

# On counterfactual inference with factors models and nearest neighbors

Raaz Dwivedi, CDT Summer School

# To think about

- Limitations of the model and how to possibly relax it
- Limitations of the analysis and how to possible sharpen it

# Recap: Latent factor model

$$R_{i,t} = R_{i,t}(A_{i,t}) \triangleq \theta_{i,t}^{(A_{i,t})} + \varepsilon_{i,t}^{(A_{i,t})}$$

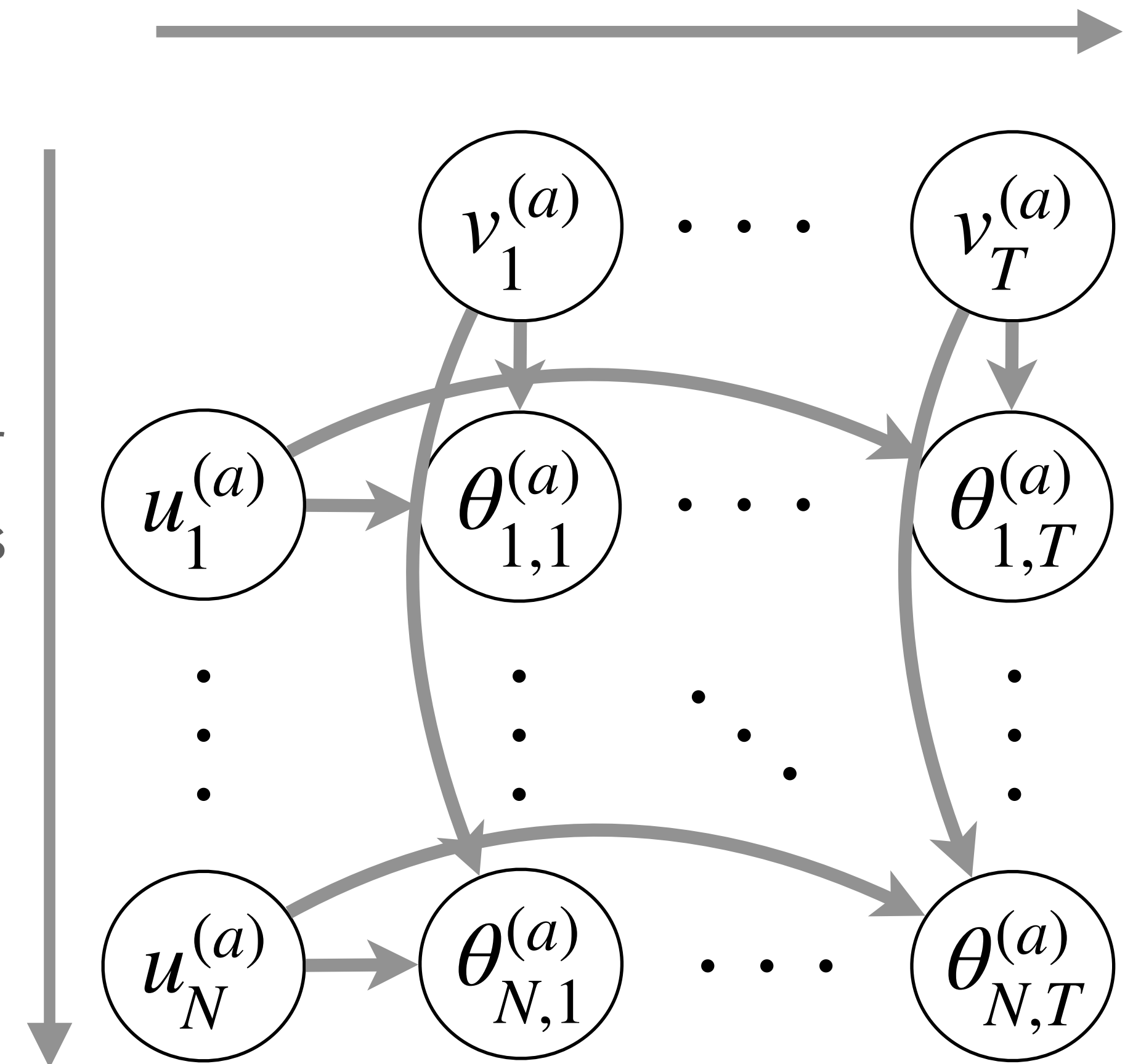
$$\theta_{i,t}^{(a)} \triangleq \langle u_i^{(a)}, v_t^{(a)} \rangle$$

user factor  
(e.g., personal  
traits)

time factor  
(e.g., societal, weather  
changes)

$N$  user  
factors

$T$  time factors

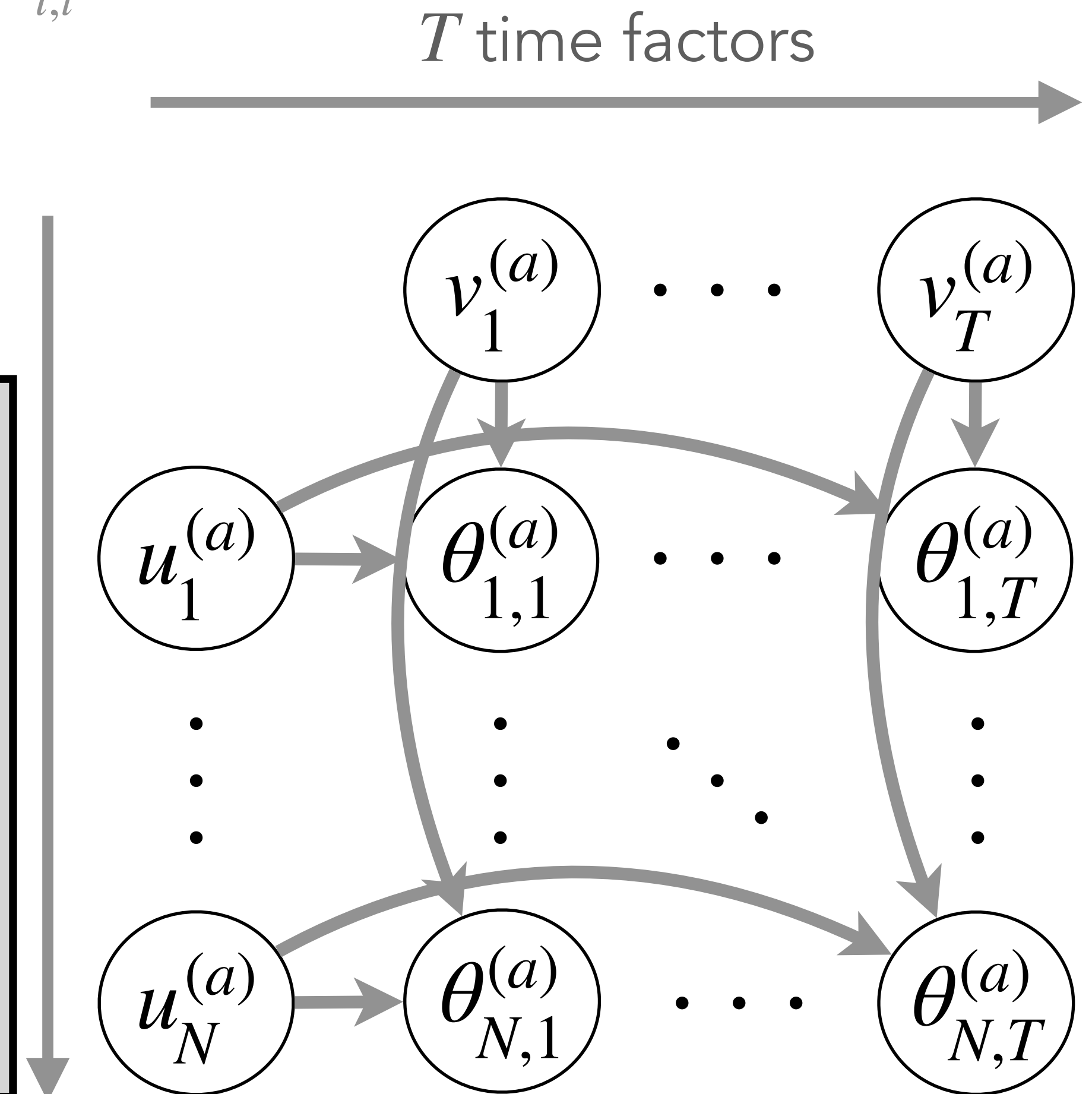


# Recap: Latent factor model

$$R_{i,t} = R_{i,t}(A_{i,t}) \triangleq \theta_{i,t}^{(A_{i,t})} + \varepsilon_{i,t}^{(A_{i,t})}$$

$$\theta_{i,t}^{(a)} \triangleq \langle u_i^{(a)}, v_t^{(a)} \rangle$$

- A generalization of mixed effects model
- No parametric assumptions on the **unknown** distributions of latent factors or noise
- Paper also considers  $\theta_{i,t}^{(a)} \triangleq f(u_i^{(a)}, v_t^{(a)})$  for **unknown** Lipschitz  $f$



# User nearest neighbors estimator for treatment $a$

$$R_{i,t} = R_{i,t}(A_{i,t}) \triangleq \theta_{i,t}^{(A_{i,t})} + \varepsilon_{i,t}^{(A_{i,t})}$$

1. Compute distance between user pairs

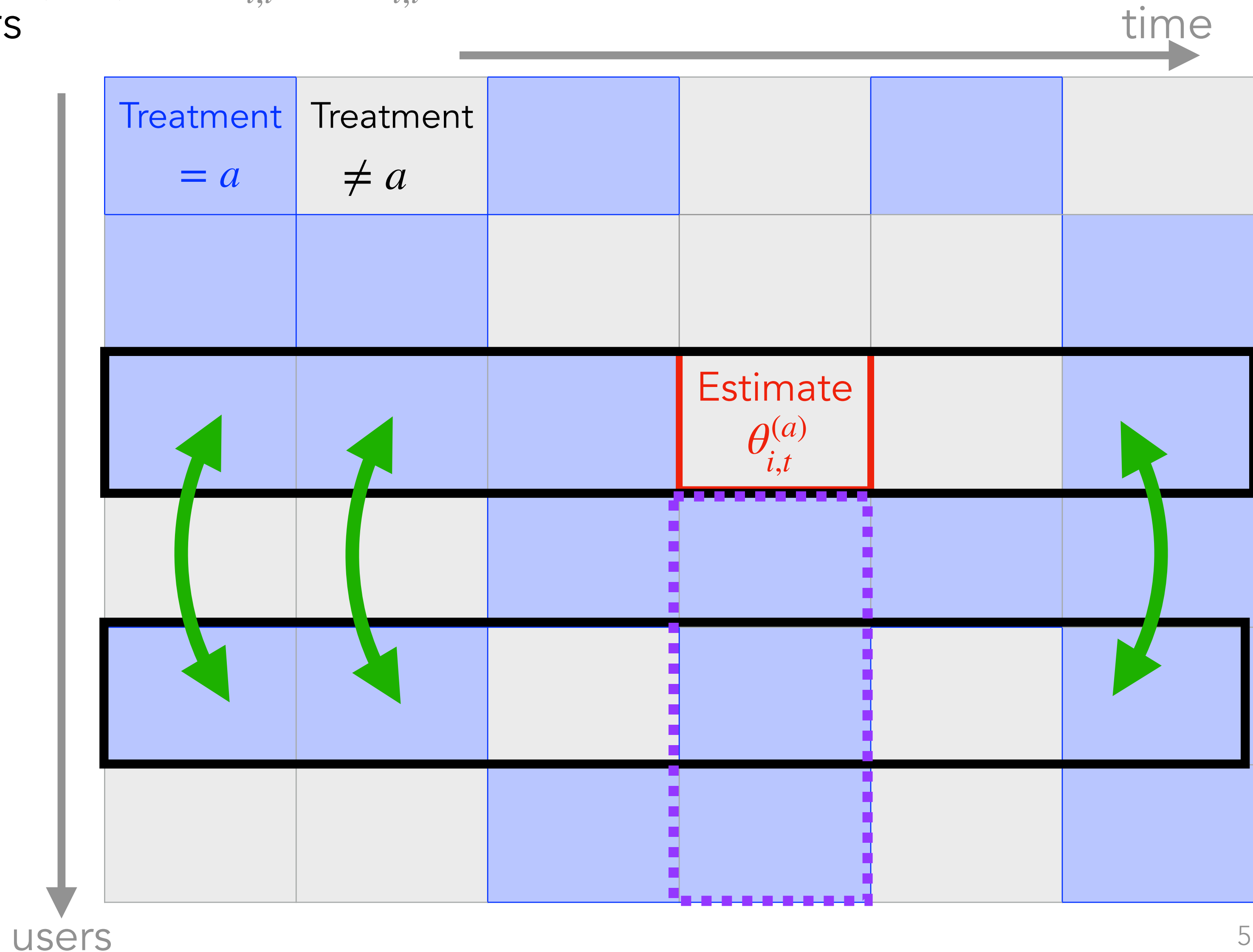
$i, j$  under treatment  $a$  **using all data**

$$\rho_{i,j}^{(a)} = \frac{\sum_{t'=1}^T (R_{i,t'} - R_{j,t'})^2 \cdot \mathbf{1}(A_{i,t'} = A_{j,t'} = a)}{\sum_{t'=1}^T \mathbf{1}(A_{i,t'} = A_{j,t'} = a)}$$

$$+ \frac{\hat{\sigma}_\rho}{\sqrt{\sum_{t'=1}^T \mathbf{1}(A_{i,t'} = A_{j,t'} = a)}}$$

2. Average over **user neighbors** treated with  $a$  at time  $t$

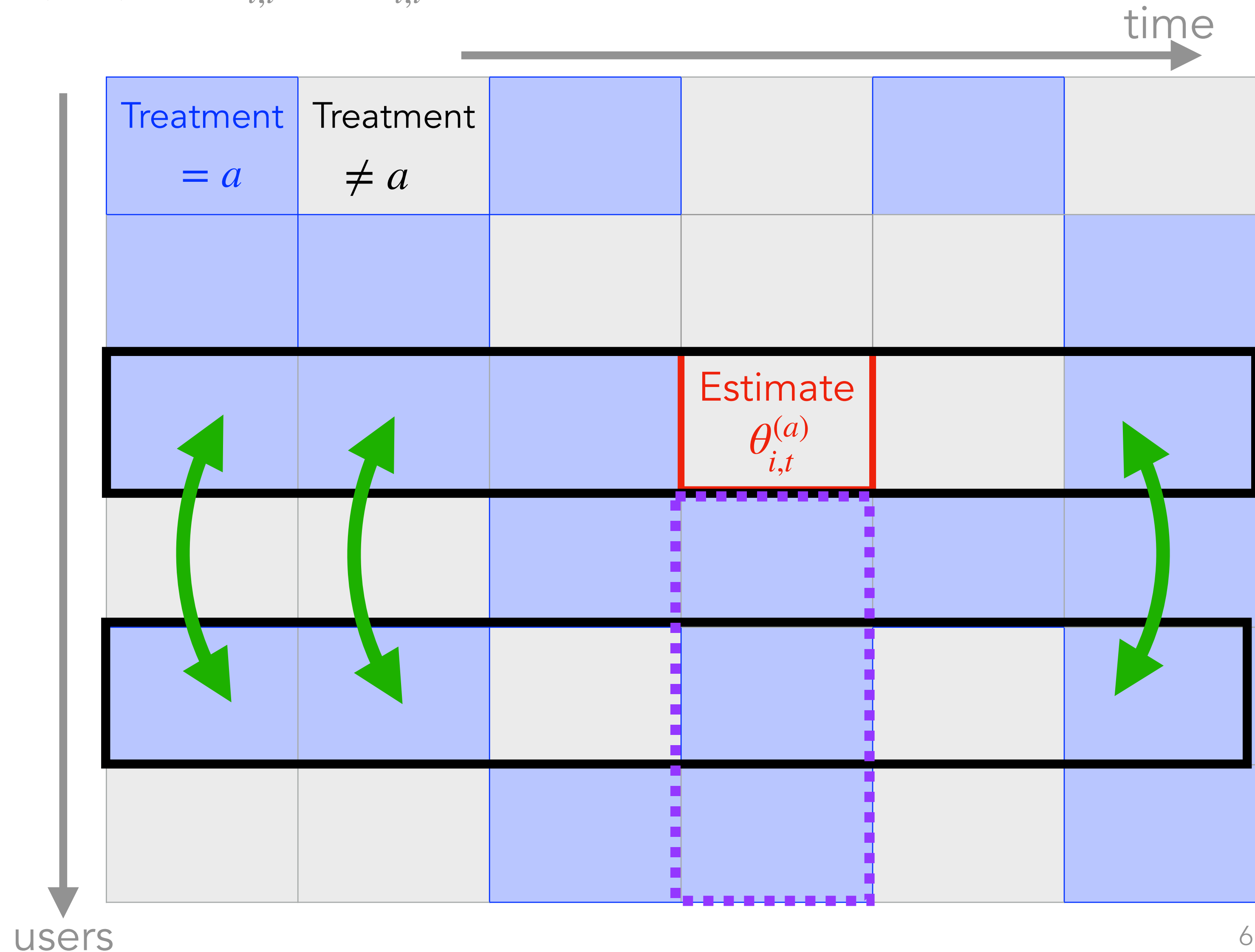
$$\hat{\theta}_{i,t,\text{user-NN}}^{(a)} = \frac{\sum_{j=1}^N R_{j,t} \mathbf{1}(\rho_{i,j}^{(a)} \leq \eta, A_{j,t} = a)}{\sum_{j=1}^N \mathbf{1}(\rho_{i,j}^{(a)} \leq \eta, A_{j,t} = a)}$$



# User nearest neighbors estimator for treatment $a$

$$R_{i,t} = R_{i,t}(A_{i,t}) \triangleq \theta_{i,t}^{(A_{i,t})} + \varepsilon_{i,t}^{(A_{i,t})}$$

