What kind of tensors are compressible?

Tianyi Shi

Cornell University

ts777@cornell.edu

July 19, 2019

Work with: Alex Townsend (Cornell University)

Tensor decomposition

- CP [Hitchcock, 1927; Cattell, 1944; Carroll & Chang, 1970; Harshman, 1970]
- Tucker [Tucker, 1963]
- **Tensor-train** [Oseledets, 11]
- · · ·

Methodologies to understand the compressibility of tensors:

Tensor decomposition

- CP [Hitchcock, 1927; Cattell, 1944; Carroll & Chang, 1970; Harshman, 1970]
- Tucker [Tucker, 1963]
- **Tensor-train** [Oseledets, 11]
- $\cdots$

Methodologies to understand the compressibility of tensors:

- Algebraic structures: $\mathcal{X}_{i,j,k} = f(x_i, y_j, z_k)$

Tensor decomposition

- CP [Hitchcock, 1927; Cattell, 1944; Carroll & Chang, 1970; Harshman, 1970]
- Tucker [Tucker, 1963]
- **Tensor-train** [Oseledets, 11]
- $\cdots$

Methodologies to understand the compressibility of tensors:

- Algebraic structures: $\mathcal{X}_{i,j,k} = f(x_i, y_j, z_k)$
- Smoothness: $f(x, y, z) \approx p_n(x, y, z)$

Tensor decomposition

- CP [Hitchcock, 1927; Cattell, 1944; Carroll & Chang, 1970; Harshman, 1970]
- Tucker [Tucker, 1963]
- **Tensor-train** [Oseledets, 11]
- $\cdots$

Methodologies to understand the compressibility of tensors:

- Algebraic structures: $\mathcal{X}_{i,j,k} = f(x_i, y_j, z_k)$
- Smoothness: $f(x, y, z) \approx p_n(x, y, z)$
- **Displacement structure**
- $\cdots$

Rank bound of tensors with displacement structure

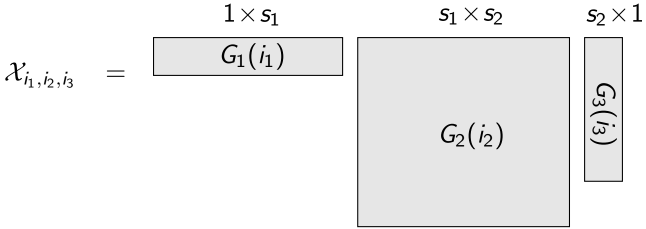
Theorem (S. & Townsend, 19)

Suppose $\mathcal{X} \times_1 A^{(1)} + \mathcal{X} \times_2 A^{(2)} + \mathcal{X} \times_3 A^{(3)} = \mathcal{G}$, where $A^{(1)}, A^{(2)}, A^{(3)}$ are Minkowski sum separated with disjoint sets E_j and F_j for $j = 1, 2$. Then, for a fixed $0 < \epsilon < 1$, we have

$$(\mathrm{rank}_\epsilon^{\mathrm{TT}}(\mathcal{X}))_j \leq k_j \nu_j, \quad \nu_j = \mathrm{rank}(G_j), \quad j = 1, 2,$$

where G_j is the jth unfolding of $\mathcal{G}$ and k_j is an integer so that $Z_{k_j}(E_j, F_j) \leq \epsilon/\sqrt{3}$.

