PARALLEL NON-NEGATIVE CP DECOMPOSITION OF DENSE TENSORS

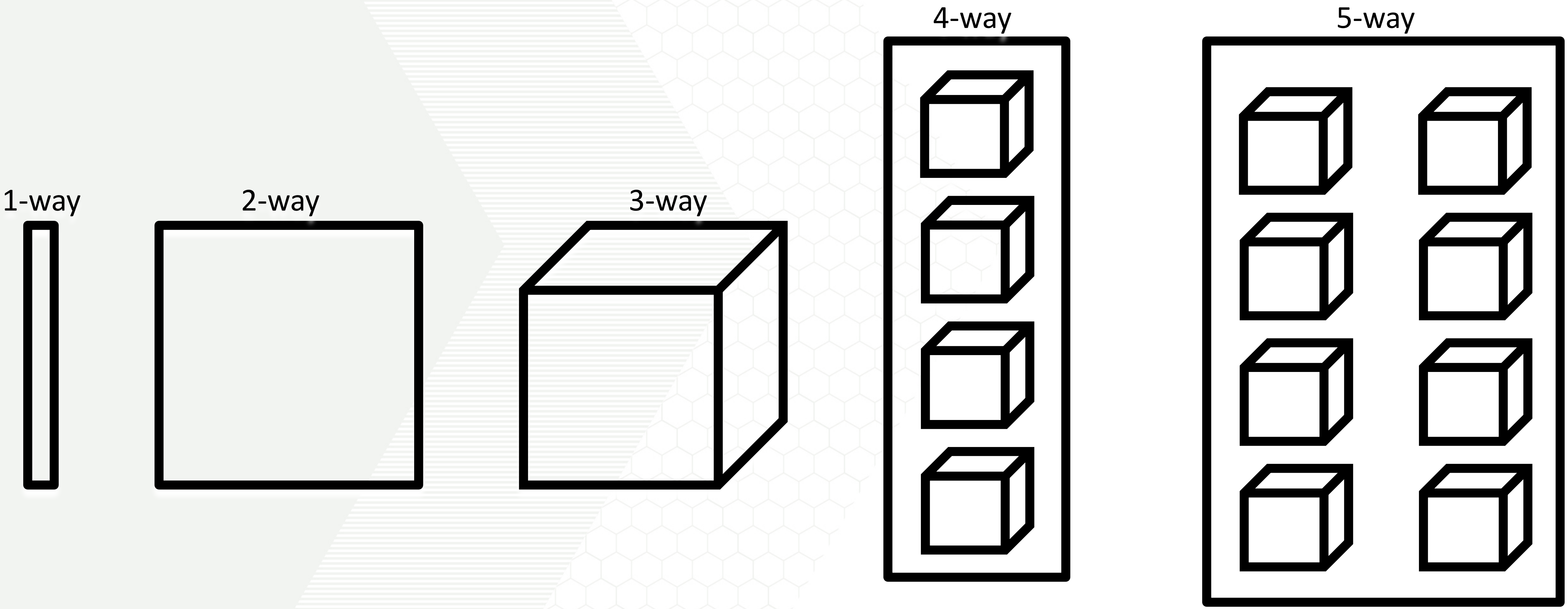
KOBY HAYASHI[1,2], GREY BALLARD[2], RAMKI KANNAN[3], SRINIVAS ESWAR[1]

