# Multiresolution Low-rank Tensor Formats

May 20, 2021

Oscar Mickelin

oscarmi@mit.edu

# Multiresolution Low-rank Tensor Formats

Mickelin, Karaman, SIAM Journal on Matrix Analysis and Applications, 41(3), 1086-1114

- Compress tensors T with multiple length-scales
- Idea: represent T as sum of compressed tensors down-sampled onto different grid length-scales

# Multiresolution Low-rank Tensor Formats

- Compress tensors T with multiple length-scales
- Idea: represent T as sum of compressed tensors down-sampled onto different grid length-scales

**Notation**

Operation to go up $\ell$ length-scales: $\mathrm{ext}_\ell(T)$

Operation to go down $\ell$ length-scales: $\mathrm{ave}_\ell(T)$

$$\mathrm{ext}_\ell(T)(i_1,\dots,i_d) = T\left(\left\lfloor \frac{i_1-1}{b_s^\ell}\right\rfloor + 1, \dots, \left\lfloor \frac{i_d-1}{b_s^\ell}\right\rfloor + 1\right)$$

$$\mathrm{ave}_\ell(S)(i_1,\dots,i_d) = \frac{1}{b_s^{\ell d}} \sum_{j_1=0}^{b_s^\ell-1} \cdots \sum_{j_d=0}^{b_s^\ell-1} S(i_1+j_1,\dots,i_d+j_d)$$

Example

$$T = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad S = \left[\underbrace{\begin{matrix} a & b \\ c & d \end{matrix}}_{S(:,:,1)} \middle| \underbrace{\begin{matrix} e & f \\ g & h \end{matrix}}_{S(:,:,2)}\right],$$

$$\mathrm{ext}_1(T) = \begin{bmatrix} a & a & b & b \\ a & a & b & b \\ c & c & d & d \\ c & c & d & d \end{bmatrix}, \quad \mathrm{ext}_1(S) = \left[\begin{array}{cccc|cccc|cccc|cccc} a & a & b & b & a & a & b & b & e & e & f & f & e & e & f & f \\ a & a & b & b & a & a & b & b & e & e & f & f & e & e & f & f \\ c & c & d & d & c & c & d & d & g & g & h & h & g & g & h & h \\ c & c & d & d & c & c & d & d & g & g & h & h & g & g & h & h \end{array}\right]$$

## Definition

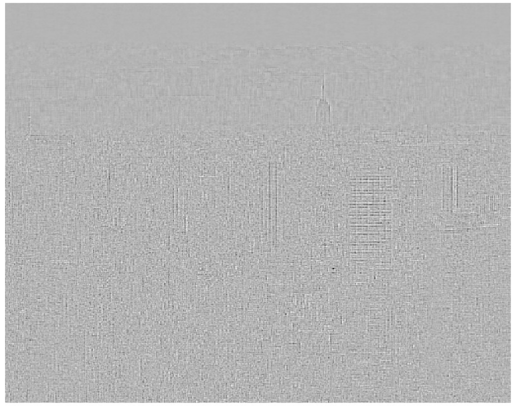
Let $\mathrm{r} = (r_0, \ldots, r_L)$ be a vector of rank bounds for each grid-scale. For any compressed tensor format $\mathcal{F}$, define the multiresolution $\mathcal{F}_{\mathrm{r}}$-format by

$$\mathrm{MS}_{\mathcal{F}_{\mathrm{r}}} = \left\{ T : T = \sum_{k=0}^{L} \mathrm{ext}_{L-k}(T_k),\, T_k \in \mathcal{F}_{r_k},\, T_k \in \mathbb{R}^{2^k \times \ldots \times 2^k} \right\}.$$

- Store as much information as possible on coarser (= cheaper) grid scales
- Tradeoff: down-sampling error vs low-rank error

