20 - 26
JUNE
2021
PORTOROŽ
SLOVENIA

8TH
EUROPEAN
CONGRESS OF
MATHEMATICS

Lana Periša

Infinite Tensor Rings

21 June 2021

This talk is funded by the Croatian Science Foundation
under the project UIP-2019-04-5200.

The paper

Roel Van Beeumen, Lana Periša, Daniel Kressner, Chao Yang -

A Flexible Power Method for Solving Infinite Dimensional Tensor Eigenvalue Problems

- https://arxiv.org/abs/2102.00146

Overview

Tensor Rings

- Finite and Infinite Tensor Rings
- Translational invariance

Motivation

- The problem
- The method

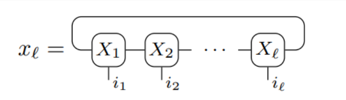
Properties of iTRs

- Normalized iTR
- Canonical form
- Rayleigh quotient
- Multiplication with iTRs

Tensor Rings

Finite Tensor Rings

Let x_ℓ be an ℓth-order tensor of size $d_1 \times d_2 \times \cdots \times d_\ell$, then its **tensor ring representation** is defined as



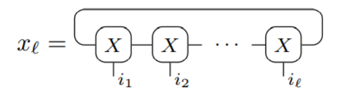
and element-wise as

$$x_\ell(i_1, i_2, \ldots, i_\ell) := Tr\left[X_1(i_1)X_2(i_2)\cdots X_\ell(i_\ell)\right],$$

where the indices i_k run from 1 to d_k, $k = 1, 2, \ldots, \ell$, and each $X_k(i_k)$ is a matrix of size $r_k \times r_{k+1}$, with $r_1 = r_{\ell+1}$.

Finite Tensor Rings

Let x_ℓ be an ℓth-order tensor of size d^ℓ, then its **translational invariant tensor ring representation** is defined as



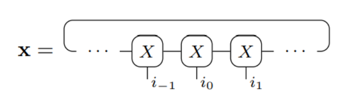
and element-wise as

$$x_\ell(i_1, i_2, \ldots, i_\ell) := Tr\left[X(i_1)X(i_2)\cdots X(i_\ell)\right],$$

where the indices i_k run from 1 to d, $k = 1, 2, \ldots, \ell$, and each $X(i_k)$ is a matrix of size $r \times r$.

Infinite Tensor Rings

A **translational invariant infinite tensor ring (iTR)** is defined as



and element-wise as

$$\mathbf{x}(\ldots, i_{-1}, i_0, i_1, \ldots) := Tr\left[\prod_{k=-\infty}^{+\infty} X(i_k)\right],$$

where all indices i_k run from 1 to d and each $X(i)$ is a matrix of size $r \times r$, with r referred to as the *rank* of $\mathbf{x}$.

Motivation

The problem

Computing the **algebraically smallest eigenvalue** of an *infinite dimensional* tensor eigenvalue problem

$$\mathbf{H}\mathbf{x} = \lambda \mathbf{x},$$

where $\mathbf{H}$ is the infinite dimensional symmetric matrix

$$\mathbf{H} = \sum_{k=-\infty}^{+\infty} \mathbf{H}_k, \quad \mathbf{H}_k = \cdots \otimes I \otimes I \otimes M_{k,k+1} \otimes I \otimes I \otimes \cdots .$$

- $I \in \mathbb{R}^{d \times d}$ identity
- $M_{k,k+1} \in \mathbb{R}^{d^2 \times d^2}$ all equal

- $\mathbf{H}$ is the infinite sum of Kronecker products of infinite number of finite matrices
- $\mathbf{H}$ is *translational invariant*
- the eigenvectors of $\mathbf{H}$ are infinite dimensional vectors

Finite case

Take $H_\ell \in \mathbb{R}^{d^\ell \times d^\ell}$ that only contains $(\ell - 1)$ terms in the summation and Kronecker products of $(\ell - 1)$ matrices.