Tensor-train decomposition



$$\operatorname{rank}^{TT}(\mathcal{X}) = (1, s_1, s_2, 1).$$

Storage:

$$\sum_{k=1}^{3} s_{k-1} s_k n_k.$$

Bound:

$$s_k \leq \operatorname{rank}(X_k), \quad (\boldsymbol{s}_\epsilon)_k \leq \operatorname{rank}_\epsilon(X_k), \qquad X_k = \operatorname{reshape}(\mathcal{X}, \prod_{s=1}^{k} n_s, \prod_{s=k+1}^{3} n_s).$$

Zolotarev number [Zolotarev, 1877]

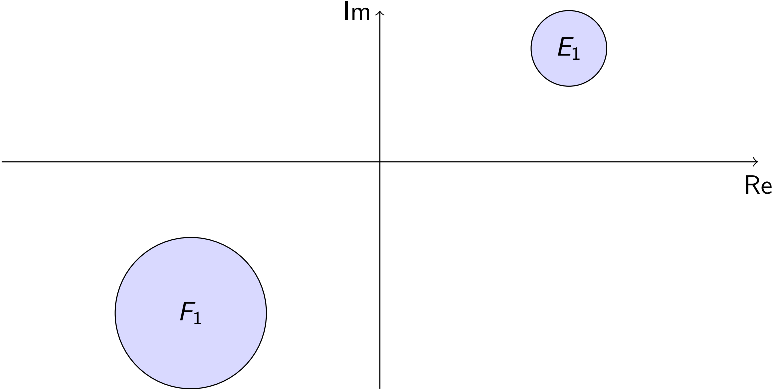
$$Z_k(E,F) := \inf_{r\in\mathcal{R}_{k,k}} \frac{\sup_{z\in E}|r(z)|}{\inf_{z\in F}|r(z)|}, \qquad k \geq 0,$$

E and F are disjoint complex sets and $\mathcal{R}_{k,k}$ is the set of irreducible rational functions of the form $p(x)/q(x)$ with polynomials p and q of degree at most k.

Zolotarev number [Zolotarev, 1877]

$$Z_k(E,F) := \inf_{r\in\mathcal{R}_{k,k}} \frac{\sup_{z\in E}|r(z)|}{\inf_{z\in F}|r(z)|}, \qquad k\geq 0,$$

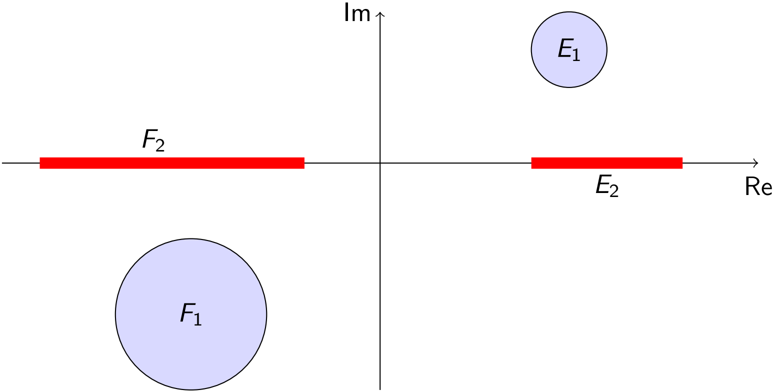
E and F are disjoint complex sets and $\mathcal{R}_{k,k}$ is the set of irreducible rational functions of the form $p(x)/q(x)$ with polynomials p and q of degree at most k.



Zolotarev number [Zolotarev, 1877]

$$Z_k(E,F) := \inf_{r \in \mathcal{R}_{k,k}} \frac{\sup_{z\in E} |r(z)|}{\inf_{z \in F} |r(z)|}, \qquad k \geq 0,$$

E and F are disjoint complex sets and $\mathcal{R}_{k,k}$ is the set of irreducible rational functions of the form $p(x)/q(x)$ with polynomials p and q of degree at most k.



