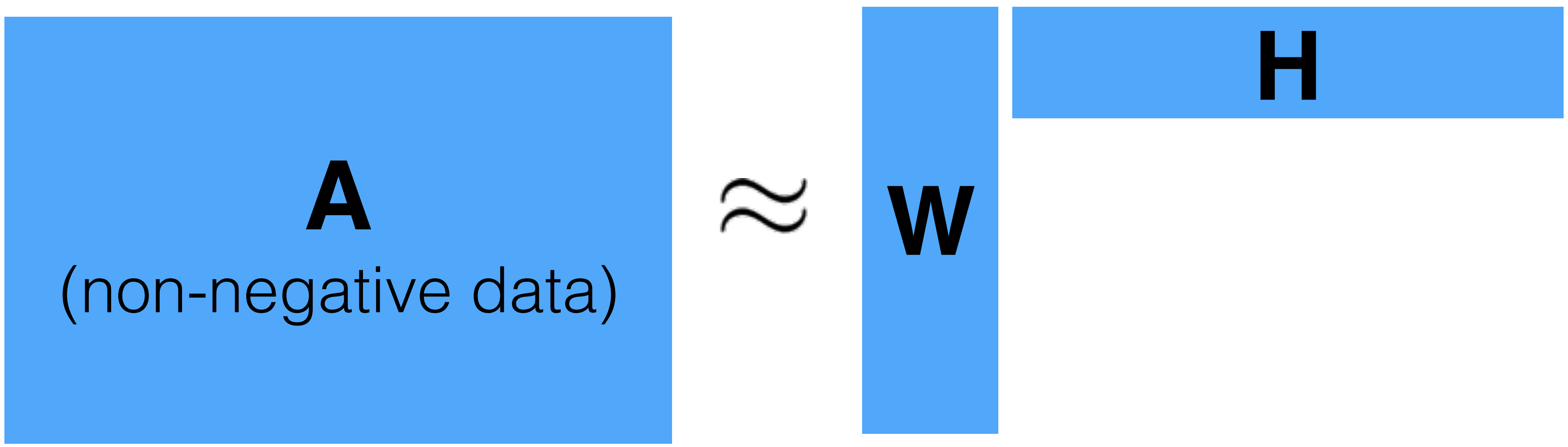
Non-Negative Matrix Factorization

$$\underset{W,H}{\text{minimize}} \quad \|A - WH\|_F^2 \qquad A \in \mathbf{R}^{n \times p}$$

$$W \in \mathbf{R}^{n \times r}$$

$$\text{subject to} \quad W \geq 0, \; H \geq 0 \qquad H \in \mathbf{R}^{r \times p}$$



