Tensors in Computations V

Lek-Heng Lim

University of Chicago

recap from lecture IV

recap from lecture IV

- tensor product defined in three ways
  - ❶ via tensor product of function spaces
  - ❷ via tensor product of more general vector spaces
  - ❸ via the universal mapping property
- ❶ is the sum of separable functions construction
- ❷ is the polyadic construction
- unfortunately no time for ❸

recap: two constructions

❶ polyadic construction: a d-tensor in $\mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_{d-1} \otimes \mathbb{V}_d$ is a 'linear combination'

$$\boldsymbol{\alpha}_1 \mathbf{e}_1 + \boldsymbol{\alpha}_2 \mathbf{e}_2 + \cdots + \boldsymbol{\alpha}_n \mathbf{e}_n$$

with $(d-1)$-tensor coefficients $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_n \in \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_{d-1}$ and $\mathbf{e}_1, \dots, \mathbf{e}_n$ basis of $\mathbb{V}_d$

❷ sum of separable functions construction: a d-tensor is a d-variate function $f : \Omega_1 \times \Omega_2 \times \cdots \times \Omega_d \to \mathbb{R}$ that is a finite sum of separable functions

$$f(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_d) = \sum_{i=1}^{r} \varphi_i(\mathbf{x}_1)\psi_i(\mathbf{x}_2) \cdots \theta_i(\mathbf{x}_d)$$

trivial case of ❷

- notation $[n] \coloneqq \{1, 2, \dots, n\}$ for any $n \in \mathbb{N}$

$$\mathbb{R}^n = \mathbb{R}^{[n]} = \{f : [n] \to \mathbb{R}\}$$
$$\mathbb{R}^{m \times n} = \mathbb{R}^{[m] \times [n]} = \{f : [m] \times [n] \to \mathbb{R}\}$$
$$\mathbb{R}^{m \times n \times p} = \mathbb{R}^{[m] \times [n] \times [p]} = \{f : [m] \times [n] \times [p] \to \mathbb{R}\}$$

- outer product of vectors $\mathbf{a} \in \mathbb{R}^m$, $\mathbf{b} \in \mathbb{R}^n$, $\mathbf{c} \in \mathbb{R}^p$

$$\mathbf{a} \otimes \mathbf{b} \coloneqq \mathbf{a}\mathbf{b}^\top = \begin{bmatrix} a_1 b_1 & \cdots & a_1 b_n \\ \vdots & \ddots & \vdots \\ a_m b_1 & \cdots & a_m b_n \end{bmatrix} \in \mathbb{R}^{m \times n}$$

$$\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c} = \left[\begin{array}{ccc|ccc|c|ccc} a_1 b_1 c_1 & \cdots & a_1 b_n c_1 & a_1 b_1 c_2 & \cdots & a_1 b_n c_2 & & a_1 b_1 c_p & \cdots & a_1 b_n c_p \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \cdots\cdots & \vdots & \ddots & \vdots \\ a_m b_1 c_1 & \cdots & a_m b_n c_1 & a_m b_1 c_2 & \cdots & a_m b_n c_2 & & a_m b_1 c_p & \cdots & a_m b_n c_p \end{array}\right] \in \mathbb{R}^{m \times n \times p}$$

goes much further

- allows us to view functions, vector fields, distributions, operators, hypermatrices, multilinear maps, tensor fields, etc, as tensors
- solution for fluid velocity $\mathbf{v}$ in the Navier–Stokes

$$\frac{\partial v_i}{\partial t} + \sum_{j=1}^{3} \frac{\partial v_i}{\partial x_j} v_j = -\frac{1}{\rho}\frac{\partial p}{\partial x_i} + \nu \sum_{j=1}^{3} \frac{\partial^2 v_i}{\partial x_j^2} + f_i, \quad i = 1, 2, 3,$$

  is a tensor $\mathbf{v} \in C^2(\mathbb{R}^3) \mathbin{\widehat{\otimes}} C^1[0, \infty) \otimes \mathbb{R}^3$
- quantum state of spin-half particle is a tensor $\Psi \in L^2(\mathbb{R}^3) \otimes \mathbb{C}^2$

$$\Psi(\mathbf{x}, \sigma) = \sum_{i=1}^{r} \psi_i(\mathbf{x}) \chi_i(\sigma)$$

  with $\psi_i \in L^2(\mathbb{R}^3)$ and $\chi_i \colon \{-\frac{1}{2}, \frac{1}{2}\} \to \mathbb{C}$

many applications

- tensor product constructions:
  - orthogonal bases
  - Riesz bases
  - frames
  - Mercer kernels
  - function spaces
  - density operators
  - multiresolution analyses
- algorithms exploting separability:
  - kernel trick
  - multipole expansion
  - Smolyak's quadrature
  - Grover quantum search
  - row–column decomposition
  - separable ODEs and integral equations
  - separable convex and integer programming
- unfortunately no time for any of these

motivation

Justifying separation of variables

Posted on **12 June 2021** by **John**

The separation of variables technique for solving partial differential equations looks like a magic trick the first time you see it. The lecturer, or author if you're more self-taught, makes an audacious assumption, like pulling a rabbit out of a hat, and it works.