Minkowski sum separation

Minkowski sum separated matrices

For normal matrices $A^{(1)}, A^{(2)}, A^{(3)}$, and disjoint sets E_j and F_j,

$$\Lambda(A^{(1)}) \subseteq E_1, \quad -(\Lambda(A^{(2)}) + \Lambda(A^{(3)})) \subseteq F_1,$$

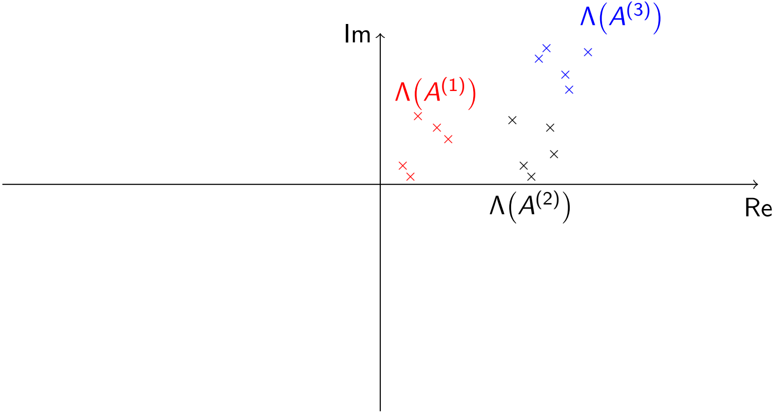
$$\Lambda(A^{(1)}) + \Lambda(A^{(2)}) \subseteq E_2, \quad -\Lambda(A^{(3)}) \subseteq F_2.$$

Minkowski sum separated matrices

$$\Lambda(A^{(1)}) \subseteq E_1, \quad -(\Lambda(A^{(2)}) + \Lambda(A^{(3)})) \subseteq F_1, \quad E_1 \cap F_1 = \varnothing,$$
$$\Lambda(A^{(1)}) + \Lambda(A^{(2)}) \subseteq E_2, \quad -\Lambda(A^{(3)}) \subseteq F_2, \quad E_2 \cap F_2 = \varnothing.$$

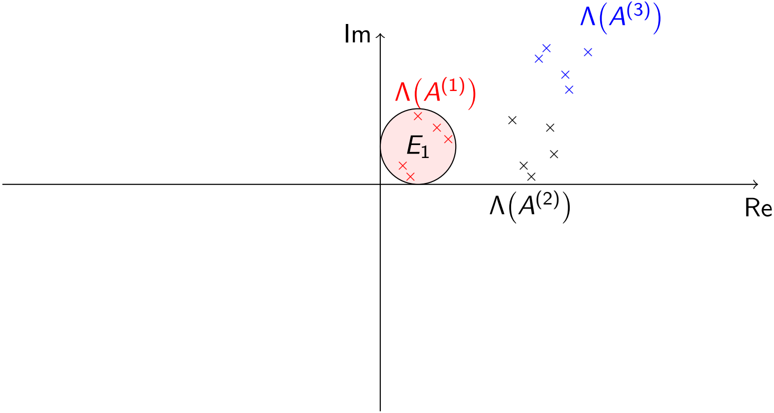
Minkowski sum separated matrices

$$\Lambda(A^{(1)}) \subseteq E_1, \quad -(\Lambda(A^{(2)}) + \Lambda(A^{(3)})) \subseteq F_1, \quad E_1 \cap F_1 = \varnothing,$$
$$\Lambda(A^{(1)}) + \Lambda(A^{(2)}) \subseteq E_2, \quad -\Lambda(A^{(3)}) \subseteq F_2, \quad E_2 \cap F_2 = \varnothing.$$