Common Technique: Alternating Minimization

while convergence criteria not satisfied **do**

$$W \leftarrow \arg\min_{\hat{W} \geq 0} \|A - \hat{W}H\|_F^2$$

$$H \leftarrow \arg\min_{\hat{H} \geq 0} \|A - W\hat{H}\|_F^2$$

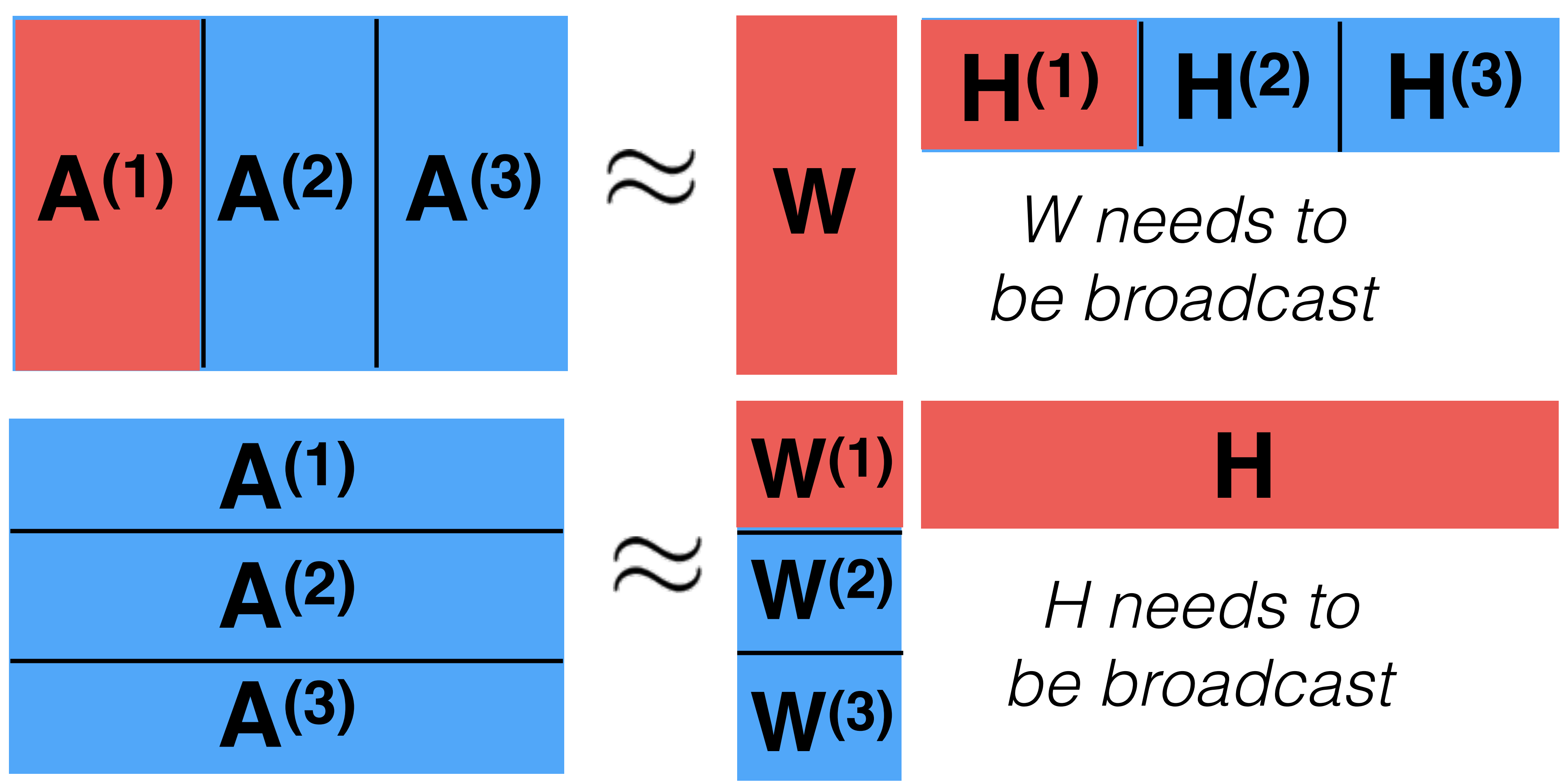
<u>Challenge:</u> How to distribute?