- H_ℓ admits Tensor Train (TT) representation

- the corresponding eigenvector can be represented by the TT

$$x_\ell(i_1, i_2, \dots, i_\ell) = X_1(i_1)X_2(i_2)\cdots X_\ell(i_\ell),$$

where $X_k(i_k)$ is an $r_k \times r_{k+1}$ matrix, with $r_1 = r_{\ell+1} = 1$, and the indices $i_k = 1, \dots, d$, for $k = 1, \dots, \ell$

- limit $\ell \to \infty$ is known in the physics literature as the *thermodynamic limit* and is important for describing macroscopic properties of quantum materials when $\mathbf{H}$ corresponds to a quantum many-body Hamiltonian

Properties of infinite case

Translational invariance property of **H** implies the *Bethe–Hulthén hypothesis*.

The elements of the eigenvector are invariant with respect to a cyclic permutation of the tensor indices, i.e.,

$$\mathbf{x}(\ldots, i_{-1}, i_0, i_1, \ldots) = \mathbf{x}(\ldots, i_0, i_1, i_2, \ldots).$$

We represent the eigenvector to be computed as a **translational invariant *infinite Tensor Ring*** (iTR)

$$\mathbf{x}(\ldots, i_{-1}, i_0, i_1, \ldots) = Tr\left[\prod_{k=-\infty}^{+\infty} X(i_k)\right],$$

where all indices $i_k = 1, \ldots, d$, and each $X(i_k)$ is a matrix of size $r \times r$.

The method

To compute the desired eigenpair we assume that the smallest eigenvalue of $\mathbf{H}$ is simple and propose to apply a **flexible power iteration** to $e^{-\mathbf{H}t}$ for some small and variable parameter t.

- Lie product formula (Suzuki–Trotter splitting):

$$e^{-\mathbf{H}t} \approx \prod_{k=-\infty}^{+\infty} e^{-\mathbf{H}_k t},$$

  can be accurate enough if t is sufficiently small
- $e^{-\mathbf{H}_k t} = \cdots \otimes I \otimes I \otimes e^{-Mt} \otimes I \otimes I \otimes \cdots$
- $\displaystyle\prod_{k=-\infty}^{+\infty} e^{-\mathbf{H}_k t} = \cdots \otimes e^{-Mt} \otimes e^{-Mt} \otimes e^{-Mt} \otimes \cdots$

(admits TT representation)

The power method ingredients

How to do this operations when **x** is an iTR?

- multiplication

$$\left(\prod_{k=-\infty}^{+\infty} e^{-\mathbf{H}_k t}\right)\mathbf{x}$$

- normalization, uniqueness
- Rayleight quotient

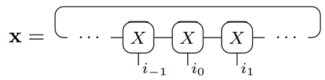
$$\mathbf{x}^T\mathbf{H}\mathbf{x} = \sum_{k=-\infty}^{+\infty} \mathbf{x}^T\mathbf{H}_k\mathbf{x}$$

- residual

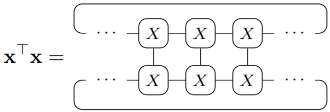
Properties of iTRs

Transfer matrix

Let **x** be an iTR

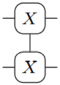


Then



and we define the **transfer matrix** T_X associated with **x** as the $r^2 \times r^2$ matrix

$$T_X := \sum_{i=1}^{d} X(i) \otimes X(i) =$$

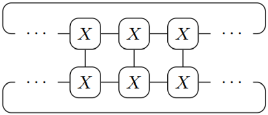


where $X(i)$ is the ith slice of X.

Normalized iTR

Let $\mathbf{x}$ be an iTR. Then $\mathbf{x}$ is **normalized**, i.e., $\mathbf{x}^T\mathbf{x} = 1$, if its corresponding transfer matrix T_X has a simple dominant eigenvalue $\eta = 1$.

$$\mathbf{x}^\top \mathbf{x} =$$

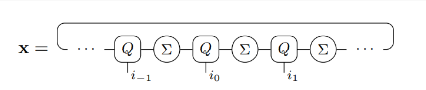


$$= \mathrm{Tr}\left[\lim_{k\to\infty} (T_X)^k\right] = \mathrm{Tr}\left[v_R v_L^\top\right] = v_L^\top v_R = 1,$$

Still, we can insert the product of any nonsingular matrix S and its inverse between two consecutive cores of an iTR and redefine each slice as $S^{-1}X(i)S$, $i = 1, \ldots, d$, so *the representation of an iTR is not unique*.