[1]GEORGIA TECH CSE

[2]WAKE FOREST UNIVERSITY

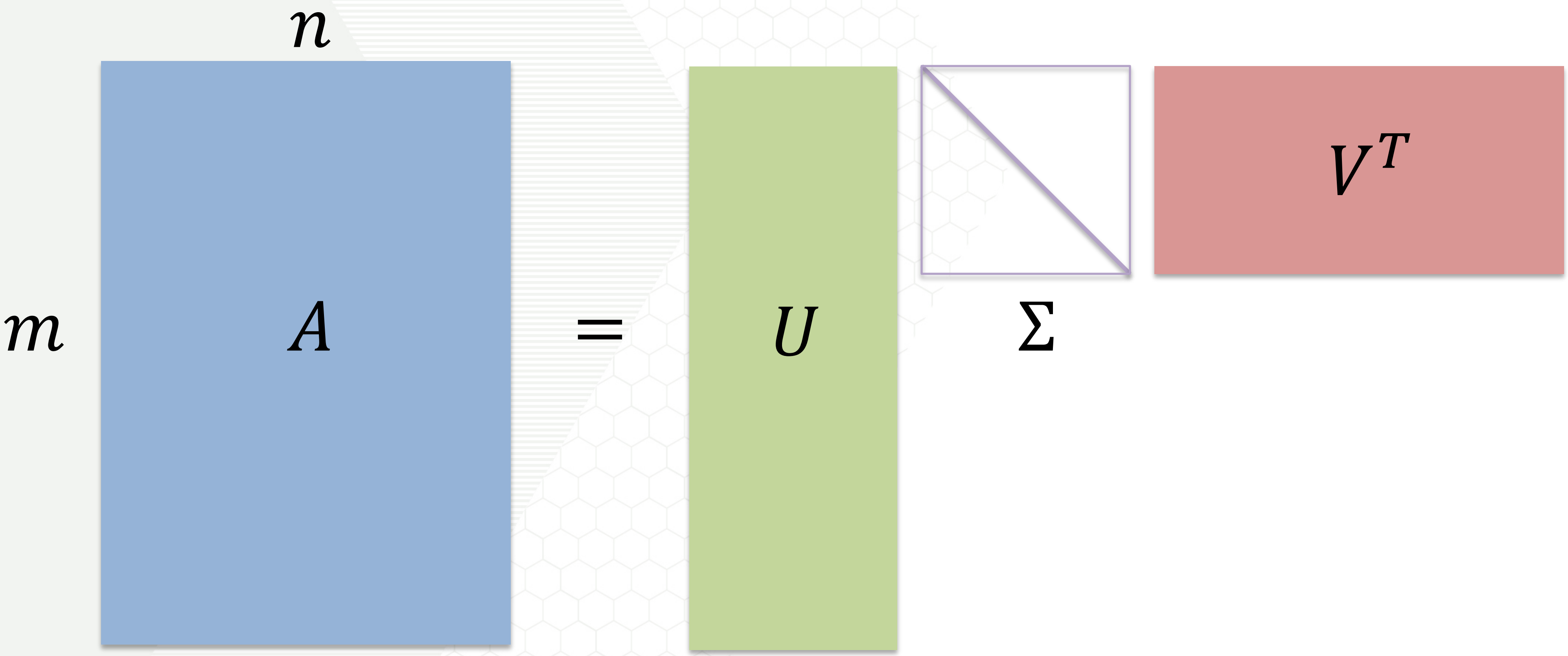
[3]OAK RIDGE NATIONAL LABS

INTRODUCTION TO TENSORS



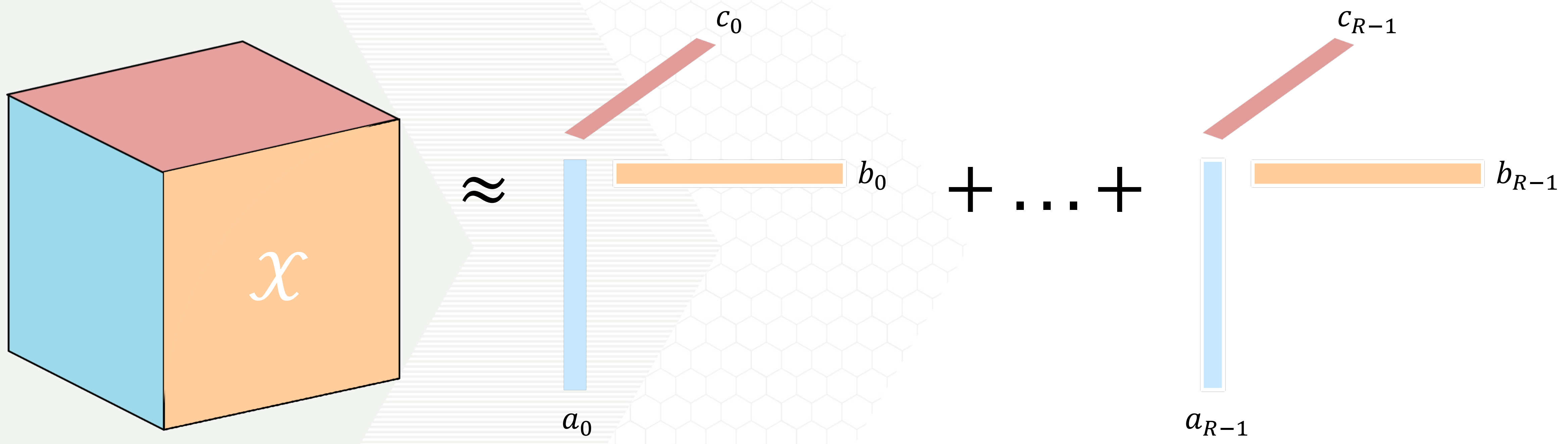
REVIEW OF THE SVD

- Singular Value Decomposition (SVD):
  - $A = U\Sigma V^T = \sum_{r=1}^{R} \sigma_r (u_r)(v_r)^T$



THE CP DECOMPOSITION

Canonical Polyadic Decomposition (CP)



$$\mathcal{X} \approx \sum_{r=0}^{R-1} a_{ir} \circ b_{jr} \circ c_{kr} = [\![A, B, C]\!]$$

CP VIA ALTERNATING LEAST SQUARES

Algorithm 5 CP_ALS

Require: $\mathcal{X}$ is an N-way tensor with dimensions $I_0 \times I_1 \times \cdots \times I_{N-1}$, $n \in [N]$, $\mathbf{U}_{(n)}$ is the n^{th} factor matrix, and a rank R

function $\mathcal{Y} = \text{CP\_ALS}(\mathcal{X}, R)$

  while stopping conditions not met **do**

    for $n \in [N]$ **do**

      $\mathbf{H} = \mathbf{U}_{(0)}^{\mathsf{T}}\mathbf{U}_{(0)} * \cdots * \mathbf{U}_{(n-1)}^{\mathsf{T}}\mathbf{U}_{(n-1)} * \mathbf{U}_{(n+1)}^{\mathsf{T}}\mathbf{U}_{(n+1)} * \cdots * \mathbf{U}_{(N-1)}^{\mathsf{T}}\mathbf{U}_{(N-1)}$

      $\mathbf{M} = \mathbf{X}_{(n)}(\mathbf{U}_{(N-1)} \odot \cdots \odot \mathbf{U}_{(n+1)} \odot \mathbf{U}_{n-1} \odot \cdots \odot \mathbf{U}_{(0)})^{\mathsf{T}}$

      solve $\mathbf{U}_{(n)} = \mathbf{M}\mathbf{H}^{\dagger}$ ▷ for nonnegativity enforce $\mathbf{U}_{(n)} > 0$

    end for

  end while

end function

Ensure: $\mathcal{Y}$ is a rank R CP Model

H MATRIX

Georgia Tech

Algorithm 5 CP_ALS

Require: $\mathcal{X}$ is an N-way tensor with dimensions $I_0 \times I_1 \times \cdots \times I_{N-1}$, $n \in [N]$, $\mathbf{U}_{(n)}$ is the n^{th} factor matrix, and a rank R

function $\mathcal{Y} = \text{CP\_ALS}(\mathcal{X}, R)$
  while stopping conditions not met **do**
    for $n \in [N]$ **do**
      $\mathbf{H} = \mathbf{U}_{(0)}^\intercal \mathbf{U}_{(0)} * \cdots * \mathbf{U}_{(n-1)}^\intercal \mathbf{U}_{(n-1)} * \mathbf{U}_{(n+1)}^\intercal \mathbf{U}_{(n+1)} * \cdots * \mathbf{U}_{(N-1)}^\intercal \mathbf{U}_{(N-1)}$
      $\mathbf{M} = \mathbf{X}_{(n)} (\mathbf{U}_{(N-1)} \odot \cdots \odot \mathbf{U}_{(n+1)} \odot \mathbf{U}_{n-1} \odot \cdots \odot \mathbf{U}_{(0)})^\intercal$
      solve $\mathbf{U}_{(n)} = \mathbf{M}\mathbf{H}^\dagger$ ▷ for nonnegativity enforce $\mathbf{U}_{(n)} > 0$
    end for
  end while
end function