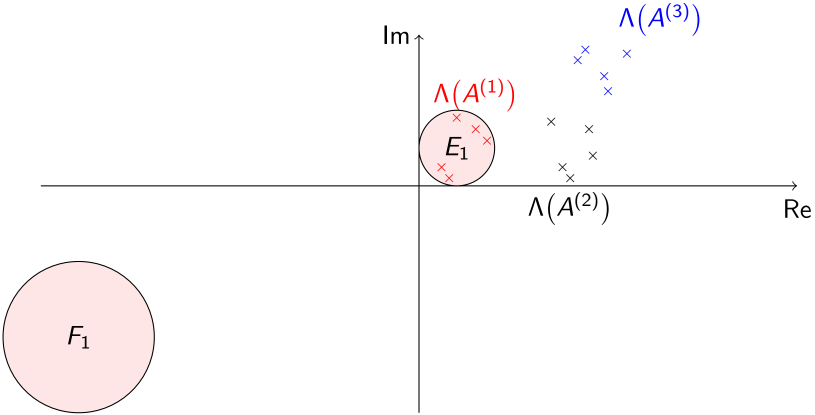
Minkowski sum separated matrices

$$\Lambda(A^{(1)}) \subseteq E_1, \quad -(\Lambda(A^{(2)}) + \Lambda(A^{(3)})) \subseteq F_1, \quad E_1 \cap F_1 = \varnothing,$$
$$\Lambda(A^{(1)}) + \Lambda(A^{(2)}) \subseteq E_2, \quad -\Lambda(A^{(3)}) \subseteq F_2, \quad E_2 \cap F_2 = \varnothing.$$



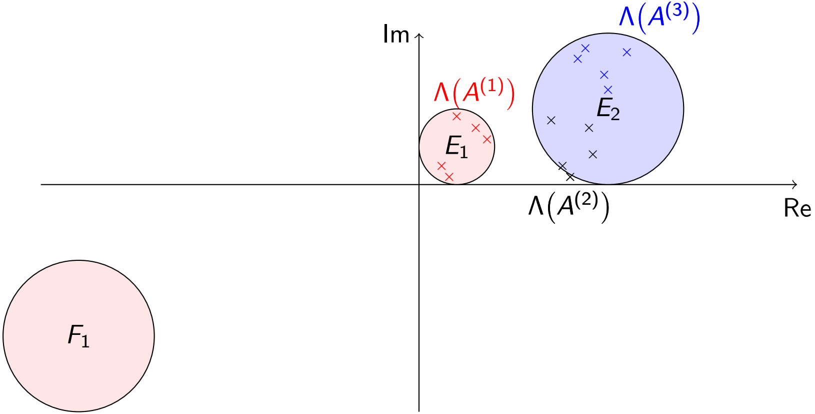
Minkowski sum separated matrices

$$\Lambda(A^{(1)}) \subseteq E_1, \quad -(\Lambda(A^{(2)}) + \Lambda(A^{(3)})) \subseteq F_1, \quad E_1 \cap F_1 = \varnothing,$$
$$\Lambda(A^{(1)}) + \Lambda(A^{(2)}) \subseteq E_2, \quad -\Lambda(A^{(3)}) \subseteq F_2, \quad E_2 \cap F_2 = \varnothing.$$



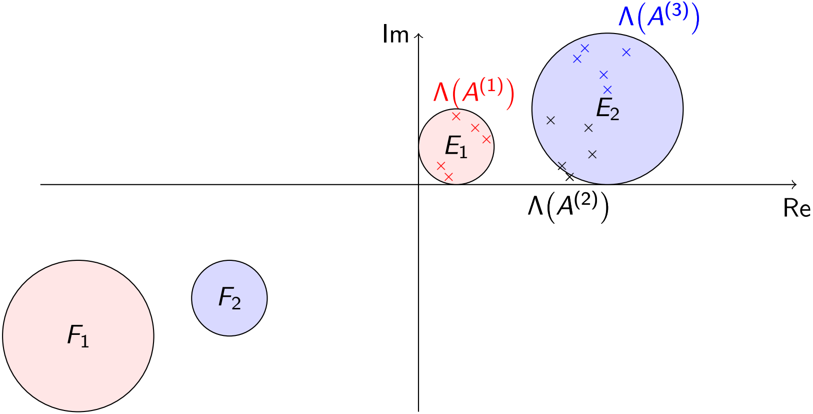
Minkowski sum separated matrices

$$\Lambda(A^{(1)}) \subseteq E_1, \quad -(\Lambda(A^{(2)}) + \Lambda(A^{(3)})) \subseteq F_1, \quad E_1 \cap F_1 = \varnothing,$$
$$\Lambda(A^{(1)}) + \Lambda(A^{(2)}) \subseteq E_2, \quad -\Lambda(A^{(3)}) \subseteq F_2, \quad E_2 \cap F_2 = \varnothing.$$