Canonical form

Let **x** be an iTR. Then its canonical form is defined as



and element-wise as

$$\mathbf{x}(\ldots, i_{-1}, i_0, i_1, \ldots) := Tr\left[\prod_{k=-\infty}^{+\infty} Q(i_k)\Sigma\right],$$

where $Q(i) \in \mathbb{R}^{r\times r}$, for $i = 1, \ldots, d$, and $\Sigma \in \mathbb{R}^{r\times r}$ is a diagonal matrix with decreasing non-negative real numbers on its diagonal and $\|\Sigma\|_F = 1$, such that the following *left and right orthogonality conditions* hold

$$\sum_{i=1}^{d} Q(i)^T\Sigma^T\Sigma Q(i) = \eta I, \qquad \sum_{i=1}^{d} Q(i)\Sigma\Sigma^T Q(i)^T = \eta I,$$

with $\eta \in \mathbb{R}$ being the dominant eigenvalue of the transfer matrix T_X.

Canonical decomposition

Algorithm 1: Canonical decomposition of iTR

Input : 3rd-order tensor X
Output: 3rd-order tensor Q and diagonal matrix Σ

1 Dominant left and right eigenvectors of transfer matrix T_X:

$$\mathrm{vec}(V_L)^\top T_X = \eta\, \mathrm{vec}(V_L)^\top, \qquad T_X\, \mathrm{vec}(V_R) = \eta\, \mathrm{vec}(V_R).$$

2 Eigendecompositions of V_L and V_R: $\quad V_L = U_L \Lambda_L U_L^\top, \quad V_R = U_R \Lambda_R U_R^\top.$

3 Form: $\quad \tilde{U}_L = U_L \Lambda_L^{1/2}, \quad \tilde{U}_R = U_R \Lambda_R^{1/2}.$

4 Singular value decomposition: $\quad V \tilde{\Sigma} W^\top = \tilde{U}_L^\top \tilde{U}_R.$

5 Form: $\quad L = W^\top \tilde{U}_R^{-1}, \quad R = \tilde{U}_L^{-\top} V.$

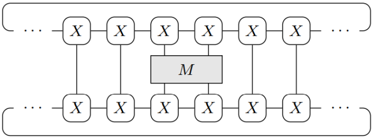
6 Canonical form: $\quad Q(i) = \|\tilde{\Sigma}\|_F\, L X(i) R, \quad \Sigma = \tilde{\Sigma}/\|\tilde{\Sigma}\|_F.$

7 (Normalization): $\quad Q(i) = Q(i)/\sqrt{\eta}.$

Rayleigh quotient

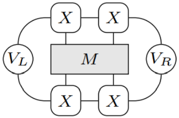
$$\mathbf{x}^\top \mathbf{H}\mathbf{x} = \sum_{k=-\infty}^{+\infty} \mathbf{x}^\top \mathbf{H}_k \mathbf{x},$$

$\mathbf{x}^\top \mathbf{H}_k \mathbf{x} =$



Let **x** be a normalized nonzero iTR and T_X its corresponding transfer matrix. Then the **Rayleigh quotient** (in an *average sense*) for a given infinite dimensional matrix **H** can be represented by

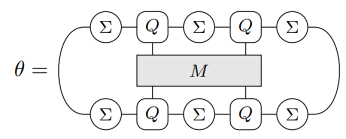
$\theta =$



where V_L and V_R are, respectively, the matricizations of the left and right dominant eigenvectors of T_X.