Ensure: $\mathcal{Y}$ is a rank R CP Model

MATRICIZED TENSOR TIMES KHATRI RAO PRODUCT

Algorithm 5 CP_ALS

Require: $\mathcal{X}$ is an N-way tensor with dimensions $I_0 \times I_1 \times \cdots \times I_{N-1}$, $n \in [N]$, $\mathbf{U}_{(n)}$ is the n^{th} factor matrix, and a rank R

function $\mathcal{Y} = \text{CP\_ALS}(\mathcal{X}, R)$
    while stopping conditions not met **do**
        for $n \in [N]$ **do**
            $\mathbf{H} = \mathbf{U}_{(0)}^{\mathsf{T}}\mathbf{U}_{(0)} * \cdots * \mathbf{U}_{(n-1)}^{\mathsf{T}}\mathbf{U}_{(n-1)} * \mathbf{U}_{(n+1)}^{\mathsf{T}}\mathbf{U}_{(n+1)} * \cdots * \mathbf{U}_{(N-1)}^{\mathsf{T}}\mathbf{U}_{(N-1)}$
            $\mathbf{M} = \mathbf{X}_{(n)}(\mathbf{U}_{(N-1)} \odot \cdots \odot \mathbf{U}_{(n+1)} \odot \mathbf{U}_{n-1} \odot \cdots \odot \mathbf{U}_{(0)})^{\mathsf{T}}$
            solve $\mathbf{U}_{(n)} = \mathbf{M}\mathbf{H}^{\dagger}$ ▷ for nonnegativity enforce $\mathbf{U}_{(n)} > 0$
        end for
    end while
end function