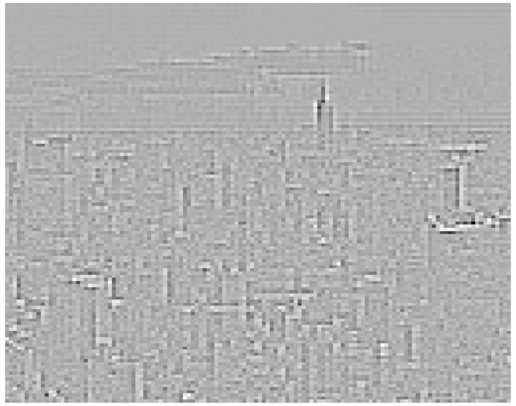
$T$

$=$



$T_L$

$+$

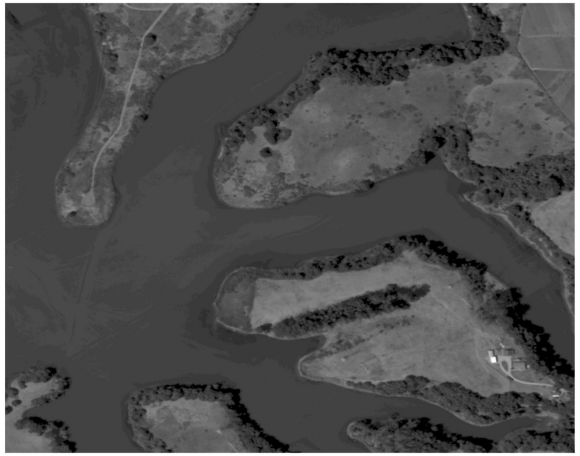


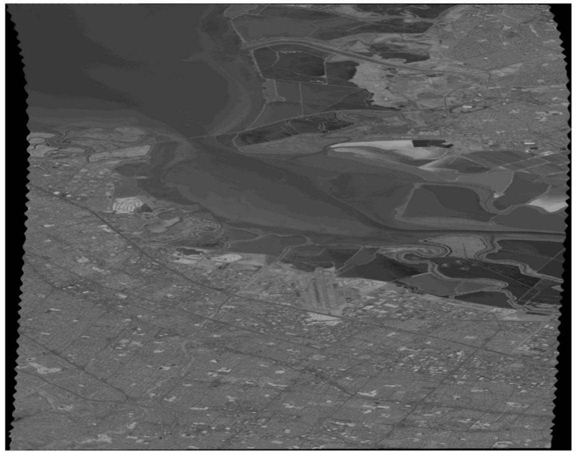
$\mathrm{ext}_1(T_{L-1})$

$+ \ldots$

# Example: hyperspectral wavelength data

Tensor $S \in \mathbb{R}^{n_1 \times n_2 \times n_3}$

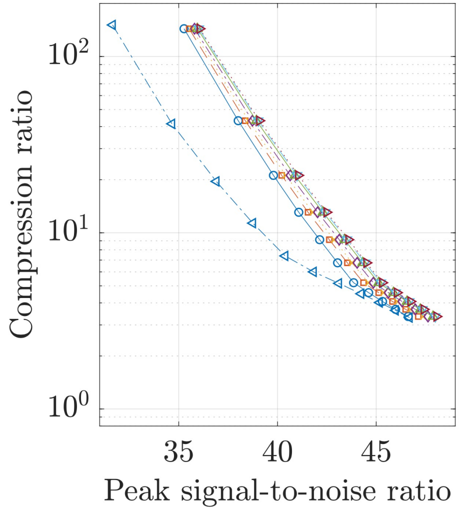




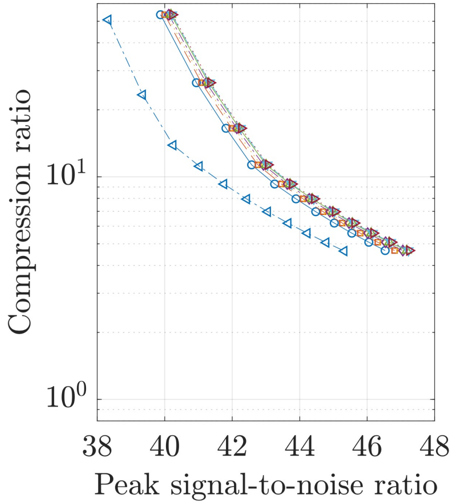
Oregon State University, SAMSON dataset