Minkowski sum separated matrices

$$\Lambda(A^{(1)}) \subseteq E_1, \quad -(\Lambda(A^{(2)}) + \Lambda(A^{(3)})) \subseteq F_1, \quad E_1 \cap F_1 = \varnothing,$$
$$\Lambda(A^{(1)}) + \Lambda(A^{(2)}) \subseteq E_2, \quad -\Lambda(A^{(3)}) \subseteq F_2, \quad E_2 \cap F_2 = \varnothing.$$



Rank bound of tensors with displacement structure (ctd.)

Theorem (S. & Townsend, 19)

Suppose $\mathcal{X} \times_1 A^{(1)} + \mathcal{X} \times_2 A^{(2)} + \mathcal{X} \times_3 A^{(3)} = \mathcal{G}$, *where* $A^{(1)}, A^{(2)}, A^{(3)}$ *are Minkowski sum separated with disjoint sets* E_j *and* F_j *for* $j = 1, 2$. *Then, for a fixed* $0 < \epsilon < 1$, *we have*

$$(\mathrm{rank}_\epsilon^{\mathrm{TT}}(\mathcal{X}))_j \le k_j \nu_j, \quad \nu_j = \mathrm{rank}(G_j), \quad j = 1, 2,$$

where G_j *is the* j*th unfolding of* $\mathcal{G}$ *and* k_j *is an integer so that* $Z_{k_j}(E_j, F_j) \le \epsilon/\sqrt{3}$.

Rank bound of tensors with displacement structure (ctd.)

Theorem (S. & Townsend, 19)

Suppose $\mathcal{X} \times_1 A^{(1)} + \mathcal{X} \times_2 A^{(2)} + \mathcal{X} \times_3 A^{(3)} = \mathcal{G}$, *where* $A^{(1)}, A^{(2)}, A^{(3)}$ *are Minkowski sum separated with disjoint sets* E_j *and* F_j *for* $j = 1, 2$. *Then, for a fixed* $0 < \epsilon < 1$, *we have*

$$(\mathrm{rank}_\epsilon^{\mathrm{TT}}(\mathcal{X}))_j \leq k_j \nu_j, \quad \nu_j = \mathrm{rank}(G_j), \quad j = 1, 2,$$

where G_j *is the* j*th unfolding of* $\mathcal{G}$ *and* k_j *is an integer so that* $Z_{k_j}(E_j, F_j) \leq \epsilon/\sqrt{3}$.

Special case

If $\Lambda(A^{(j)}) \subseteq [a, b]$ for $0 < a < b < \infty$, and $\gamma_j = \frac{(3a+j(b-a))(3b-j(b-a))}{9ab}$, then

$$(\mathrm{rank}_\epsilon^{\mathrm{TT}}(\mathcal{X}))_j \leq k_j \nu_j, \qquad k_j = \left\lceil \frac{\log(16\gamma_j)\log(4\sqrt{3}/\epsilon)}{\pi^2} \right\rceil.$$

Solving 3D Poisson equation

$$-(u_{xx} + u_{yy} + u_{zz}) = f \text{ on } \Omega = [-1, 1]^3, \qquad u|_{\partial\Omega} = 0.$$

Ultraspherical spectral discretization [Fortunato & Townsend, 17]:

$$u = (1 - x^2)(1 - y^2)(1 - z^2) \sum_{i=0}^{m} \sum_{j=0}^{n} \sum_{k=0}^{p} \mathcal{X}_{i,j,k} \tilde{C}_i^{(3/2)}(x) \tilde{C}_j^{(3/2)}(y) \tilde{C}_k^{(3/2)}(z),$$

$$f = \sum_{i=0}^{m} \sum_{j=0}^{n} \sum_{k=0}^{p} \mathcal{F}_{i,j,k} \tilde{C}_i^{(3/2)}(x) \tilde{C}_j^{(3/2)}(y) \tilde{C}_k^{(3/2)}(z),$$

