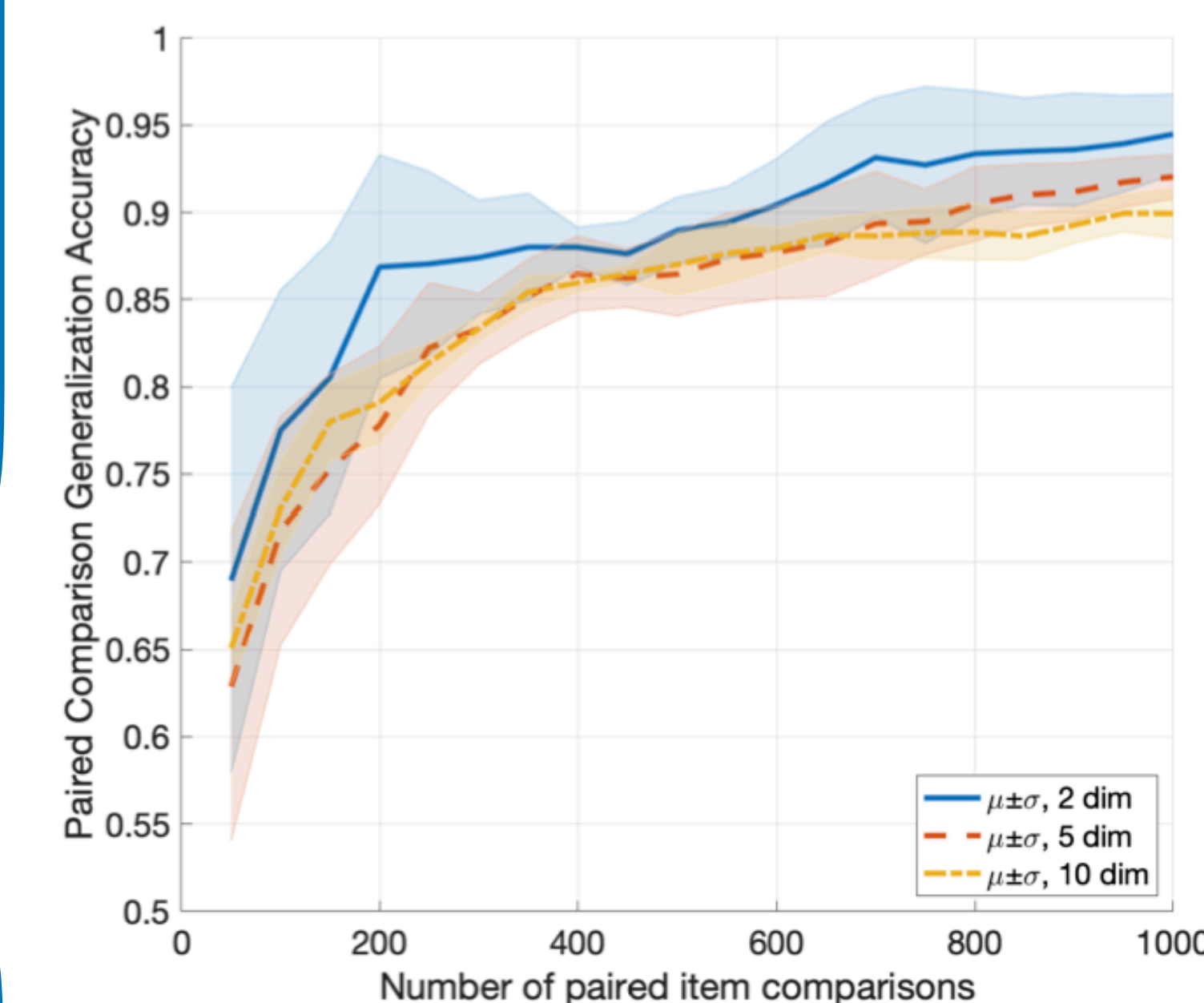
Trace regularization encourages low rank solutions and $\lambda > 0$ is a user specified parameter. To recover a vector estimate $\hat{p}_i$, we take the first d entries of the last column of $\hat{P}_i$