NASA JPL, AVIRIS Data

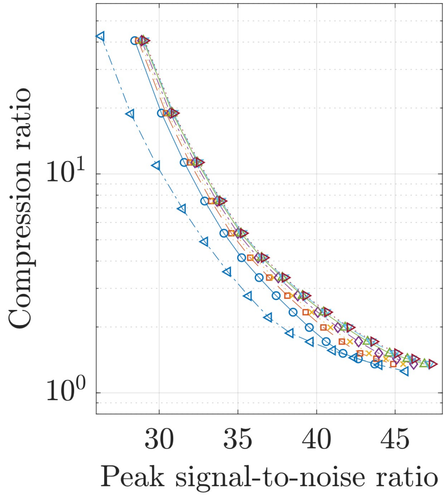
US Army Corps of Engineers, HyperCube



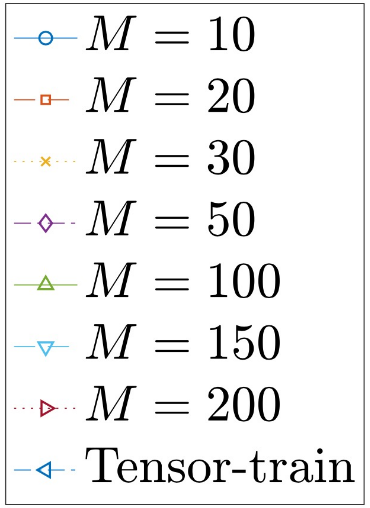

0.82 GB




1.26 GB




67 MB

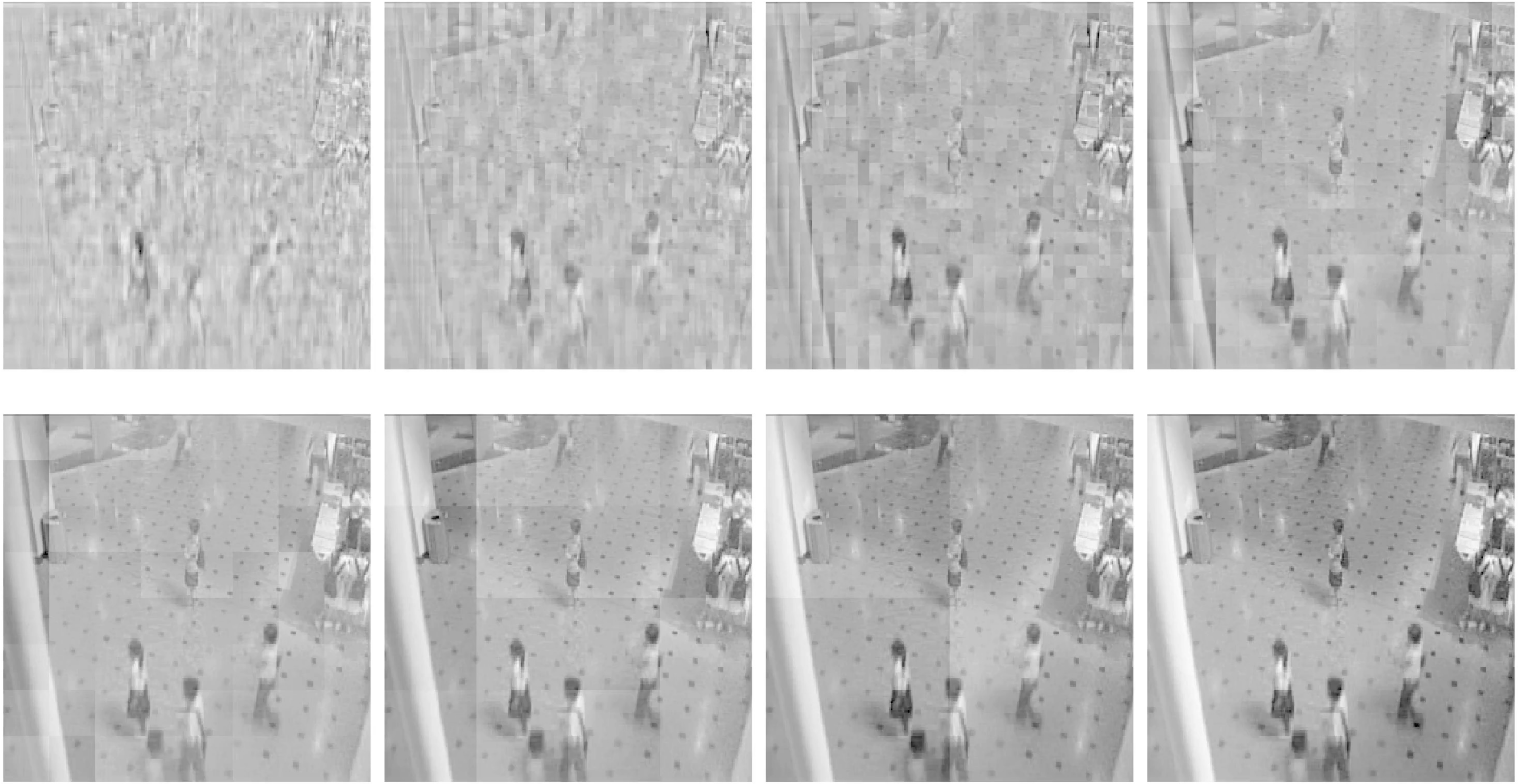



$M$ = number of iterations

# Example: video data

Multiscale

Tensor-train



# Theoretical difficulties

- Multiresolution format not closed – even for closed base formats!