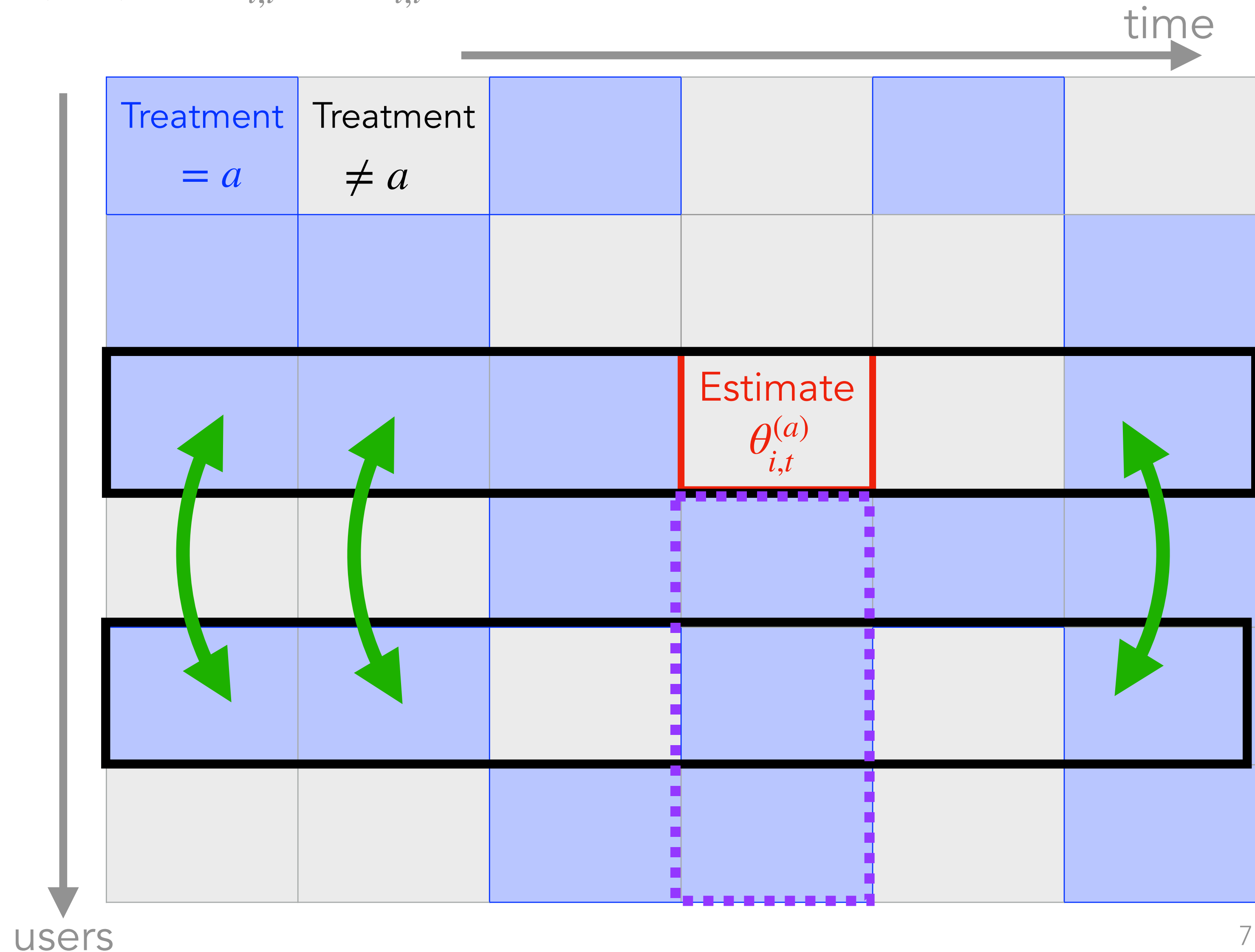
Note that the estimator is agnostic to sampling policy or the generative model for the latent factors



# User nearest neighbors estimator for treatment $a$

$$R_{i,t} = R_{i,t}(A_{i,t}) \triangleq \theta_{i,t}^{(A_{i,t})} + \varepsilon_{i,t}^{(A_{i,t})}$$

Something should bother you about this estimator when treatments are sampled sequentially !!!



# User-NN: A non-asymptotic guarantee for each $(i, t, a)$

**Informal theorem:** [Dwivedi-Tian-Tomkins-Klasnja-Murphy-Shah '22a]

For suitably chosen  $\eta$  & **under regularity conditions on latent factors**

- iid latent factors, sub-Gaussian noise
- sequentially adaptive policies with conditionally independent treatments across users that choose  $a$  with probability  $\geq p^\dagger$



# User-NN: A non-asymptotic guarantee for each $(i, t, a)$

**Informal theorem:** [Dwivedi-Tian-Tomkins-Klasnja-Murphy-Shah '22a]

For suitably chosen  $\eta$  & **under regularity conditions on latent factors**

- iid latent factors, sub-Gaussian noise
- sequentially adaptive policies with conditionally independent treatments across users that choose  $a$  with probability  $\geq p^\dagger$

for each user  $i$  at each time  $t$ , with high probability

$$(\hat{\theta}_{i,t,\text{user-NN}}^{(a)} - \theta_{i,t}^{(a)})^2 \lesssim \frac{1}{\sqrt{p^2 T}} + \frac{1}{p \cdot \#Neighbors}$$

( $\dagger$  Thus  $p$  can not decay faster than  $\gtrsim T^{-1/2}$ )

# User-NN: A non-asymptotic guarantee for each $(i, t, a)$

**Informal theorem:** [Dwivedi-Tian-Tomkins-Klasnja-Murphy-Shah '22a]

For suitably chosen  $\eta$  & under regularity conditions

- iid latent factors, sub-Gaussian noise
- sequentially adaptive policies with conditionally independent treatments across users that choose  $a$  with probability  $\geq p^\dagger$

for each user  $i$  at each time  $t$ , with high probability

$$(\hat{\theta}_{i,t,\text{user-NN}}^{(a)} - \theta_{i,t}^{(a)})^2 \lesssim \frac{1}{\sqrt{p^2 T}} + \frac{1}{pN/M}$$

User factor distribution



(user factors  $\sim$  uniform over a finite set of size  $M$ )

# User-NN: A non-asymptotic guarantee for each $(i, t, a)$

**Informal theorem:** [Dwivedi-Tian-Tomkins-Klasnja-Murphy-Shah '22a]

For suitably chosen  $\eta$  & under regularity conditions

- iid latent factors, sub-Gaussian noise
- sequentially adaptive policies with conditionally independent treatments across users that choose  $a$  with probability  $\geq p^\dagger$

for each user  $i$  at each time  $t$ , with high probability

$$(\hat{\theta}_{i,t,\text{user-NN}}^{(a)} - \theta_{i,t}^{(a)})^2 \lesssim \frac{1}{\sqrt{p^2 T}} + \frac{1}{pN/M}$$

$$(\hat{\theta}_{i,t,\text{user-NN}}^{(a)} - \theta_{i,t}^{(a)})^2 \lesssim \frac{1}{\sqrt{p^2 T}} + \frac{1}{(pN)^{2/(d+2)}}$$

User factor distribution



(user factors  $\sim$  uniform over a finite set of size  $M$ )

(user factors  $\sim$  Uniform in  $[-1,1]^d$ )

# User-NN guarantees

- Asymptotic **confidence intervals** as  $N, T \rightarrow \infty$ :

$$\hat{\theta}_{i,t,\text{user-NN}}^{(a)} \pm \frac{1.96 \hat{\sigma}}{\sqrt{\# \text{neighbors}_{i,t,a}}}$$



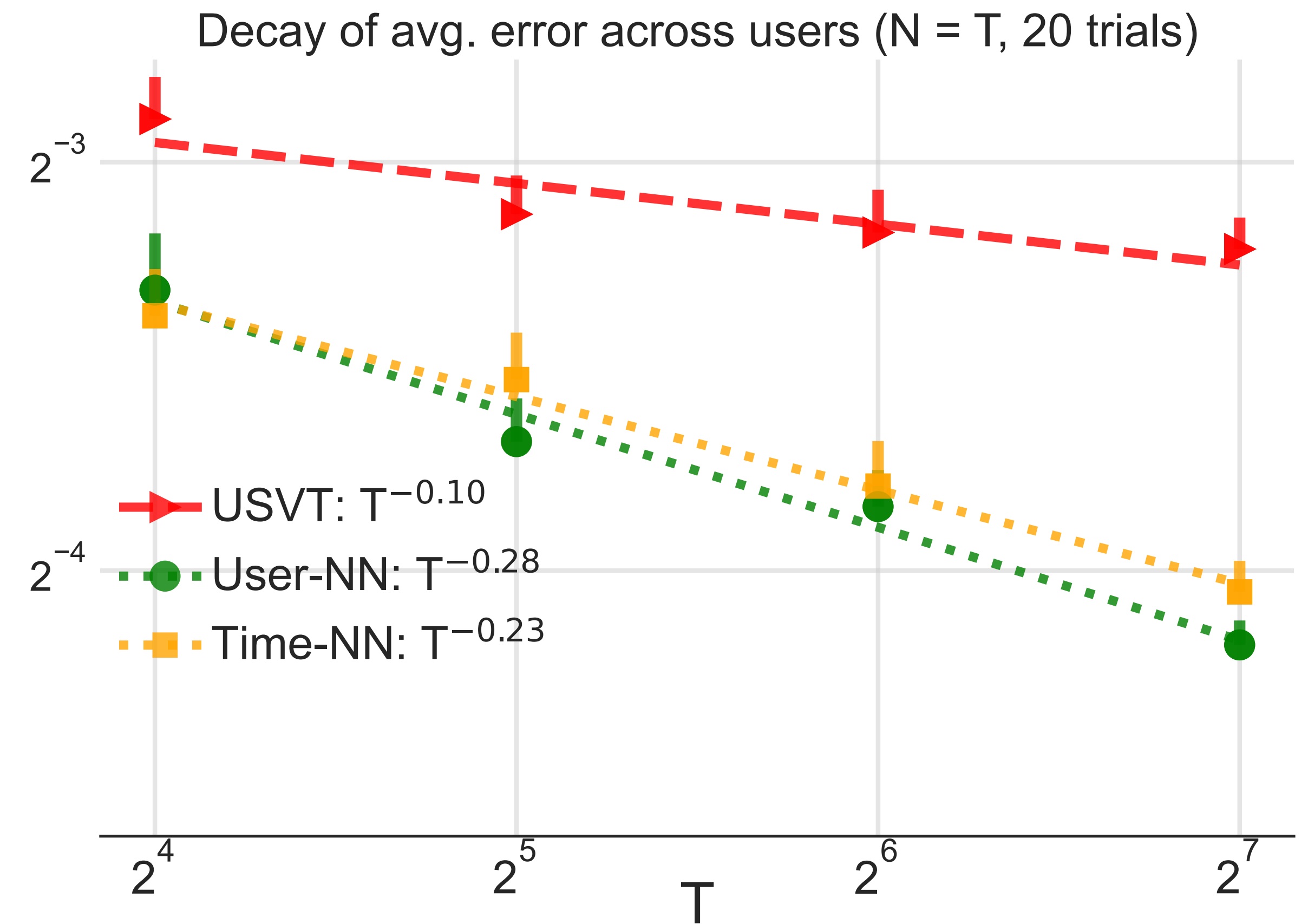
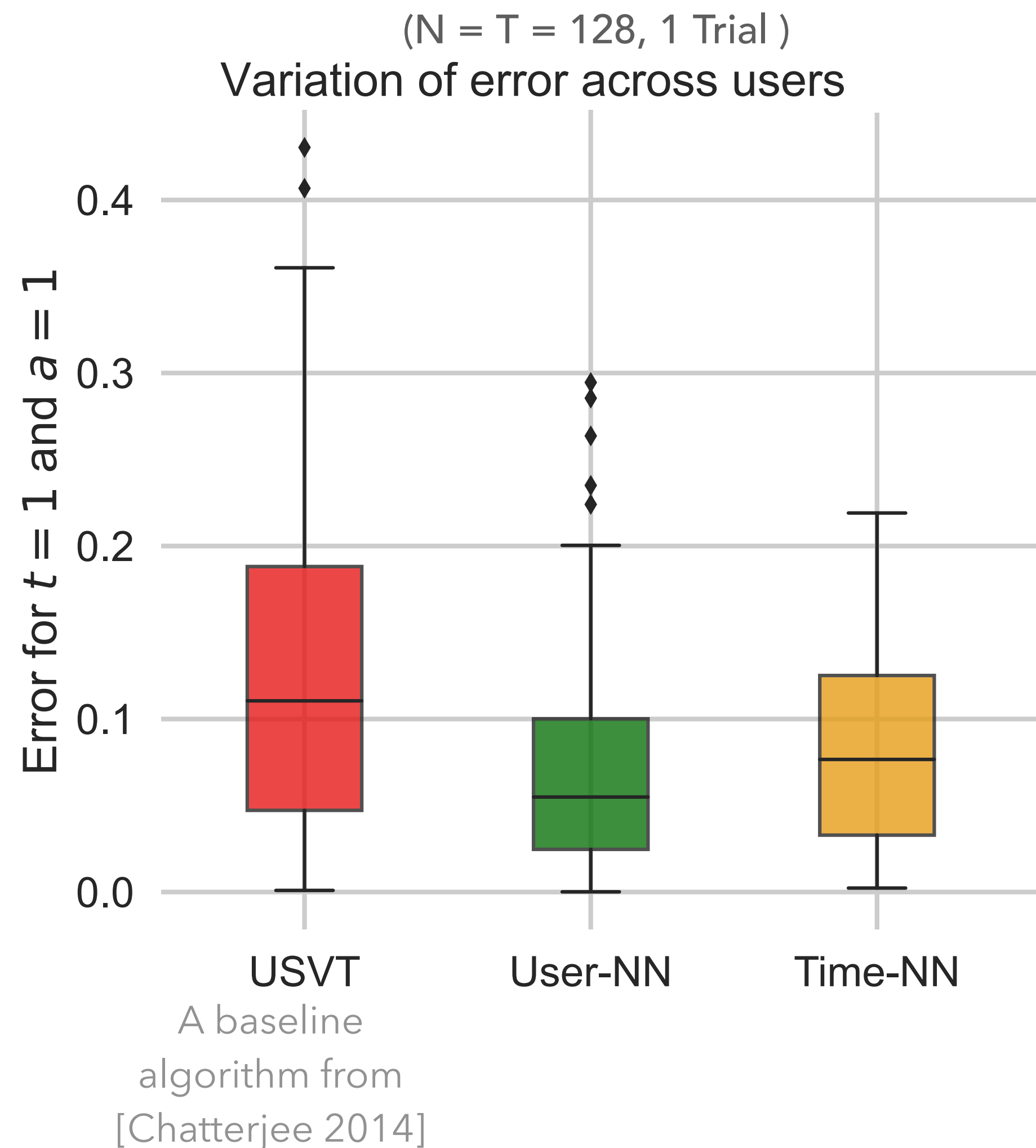
Confidence intervals for treatment effect  $\theta_{i,t}^{(1)} - \theta_{i,t}^{(0)}$

$$| \hat{\theta}_{i,t,\text{user-NN}}^{(a)} - \theta_{i,t}^{(a)} | = \tilde{O} \left( \frac{1}{T^{1/4}} + \frac{1}{\sqrt{N}} \right)$$

$$| \hat{\theta}_{i,t,\text{time-NN}}^{(a)} - \theta_{i,t}^{(a)} | = \tilde{O} \left( \frac{1}{N^{1/4}} + \frac{1}{\sqrt{T}} \right)$$

# Simulation results

Uniform latent factors on  $[-0.5, 0.5]^4$ , Gaussian noise, pooled  $\varepsilon$ -greedy policy ( $\varepsilon = 0.5$ )



USVT: A baseline algorithm from [Chatterjee 2014]

# We prove a general error bound for user NN (with actions sampled by learning policies)

$$(\hat{\theta}_{i,t,\eta}^{(a)} - \theta_{i,t}^{(a)})^2 \lesssim \frac{1}{\lambda_{\star}^2} \left( \underbrace{\eta - 2\sigma^2}_{\eta'} + \underbrace{\frac{C}{\sqrt{p^2 T}}}_{e_T} \right) + \frac{\sigma^2}{p N_{i,\eta'-e_T}} + c_{noise} \left[ \frac{N_{i,\eta'+e_T} - N_{i,\eta'-e_T}}{p N_{i,\eta'-e_T}} \right]^2$$

$\eta'$   
NN bias  
due to threshold

$e_T$   
Error in  
NN distance

NN noise  
variance

NN bias  
inflation due to  
**learning** policy

$$\lambda_{\star} \triangleq \lambda_{\min}(\Sigma_v) \quad \text{where } \Sigma_v = \mathbb{E}[v_t v_t^\top]$$

$$N_{i,\gamma} \triangleq |\{j \neq i : (u_i - u_j)^\top \Sigma_v (u_i - u_j) \leq \gamma\}|$$