Ensure: $\mathcal{Y}$ is a rank R CP Model

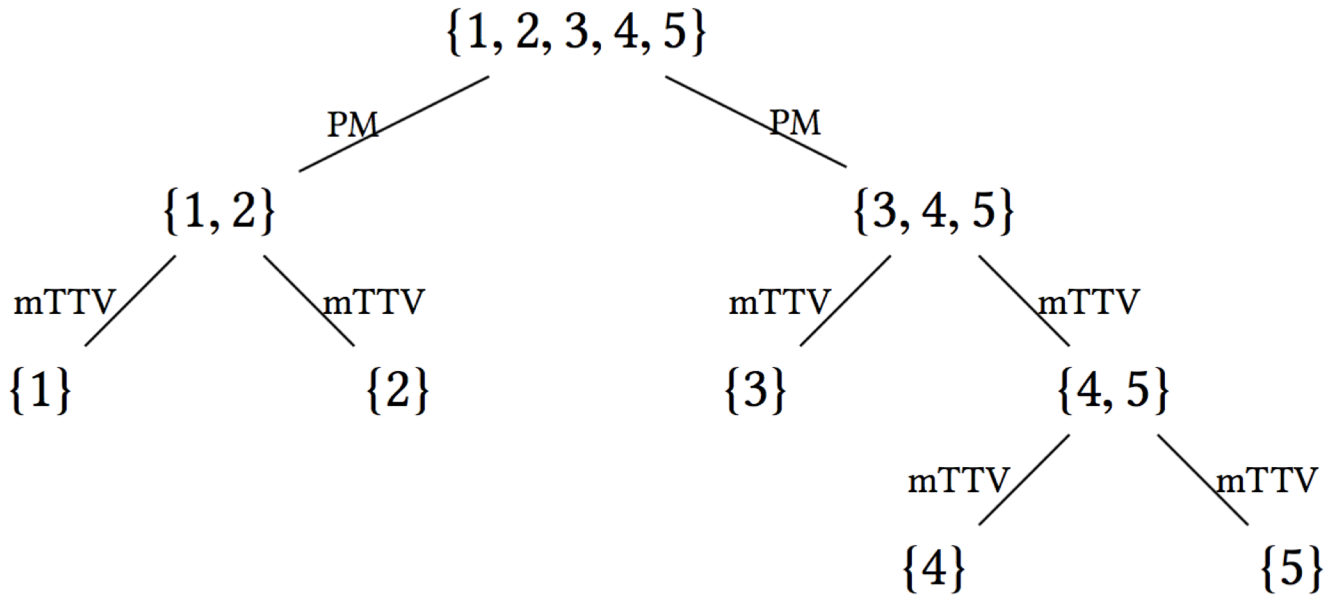
MTTKRP

Dimension Tree Optimization for N-MTTKRPS

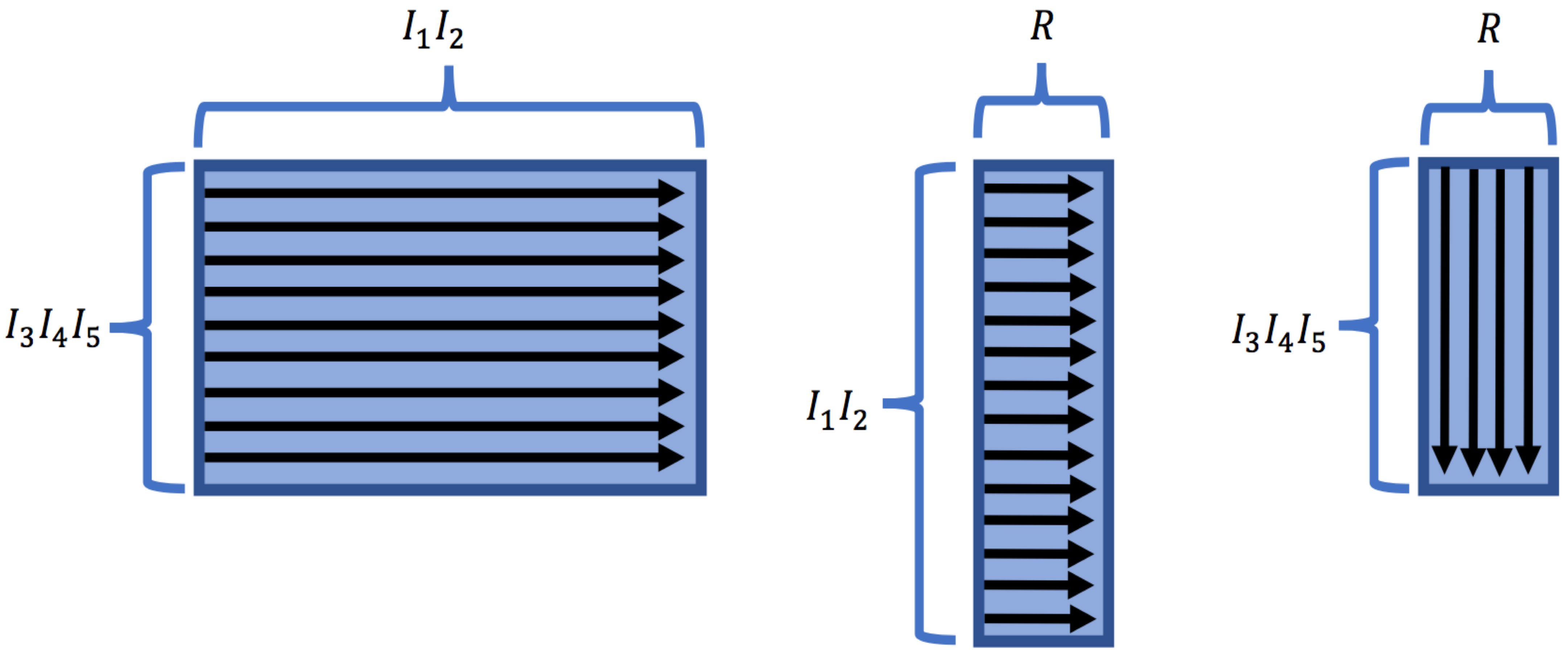
- Node = set of matrix and tensor
- Edge = computation
- Leaf Node = completed MTTKRP

Translations

- PM = Partial MTTKRP = GEMM
- mTTV = multiple Tensor Times Vector = multiple GEMV