Experiments

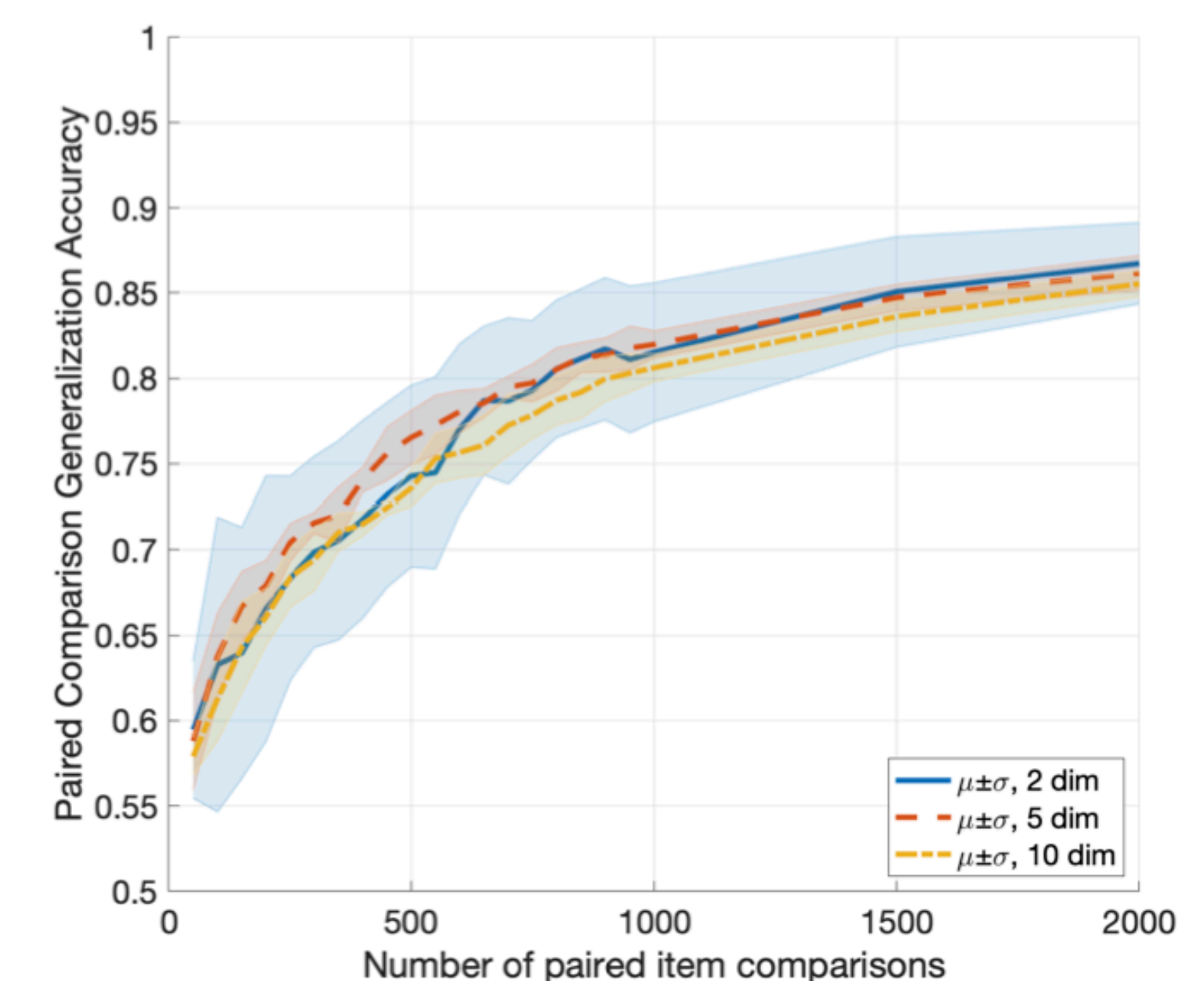
We characterize performance using

- Paired comparison generalization accuracy (PCGA):** Fraction of paired comparisons induced by learned items that are consistent with ground truth paired comparisons
- Averaged norm-square error (ANSE):**