# Steps towards deriving the general bound



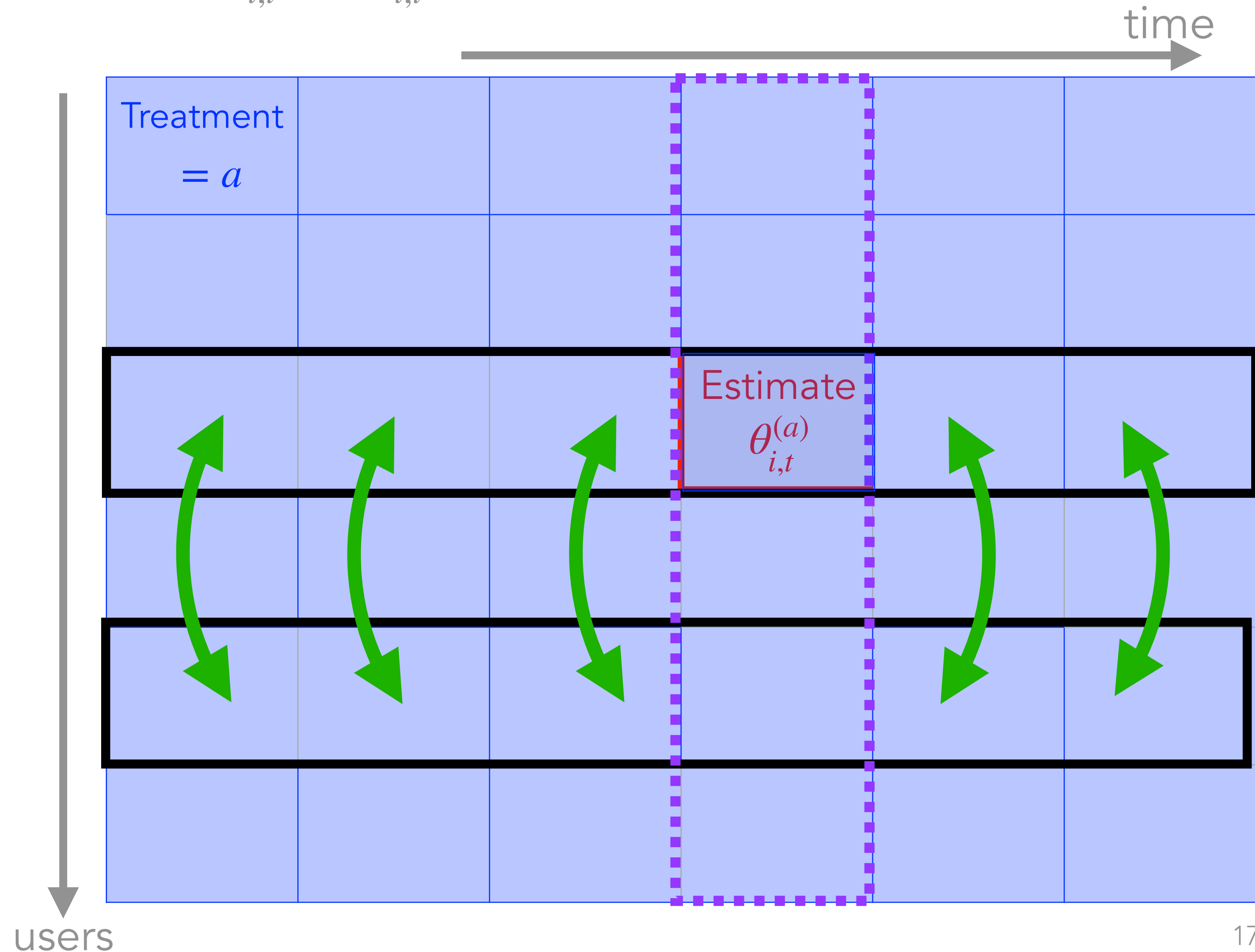
# User-NN with data split (no missingness)

$$R_{i,t} = R_{i,t}(A_{i,t}) \triangleq \theta_{i,t}^{(A_{i,t})} + \varepsilon_{i,t}^{(A_{i,t})}$$

$$\rho_{i,j}^{(a)} = \frac{\sum_{t' \neq t} (R_{i,t'} - R_{j,t'})^2}{T - 1}$$

$$\hat{\theta}_{i,t,\text{user-NN}}^{(a)} = \frac{\sum_{j=1}^N R_{j,t} \mathbf{1}(\rho_{i,j}^{(a)} \leq \eta)}{\sum_{j=1}^N \mathbf{1}(\rho_{i,j}^{(a)} \leq \eta)}$$

Why may we want to do a data split?



# Proof outline for user-NN (no missingness)

- Simple case: Always assign  $A_{j,t} = a$  and  $\theta_{i,t}^{(a)} \triangleq u_i v_t$

$$R_{i,t} = R_{i,t}(A_{i,t}) \triangleq \theta_{i,t}^{(A_{i,t})} + \varepsilon_{i,t}^{(A_{i,t})}$$

# Proof outline for user-NN (no missingness)

- Simple case: Always assign  $A_{j,t} = a$  and  $\theta_{i,t}^{(a)} \triangleq u_i v_t$

- $$\begin{aligned}\hat{\theta}_{i,t,\text{user-NN}}^{(a)} &= \frac{\sum_{j \in \text{user nbrs}} R_{j,t}}{|\text{user nbrs}|} = \frac{\sum_{j \in \text{user nbrs}} \theta_{j,t}^{(a)} + \varepsilon_{j,t}^{(a)}}{|\text{user nbrs}|} \\ &= \frac{\sum_{j \in \text{user nbrs}} u_j}{|\text{user nbrs}|} v_t + \frac{\sum_{j \in \text{user nbrs}} \varepsilon_{j,t}^{(a)}}{|\text{user nbrs}|}\end{aligned}$$

$$R_{i,t} = R_{i,t}(A_{i,t}) \triangleq \theta_{i,t}^{(A_{i,t})} + \varepsilon_{i,t}^{(A_{i,t})}$$

# Proof outline for user-NN (no missingness)

- Simple case: Always assign  $A_{j,t} = a$  and  $\theta_{i,t}^{(a)} \triangleq u_i v_t$

- $$\begin{aligned}\hat{\theta}_{i,t,\text{user-NN}}^{(a)} &= \frac{\sum_{j \in \text{user nbrs}} R_{j,t}}{|\text{user nbrs}|} = \frac{\sum_{j \in \text{user nbrs}} \theta_{j,t}^{(a)} + \varepsilon_{j,t}^{(a)}}{|\text{user nbrs}|} \\ &= \underbrace{\frac{\sum_{j \in \text{user nbrs}} u_j}{|\text{user nbrs}|}}_{\hat{u}_i} v_t + \underbrace{\frac{\sum_{j \in \text{user nbrs}} \varepsilon_{j,t}^{(a)}}{|\text{user nbrs}|}}_{\bar{\varepsilon}_t}\end{aligned}$$

$$\text{user-nbrs} = \{j : \rho_{i,j}^{(a)} \leq \eta\}$$

# Proof outline for user-NN (no missingness)

- Simple case: Always assign  $A_{j,t} = a$  and  $\theta_{i,t}^{(a)} \triangleq u_i v_t$

$$\begin{aligned}
 \bullet \quad \hat{\theta}_{i,t,\text{user-NN}}^{(a)} &= \frac{\sum_{j \in \text{user nbrs}} R_{j,t}}{|\text{user nbrs}|} = \frac{\sum_{j \in \text{user nbrs}} \theta_{j,t}^{(a)} + \varepsilon_{j,t}^{(a)}}{|\text{user nbrs}|} \\
 &= \underbrace{\frac{\sum_{j \in \text{user nbrs}} u_j}{|\text{user nbrs}|}}_{\hat{u}_i} v_t + \underbrace{\frac{\sum_{j \in \text{user nbrs}} \varepsilon_{j,t}^{(a)}}{|\text{user nbrs}|}}_{\bar{\varepsilon}_t}
 \end{aligned}$$

$$\bullet \quad |\theta_{i,t}^{(a)} - \hat{\theta}_{i,t,\text{user-NN}}^{(a)}| = |u_i v_t - \hat{\theta}_{i,t,\text{user-NN}}^{(a)}| \leq |u_i v_t - \hat{u}_i v_t| + |\bar{\varepsilon}_t|$$

# Re-expressing the distance without data-split

- Simple case: Always assign  $A_{j,t} = a$  and  $\theta_{i,t}^{(a)} \triangleq u_i v_t$

- $$\rho_{i,j}^{(a)} = \frac{\sum_{t'=1}^T (R_{i,t'} - R_{j,t'})^2}{T} = \frac{1}{T} \sum_{t'=1}^T \left[ (u_i v_{t'} - u_j v_{t'})^2 + (\varepsilon_{i,t'} - \varepsilon_{j,t'})^2 + 2(u_i v_{t'} - u_j v_{t'}) (\varepsilon_{i,t'} - \varepsilon_{j,t'}) \right]$$
$$= (u_i - u_j)^2 \frac{\sum_{t'=1}^T v_{t'}^2}{T} + \frac{\sum_{t'=1}^T (\varepsilon_{i,t'} - \varepsilon_{j,t'})^2}{T} + 2(u_i - u_j) \frac{\sum_{t'=1}^T v_{t'}' (\varepsilon_{i,t'} - \varepsilon_{j,t'})}{T}$$

# Re-expressing the distance without data-split

- $$\rho_{i,j}^{(a)} = (u_i - u_j)^2 \frac{\sum_{t'=1}^T v_{t'}^2}{T} + \frac{\sum_{t'=1}^T (\varepsilon_{i,t'} - \varepsilon_{j,t'})^2}{T} + 2(u_i - u_j) \frac{\sum_{t'=1}^T v_{t'}'(\varepsilon_{i,t'} - \varepsilon_{j,t'})}{T}$$

- $\rho_{i,j}^{(a)}$  depends on noise at time  $t$   
 $\implies$  user-nbrs =  $\{j : \rho_{i,j}^{(a)} \leq \eta\}$  are correlated with noise at time  $t$

- $$\bar{\varepsilon}_t = \frac{\sum_{j \in \text{user nbrs}} \varepsilon_{j,t}^{(a)}}{|\text{user nbrs}|}$$

-- need not behave as mean of iid noise and need not decay to 0

# Re-expressing the distance without data-split

- $$\rho_{i,j}^{(a)} = (u_i - u_j)^2 \frac{\sum_{t'=1}^T v_{t'}^2}{T} + \frac{\sum_{t'=1}^T (\varepsilon_{i,t'} - \varepsilon_{j,t'})^2}{T} + 2(u_i - u_j) \frac{\sum_{t'=1}^T v_{t'}'(\varepsilon_{i,t'} - \varepsilon_{j,t'})}{T}$$

- $\rho_{i,j}^{(a)}$  depends on noise at time  $t$   
 $\implies$  user-nbrs =  $\{j : \rho_{i,j}^{(a)} \leq \eta\}$  are correlated with noise at time  $t$

- $$\bar{\varepsilon}_t = \frac{\sum_{j \in \text{user nbrs}} \varepsilon_{j,t}^{(a)}}{|\text{user nbrs}|}$$

Data splitting, or, excluding time  $t$  data in defining the distance  $\rho_{i,j}^{(a)}$  breaks this dependence and the noise at time  $t$  is not correlated with user-nbrs.

-- need not behave as mean of iid noise and need not decay to 0



# What do these three terms concentrate on?

$$\bullet \rho_{i,j}^{(a)} = (u_i - u_j)^2 \frac{\sum_{t' \neq t} v_{t'}^2}{T-1} + \frac{\sum_{t' \neq t} (\varepsilon_{i,t'} - \varepsilon_{j,t'})^2}{T-1} + 2(u_i - u_j) \frac{\sum_{t' \neq t} v_{t'}' (\varepsilon_{i,t'} - \varepsilon_{j,t'})}{T-1}$$

# Inverting the distance to get a control on $|u_i - u_j|$

- Assume  $v_t, \varepsilon$  are bounded and  $\mathbb{E}[v_t^2] = v_\star^2$  then

- $|\rho_{i,j}^{(a)} - (u_i - u_j)^2 v_\star^2 - 2\sigma^2| \leq \frac{C\sqrt{\log(c/\delta)}}{\sqrt{T}}$  with probability  $1 - \delta$

- Treat  $\delta$  as a constant

- Rearranging terms  $|u_i - u_j|^2 \leq \frac{1}{v_\star^2} \left( \rho_{i,j}^{(a)} - 2\sigma^2 + \frac{C}{\sqrt{T}} \right)$

Univariate factors: A general error bound for user NN when  $A_{j,t} \equiv a$  and we use data split, ignore time  $t$  data while computing distance)

$$(\hat{\theta}_{i,t,\eta}^{(a)} - \theta_{i,t}^{(a)})^2 \lesssim \underbrace{\frac{1}{v_{\star}^2} \left( \underbrace{\eta - 2\sigma^2}_{\eta'} + \underbrace{\frac{C}{\sqrt{T-1}}}_{e_T} \right)}_{\text{NN bias due to threshold}} + \underbrace{\frac{\sigma^2}{N_{i,\eta'-e_T}}}_{\text{NN noise variance}}$$

$$v_{\star}^2 = \mathbb{E}[v_{t'}^2]$$

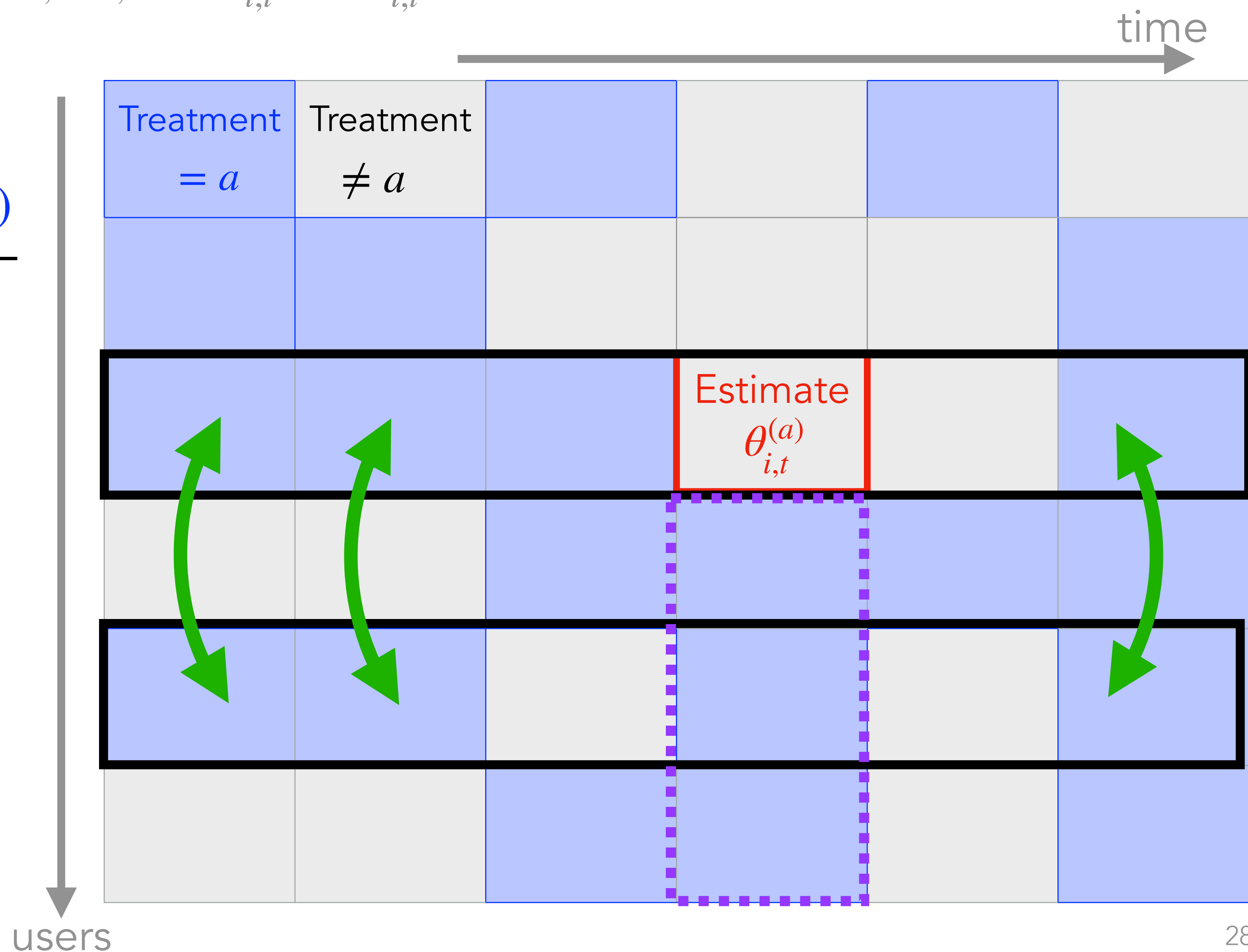
$$N_{i,\gamma} \triangleq |\{j \neq i : (u_i - u_j)^2 v_{\star}^2 \leq \gamma\}|$$

# User-NN with data split under **pure exploration** $\mathbb{P}(A_{i,t} = a) = p$

$$R_{i,t} = R_{i,t}(A_{i,t}) \triangleq \theta_{i,t}^{(A_{i,t})} + \varepsilon_{i,t}^{(A_{i,t})}$$

$$\rho_{i,j}^{(a)} = \frac{\sum_{t' \neq t} (R_{i,t'} - R_{j,t'})^2 \cdot \mathbf{1}(A_{i,t'} = A_{j,t'} = a)}{\sum_{t' \neq t} \mathbf{1}(A_{i,t'} = A_{j,t'} = a)}$$

$$\hat{\theta}_{i,t,\text{user-NN}}^{(a)} = \frac{\sum_{j=1}^N R_{j,t} \mathbf{1}(\rho_{i,j}^{(a)} \leq \eta, A_{j,t} = a)}{\sum_{j=1}^N \mathbf{1}(\rho_{i,j}^{(a)} \leq \eta, A_{j,t} = a)}$$



# Univariate factors + pure exploration policy $\mathbb{P}(A_{i,t} = a) = p$ : A general error bound for user NN with data split

$$(\hat{\theta}_{i,t,\eta}^{(a)} - \theta_{i,t}^{(a)})^2 \lesssim \frac{1}{v_{\star}^2} \left( \underbrace{\eta - 2\sigma^2}_{\eta'} + \underbrace{\frac{C}{\sqrt{p^2(T-1)}}}_{e_T} \right) + \frac{\sigma^2}{pN_{i,\eta'-e_T}}$$

$\eta'$                        $e_T$                      

NN bias  
due to threshold      Error in  
NN distance              NN noise  
variance

$$v_{\star}^2 = \mathbb{E}[v_{t'}^2]$$

$$N_{i,\gamma} \triangleq |\{j \neq i : (u_i - u_j)^2 v_{\star}^2 \leq \gamma\}|$$