Solving 3D Poisson equation

$$-(u_{xx} + u_{yy} + u_{zz}) = f \text{ on } \Omega = [-1, 1]^3, \qquad u|_{\partial\Omega} = 0.$$

Ultraspherical spectral discretization [Fortunato & Townsend, 17]:

$$u = (1 - x^2)(1 - y^2)(1 - z^2) \sum_{i=0}^{m} \sum_{j=0}^{n} \sum_{k=0}^{p} \mathcal{X}_{i,j,k} \tilde{C}_i^{(3/2)}(x) \tilde{C}_j^{(3/2)}(y) \tilde{C}_k^{(3/2)}(z),$$

$$f = \sum_{i=0}^{m} \sum_{j=0}^{n} \sum_{k=0}^{p} \mathcal{F}_{i,j,k} \tilde{C}_i^{(3/2)}(x) \tilde{C}_j^{(3/2)}(y) \tilde{C}_k^{(3/2)}(z),$$

$$\mathcal{X} \times_1 A^{-1} + \mathcal{X} \times_2 A^{-1} + \mathcal{X} \times_3 A^{-1} = \mathcal{G}, \quad \Lambda(A) \subseteq [-1,\ -1/(30n^4)],$$
$$\mathcal{G} = \mathcal{F} \times_1 M^{-1} \times_2 M^{-1} \times_3 M^{-1},$$

Solving 3D Poisson equation

$$-(u_{xx} + u_{yy} + u_{zz}) = f \text{ on } \Omega = [-1,1]^3, \qquad u|_{\partial\Omega} = 0.$$

Ultraspherical spectral discretization [Fortunato & Townsend, 17]:

$$u = (1-x^2)(1-y^2)(1-z^2) \sum_{i=0}^{m} \sum_{j=0}^{n} \sum_{k=0}^{p} \mathcal{X}_{i,j,k} \tilde{C}_i^{(3/2)}(x) \tilde{C}_j^{(3/2)}(y) \tilde{C}_k^{(3/2)}(z),$$

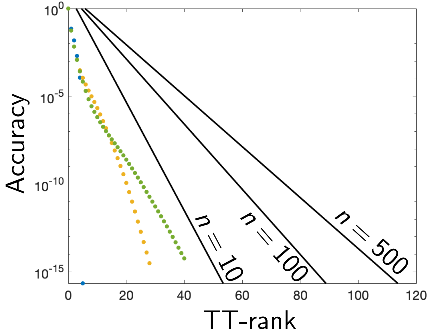
$$f = \sum_{i=0}^{m} \sum_{j=0}^{n} \sum_{k=0}^{p} \mathcal{F}_{i,j,k} \tilde{C}_i^{(3/2)}(x) \tilde{C}_j^{(3/2)}(y) \tilde{C}_k^{(3/2)}(z),$$

$$\mathcal{X} \times_1 A^{-1} + \mathcal{X} \times_2 A^{-1} + \mathcal{X} \times_3 A^{-1} = \mathcal{G}, \quad \Lambda(A) \subseteq [-1,\ -1/(30n^4)],$$
$$\mathcal{G} = \mathcal{F} \times_1 M^{-1} \times_2 M^{-1} \times_3 M^{-1},$$

$$(\mathrm{rank}_\epsilon^{\mathrm{TT}}(\mathcal{X}))_j \le s_j, \quad s_j = \mathcal{O}(\nu_j \log(n) \log(1/\epsilon)).$$

Solving 3D Poisson equation

When $f = 1$,



3D Poisson solver

- Constructive bound proof
- ADI-based algorithm
- Solve in TT format if unfoldings of $\mathcal{F}$ is low rank

Alternating Direction Implicit (ADI) method and factored ADI

Solving matrix Sylvester equation:

$$AX - XB = F, \quad A \in \mathbb{C}^{m\times m},\ B \in \mathbb{C}^{n\times n},$$

ADI [Wachspress, 08]

- Find vectors of shift parameters $\boldsymbol{p}$, and $\boldsymbol{q}$.
- Solve $(A - q_i I)X_{i+1/2} = F + X_i(B - q_i I)$.
- Solve $X_{i+1}(B - p_i I) = (A - p_i I)X_{i+1/2} - F$.