Canonical Rayleigh quotient

Let **x** be a normalized iTR in canonical form with Q and Σ being the tensor and matrix from the canonical representation. Then the **Rayleigh quotient** associated with the infinite dimensional matrix **H** can be represented by

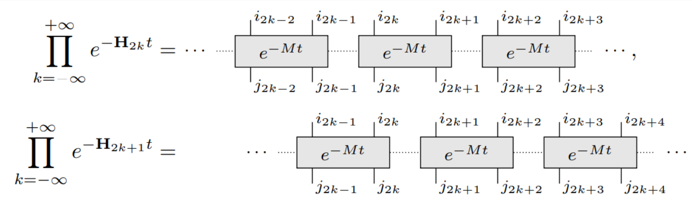


Multiplication with iTR

$$\mathbf{H} = \left(\sum_{k=-\infty}^{+\infty} \mathbf{H}_{2k}\right) + \left(\sum_{k=-\infty}^{+\infty} \mathbf{H}_{2k+1}\right) =: \ \mathbf{H}_e + \mathbf{H}_o$$

Using the Suzuki–Trotter splitting twice:

$$e^{-\mathbf{H}t} = e^{-(\mathbf{H}_e+\mathbf{H}_o)t} \approx \left(\prod_{k=-\infty}^{+\infty} e^{-\mathbf{H}_{2k}t}\right)\left(\prod_{k=-\infty}^{+\infty} e^{-\mathbf{H}_{2k+1}t}\right)$$

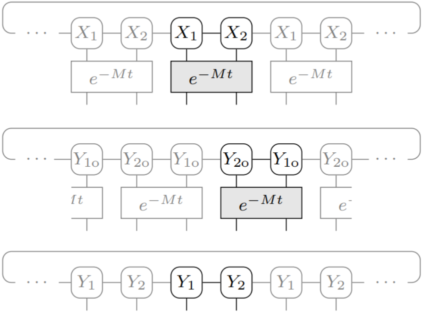


The matrix–vector operation $\mathbf{y} = e^{-\mathbf{H}t}\mathbf{x}$

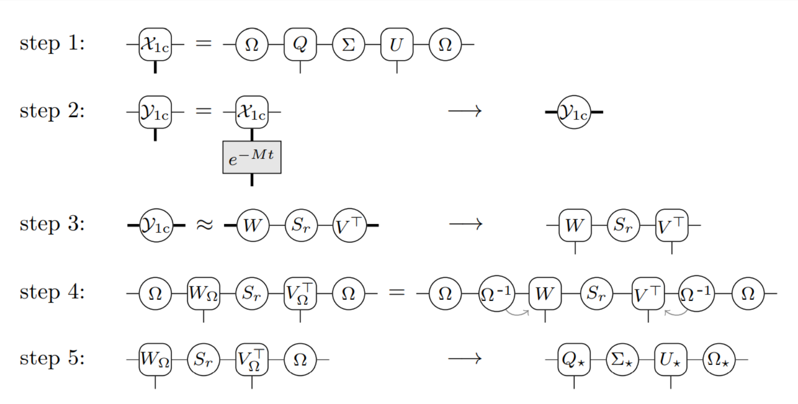
$$\left(\prod_{k=-\infty}^{+\infty} e^{-\mathbf{H}_{2k+1}t}\right)\mathbf{x} =$$

$$\left(\prod_{k=-\infty}^{+\infty} e^{-\mathbf{H}_{2k}t}\right)\tilde{\mathbf{y}}_{\text{o}} =$$

$$\tilde{\mathbf{y}} =$$



The multiplication details



Recap

- An iTR can be seen as the infinite limit of finite size tensor ring.
- Due to the translational invariance, we only need to store, and work, with d matrices of size $r \times r$.
- Most operations can be efficiently implemented involving only tensors of size $r \times d \times r$.
- Special structure of $\mathbf{H}$ allows us to split it in even and odd terms.
- We are able to multiply a matrix exponential with an iTR and keep the product in iTR form.
- We keep the rank low by doing the truncated SVD.

Thank you! Questions?