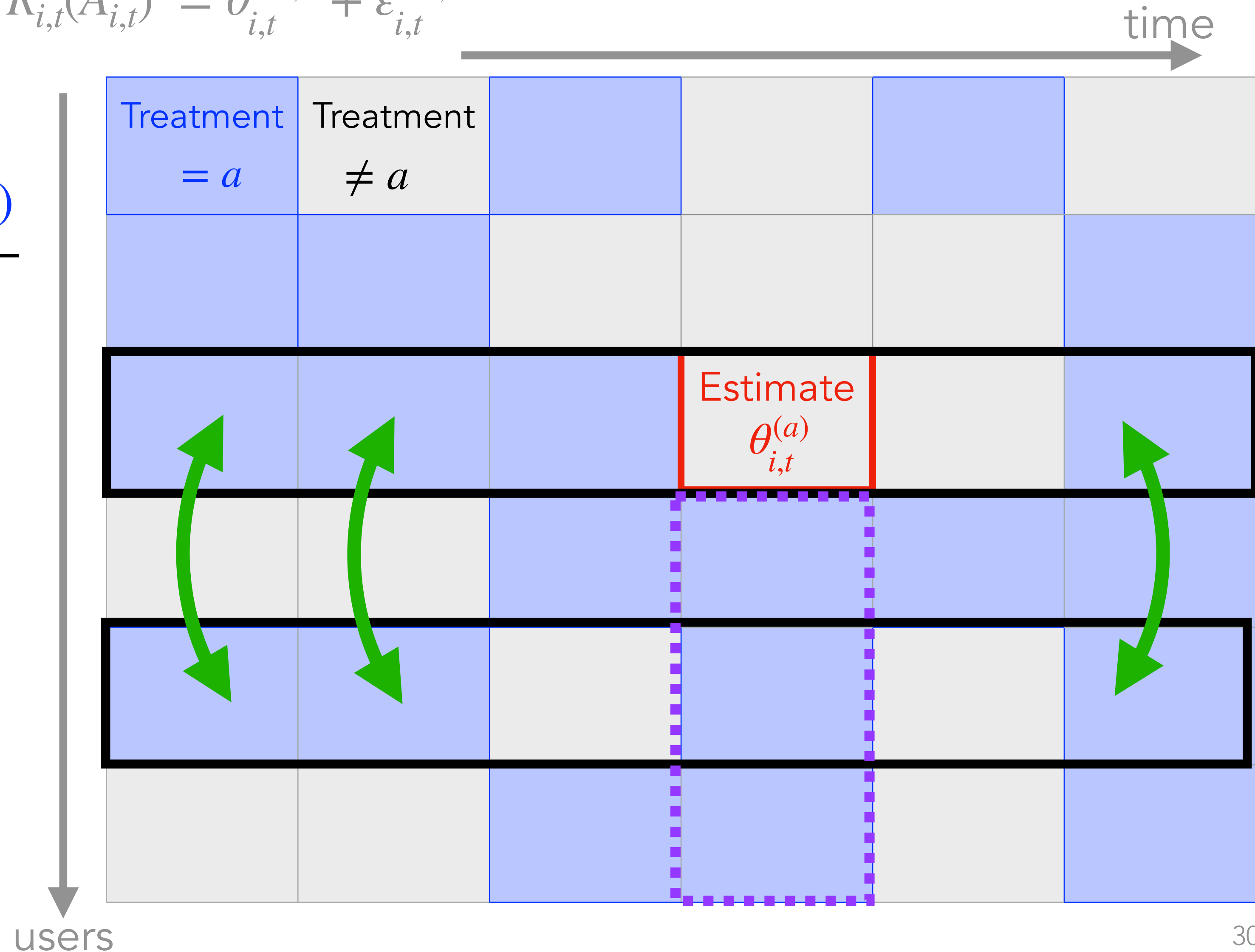
# User-NN with data split:

Will this trivially work when policy is sequential/learning?

$$R_{i,t} = R_{i,t}(A_{i,t}) \triangleq \theta_{i,t}^{(A_{i,t})} + \varepsilon_{i,t}^{(A_{i,t})}$$

$$\rho_{i,j}^{(a)} = \frac{\sum_{t' \neq t} (R_{i,t'} - R_{j,t'})^2 \cdot \mathbf{1}(A_{i,t'} = A_{j,t'} = a)}{\sum_{t' \neq t} \mathbf{1}(A_{i,t'} = A_{j,t'} = a)}$$

$$\hat{\theta}_{i,t,\text{user-NN}}^{(a)} = \frac{\sum_{j=1}^N R_{j,t} \mathbf{1}(\rho_{i,j}^{(a)} \leq \eta, A_{j,t} = a)}{\sum_{j=1}^N \mathbf{1}(\rho_{i,j}^{(a)} \leq \eta, A_{j,t} = a)}$$



# User-NN with data split:

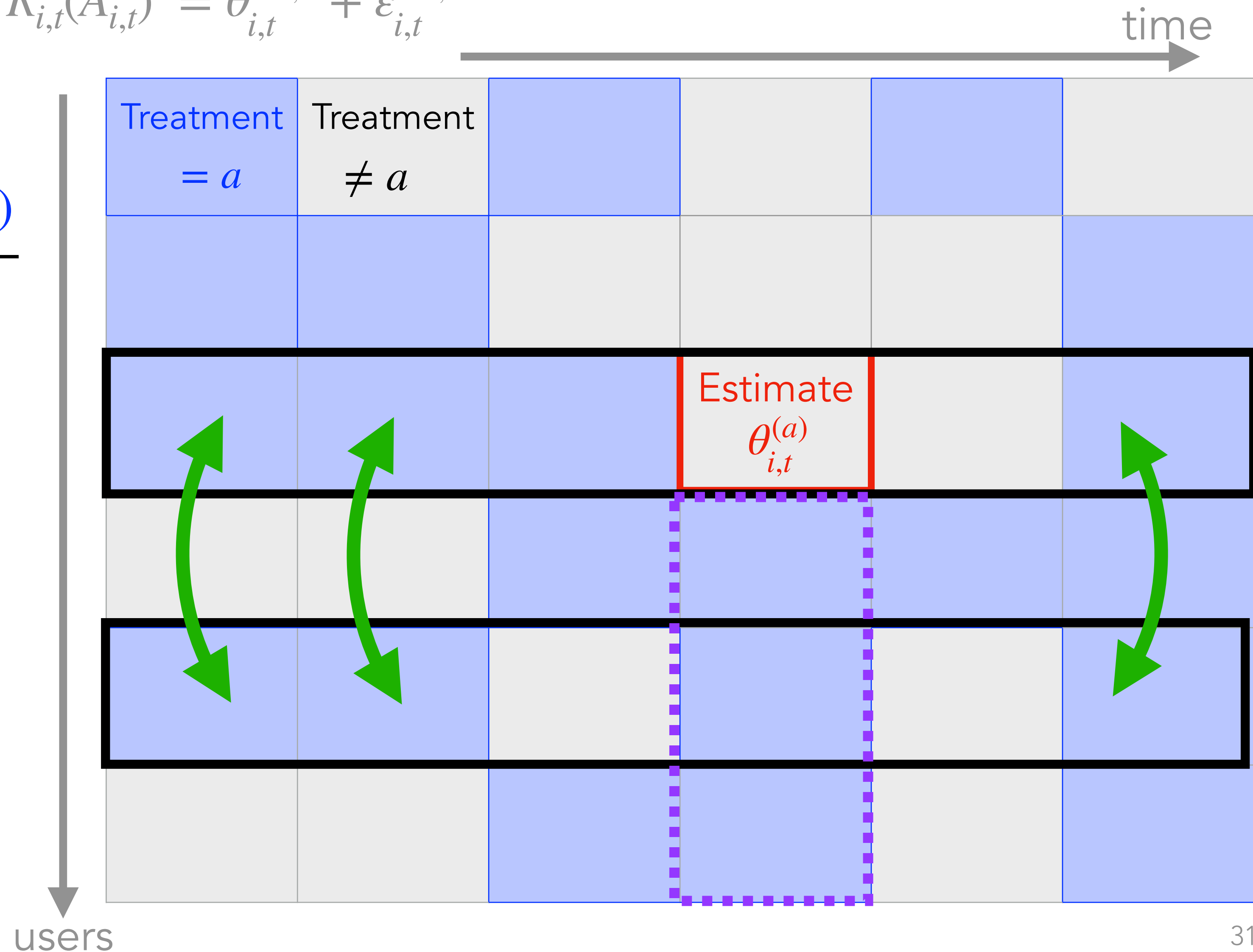
## Will this trivially work when policy is sequential/learning?

$$R_{i,t} = R_{i,t}(A_{i,t}) \triangleq \theta_{i,t}^{(A_{i,t})} + \varepsilon_{i,t}^{(A_{i,t})}$$

$$\rho_{i,j}^{(a)} = \frac{\sum_{t' < t} (R_{i,t'} - R_{j,t'})^2 \cdot \mathbf{1}(A_{i,t'} = A_{j,t'} = a)}{\sum_{t' < t} \mathbf{1}(A_{i,t'} = A_{j,t'} = a)}$$

$$\hat{\theta}_{i,t,\text{user-NN}}^{(a)} = \frac{\sum_{j=1}^N R_{j,t} \mathbf{1}(\rho_{i,j}^{(a)} \leq \eta, A_{j,t} = a)}{\sum_{j=1}^N \mathbf{1}(\rho_{i,j}^{(a)} \leq \eta, A_{j,t} = a)}$$

**How about this?**



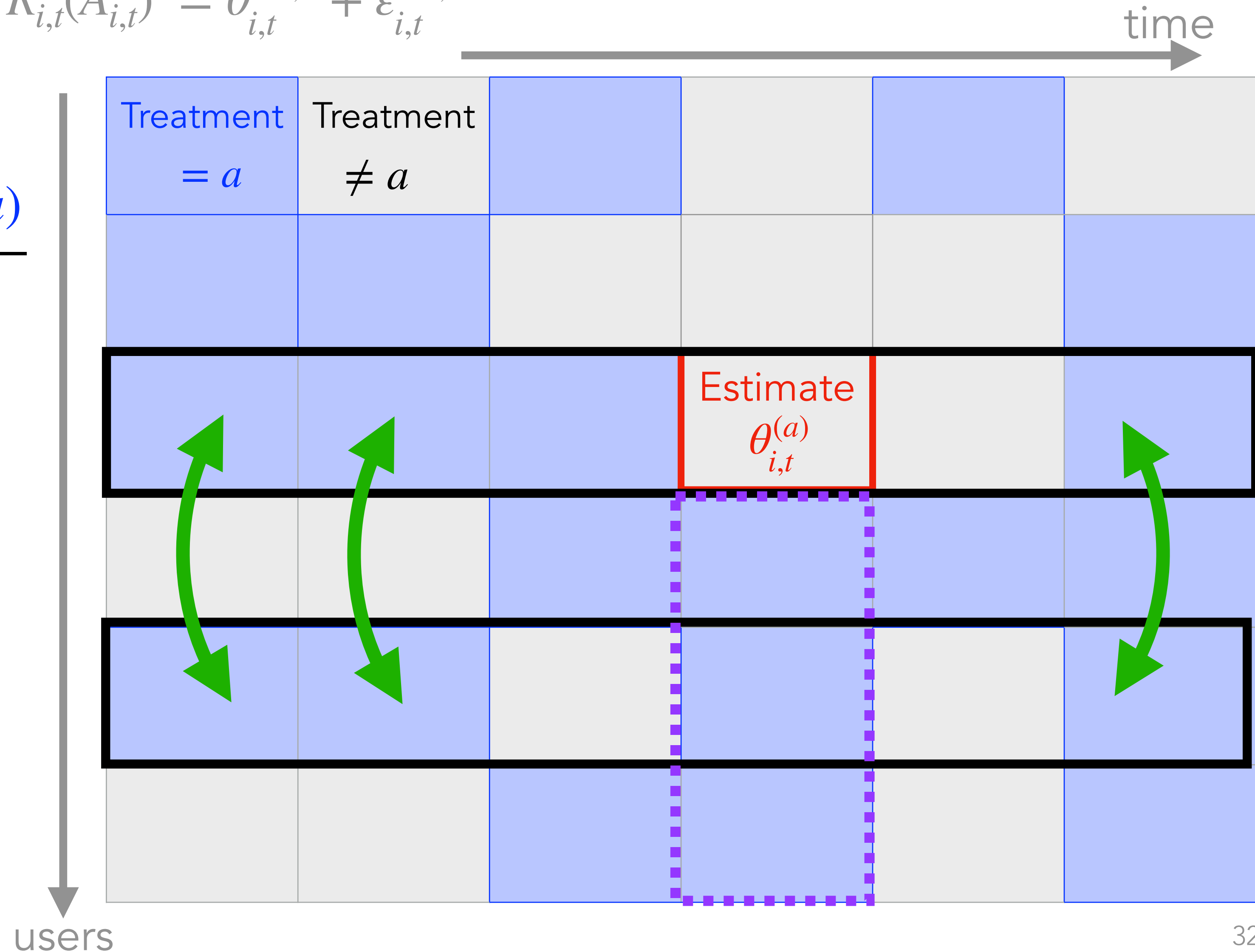
# Back to our original user-NN: No data-split + learning policy

$$R_{i,t} = R_{i,t}(A_{i,t}) \triangleq \theta_{i,t}^{(A_{i,t})} + \varepsilon_{i,t}^{(A_{i,t})}$$

$$\rho_{i,j}^{(a)} = \frac{\sum_{t'=1}^T (R_{i,t'} - R_{j,t'})^2 \cdot \mathbf{1}(A_{i,t'} = A_{j,t'} = a)}{\sum_{t'=1}^T \mathbf{1}(A_{i,t'} = A_{j,t'} = a)}$$

$$\hat{\theta}_{i,t,\text{user-NN}}^{(a)} = \frac{\sum_{j=1}^N R_{j,t} \mathbf{1}(\rho_{i,j}^{(a)} \leq \eta, A_{j,t} = a)}{\sum_{j=1}^N \mathbf{1}(\rho_{i,j}^{(a)} \leq \eta, A_{j,t} = a)}$$

How about this?





# A new sandwich argument for user-nbrs = $\{j : \rho_{i,j}^{(a)} \leq \eta\}$

- First using Azuma-Hoeffding concentration, we show that

$$|\rho_{i,j}^{(a)} - (u_i - u_j)^2 v_{\star}^2 - 2\sigma^2| \leq \frac{C}{\sqrt{p^2 T}}$$

with high probability for a learning policy with exploration  $p$

# A new sandwich argument for user-nbrs = $\{j : \rho_{i,j}^{(a)} \leq \eta\}$

- First using Azuma-Hoeffding concentration, we show that

$$|\rho_{i,j}^{(a)} - (u_i - u_j)^2 v_{\star}^2 - 2\sigma^2| \leq \frac{C}{\sqrt{p^2 T}}$$

with high probability for a learning policy with exploration  $p$

- Define  $Q_1 \triangleq \{j : (u_i - u_j)^2 v_{\star}^2 \leq (\eta - 2\sigma^2 + \frac{C}{\sqrt{p^2 T}})\}$

$$\text{and } Q_2 \triangleq \{j : (u_i - u_j)^2 v_{\star}^2 \leq (\eta - 2\sigma^2 - \frac{C}{\sqrt{p^2 T}})\}$$

# A new sandwich argument for user-nbrs = $\{j : \rho_{i,j}^{(a)} \leq \eta\}$

- First using Azuma-Hoeffding concentration, we show that

$$|\rho_{i,j}^{(a)} - (u_i - u_j)^2 v_{\star}^2 - 2\sigma^2| \leq \frac{C}{\sqrt{p^2 T}}$$

with high probability for a learning policy with exploration  $p$

- Define  $Q_1 \triangleq \{j : (u_i - u_j)^2 v_{\star}^2 \leq (\eta - 2\sigma^2 + \frac{C}{\sqrt{p^2 T}})\}$

$$\text{and } Q_2 \triangleq \{j : (u_i - u_j)^2 v_{\star}^2 \leq (\eta - 2\sigma^2 - \frac{C}{\sqrt{p^2 T}})\}$$

- Then we have  $Q_2 \subseteq \text{User-nbrs} \subseteq Q_1$

# Applying the sandwich argument for neighbors

- $$\bar{\varepsilon}_t^2 = \frac{(\sum_{j \in \text{user nbrs}} \varepsilon_{j,t}^{(a)} \mathbf{1}(A_{j,t} = a))^2}{(\text{user nbrs with } A_{j,t} = a)^2}$$

# Applying the sandwich argument for neighbors

- $\bar{\varepsilon}_t^2 = \frac{(\sum_{j \in \text{user nbrs}} \varepsilon_{j,t}^{(a)} \mathbf{1}(A_{j,t} = a))^2}{(\text{user nbrs with } A_{j,t} = a)^2}$

$$= \frac{(\sum_{j \in Q_1} \varepsilon_{j,t}^{(a)} \mathbf{1}(A_{j,t} = a) + \sum_{j \in \text{user nbrs} \setminus Q_1} \varepsilon_{j,t}^{(a)} \mathbf{1}(A_{j,t} = a))^2}{(\text{user nbrs with } A_{j,t} = a)^2}$$

**Lower part of the sandwich**

# Applying the sandwich argument for neighbors

- $\bar{\varepsilon}_t^2 = \frac{(\sum_{j \in \text{user nbrs}} \varepsilon_{j,t}^{(a)} \mathbf{1}(A_{j,t} = a))^2}{(\text{user nbrs with } A_{j,t} = a)^2}$

$$= \frac{(\sum_{j \in Q_1} \varepsilon_{j,t}^{(a)} \mathbf{1}(A_{j,t} = a) + \sum_{j \in \text{user nbrs} \setminus Q_1} \varepsilon_{j,t}^{(a)} \mathbf{1}(A_{j,t} = a))^2}{(\text{user nbrs with } A_{j,t} = a)^2} \quad \text{Lower part of the sandwich}$$

$$\leq 2 \frac{(\sum_{j \in Q_1} \varepsilon_{j,t}^{(a)} \mathbf{1}(A_{j,t} = a))^2}{(\text{users} \in Q_1 \text{ with } A_{j,t} = a)^2} + 2c_{\text{noise}} \frac{(\sum_{j \in Q_2 \setminus Q_1} \mathbf{1}(A_{j,t} = a))^2}{(\text{users} \in Q_1 \text{ with } A_{j,t} = a)^2}$$

Upper part of the sandwich

# Applying the sandwich argument for neighbors

$$\begin{aligned}
 \bullet \quad \bar{\varepsilon}_t^2 &= \frac{(\sum_{j \in \text{user nbrs}} \varepsilon_{j,t}^{(a)} \mathbf{1}(A_{j,t} = a))^2}{(\text{user nbrs with } A_{j,t} = a)^2} \\
 &= \frac{(\sum_{j \in Q_1} \varepsilon_{j,t}^{(a)} \mathbf{1}(A_{j,t} = a) + \sum_{j \in \text{user nbrs} \setminus Q_1} \varepsilon_{j,t}^{(a)} \mathbf{1}(A_{j,t} = a))^2}{(\text{user nbrs with } A_{j,t} = a)^2} \\
 &\leq 2 \frac{(\sum_{j \in Q_1} \varepsilon_{j,t}^{(a)} \mathbf{1}(A_{j,t} = a))^2}{(\text{users} \in Q_1 \text{ with } A_{j,t} = a)^2} + 2c_{\text{noise}} \frac{(\sum_{j \in Q_2 \setminus Q_1} \mathbf{1}(A_{j,t} = a))^2}{(\text{users} \in Q_1 \text{ with } A_{j,t} = a)^2} \\
 &\lesssim \frac{\sigma^2}{pN_{i,\eta'-e_T}} + c_{\text{noise}} \left[ \frac{N_{i,\eta'+e_T} - N_{i,\eta'-e_T}}{pN_{i,\eta'-e_T}} \right]^2
 \end{aligned}$$