Alternating Direction Implicit (ADI) method and factored ADI

Solving matrix Sylvester equation:

$$AX - XB = F, \quad A \in \mathbb{C}^{m\times m},\ B \in \mathbb{C}^{n\times n},$$

ADI [Wachspress, 08]

- Find vectors of shift parameters $\boldsymbol{p}$, and $\boldsymbol{q}$.
- Solve $(A - q_i I)X_{i+1/2} = F + X_i(B - q_i I)$.
- Solve $X_{i+1}(B - p_i I) = (A - p_i I)X_{i+1/2} - F$.

f-ADI [Benner, Li, & Truhar, 09]

If $F = MN^*$, $M \in \mathbb{C}^{m\times\nu}$, $N \in \mathbb{C}^{n\times\nu}$, and ν smaller than m and n, solve for $X = ZDY^*$.

TT-SVD and f-ADI based Poisson solver (rough sketch)

TT-SVD is an algorithm that calculates the TT decomposition of a given tensor "train" by "train":

TT-SVD and f-ADI based Poisson solver (rough sketch)

TT-SVD is an algorithm that calculates the TT decomposition of a given tensor "train" by "train":

Compute SVD of $X_1 = U_1 D_1 V_1^*$ and use U_1 as first "train".

TT-SVD and f-ADI based Poisson solver (rough sketch)

TT-SVD is an algorithm that calculates the TT decomposition of a given tensor "train" by "train":

Compute SVD of $X_1 = U_1 D_1 V_1^*$ and use U_1 as first "train".

Use f-ADI to solve only for orthogonal row space U_1 of X_1, which satisfies $A^{-1}X_1 + X_1(I \otimes A^{-1} + A^{-1} \otimes I)^T = G_1 = M_1 N_1^*$.

TT-SVD and f-ADI based Poisson solver (rough sketch)

TT-SVD is an algorithm that calculates the TT decomposition of a given tensor "train" by "train":

Compute SVD of $X_1 = U_1 D_1 V_1^*$ and use U_1 as first "train".

Use f-ADI to solve only for orthogonal row space U_1 of X_1, which satisfies $A^{-1}X_1 + X_1(I \otimes A^{-1} + A^{-1} \otimes I)^T = G_1 = M_1 N_1^*$.

Compute SVD of a reshaping of $D_1 V_1^*$ to get second and third "train".

TT-SVD and f-ADI based Poisson solver (rough sketch)

TT-SVD is an algorithm that calculates the TT decomposition of a given tensor "train" by "train":

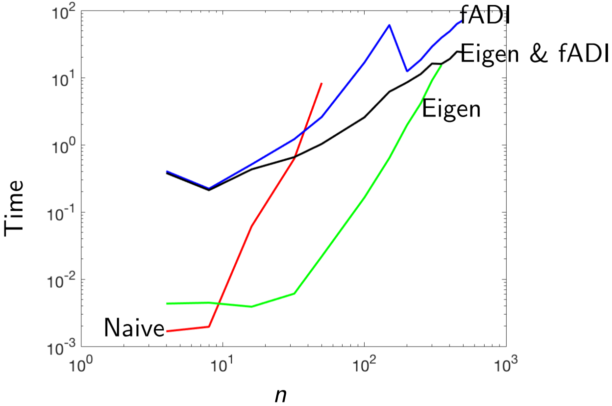
Compute SVD of $X_1 = U_1 D_1 V_1^*$ and use U_1 as first "train".