Example

$$T^{(n)} = \begin{bmatrix} n & n \\ n & n \end{bmatrix} - \begin{bmatrix} \sqrt{n+1} \\ \sqrt{n-1} \end{bmatrix} \begin{bmatrix} \sqrt{n+1} & \sqrt{n-1} \end{bmatrix}, \quad \mathbf{r} = (1,1)$$

$$T^{(n)} \to T := \left[\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right],$$

$$T + a\left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right] = \left[\begin{smallmatrix} a-1 & a \\ a & a+1 \end{smallmatrix}\right] \quad \text{rank } 2$$

- Propagates to e.g., canonical decomposition and tensor-train format

# Theoretical remedies

- “Stable” sequences still converge for closed base formats

DEFINITION 4.2. *A sequence of tensors* $T^{(n)}$ *in* $MS_{\mathcal{F}_{\mathbf{r}}}$ *with*

$$T^{(n)} = \sum_{k=0}^{L} ext_{L-k}(T_k^{(n)}) \tag{4.2}$$

*is called stable if there is a constant* $C < \infty$ *such that* $\|T_k^{(n)}\| \leqslant C\|T^{(n)}\|$ *for each* $k = 0, \ldots, L$ *and* $n$.

THEOREM 4.3. *The format* $\mathcal{F}$ *is closed if and only if, for all possible rank vectors* $\mathbf{r}$*, all stable, convergent sequences in* $MS_{\mathcal{F}_{\mathbf{r}}}$ *converge to a tensor in* $MS_{\mathcal{F}_{\mathbf{r}}}$.

# Compressing tensors into format

Non-convex problem.

1. Start with initial approximation $\sum_{k=0}^{L} \mathrm{ext}_{L-k}(T_k^{(0)})$ to $T$
2. For $k = 1$ : max iterations, improve the approximation on the $k$:th scale by the update equation

$$T_k^{(n)} = \operatorname*{argmin}_{S \in \mathcal{F}_{r_k}} \| T - \sum_{\ell<k} \mathrm{ext}_{L-\ell}(T_\ell^{(n)}) - \sum_{\ell>k} \mathrm{ext}_{L-\ell}(T_\ell^{(n-1)}) - \mathrm{ext}_{L-k}(S) \| .$$

Can be shown to reduce to an optimal tensor approximation problem on each scale

# Compressing tensors into format

Non-convex problem.