# Draw illustration



Our bound: Obtained by tuning a general error bound for user NN for sequential pooled policies over  $\eta$

$$(\hat{\theta}_{i,t,\eta}^{(a)} - \theta_{i,t}^{(a)})^2 \lesssim \frac{1}{\lambda_{\star}^2} \left( \underbrace{\eta - 2\sigma^2}_{\substack{\eta' \\ \text{NN bias} \\ \text{due to threshold}}} + \underbrace{\frac{C}{\sqrt{p^2 T}}}_{\substack{e_T \\ \text{Error in} \\ \text{NN distance}}} \right) + \underbrace{\frac{\sigma^2}{p N_{i,\eta' - e_T}}}_{\substack{\text{NN noise} \\ \text{variance}}} + c_{\text{noise}} \left[ \frac{N_{i,\eta' + e_T} - N_{i,\eta' - e_T}}{p N_{i,\eta' - e_T}} \right]^2$$

NN bias  
inflation due to  
**learning** policy

$$\lambda_{\star} \triangleq \lambda_{\min}(\Sigma_v) \quad \text{where } \Sigma_v = \mathbb{E}[v_t v_t^\top]$$

$$N_{i,\gamma} \triangleq |\{j \neq i : (u_i - u_j)^\top \Sigma_v (u_i - u_j) \leq \gamma\}|$$

# Proof summary for user-NN

$$\begin{aligned}
 \bullet \quad \hat{\theta}_{i,t,\text{user-NN}}^{(a)} &= \frac{\sum_{j \in \text{user nbrs}} R_{j,t}}{|\text{user nbrs}|} = \frac{\sum_{j \in \text{user nbrs}} \theta_{j,t}^{(a)} + \varepsilon_{j,t}^{(a)}}{|\text{user nbrs}|} \\
 &= \underbrace{\frac{\sum_{j \in \text{user nbrs}} u_j}{|\text{user nbrs}|}}_{\hat{u}_i} v_t + \underbrace{\frac{\sum_{j \in \text{user nbrs}} \varepsilon_{j,t}^{(a)}}{|\text{user nbrs}|}}_{\bar{\varepsilon}_t}
 \end{aligned}$$

$$\bullet \quad |\theta_{i,t}^{(a)} - \hat{\theta}_{i,t,\text{user-NN}}^{(a)}| = |u_i v_t - \hat{\theta}_{i,t,\text{user-NN}}^{(a)}| \leq |u_i v_t - \hat{u}_i v_t| + |\bar{\varepsilon}_t|$$

# Summary of the proof sketch for unit or time nearest neighbors

- $|u_i v_t - \hat{\theta}_{i,t,\text{user-NN}}^{(a)}| \leq |u_i v_t - \hat{u}_i v_t| + |\bar{\epsilon}_t| = O(|u_i - \hat{u}_i|)$
- $|u_i v_t - \hat{\theta}_{i,t,\text{time-NN}}^{(a)}| \leq |u_i v_t - u_i \hat{v}_t| + |\bar{\epsilon}_i| = O(|v_t - \hat{v}_t|)$
- Can we combine both to improve the quality?

Can we make the error rates symmetric in N and T?

$$| \hat{\theta}_{i,t,\text{user-NN}}^{(a)} - \theta_{i,t}^{(a)} | = \tilde{O} \left( \frac{1}{T^{1/4}} + \frac{1}{\sqrt{N}} \right)$$



$$| \text{??} - \theta_{i,t}^{(a)} | = \tilde{O} \left( \frac{1}{\sqrt{T}} + \frac{1}{\sqrt{N}} \right)$$



$$| \hat{\theta}_{i,t,\text{time-NN}}^{(a)} - \theta_{i,t}^{(a)} | = \tilde{O} \left( \frac{1}{N^{1/4}} + \frac{1}{\sqrt{T}} \right)$$

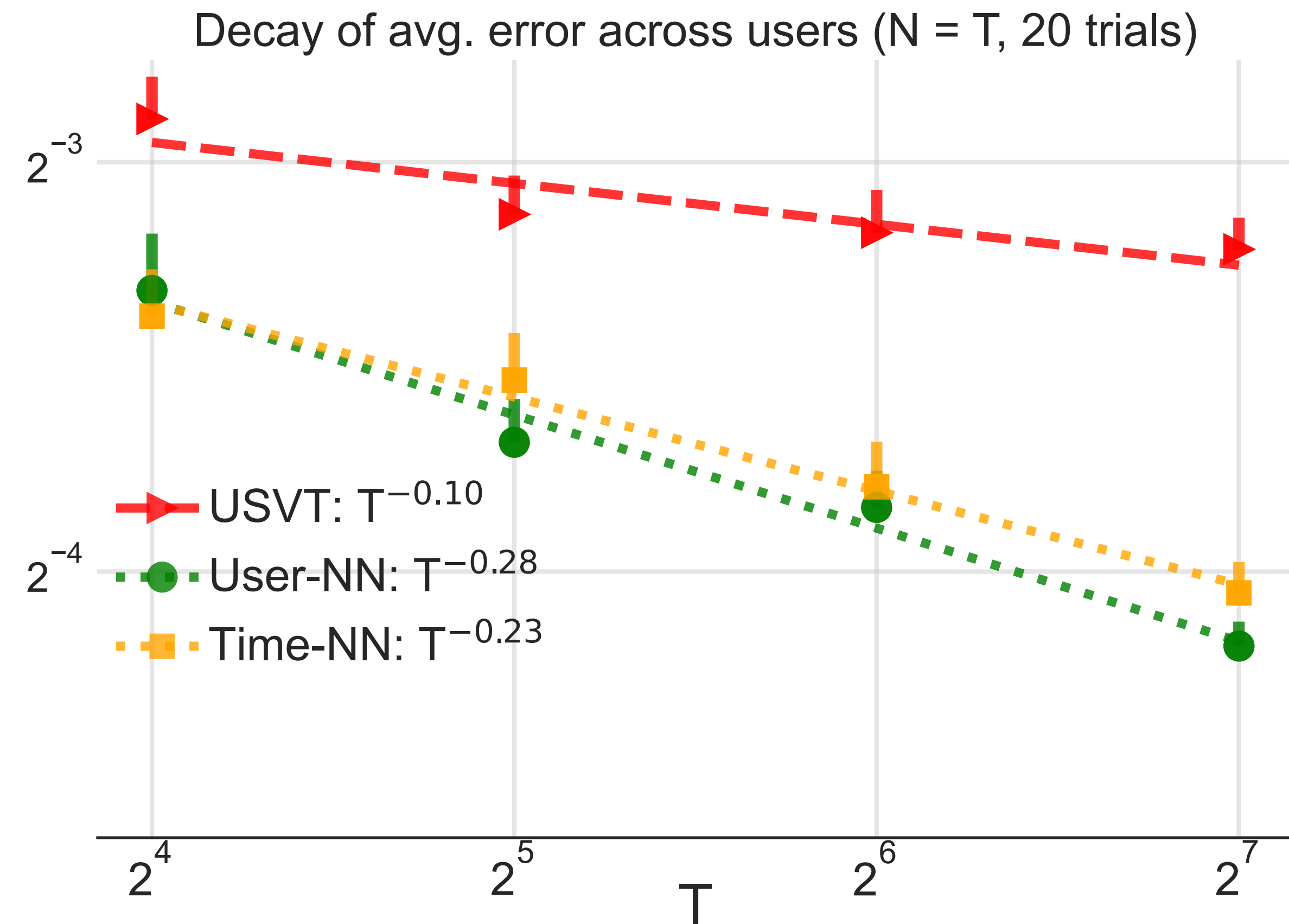
# Can we make the error rates symmetric in N and T?

$$|\hat{\theta}_{i,t,\text{user-NN}}^{(a)} - \theta_{i,t}^{(a)}| = \tilde{O}\left(\frac{1}{T^{1/4}} + \frac{1}{\sqrt{N}}\right)$$

$$|\text{??} - \theta_{i,t}^{(a)}| = \tilde{O}\left(\frac{1}{\sqrt{T}} + \frac{1}{\sqrt{N}}\right)$$

$$|\hat{\theta}_{i,t,\text{time-NN}}^{(a)} - \theta_{i,t}^{(a)}| = \tilde{O}\left(\frac{1}{N^{1/4}} + \frac{1}{\sqrt{T}}\right)$$

Uniform factors on  $[-0.5, 0.5]^4$ , Gaussian noise,  
pooled  $\varepsilon$ -greedy policy ( $\varepsilon = 0.5$ )



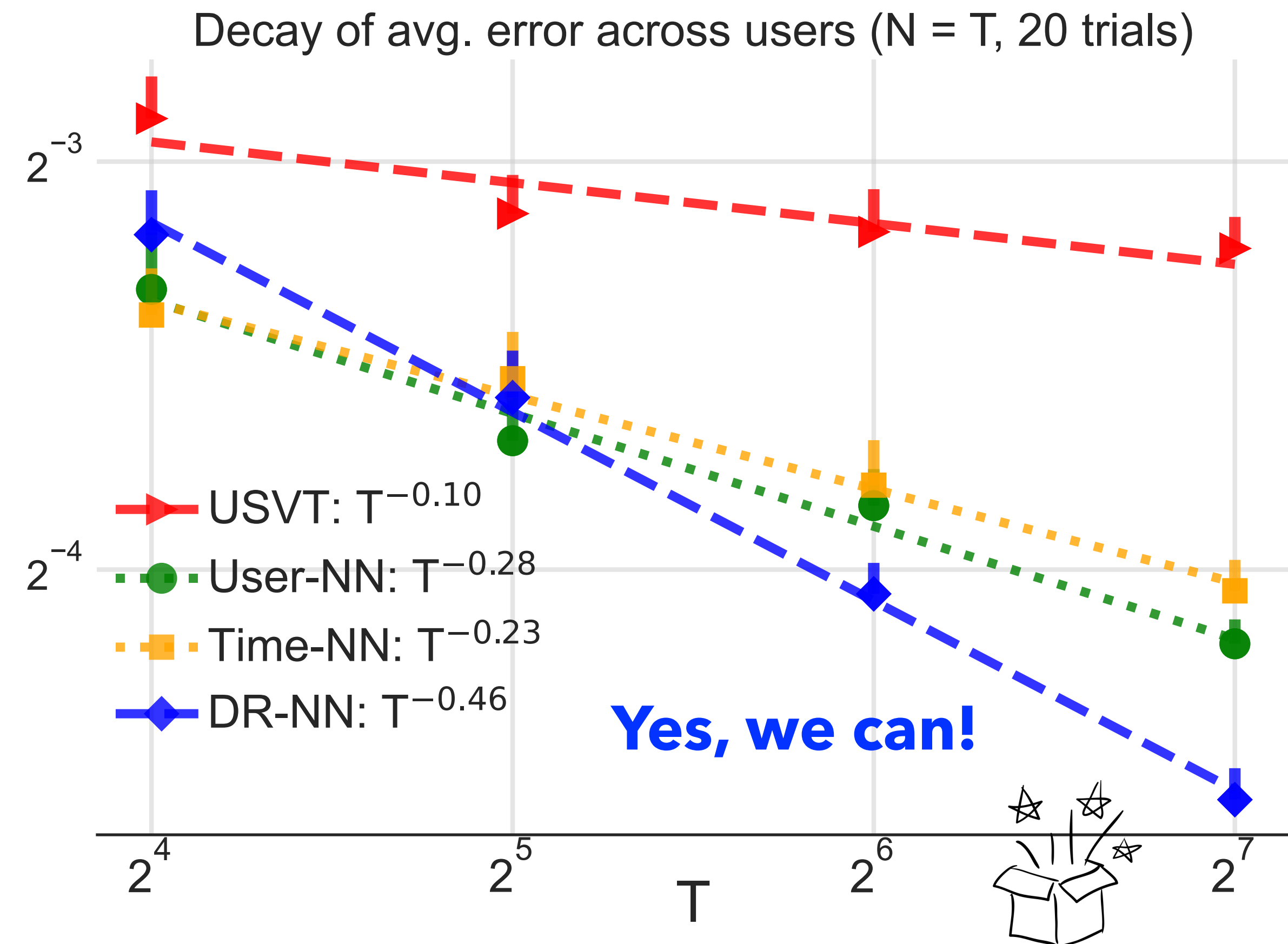
# Can we make the error rates symmetric in N and T?

Uniform factors on  $[-0.5, 0.5]^4$ , Gaussian noise,  
pooled  $\varepsilon$ -greedy policy ( $\varepsilon = 0.5$ )

$$|\hat{\theta}_{i,t,\text{DR-NN}}^{(a)} - \theta_{i,t}^{(a)}| = \tilde{O}\left(\frac{1}{\sqrt{T}} + \frac{1}{\sqrt{N}}\right)$$

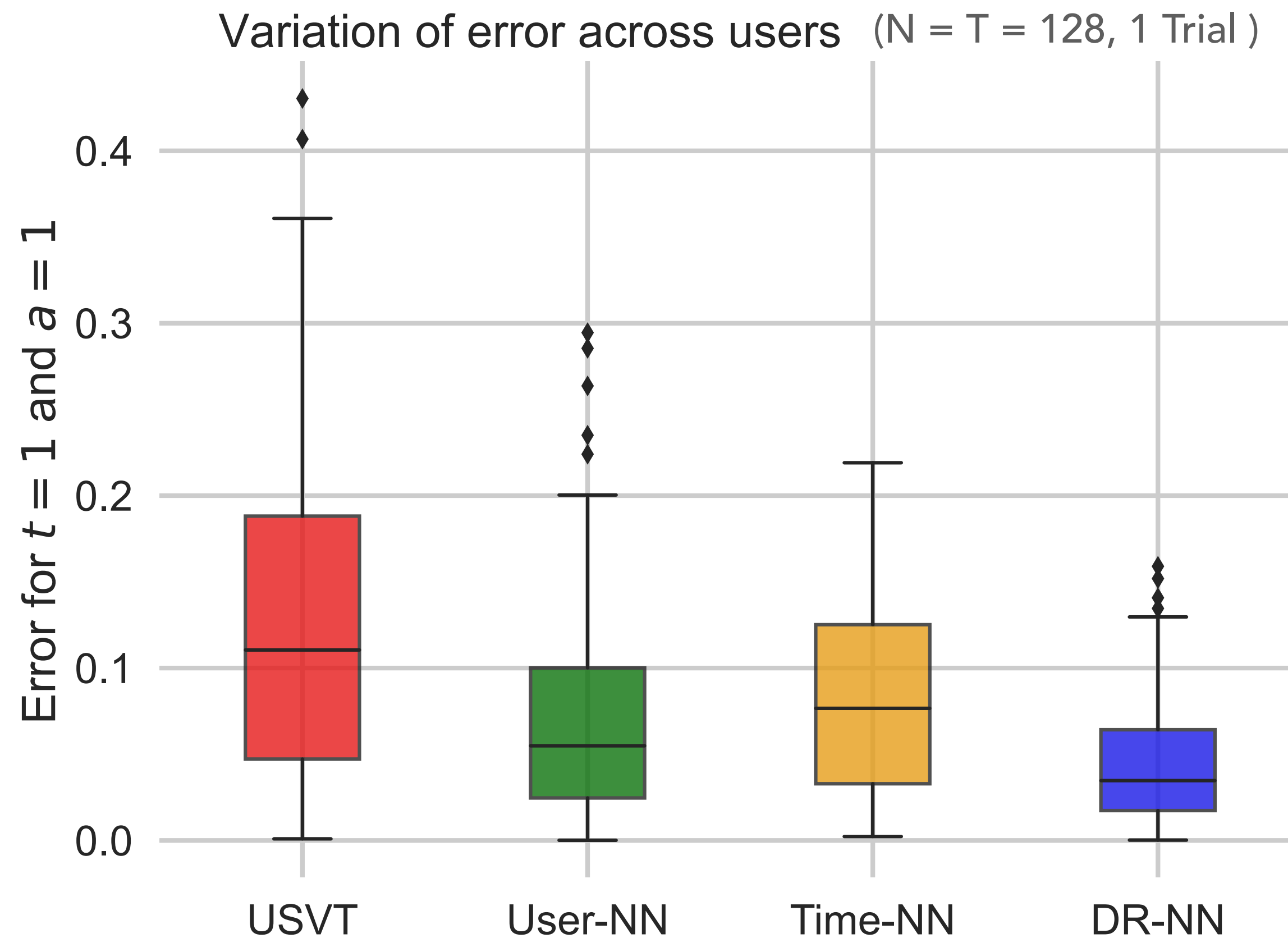
With a **suitable variant** of nearest neighbors

[**Dwivedi**-Tian-Tomkins-Klasnja-Murphy-Shah '22b]