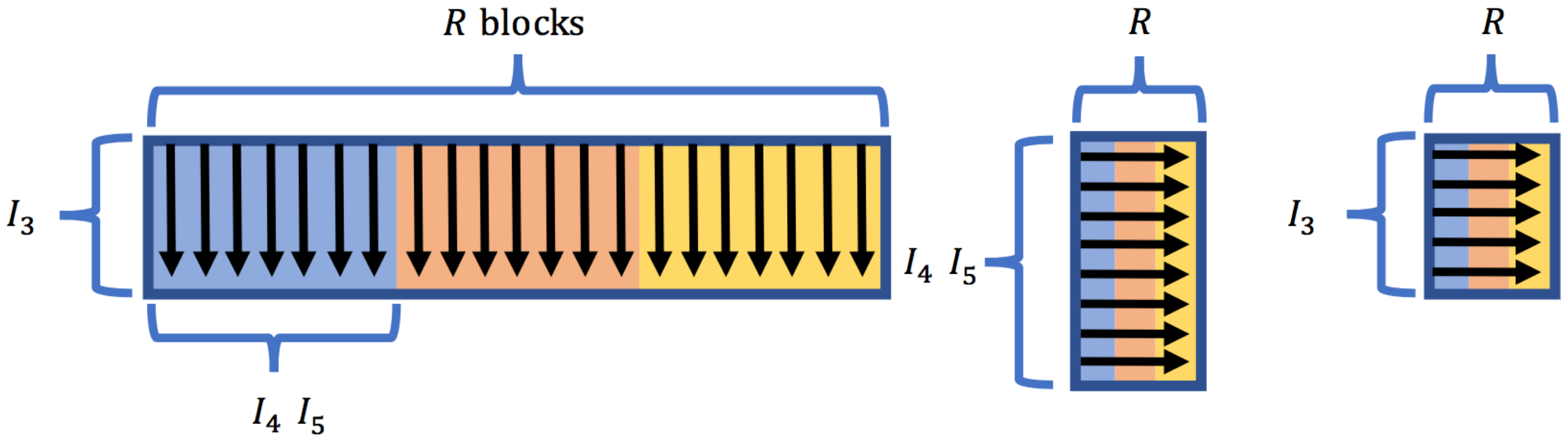


PARTIAL MTTKRP = GEMM



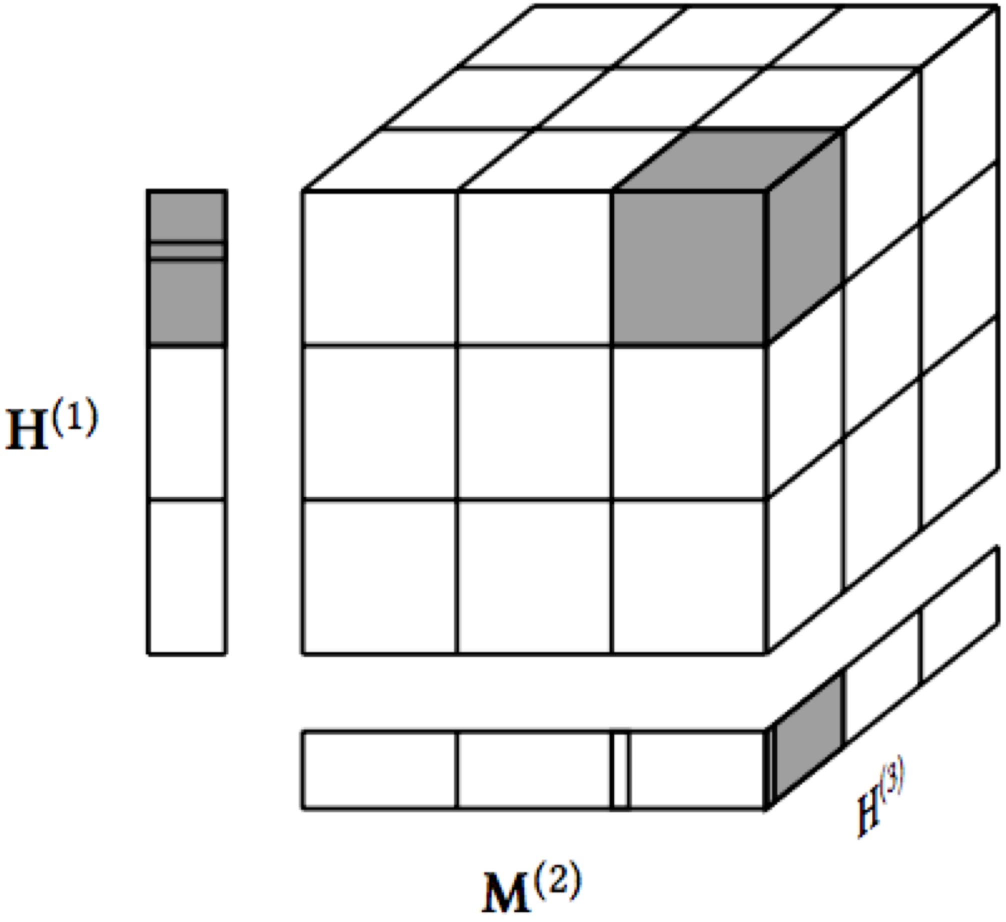
$$X^T_{(1:2)} \cdot K_{1:2} = T_{(1:3)}$$