Use f-ADI to solve only for orthogonal row space U_1 of X_1, which satisfies $A^{-1} X_1 + X_1 (I \otimes A^{-1} + A^{-1} \otimes I)^T = G_1 = M_1 N_1^*$.

Compute SVD of a reshaping of $D_1 V_1^*$ to get second and third "train".

Use f-ADI to solve for both orthogonal row and column spaces of C_2, which satisfies
$(I \otimes (U_1^* A^{-1} U_1) + A^{-1} \otimes I) C_2 + C_2 (A^{-1})^T = (I \otimes U_1^*) G_2 = (I \otimes U_1^*) M_2 N_2^*$.

$$f = -2(1-y^2)(1-z^2) - 2(1-x^2)(1-z^2) - 2(1-x^2)(1-y^2).$$



Overall complexity

- Converting between Chebyshev and $\tilde{C}^{(3/2)}$ takes $\mathcal{O}(n^2(\log n)^2\log(1/\epsilon))$ complexity [Townsend, Webb, & Olver, 17].
- Solving for trains using fADI and ADI takes $\mathcal{O}(n^2(\log n)^2(\log(1/\epsilon))^2)$ complexity.

Summary

- Several methodologies guarantee compressibility of certain tensors in various formats.
- Super fast spectrally accurate Poisson equation solver.

Ongoing work

- Make the fast Poisson solver open-source codes.
- What about

$$\sum_{j=1}^{d} A_j X B_j^T = F,$$

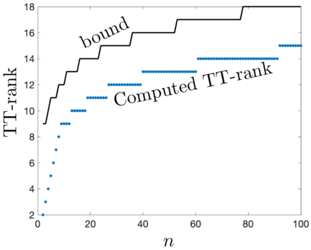
$$\sum_{j=1}^{d} \mathcal{X} \times_1 A_j \times_2 B_j \times_3 C_j = \mathcal{F}.$$

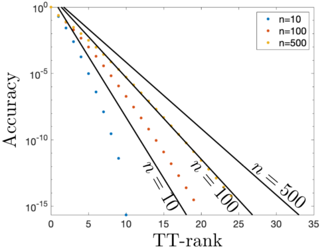
Bonus: Hilbert tensor

$$\mathcal{H}_{i,j,k} = \frac{1}{i+j+k-2}, \qquad 1 \le i,j,k \le n.$$
$$\mathcal{H} \times_1 D + \mathcal{H} \times_2 D + \mathcal{H} \times_3 D = \mathcal{S},$$

$\mathcal{S}$ is the tensor of all ones and D is a diagonal matrix with $D_{ii} = i - \frac{2}{3}$.

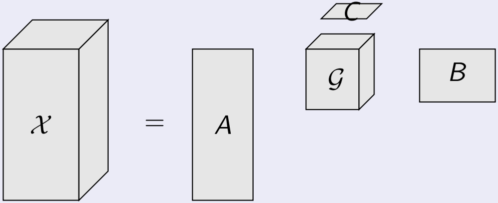
$$\mathrm{rank}_\epsilon^{\mathrm{TT}}(\mathcal{H}) \le_{\mathrm{lex}} (1, s_1, s_1, 1), \qquad s_1 = \left\lceil \frac{1}{\pi^2} \log\left(\frac{16n(2n-1)}{3n-2}\right) \log\left(\frac{4\sqrt{3}}{\epsilon}\right) \right\rceil.$$





Bonus: Solving 3D Poisson in Orthogonal Tucker format

Tucker decomposition [Tucker, 1963]



HOSVD [De Lathauwer, De Moor, & Vanderwalle, 00] and fADI-based Poisson solver

- Use fADI to solve for A, B, and C as orthogonal row spaces of matricizations of $\mathcal{X}$.
- Solve a smaller tensor Sylvester equation of $\mathcal{G}$.