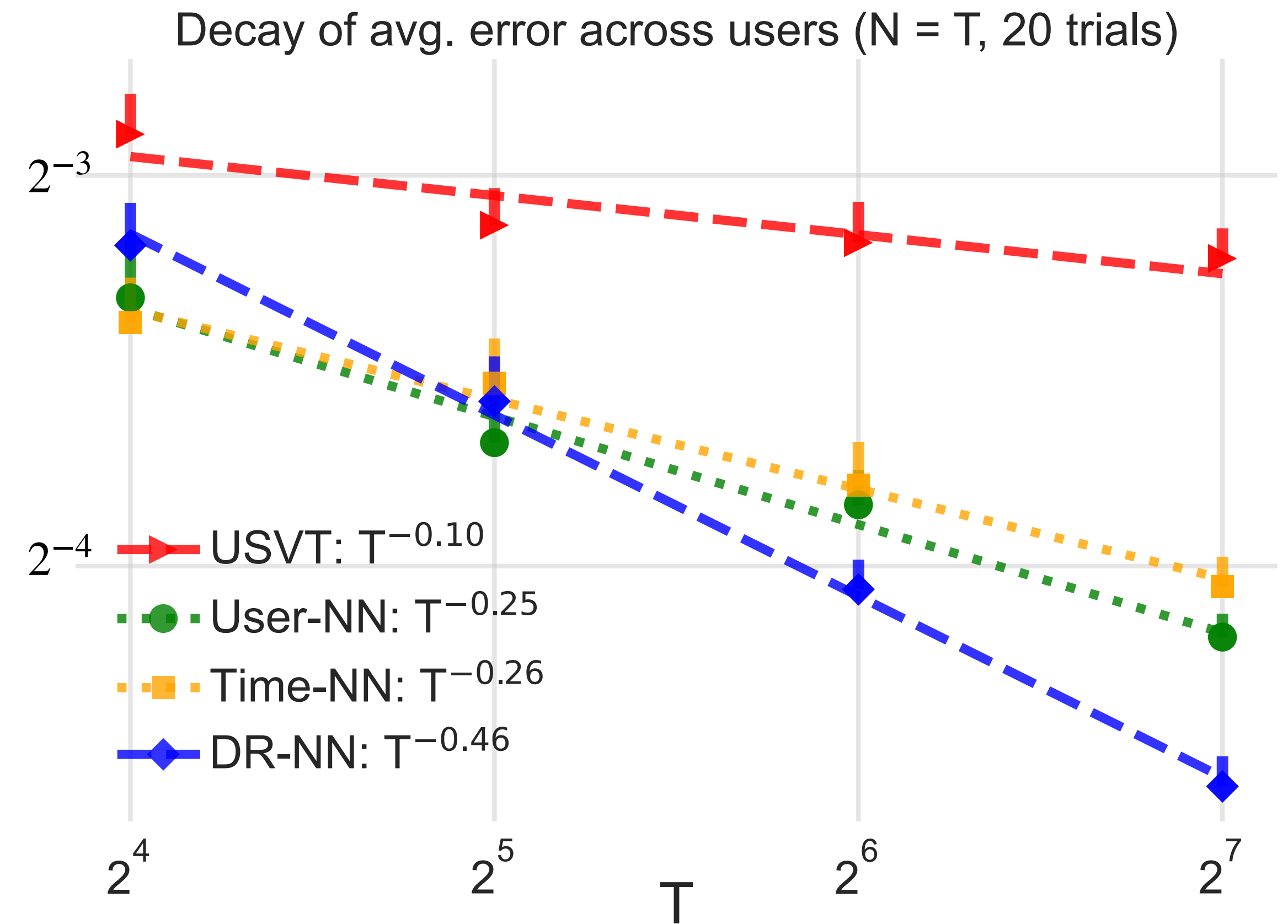
# Simulation results

Uniform latent factors on  $[-0.5, 0.5]^4$ , Gaussian noise, pooled  $\varepsilon$ -greedy policy ( $\varepsilon = 0.5$ )



A baseline  
algorithm from  
[Chatterjee 2014]

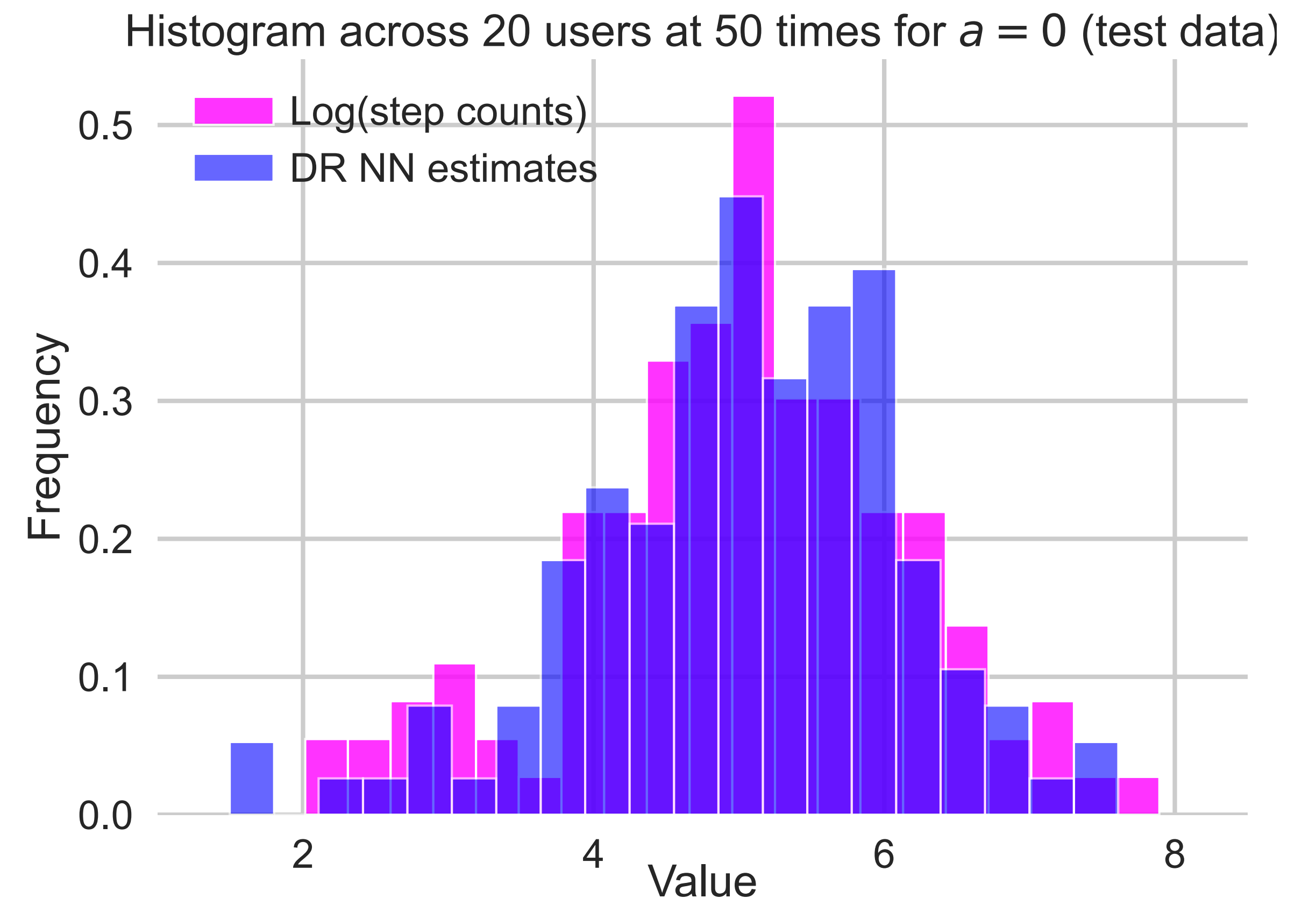
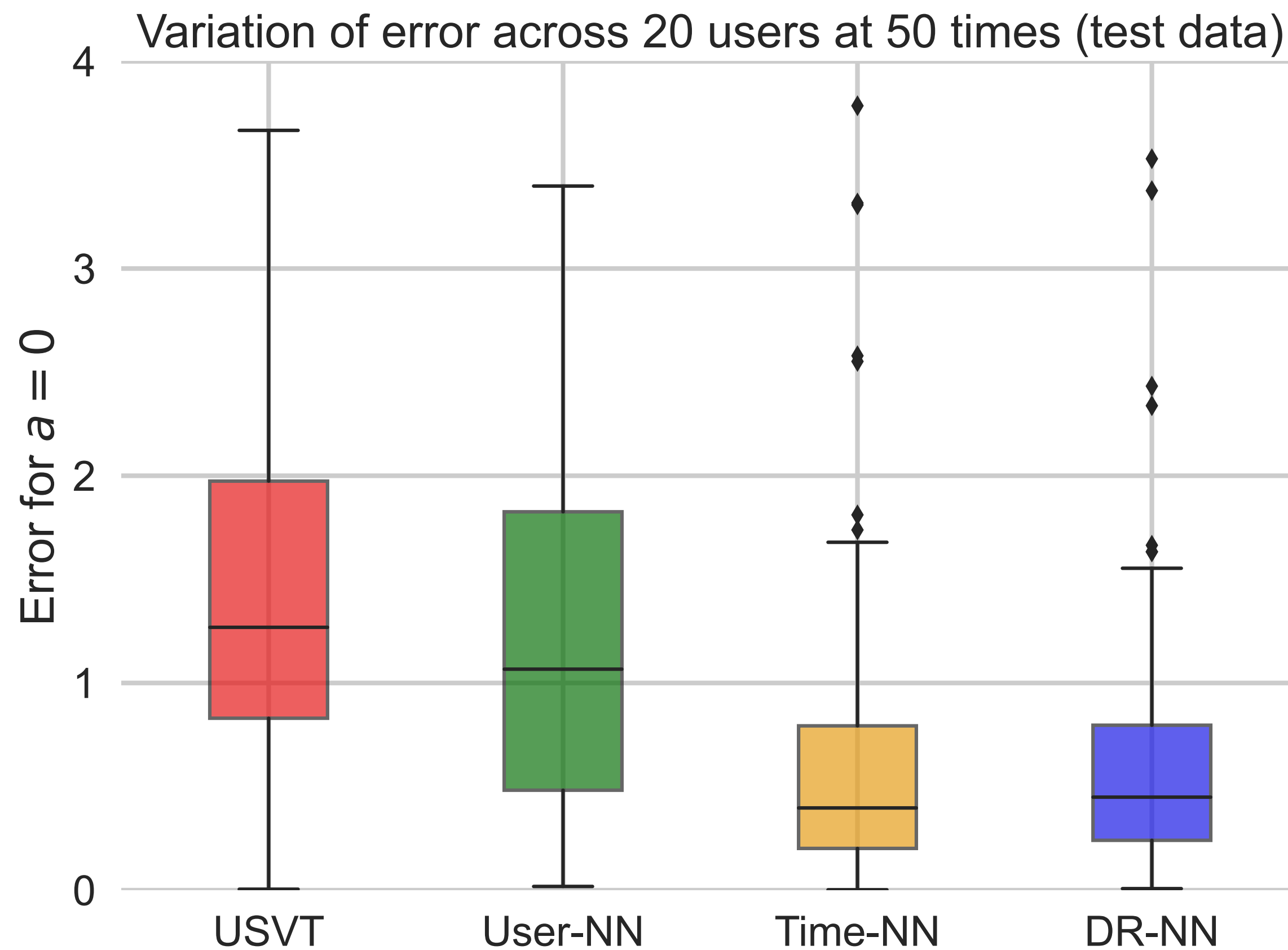
**DR-NN error  $\ll \min$  { user-NN error, time-NN error }**



# Personalized HeartSteps results



Treatments assigned with Thompson sampling independently for 91 users for 90 days, 5 times a day



**DR-NN error  $\approx \min$  { user-NN error, time-NN error }**



# In the search of improved estimator...

- Let's ignore the noise term and consider one nearest neighbor. "j" is a user neighbor so that  $\hat{u}_i = u_j$  and "t'" is time neighbor so that  $\hat{v}_t = v_{t'}$

- $\hat{\theta}_{i,t,\text{user-NN}}^{(a)} = R_{j,t} = u_j v_t$  and  $\hat{\theta}_{i,t,\text{time-NN}}^{(a)} = R_{i,t'} = u_i v_{t'}$

- Can we combine to improve?

- Average the two estimates:  $\frac{u_j v_t + u_i v_{t'}}{2} = \frac{R_{j,t} + R_{i,t'}}{2}$

- Use both neighbors: Outcome of user  $j$  at time  $t'$ :  $u_j v_{t'} = R_{j,t'}$

# Discussion questions

- What are the limitations of the factor model and the assumptions made for stating the non-asymptotic guarantee? Can you try to weaken these assumptions, to include states, delayed effects?
- Given all these counterfactual estimates, what kind of quantities could you investigate? How would you use them for between study analyses or to help the design of next study?
- **Hard:** Would the “averaged/merged” combination strategy significantly improve the performance?
  - Can you think of other ways to improve the NN estimator for the current model or more generally?

# Error of the “averaged” estimate

$$\begin{aligned} \left| u_i v_t - \frac{u_j v_t + u_i v_{t'}}{2} \right| &= \frac{|u_i v_t - u_j v_t + u_i v_t - u_i v_{t'}|}{2} \leq \frac{|u_i v_t - u_j v_t| + |u_i v_t - u_i v_{t'}|}{2} \\ &= \frac{|u_i - u_j| |v_t| + |u_i| |v_t - v_{t'}|}{2} \\ &= \frac{O(|u_i - u_j|) + O(|v_t - v_{t'}|)}{2} \\ &= \frac{1}{2} (\text{User-NN Error} + \text{Time-NN Error}) \\ &\approx \max \{ \text{User-NN Error}, \text{Time-NN Error} \} \end{aligned}$$

# Error of the “merged” estimate

$$|u_i v_t - u_j v_{t'}| = |u_i v_t - u_j v_t + u_j v_t - u_j v_{t'}| \leq |u_i v_t - u_j v_t| + |u_j v_t - u_j v_{t'}|$$

$$= |u_i - u_j| |v_t| + |u_j| |v_t - v_{t'}|$$

$$= O(|u_i - u_j|) + O(|v_t - v_{t'}|)$$

$$= \text{User-NN Error} + \text{Time-NN Error}$$

$$\approx \max \{ \text{User-NN Error}, \text{Time-NN Error} \}$$

# What do we desire?

- **Convert + to ×:**  $|u_i v_t - ??| = |u_i - u_j| \times |v_t - v_{t'}|$   
 $= \text{User-NN Error} \times \text{Time-NN Error}$   
or **max to min:**  $\approx \min \{ \text{User-NN Error}, \text{Time-NN Error} \}$

# What should be our estimator? Let's expand the RHS...

$$\cancel{u_i v_t} - \text{??} = (u_i - u_j) \times (v_t - v_{t'})$$

$$= \cancel{u_i v_t} - u_j v_t - u_i v_{t'} + u_j v_{t'}$$

$$\Rightarrow \text{??} = u_j v_t + u_i v_{t'} - u_j v_{t'}$$

$$R_{j,t} + R_{i,t'} - R_{j,t'}$$

This is our **improved** nearest neighbors estimator!

$$\cancel{u_i v_t} - \text{??} = (u_i - u_j) \times (v_t - v_{t'})$$

$$= \cancel{u_i v_t} - u_j v_t - u_i v_{t'} + u_j v_{t'}$$

$$\Rightarrow \text{??} = u_j v_t + u_i v_{t'} - u_j v_{t'}$$

$$\hat{\theta}_{i,t,\text{DR-NN}}^{(a)} = \frac{\sum_{j,t'} (R_{j,t} + R_{i,t'} - R_{j,t'}) \mathbf{1}_{i,t,j,t'}}{\sum_{j,t'} \mathbf{1}_{i,t,j,t'}}$$



$$\mathbf{1}_{i,t,j,t'} = \mathbf{1}(\rho_{i,j}^{(a)} \leq \eta, \rho_{t,t'}^{(a)} \leq \eta', A_{j,t} = A_{i,t'} = A_{j,t'} = a)$$

# This is our **improved** nearest neighbors estimator!

$$\cancel{u_i v_t} - \textcircled{??} = (u_i - u_j) \times (v_t - v_{t'})$$

$$= \cancel{u_i v_t} - u_j v_t - u_i v_{t'} + u_j v_{t'}$$

$$\Rightarrow \textcircled{??} = u_j v_t + u_i v_{t'} - u_j v_{t'}$$

**DR-NN error**  $\approx$  **user-NN error**  $\times$  **time-NN error**

$\lesssim$  **min{user-NN error, time-NN error}**

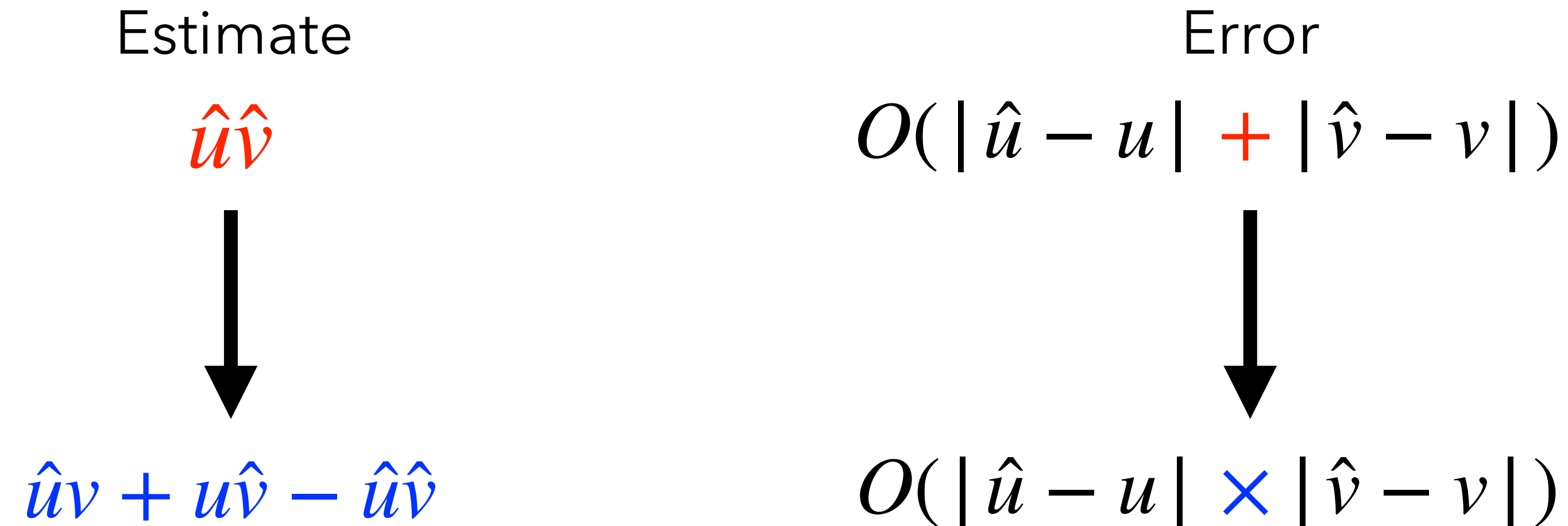
**Doubly robust to heterogeneity in user factors & time factors**

Double robustness, double machine learning...