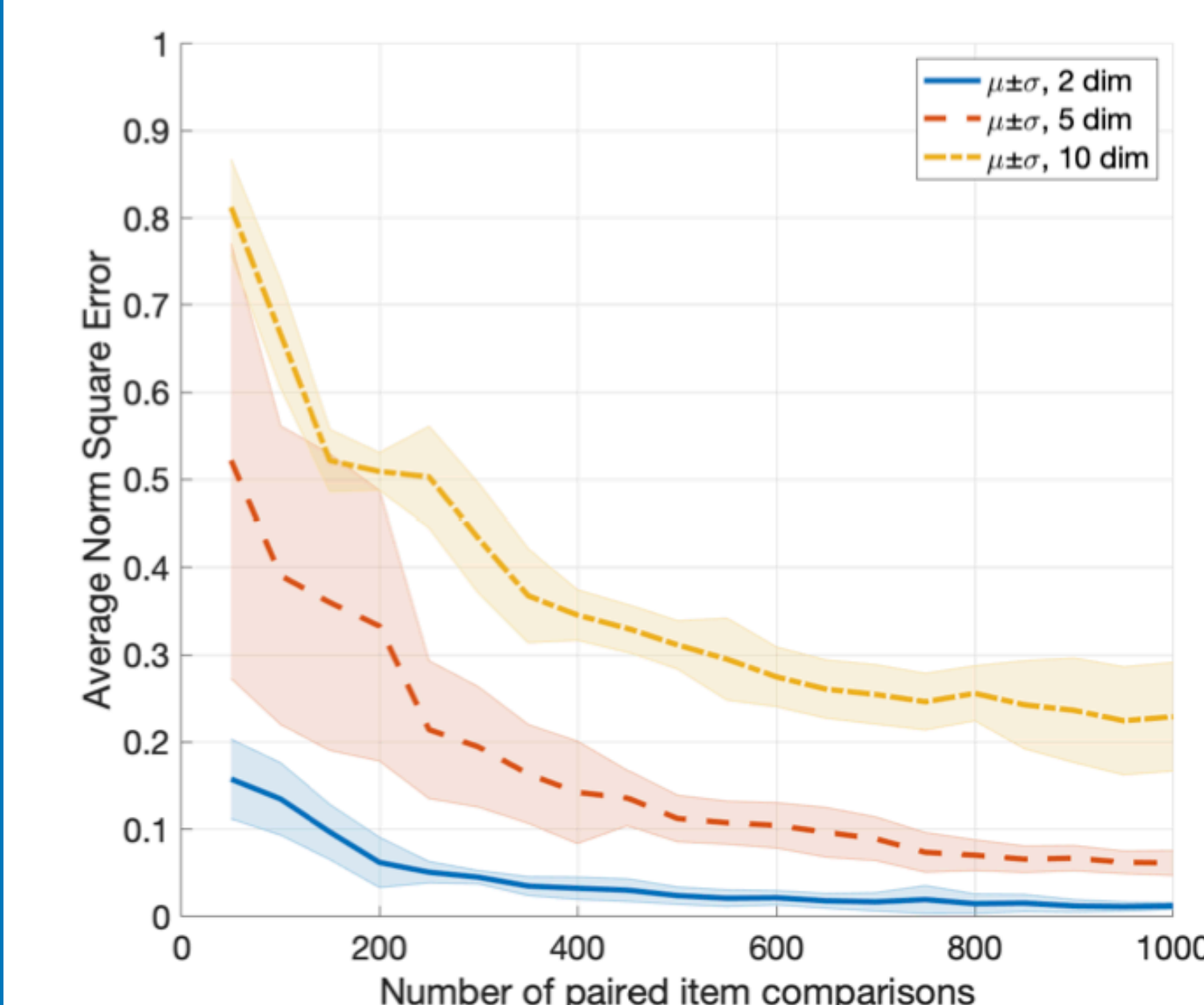
$$\frac{1}{n} \sum_{i=1}^n \|\hat{p}_i - p_i\|_2^2$$