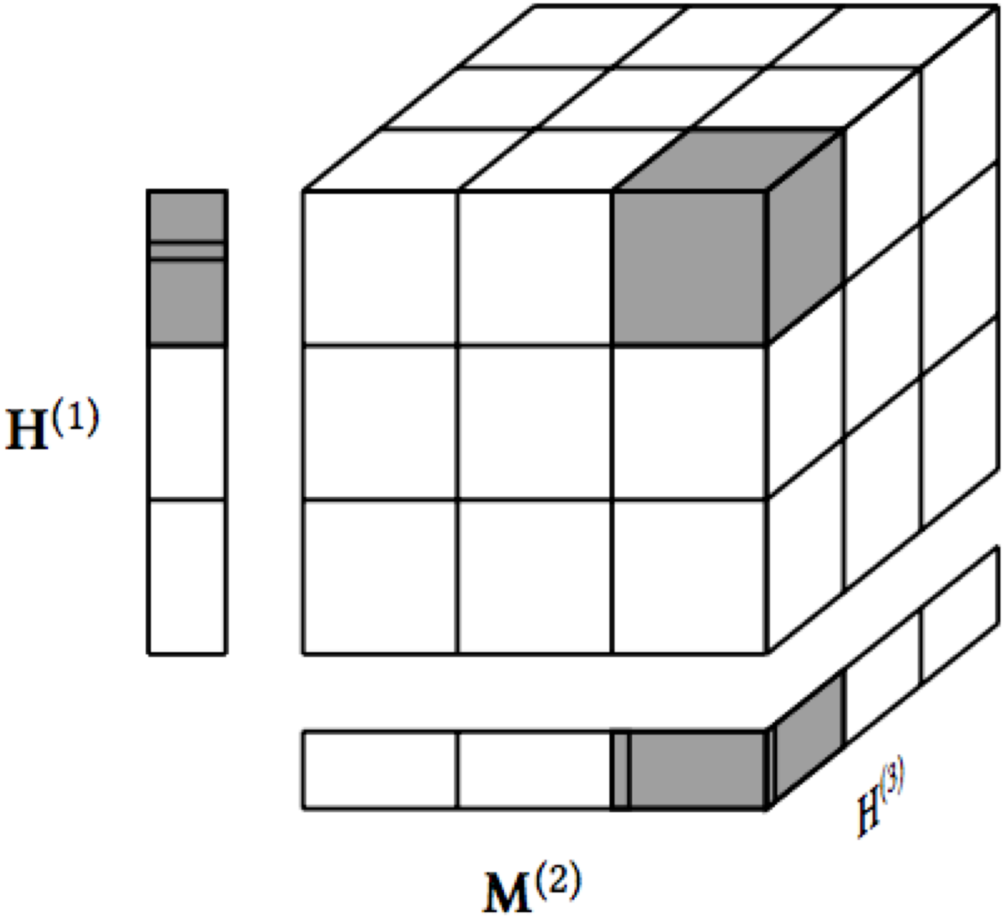
MULTIPLE TENSOR TIMES VECTOR = GEMV



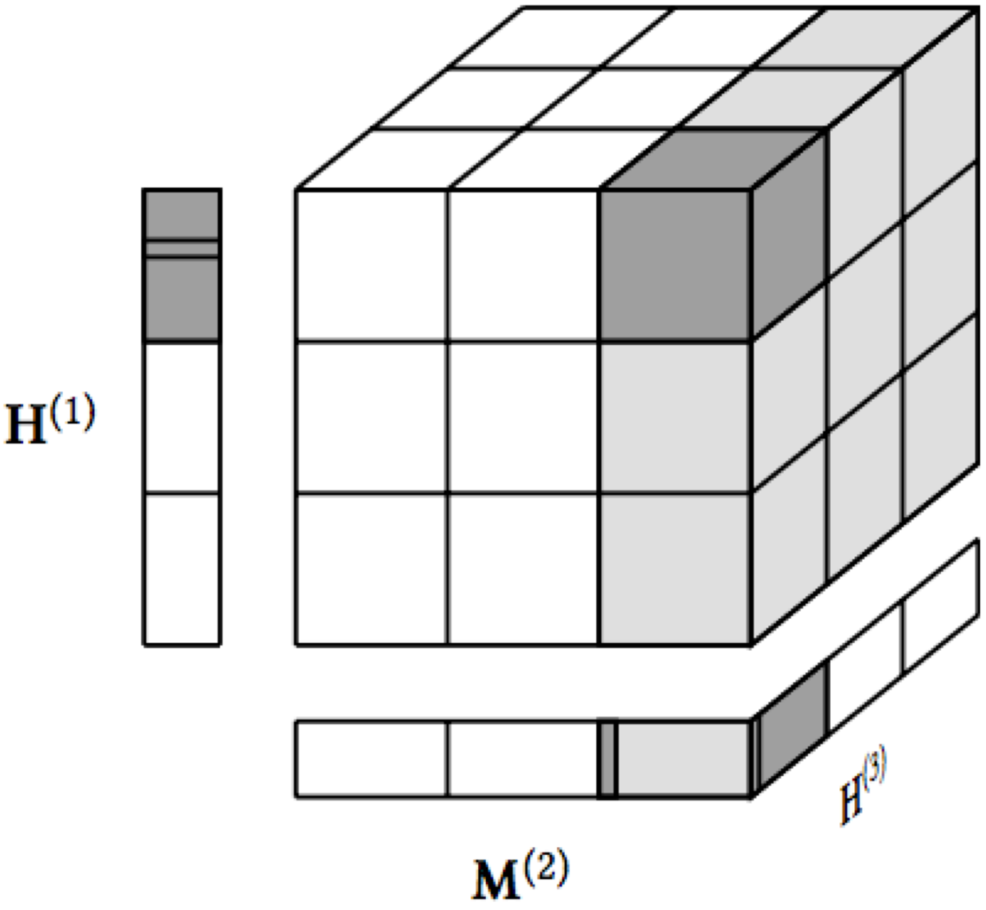
$$T_{(1)}[r] \quad \cdot \quad K_{4:5}(:,r) = M^{(3)}(:,r)$$