Optimizing H parallelizes *across <u>columns</u>* of A

Optimizing W parallelizes *across <u>rows</u>* of A

<u>Naive Algorithm:</u>

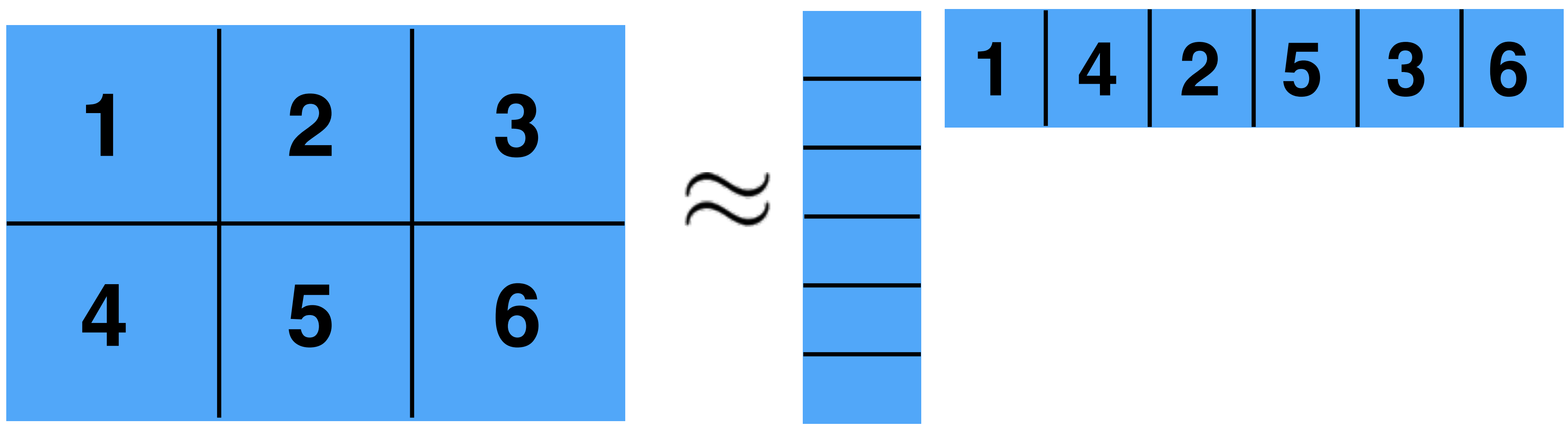
Partition A in blocks row-wise and col-wise



Improved Algorithm: Kannan et al. (2016)

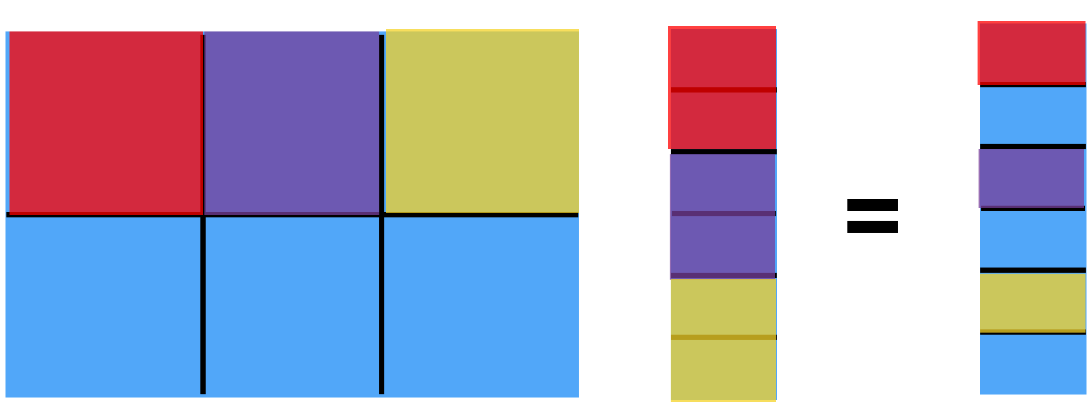
$$\text{minimize}_{\hat{W}} \; \|A - \hat{W}H\|_F^2 \iff \text{minimize}_{\hat{W}} \; \|AH^T - \hat{W}HH^T\|_F^2$$

Partition A in a grid, partition W and H as before

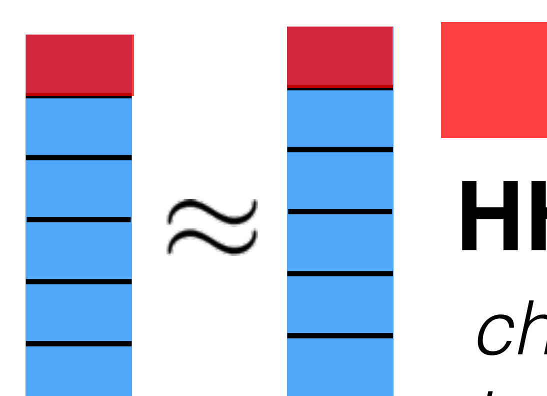


Compute AH^T

gather -> reduce -> scatter



Update W



cheap to broadcast HH^T