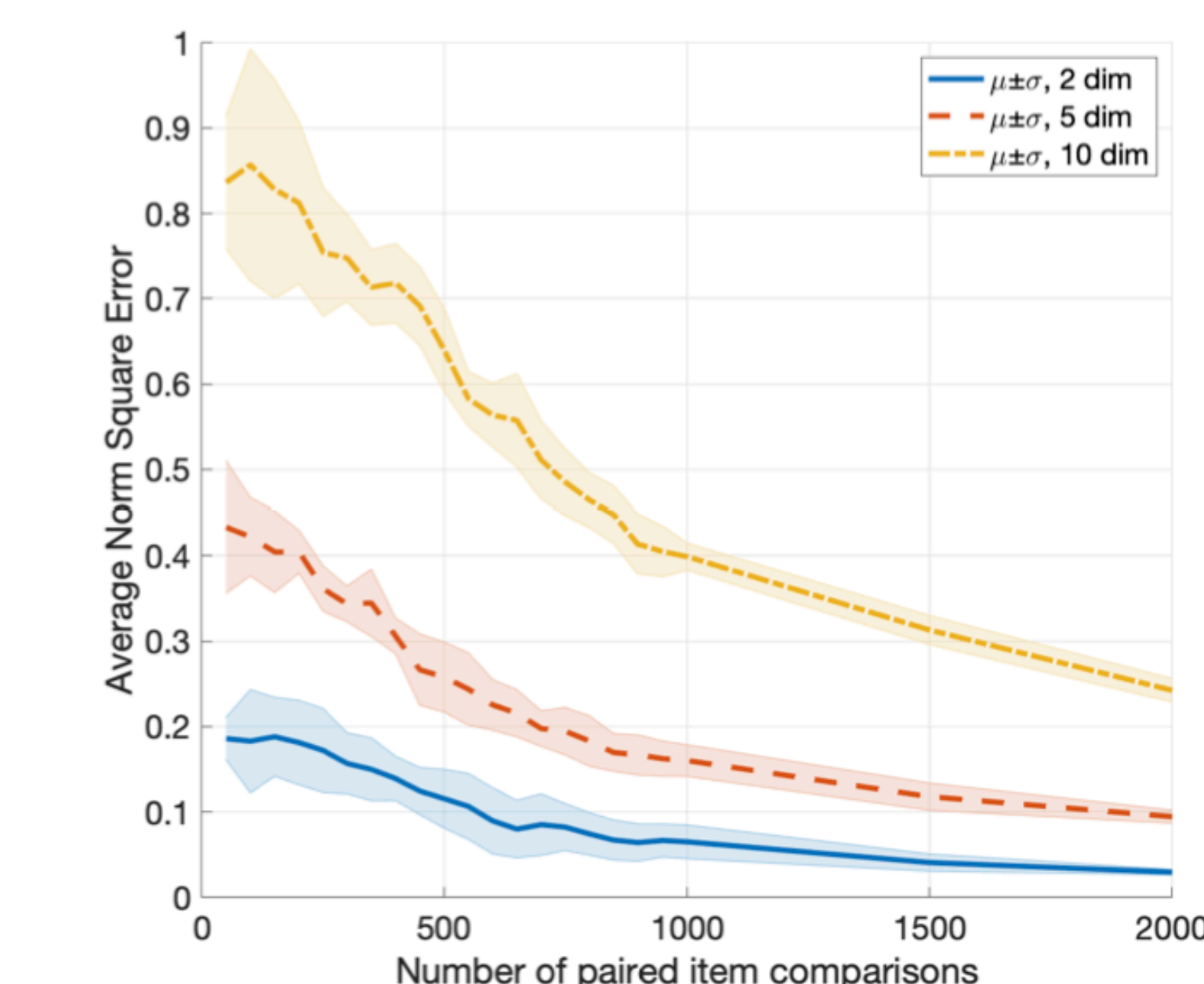
PCGA (5 items)



PCGA (20 items)



ANSE (5 items)



ANSE (20 items)

$\ell = 100$ users generated uniformly on $[-1, 1]^d$ and $n = 5, 20$ items generated uniformly on $[-0.5, 0.5]^d$