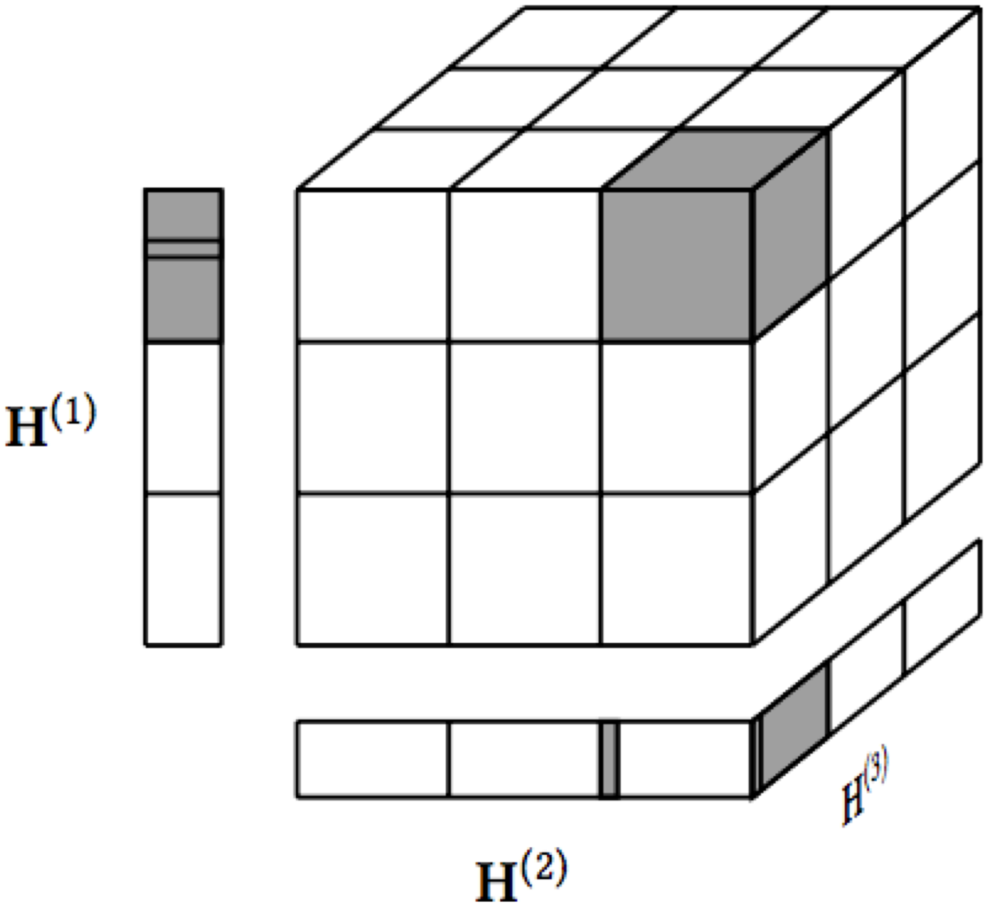
DISTRIBUTED MODEL



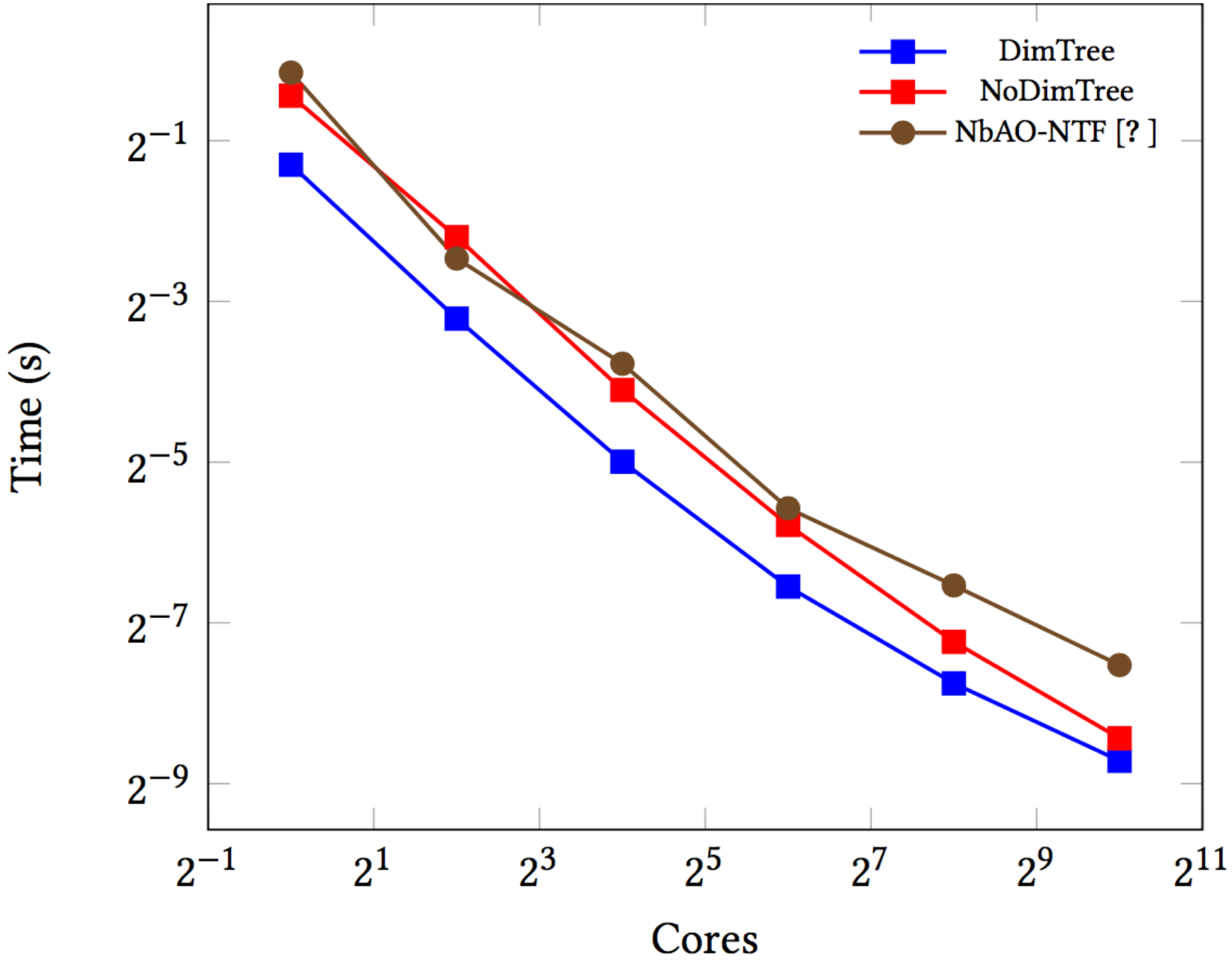


DISTRIBUTED MODEL





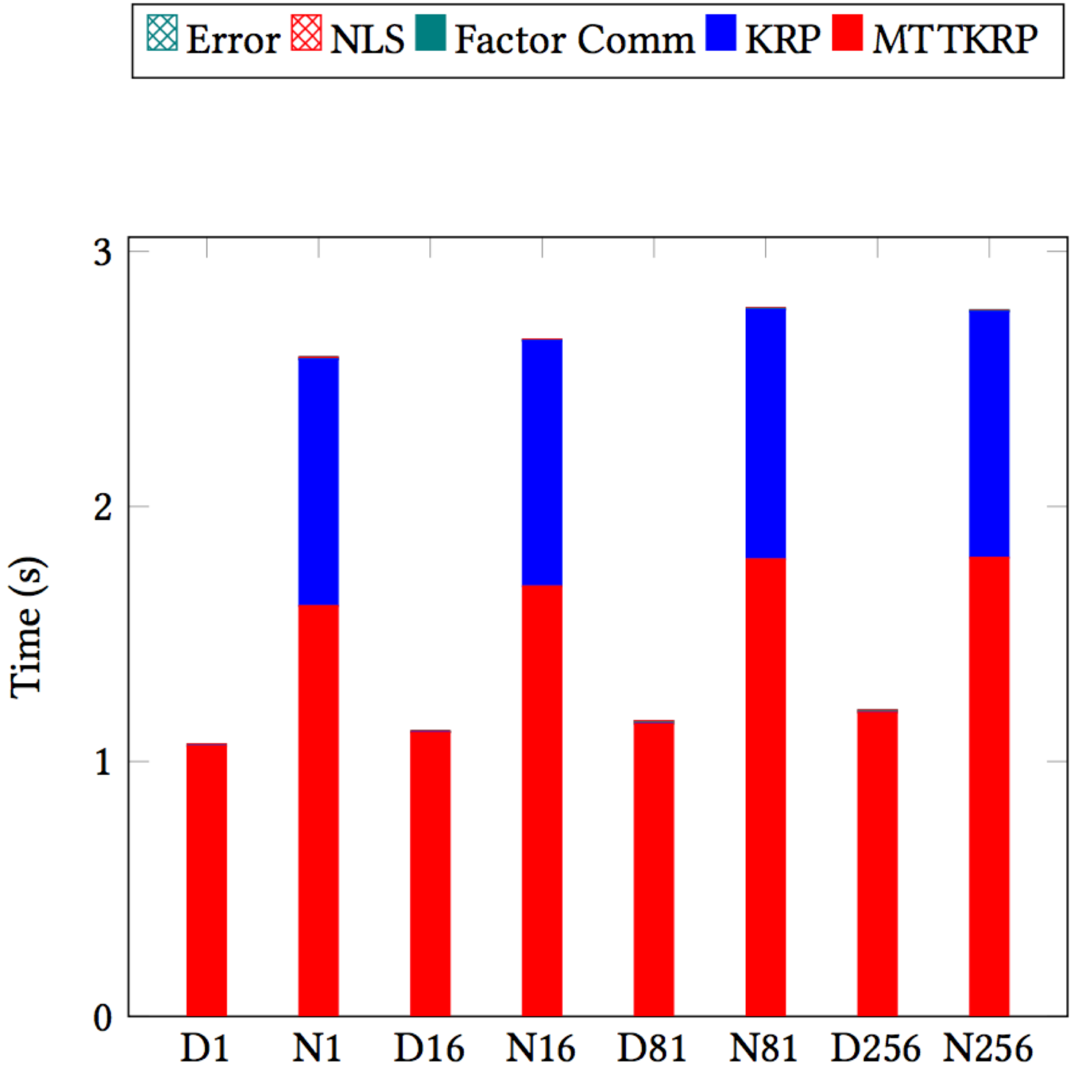
STRONG SCALING



- Tensor = 1023×1344×33
- Proc grid = p×p×1
- $p = \{1,2,4,8,16,32\}$

WEAK SCALING

- 4-way tensor
- Low Rank = 32
- $p \times p \times p \times p : p = \{1,2,3,4\}$
- Tensor is $128p \times 128p \times 128p \times 128p$



LINEAR SYSTEM SOLVE

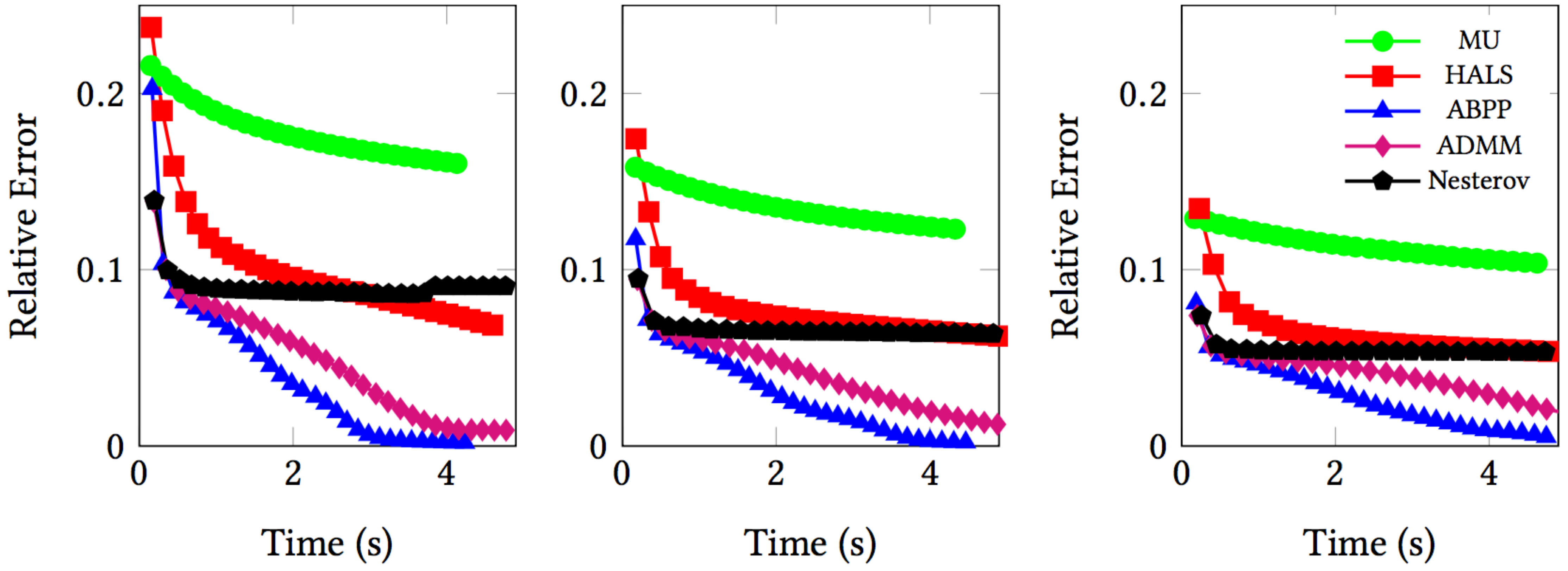
Algorithm 5 CP_ALS

Require: $\mathcal{X}$ is an N-way tensor with dimensions $I_0 \times I_1 \times \cdots \times I_{N-1}$, $n \in [N]$, $\mathbf{U}_{(n)}$ is the n^{th} factor matrix, and a rank R

function $\mathcal{Y} = \text{CP\_ALS}(\mathcal{X}, R)$
    while stopping conditions not met **do**
        for $n \in [N]$ **do**
            $\mathbf{H} = \mathbf{U}_{(0)}^{\mathsf{T}}\mathbf{U}_{(0)} * \cdots * \mathbf{U}_{(n-1)}^{\mathsf{T}}\mathbf{U}_{(n-1)} * \mathbf{U}_{(n+1)}^{\mathsf{T}}\mathbf{U}_{(n+1)} * \cdots * \mathbf{U}_{(N-1)}^{\mathsf{T}}\mathbf{U}_{(N-1)}$
            $\mathbf{M} = \mathbf{X}_{(n)}(\mathbf{U}_{(N-1)} \odot \cdots \odot \mathbf{U}_{(n+1)} \odot \mathbf{U}_{n-1} \odot \cdots \odot \mathbf{U}_{(0)})^{\mathsf{T}}$
            solve $\mathbf{U}_{(n)} = \mathbf{M}\mathbf{H}^{\dagger}$ ▷ for nonnegativity enforce $\mathbf{U}_{(n)} > 0$
        end for
    end while
end function

Ensure: $\mathcal{Y}$ is a rank R CP Model

LINEAR SOLVE TESTS



(a) Low rank $k = 32$ (b) Low rank $k = 64$ (c) Low rank $k = 96$

END

- Manipulating BLAS for tensor operations is clunky
- What is the right way to think about/present tensor operations?
  - MTTKRP, tensor contraction with low rank tensor? Matrix Multiply?

Parallel Nonnegative CP Decomposition of Dense Tensors. 25th IEEE In- ternational Conference on High Performance Computing, Data, and Analytics, HiPC 2018, with Grey Ballard and Ramakrishnan Kannan.