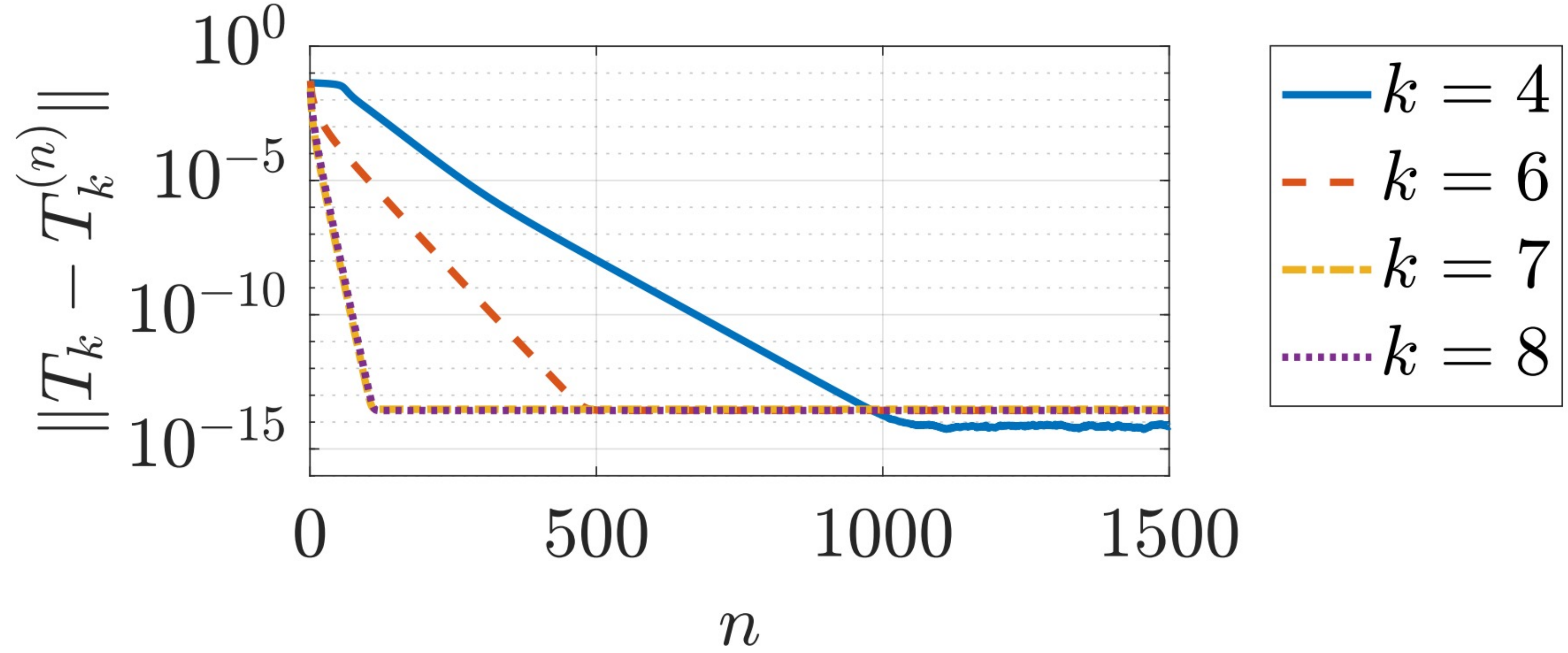
1. Start with initial approximation $\sum_{k=0}^{L} \mathrm{ext}_{L-k}(T_k^{(0)})$ to $T$
2. For $k = 1$ : max iterations, improve the approximation on the $k$:th scale by the update equation

$$T_k^{(n)} = \underset{S \in \mathcal{F}_{r_k}}{\operatorname{argmin}} \| T - \sum_{\ell < k} \mathrm{ext}_{L-\ell}(T_\ell^{(n)}) - \sum_{\ell > k} \mathrm{ext}_{L-\ell}(T_\ell^{(n-1)}) - \mathrm{ext}_{L-k}(S) \|.$$

Can be shown to reduce to an optimal tensor approximation problem on each scale

$$\begin{aligned}
\operatorname{argmin}_S \|A - \mathrm{ext}_{L-k}(S)\|^2 &= \operatorname{argmin}_S \|A\|^2 - 2\langle A, \mathrm{ext}_{L-k}(S)\rangle + \|\mathrm{ext}_{L-k}(S)\|^2 \\
&= \operatorname{argmin}_S \|A\|^2 - 2 b_s^{d(L-k)} \langle \mathrm{ave}_{L-k}(A), S\rangle + b_s^{d(L-k)} \|S\|^2 \\
&= \operatorname{argmin}_S \|A\|^2 - 2\langle \mathrm{ave}_{L-k}(A), S\rangle + \|S\|^2 \\
&= \operatorname{argmin}_S \|\mathrm{ave}_{L-k}(A) - S\|^2
\end{aligned}$$

# Compressing tensors into format




Can be shown to reduce to an optimal tensor approximation problem on each scale

Local convergence guarantees with linear convergence

# Thank you for listening!

## Conclusion

- Simple black-box augmentation of tensor formats for data with multiple length scales
- Interesting theoretical and numerical properties

## Future work

- Adaptive rank allocation schemes
- Sketching algorithms

Happy to discuss! oscarmi@mit.edu

# Theoretical remedies

THEOREM 4.3. *The format $\mathcal{F}$ is closed if and only if, for all possible rank vectors $\mathbf{r}$, all stable, convergent sequences in $MS_{\mathcal{F}_{\mathbf{r}}}$ converge to a tensor in $MS_{\mathcal{F}_{\mathbf{r}}}$.*

*Proof:*

$\Leftarrow$: take $\mathbf{r} = (0, \dots, 0, r)$

$\Rightarrow$: take $T^{(n)}$ in $MS_{\mathcal{F}_{\mathbf{r}}}$, $\quad T^{(n)} \to T$

For large $n$, $\|T^{(n)}\| \leq \|T\| + 1$, so $\|T_k^{(n)}\| \leq C\|T\| + C$

Convergent subsequences $T_k^{(n_j)}$ converge in $\mathcal{F}_{r_k}$, by closedness $\square$