[... Cassel+ '77, Robinson '88, Särndal+ '89, Robins+ '94, '95, '08, '09, Newey+ '94, '18, Bickel+ '98, van der Laan+ '03, Lunceford+ '04, Davidian+ '05, Li+ '11, Jiang+ '15, Chernozhukov+ '18, Hirshberg+ '18, Diaz '19, Arkhangelsky+ '21, Dorn+ '21 ...]



# Simplified view of doubly robust estimator for $uv$



Problem setting	$u$	$v$
Factor model (this talk)	user factor	time factor
Observational studies (Causal inference)	propensity function	mean outcome function
Off policy evaluation (Reinforcement learning)	importance ratio	reward function

Double robustness, double machine learning...

[... Cassel+ '77, Robinson '88, Särndal+ '89, Robins+ '94, '95, '08, '09, Newey+ '94, '18, Bickel+ '98, van der Laan+ '03, Lunceford+ '04, Davidian+ '05, Li+ '11, Jiang+ '15, Chernozhukov+ '18, Hirshberg+ '18, Diaz '19, Arkhangelsky+ '21, Dorn+ '21 ...]

# **Appendix [not covered in lecture]: Further details on derivations**

# Disclaimer

- $c, C$  are universal constants that might take a different value in each appearance

# Proof sketch for user-NN

- Simple case: Always assign  $A_{j,t} = a$  and  $\theta_{i,t}^{(a)} \triangleq u_i v_t$

$$\begin{aligned} \bullet \quad \hat{\theta}_{i,t,\text{user-NN}}^{(a)} &= \frac{\sum_{j \in \text{user nbrs}} R_{j,t}}{|\text{user nbrs}|} = \frac{\sum_{j \in \text{user nbrs}} \theta_{j,t}^{(a)} + \varepsilon_{j,t}^{(a)}}{|\text{user nbrs}|} \\ &= \frac{\sum_{j \in \text{user nbrs}} u_j}{|\text{user nbrs}|} v_t + \frac{\sum_{j \in \text{user nbrs}} \varepsilon_{j,t}^{(a)}}{|\text{user nbrs}|} \end{aligned}$$

$$\bullet \quad |u_i v_t - \hat{u}_i v_t| \leq \max_{j \in \text{user nbrs}} |u_i - u_j| |v_t| \lesssim \sqrt{\eta - 2\sigma^2} + \frac{1}{T^{1/4}}$$

$$\bullet \quad \bar{\varepsilon}_t = \frac{\sum_{j \in \text{user nbrs}} \varepsilon_{j,t}^{(a)}}{|\text{user nbrs}|} \lesssim \frac{\sigma}{\sqrt{|\text{user nbrs}|}} = \frac{\sigma}{\sqrt{N_\eta}}$$

Our goal: Control  $\max_{j \in \text{user nbrs}} |u_i - u_j| |v_t|$

- $|v_t|$  is bounded so suffices to bound  $\max_{j \in \text{user nbrs}} |u_i - u_j|$
- user neighbours =  $\{\rho_{i,j}^{(a)} \leq \eta\}$

- $$\rho_{i,j}^{(a)} = \frac{\sum_{t'=1}^T (R_{i,t'} - R_{j,t'})^2 \cdot \mathbf{1}(A_{i,t'} = A_{j,t'} = a)}{\sum_{t'=1}^T \mathbf{1}(A_{i,t'} = A_{j,t'} = a)} + \frac{\hat{\sigma}_\rho}{\sqrt{\sum_{t'=1}^T \mathbf{1}(A_{i,t'} = A_{j,t'} = a)}}$$

- In theory, we ignore the second term

# Controlling the bias via concentration of distance

- Simple case: Always assign  $A_{j,t} = a$  and  $\theta_{i,t}^{(a)} \triangleq u_i v_t$

- $$\rho_{i,j}^{(a)} = \frac{\sum_{t'=1}^T (R_{i,t'} - R_{j,t'})^2}{T}$$

# Re-expressing the distance

- Simple case: Always assign  $A_{j,t} = a$  and  $\theta_{i,t}^{(a)} \triangleq u_i v_t$

- $$\begin{aligned}\rho_{i,j}^{(a)} &= \frac{\sum_{t'=1}^T (R_{i,t'} - R_{j,t'})^2}{T} \\ &= \frac{1}{T} \sum_{t'=1}^T \left[ (u_i v_{t'} - u_j v_{t'})^2 + (\varepsilon_{i,t'} - \varepsilon_{j,t'})^2 + 2(u_i v_{t'} - u_j v_{t'})(\varepsilon_{i,t'} - \varepsilon_{j,t'}) \right] \\ &= (u_i - u_j)^2 \frac{\sum_{t'=1}^T v_{t'}^2}{T} + \frac{\sum_{t'=1}^T (\varepsilon_{i,t'} - \varepsilon_{j,t'})^2}{T} + 2(u_i - u_j) \frac{\sum_{t'=1}^T v_{t'}'(\varepsilon_{i,t'} - \varepsilon_{j,t'})}{T}\end{aligned}$$

# Re-expressing the distance: Collecting into three terms

$$\bullet \rho_{i,j}^{(a)} = (u_i - u_j)^2 \frac{\sum_{t'=1}^T v_{t'}^2}{T} + \frac{\sum_{t'=1}^T (\varepsilon_{i,t'} - \varepsilon_{j,t'})^2}{T} + 2(u_i - u_j) \frac{\sum_{t'=1}^T v_{t'}'(\varepsilon_{i,t'} - \varepsilon_{j,t'})}{T}$$



# What do these three terms concentrate on?

$$\bullet \rho_{i,j}^{(a)} = (u_i - u_j)^2 \frac{\sum_{t'=1}^T v_{t'}^2}{T} + \frac{\sum_{t'=1}^T (\varepsilon_{i,t'} - \varepsilon_{j,t'})^2}{T} + 2(u_i - u_j) \frac{\sum_{t'=1}^T v_{t'}'(\varepsilon_{i,t'} - \varepsilon_{j,t'})}{T}$$

Recall our assumptions –

$v_t$  are iid

$\varepsilon_{i,t}$  are iid zero mean with variance  $\sigma$

$v_t$  and  $\varepsilon_{i,t}$  are independent of each other

# What do these three terms concentrate on?

$$\bullet \rho_{i,j}^{(a)} = (u_i - u_j)^2 \frac{\sum_{t'=1}^T v_{t'}^2}{T} + \frac{\sum_{t'=1}^T (\varepsilon_{i,t'} - \varepsilon_{j,t'})^2}{T} + 2(u_i - u_j) \frac{\sum_{t'=1}^T v_{t'}'(\varepsilon_{i,t'} - \varepsilon_{j,t'})}{T}$$

$$\bullet \left| \frac{\sum_{t'=1}^T v_{t'}^2}{T} - ? \right| \lesssim ?$$

$$\bullet \left| \frac{\sum_{t'=1}^T (\varepsilon_{i,t'} - \varepsilon_{j,t'})^2}{T} - ? \right| \lesssim ?$$

$$\bullet \left| \frac{\sum_{t'=1}^T v_{t'}'(\varepsilon_{i,t'} - \varepsilon_{j,t'})}{T} - ? \right| \lesssim ?$$

Recall our assumptions –

$v_t$  are iid

$\varepsilon_{i,t}$  are iid zero mean with variance  $\sigma$

$v_t$  and  $\varepsilon_{i,t}$  are independent of each other

# Tools for concentration

- **Markov's inequality:** Let  $X_1, X_2, \dots, X_T$  be iid random variables with mean  $\mu$  and variance  $\text{Var}(X)$ , then

$$\mathbb{P} \left[ \left| \frac{\sum_{i=1}^T X_i}{T} - \mu \right| \leq \sqrt{\frac{\text{Var}(X)}{\delta T}} \right] \geq 1 - \delta$$

# Tools for concentration

- **Markov's inequality:** Let  $X_1, X_2, \dots, X_T$  be iid random variables with mean  $\mu$  and variance  $\text{Var}(X)$ , then

$$\mathbb{P} \left[ \left| \frac{\sum_{i=1}^T X_i}{T} - \mu \right| \leq \sqrt{\frac{\text{Var}(X)}{\delta T}} \right] \geq 1 - \delta$$

- **Chernoff-Hoeffding bound:** If  $X_i$  have mean  $\mu$  and are  $\gamma$ -sub-Gaussian, i.e.,  $\mathbb{E}[e^{t(X-\mu)}] \leq e^{t^2\gamma^2/2}$  then

$$\mathbb{P} \left[ \left| \frac{\sum_{i=1}^T X_i}{T} - \mu \right| \leq \gamma \sqrt{2 \log(1/\delta)} \right] \geq 1 - \delta$$

- Useful fact if  $|X_i| \leq c$ , then we can use  $\gamma = c$

# What do these three terms concentrate around?

**Their means!**

- $\rho_{i,j}^{(a)} = (u_i - u_j)^2 \frac{\sum_{t'=1}^T v_{t'}^2}{T} + \frac{\sum_{t'=1}^T (\varepsilon_{i,t'} - \varepsilon_{j,t'})^2}{T} + 2(u_i - u_j) \frac{\sum_{t'=1}^T v_{t'}'(\varepsilon_{i,t'} - \varepsilon_{j,t'})}{T}$

- $\left| \frac{\sum_{t'=1}^T v_{t'}^2}{T} - \mathbb{E}[v_{t'}^2] \right| \lesssim \frac{c\sqrt{\text{Var}(v_{t'}^2)}}{\sqrt{\delta T}} \text{ or } \frac{c\sqrt{\text{Sub-Gauss}(v_{t'}^2) \cdot \log(1/\delta)}}{\sqrt{T}}$

# What do these three terms concentrate around?

Their means!

$$\bullet \rho_{i,j}^{(a)} = (u_i - u_j)^2 \frac{\sum_{t'=1}^T v_{t'}^2}{T} + \frac{\sum_{t'=1}^T (\varepsilon_{i,t'} - \varepsilon_{j,t'})^2}{T} + 2(u_i - u_j) \frac{\sum_{t'=1}^T v_{t'}'(\varepsilon_{i,t'} - \varepsilon_{j,t'})}{T}$$

$$\bullet \left| \frac{\sum_{t'=1}^T v_{t'}^2}{T} - \mathbb{E}[v_{t'}^2] \right| \lesssim \frac{c\sqrt{\text{Var}(v_{t'}^2)}}{\sqrt{\delta T}} \text{ or } \frac{c\sqrt{\text{Sub-Gauss}(v_{t'}^2) \cdot \log(1/\delta)}}{\sqrt{T}}$$

$$\bullet \left| \frac{\sum_{t'=1}^T (\varepsilon_{i,t'} - \varepsilon_{j,t'})^2}{T} - 2\sigma^2 \right| \lesssim \frac{c\sqrt{2\text{Var}(\varepsilon^2)}}{\sqrt{\delta T}} \text{ or } \frac{c\sqrt{\text{Sub-Gauss}(\varepsilon^2) \cdot \log(1/\delta)}}{\sqrt{T}}$$

# What do these three terms concentrate around?

Their means!

- $$\rho_{i,j}^{(a)} = (u_i - u_j)^2 \frac{\sum_{t'=1}^T v_{t'}^2}{T} + \frac{\sum_{t'=1}^T (\varepsilon_{i,t'} - \varepsilon_{j,t'})^2}{T} + 2(u_i - u_j) \frac{\sum_{t'=1}^T v_{t'}'(\varepsilon_{i,t'} - \varepsilon_{j,t'})}{T}$$

- $$\left| \frac{\sum_{t'=1}^T v_{t'}^2}{T} - \mathbb{E}[v_{t'}^2] \right| \lesssim \frac{c\sqrt{\text{Var}(v_{t'}^2)}}{\sqrt{\delta T}} \text{ or } \frac{c\sqrt{\text{Sub-Gauss}(v_{t'}^2) \cdot \log(1/\delta)}}{\sqrt{T}}$$

- $$\left| \frac{\sum_{t'=1}^T (\varepsilon_{i,t'} - \varepsilon_{j,t'})^2}{T} - 2\sigma^2 \right| \lesssim \frac{c\sqrt{2\text{Var}(\varepsilon^2)}}{\sqrt{\delta T}} \text{ or } \frac{c\sqrt{\text{Sub-Gauss}(\varepsilon^2) \cdot \log(1/\delta)}}{\sqrt{T}}$$

- $$\left| \frac{\sum_{t'=1}^T v_{t'}'(\varepsilon_{i,t'} - \varepsilon_{j,t'})}{T} - 0 \right| \lesssim \frac{c\sqrt{2\text{Var}(v_t \varepsilon)}}{\sqrt{\delta T}} \text{ or } \frac{c\sqrt{\text{Sub-Gauss}(v_t \varepsilon) \cdot \log(1/\delta)}}{\sqrt{T}}$$

# Inverting the distance to get a control on $|u_i - u_j|$

- Assume  $v_t, \varepsilon$  are bounded and  $\mathbb{E}[v_t^2] = v_\star^2$  then

- $|\rho_{i,j}^{(a)} - (u_i - u_j)^2 v_\star^2 - 2\sigma^2| \leq \frac{C\sqrt{\log(c/\delta)}}{\sqrt{T}}$  with probability  $1 - \delta$

- Treat  $\delta$  as a constant

- Rearranging terms  $|u_i - u_j|^2 \leq \frac{1}{v_\star^2} \left( \rho_{i,j}^{(a)} - 2\sigma^2 + \frac{C}{\sqrt{T}} \right)$



# Inverting the distance to get a control on $|u_i - u_j|$

- Assume  $v_t, \varepsilon$  are bounded and  $\mathbb{E}[v_t^2] = v_\star^2$  then

- $|\rho_{i,j}^{(a)} - (u_i - u_j)^2 v_\star^2 - 2\sigma^2| \leq \frac{C\sqrt{\log(c/\delta)}}{\sqrt{T}}$  with probability  $1 - \delta$

- Treat  $\delta$  as a constant

- Rearranging terms  $|u_i - u_j|^2 \leq \frac{1}{v_\star^2} \left( \rho_{i,j}^{(a)} - 2\sigma^2 + \frac{C}{\sqrt{T}} \right)$

- So if  $\rho_{i,j}^{(a)} \leq \eta \implies |u_i - u_j| \lesssim \sqrt{\eta - 2\sigma^2} + \frac{C}{T^{1/4}}$  if  $v_\star^2 > 0$ .

**But how many users would satisfy  $\rho_{i,j}^{(a)} \leq \eta$ ?**

- $N_{i,r} \triangleq |\{j \neq i : (u_i - u_j)^2 v_\star^2 \leq r\}|$

- $|\text{User-nbrs}| = |\rho_{i,j}^{(a)} \leq \eta| \geq N_{i,\gamma} \text{ for } \gamma = \eta - 2\sigma^2 - \frac{C\sqrt{\log(c/\delta)}}{\sqrt{T}}$

# But how many users would satisfy $\rho_{i,j}^{(a)} \leq \eta$ ?

- $N_{i,r} \triangleq |\{j \neq i : (u_i - u_j)^2 v_\star^2 \leq r\}|$

- $|\text{User-nbrs}| = |\rho_{i,j}^{(a)} \leq \eta| \geq N_{i,\gamma} \text{ for } \gamma = \eta - 2\sigma^2 - \frac{C\sqrt{\log(c/\delta)}}{\sqrt{T}}$

- Why do we care? Variance

- $|\bar{\epsilon}_t| = \frac{|\sum_{j \in \text{user nbrs}} \epsilon_{j,t}^{(a)}|}{|\text{user nbrs}|} \lesssim \frac{\sigma}{\sqrt{N_{i,\gamma}}}$

# Univariate factors:

## A general error bound for user NN

$$(\hat{\theta}_{i,t,\eta}^{(a)} - \theta_{i,t}^{(a)})^2 \lesssim \underbrace{\frac{1}{v_{\star}^2} \left( \underbrace{\eta - 2\sigma^2}_{\eta'} + \underbrace{\frac{C}{\sqrt{T}}}_{e_T} \right)}_{\text{NN bias due to threshold}} + \underbrace{\frac{\sigma^2}{N_{i,\eta'-e_T}}}_{\text{NN noise variance}}$$

$$N_{i,\gamma} \triangleq |\{j \neq i : (u_i - u_j)^2 v_{\star}^2 \leq \gamma\}|$$

# Univariate factors + constant policy $\mathbb{P}(A_{i,t} = a) = :$ A general error bound for user NN

$$(\hat{\theta}_{i,t,\eta}^{(a)} - \theta_{i,t}^{(a)})^2 \lesssim \frac{1}{v_{\star}^2} \left( \underbrace{\eta - 2\sigma^2}_{\eta'} + \frac{C}{\sqrt{p^2 T}} \right) + \frac{\sigma^2}{p N_{i,\eta' - e_T}}$$

$\eta'$

$e_T$

NN bias  
due to threshold

Error in  
NN distance

NN noise  
variance

$$N_{i,\gamma} \triangleq |\{j \neq i : (u_i - u_j)^2 v_{\star}^2 \leq \gamma\}|$$

Multivariate factors + **learning policy** with  $\mathbb{P}(A_{i,t} = a \mid \text{History}_{t-1}) \geq p$ :  
A general error bound for user NN

$$(\hat{\theta}_{i,t,\eta}^{(a)} - \theta_{i,t}^{(a)})^2 \lesssim \underbrace{\frac{1}{v_{\star}^2} \left( \underbrace{\eta - 2\sigma^2}_{\eta'} + \underbrace{\frac{C}{\sqrt{p^2 T}}}_{e_T} \right)}_{\substack{\text{NN bias} \\ \text{due to threshold}}} + \underbrace{\frac{\sigma^2}{p N_{i,\eta'-e_T}}}_{\substack{\text{NN noise} \\ \text{variance}}} + \underbrace{c_{\text{noise}} \left[ \frac{N_{i,\eta'+e_T} - N_{i,\eta'-e_T}}{p N_{i,\eta'-e_T}} \right]^2}_{\substack{\text{NN bias} \\ \text{inflation due to} \\ \textbf{learning} \text{ policy}}}$$

$$N_{i,\gamma} \triangleq |\{j \neq i : (u_i - u_j)^2 v_{\star}^2 \leq \gamma\}|$$

# Scalings of $N_{i,\gamma}$

- When factors are sampled independently and uniformly from a discrete set  $\{\Delta, \dots, (M-1)\Delta\}$ 
  - $N_{i,r} \geq cN/M$  for any  $r \geq 0$  if  $v_\star > 0$ .
- When factors are sampled independently and uniformly from a continuous set  $[0,1]$ 
  - $N_{i,r} \geq c\sqrt{r/v_\star}$  for any  $r \geq 0$ .
- **HW:** You can now tune  $\eta$  to get refined error bounds.

# Multivariate factors: Bias analysis

- Assume  $v_t, \varepsilon$  are bounded and  $\mathbb{E}[v_t v_t^\top] = \Sigma_v$  then
- $|\rho_{i,j}^{(a)} - (u_i - u_j)^\top \Sigma_v (u_i - u_j) - 2\sigma^2| \leq \frac{C\sqrt{\log(c/\delta)}}{\sqrt{T}}$  with probability  $1 - \delta$
- Treat  $\delta$  as a constant
- Rearranging terms  $\|u_i - u_j\|_2^2 \leq \frac{1}{\lambda_\star^2} \left( \rho_{i,j}^{(a)} - 2\sigma^2 + \frac{C}{\sqrt{T}} \right)$  where  $\lambda_\star = \lambda_{\min}(\Sigma_v)$
- So if  $\rho_{i,j}^{(a)} \leq \eta \implies \|u_i - u_j\|_2 \lesssim \sqrt{\eta - 2\sigma^2} + \frac{C}{T^{1/4}}$  if  $\lambda_\star^2 > 0$ .



# Multivariate factors: Variance analysis

- $N_{i,r} \triangleq |\{j \neq i : (u_i - u_j)^\top \Sigma_v (u_i - u_j) \leq r\}|$
- $|\text{User-nbrs}| = |\rho_{i,j}^{(a)} \leq \eta| \geq N_{i,\gamma} \text{ for } \gamma = \eta - 2\sigma^2 - \frac{C\sqrt{\log(c/\delta)}}{\sqrt{T}}$
- Why do we care? Variance
  - $|\bar{\epsilon}_t| = \frac{|\sum_{j \in \text{user nbrs}} \epsilon_{j,t}^{(a)}|}{|\text{user nbrs}|} \lesssim \frac{\sigma}{\sqrt{N_{i,\gamma}}}$

# Multivariate factors:

A general error bound for user NN when  $A_{i,t}$  is always  $a$

$$(\hat{\theta}_{i,t,\eta}^{(a)} - \theta_{i,t}^{(a)})^2 \lesssim \underbrace{\frac{1}{\lambda_{\star}^2} \left( \underbrace{\eta - 2\sigma^2}_{\eta'} + \underbrace{\frac{C}{\sqrt{T}}}_{e_T} \right)}_{\text{NN bias due to threshold}} + \underbrace{\frac{\sigma^2}{N_{i,\eta'-e_T}}}_{\text{NN noise variance}}$$

$$N_{i,\gamma} \triangleq |\{j \neq i : (u_i - u_j)^\top \Sigma_v (u_i - u_j) \leq \gamma\}|$$

# Multivariate factors + learning policy with exploration $p$ :

## A general error bound for user NN

$$(\hat{\theta}_{i,t,\eta}^{(a)} - \theta_{i,t}^{(a)})^2 \lesssim \frac{1}{\lambda_{\star}^2} \left( \underbrace{\eta}_{\eta'} - 2\sigma^2 + \frac{C}{\sqrt{p^2 T}} \right) + \frac{\sigma^2}{p N_{i,\eta'-e_T}} + c_{\text{noise}} \left[ \frac{N_{i,\eta'+e_T} - N_{i,\eta'-e_T}}{p N_{i,\eta'-e_T}} \right]^2$$

$\eta'$

NN bias  
due to threshold

$e_T$

Error in  
NN distance

NN noise  
variance

NN bias  
inflation due to  
**learning** policy

$$N_{i,\gamma} \triangleq |\{j \neq i : (u_i - u_j)^\top \Sigma_v (u_i - u_j) \leq \gamma\}|$$

# Constant policy: Bias analysis

- Assume  $A_{j,t}$  are iid Bernoulli random variables  $p$  – constant MRT – Like in HeartSteps V1
- Let  $a = 1$ , then what is the distribution of  $B_{i,j,t'} \triangleq \mathbf{1}(A_{i,t'}=A_{j,t'}=a)$ ?

$$\begin{aligned} \bullet \rho_{i,j}^{(a)} &= \frac{\sum_{t'=1}^T (R_{i,t'} - R_{j,t'})^2 \cdot B_{i,j,t'}}{\sum_{t'=1}^T B_{i,j,t'}} \\ &= \frac{\sum_{t'=1}^T \left[ (u_i v_{t'} - u_j v_{t'})^2 + (\varepsilon_{i,t'} - \varepsilon_{j,t'})^2 + 2(u_i v_{t'} - u_j v_{t'})(\varepsilon_{i,t'} - \varepsilon_{j,t'}) \right] \cdot B_{i,j,t'}}{\sum_{t'=1}^T B_{i,j,t'}} \end{aligned}$$

# Bias analysis: The denominator changes

- $$\left| (u_i - u_j)^2 \frac{\sum_{t'=1}^T v_{t'}^2 \cdot B_{i,j,t'}}{\sum_{t'=1}^T B_{i,j,t'}} - (u_i - u_j)^2 \mathbb{E}[v_{t'}^2] \right| \lesssim \frac{c_v^2 \sqrt{\log(1/\delta)}}{\sqrt{T_{i,j}}}$$

- $$\left| \frac{\sum_{t'=1}^T (\varepsilon_{i,t'} - \varepsilon_{j,t'})^2 \cdot B_{i,j,t'}}{\sum_{t'=1}^T B_{i,j,t'}} - 2\sigma^2 \right| \lesssim \frac{c_\varepsilon \sqrt{\log(1/\delta)}}{\sqrt{T_{i,j}}}$$

- $$\left| \frac{T_{i,j}}{T} - p^2 \right| \lesssim \frac{c \sqrt{\log(1/\delta)}}{\sqrt{T}}$$

- A better bound available:  $T_{i,j} \geq cp^2T$  with probability  $\geq 1 - e^{-cp^2T}$ .

# Bias analysis: The denominator changes

- Assume  $v_t, \varepsilon$  are bounded and  $\mathbb{E}[v_t v_t^\top] = \Sigma_v$  then
- Hence  $|\rho_{i,j}^{(a)} - (u_i - u_j)^\top \Sigma_v (u_i - u_j) - 2\sigma^2| \leq \frac{C\sqrt{\log(c/\delta)}}{\sqrt{p^2 T}}$  with probability  $1 - \delta$
- Treat  $\delta$  as a constant
- Rearranging terms  $\|u_i - u_j\|_2^2 \leq \frac{1}{\lambda_\star^2} \left( \rho_{i,j}^{(a)} - 2\sigma^2 + \frac{C}{\sqrt{p^2 T}} \right)$  where  $\lambda_\star = \lambda_{\min}(\Sigma_v)$
- So if  $\rho_{i,j}^{(a)} \leq \eta \implies \|u_i - u_j\|_2 \lesssim \sqrt{\eta - 2\sigma^2} + \frac{C}{p^{1/2} T^{1/4}}$  if  $\lambda_\star^2 > 0$ .

# Constant policy: Variance analysis: denominator changes

- $N_{i,r} \triangleq |\{j \neq i : (u_i - u_j)^\top \Sigma_v (u_i - u_j) \leq r\}|$
- $|\text{User-nbrs}| = |\rho_{i,j}^{(a)} \leq \eta| \geq N_{i,\gamma} \text{ for } \gamma = \eta - 2\sigma^2 - \frac{C\sqrt{\log(c/\delta)}}{\sqrt{p^2 T}}$
- Why do we care? Variance
  - $|\bar{\epsilon}_t| = \frac{|\sum_{j \in \text{user nbrs}} \epsilon_{j,t}^{(a)} \mathbf{1}(A_{j,t} = a)|}{|\text{user nbrs with } A_{j,t} = a|} \lesssim \frac{\sigma}{\sqrt{p N_{i,\gamma}}}$

# Multivariate factors + constant policy:

## A general error bound for user NN

$$(\hat{\theta}_{i,t,\eta}^{(a)} - \theta_{i,t}^{(a)})^2 \lesssim \frac{1}{\lambda_{\star}^2} \left( \underbrace{\eta - 2\sigma^2}_{\eta'} + \frac{C}{\sqrt{p^2 T}} \right) + \frac{\sigma^2}{p N_{i,\eta' - e_T}}$$

$\eta'$

$e_T$

NN bias  
due to threshold

Error in  
NN distance

NN noise  
variance

$$N_{i,\gamma} \triangleq |\{j \neq i : (u_i - u_j)^\top \Sigma_v (u_i - u_j) \leq \gamma\}|$$



# Multivariate factors + learning policy with exploration $p$ :

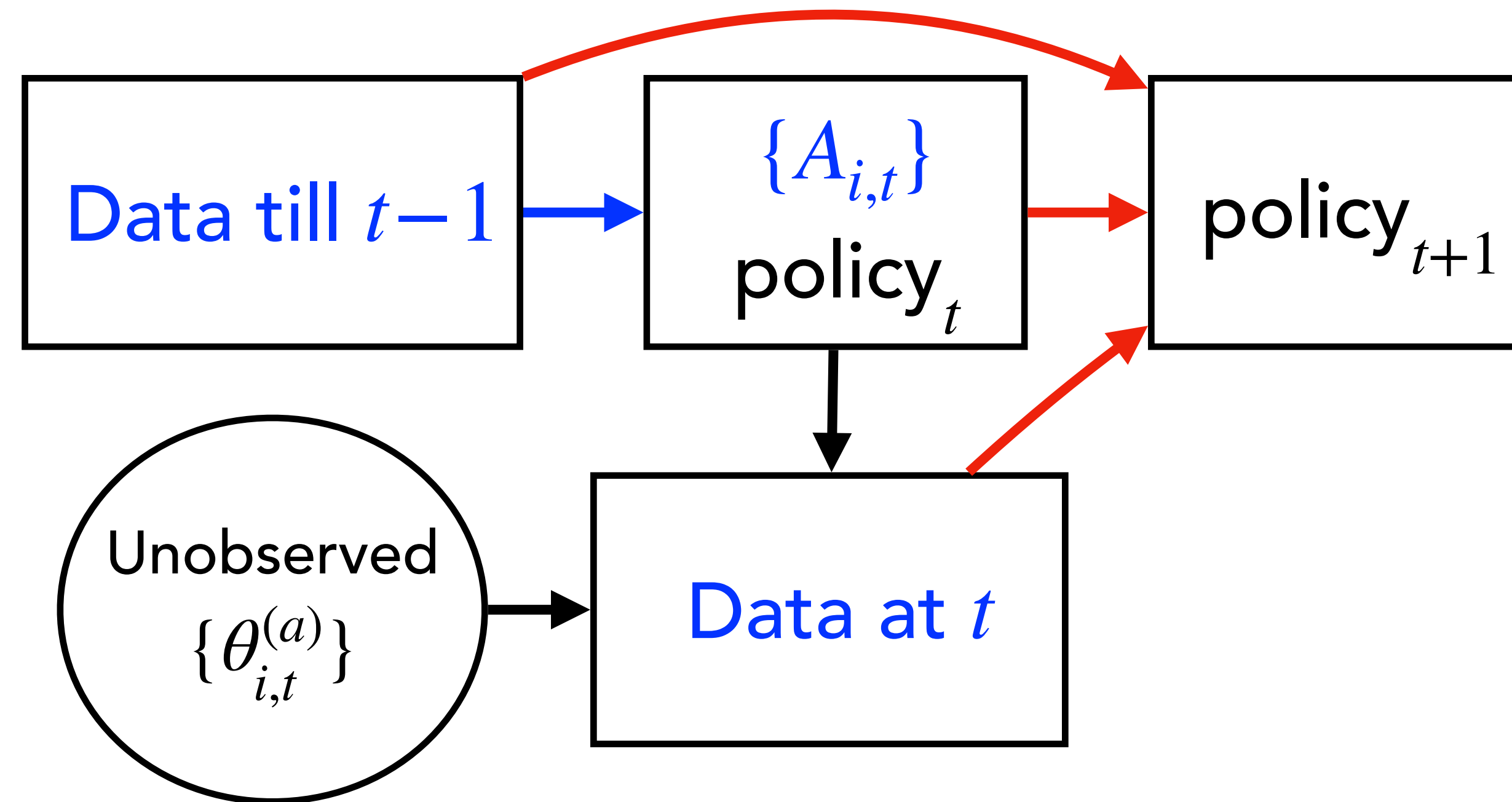
## A general error bound for user NN

$$(\hat{\theta}_{i,t,\eta}^{(a)} - \theta_{i,t}^{(a)})^2 \lesssim \frac{1}{\lambda_{\star}^2} \left( \underbrace{\eta - 2\sigma^2}_{\substack{\eta' \\ \text{NN bias} \\ \text{due to threshold}}} + \underbrace{\frac{C}{\sqrt{p^2 T}}}_{\substack{e_T \\ \text{Error in} \\ \text{NN distance}}} \right) + \underbrace{\frac{\sigma^2}{p N_{i,\eta'-e_T}}}_{\substack{\text{NN noise} \\ \text{variance}}} + \underbrace{c_{\text{noise}} \left[ \frac{N_{i,\eta'+e_T} - N_{i,\eta'-e_T}}{p N_{i,\eta'-e_T}} \right]^2}_{\substack{\text{NN bias} \\ \text{inflation due to} \\ \text{learning policy}}}$$

$$N_{i,\gamma} \triangleq |\{j \neq i : (u_i - u_j)^\top \Sigma_v (u_i - u_j) \leq \gamma\}|$$

# Learning policy:

## Sequential dependence between observations



# Learning policy: Bias analysis

## Similar except now with Martingales

- Still goes through using “Azuma-Hoeffding bounds” and careful Martingale construction

- $$\left| (u_i - u_j)^2 \frac{\sum_{t'=1}^T v_{t'}^2 \cdot B_{i,j,t'}}{\sum_{t'=1}^T B_{i,j,t'}} - (u_i - u_j)^2 \mathbb{E}[v_{t'}^2] \right| \lesssim \frac{c_v^2 \sqrt{\log(1/\delta)}}{\sqrt{T_{i,j}}}$$

- $$\left| \frac{\sum_{t'=1}^T (\varepsilon_{i,t'} - \varepsilon_{j,t'})^2 \cdot B_{i,j,t'}}{\sum_{t'=1}^T B_{i,j,t'}} - 2\sigma^2 \right| \lesssim \frac{c_\varepsilon \sqrt{\log(1/\delta)}}{\sqrt{T_{i,j}}}$$

- $$\left| \frac{T_{i,j}}{T} - p^2 \right| \lesssim \frac{c \sqrt{\log(1/\delta)}}{\sqrt{T}}$$

- A better bound available:  $T_{i,j} \geq cp^2T$  with probability  $\geq 1 - e^{-cp^2T}$ .

# Learning policy: Bias bounds

## Essentially same as the MRT bound

- If  $\rho_{i,j}^{(a)} \leq \eta \quad \Rightarrow \quad \|u_i - u_j\|_2 \lesssim \frac{1}{\lambda_\star} (\sqrt{\eta - 2\sigma^2} + \frac{C}{p^{1/2}T^{1/4}})$  if  $\lambda_\star^2 > 0$ .

# Learning policy: Variance analysis — Non-trivial changes

- $N_{i,r} \triangleq |\{j \neq i : (u_i - u_j)^\top \Sigma_v (u_i - u_j) \leq r\}|$

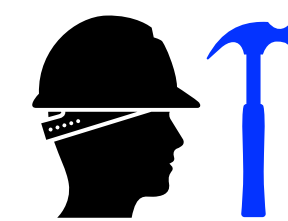
- $|\text{User-nbrs}| = |\rho_{i,j}^{(a)} \leq \eta| \geq N_{i,\gamma} \text{ for } \gamma = \eta - 2\sigma^2 - \frac{C\sqrt{\log(c/\delta)}}{\sqrt{p^2 T}}$

- Why do we care? Variance

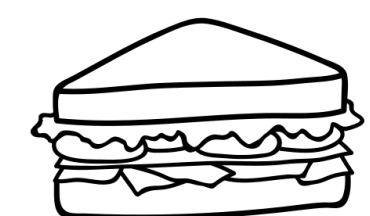
- $|\bar{\epsilon}_t| = \frac{|\sum_{j \in \text{user nbrs}} \epsilon_{j,t}^{(a)} \mathbf{1}(A_{j,t} = a)|}{|\text{user nbrs with } A_{j,t} = a|}$



noise at  $t$  **correlated** with user neighbors (**learning policy**)



Martingale concentration, **new sandwich argument** for NN



# Learning policy: Variance bounds — Has a “bias” like term

- $$\bar{\varepsilon}_t^2 = \left( \frac{|\sum_{j \in \text{user nbrs}} \varepsilon_{j,t}^{(a)} \mathbf{1}(A_{j,t} = a)|}{|\text{user nbrs with } A_{j,t} = a|} \right)^2$$

Martingale concentration, **new sandwich argument** for NN

$$\lesssim \frac{\sigma^2}{p N_{i,\eta'-e_T}} + c_{\text{noise}} \left[ \frac{N_{i,\eta'+e_T} - N_{i,\eta'-e_T}}{p N_{i,\eta'-e_T}} \right]^2$$

# Multivariate factors + learning policy with exploration $p$ :

## A general error bound for user NN

$$(\hat{\theta}_{i,t,\eta}^{(a)} - \theta_{i,t}^{(a)})^2 \lesssim \frac{1}{\lambda_{\star}^2} \left( \underbrace{\eta}_{\eta'} - 2\sigma^2 + \frac{C}{\sqrt{p^2 T}} \right) + \frac{\sigma^2}{p N_{i,\eta'-e_T}} + c_{\text{noise}} \left[ \frac{N_{i,\eta'+e_T} - N_{i,\eta'-e_T}}{p N_{i,\eta'-e_T}} \right]^2$$

$\eta'$

NN bias  
due to threshold

$e_T$

Error in  
NN distance

NN noise  
variance

NN bias  
inflation due to  
**learning** policy

$$N_{i,\gamma} \triangleq |\{j \neq i : (u_i - u_j)^\top \Sigma_v (u_i - u_j) \leq \gamma\}|$$