For example, you might first see the heat equation

$$u_t = c^2 u_{xx}.$$

The professor asks you to assume the solution has the form

$$u(x, t) = X(x)\, T(t).$$

i.e. the solution can be separated into the product of a function of x alone and a function of t alone.

Following that you might see Laplace's equation on a rectangle

$$u_{xx} + u_{yy} = 0$$

with the analogous assumption that

$$u(x, y) = X(x)\, Y(y),$$

common mistake

- incorrect: "separation of variables works because sums of separable functions are dense in the function space"
- take wave equation for example

$$\frac{\partial^2 f}{\partial x^2} - \frac{\partial^2 f}{\partial t^2} = 0$$

  and $f(x, t) = \varphi(x)\psi(t)$ solves it
- $\xi = x - t$, $\eta = x + t$ turns it into

$$\frac{\partial^2 f}{\partial \xi \, \partial \eta} = 0$$

  and $f(\xi, \eta) = \varphi(\xi)\psi(\eta)$ leads nowhere
- sums of separable functions are dense in (ξ, η) coordinates as they are in (x, t) coordinates so this cannot be the reason
- whether it works or not depends on choice of coordinates

justifying separation-of-variables

additive and multiplicative separability

- $f : \Omega_1 \times \Omega_2 \times \cdots \times \Omega_d \to \mathbb{R}$ or $\mathbb{C}$
- additive separability

$$f(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_d) = f_1(\mathbf{x}_1) + f_2(\mathbf{x}_2) + \cdots + f_d(\mathbf{x}_d)$$

  for some $f_i : \Omega_i \to \mathbb{R}$

- multiplicative separability

$$f(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_d) = f_1(\mathbf{x}_1) f_2(\mathbf{x}_2) \cdots f_d(\mathbf{x}_d)$$

  for some $f_i : \Omega_i \to \mathbb{R}$

- both forms intimately related but we focus on the latter first

in terms of tensors

- tensor product of functions $\varphi : \Omega_1 \to \mathbb{R}$, $\psi : \Omega_2 \to \mathbb{R}, \ldots, \theta : \Omega_d \to \mathbb{R}$

$$\varphi \otimes \psi \otimes \cdots \otimes \theta : \Omega_1 \times \Omega_2 \times \cdots \times \Omega_d \to \mathbb{R}$$

  defined by

$$[\varphi \otimes \psi \otimes \cdots \otimes \theta](\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_d) := \varphi(\mathbf{x}_1)\psi(\mathbf{x}_2)\cdots\theta(\mathbf{x}_d)$$

- vector space of real-valued functions

$$\mathbb{V}_i := \{f : \Omega_i \to \mathbb{R}\}$$

- tensor product of vector spaces

$$\mathbb{V}_1 \otimes \mathbb{V}_2 \otimes \cdots \otimes \mathbb{V}_d := \left\{ \sum_{i=1}^{r} \varphi_i \otimes \psi_i \otimes \cdots \otimes \theta_i : \varphi_i \in \mathbb{V}_1, \psi_i \in \mathbb{V}_2, \ldots \theta_i \in \mathbb{V}_d,\ r \in \mathbb{N} \right\}$$

- by definition $\mathbb{V}_1 \otimes \mathbb{V}_2 \otimes \cdots \otimes \mathbb{V}_d$ comprises all finite-rank tensors

operators

- $\Phi_1 \colon \mathbb{V}_1 \to \mathbb{W}_1, \dots, \Phi_d \colon \mathbb{V}_d \to \mathbb{W}_d$ linear operators
- Kronecker product of operators

$$[\Phi_1 \otimes \cdots \otimes \Phi_d](\mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_d) \coloneqq \Phi_1(\mathbf{v}_1) \otimes \cdots \otimes \Phi_d(\mathbf{v}_d)$$

  and extend linearly to all of $\mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_d$
- an operator $\Phi \colon \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_d \to \mathbb{V}_1 \otimes \cdots \otimes \mathbb{V}_d$ is separable if

$$\Phi = \Phi_1 \otimes I_2 \otimes \cdots \otimes I_d + I_1 \otimes \Phi_2 \otimes \cdots \otimes I_d + \cdots + I_1 \otimes I_2 \otimes \cdots \otimes \Phi_d$$

  where I_k is identity operator on $\mathbb{V}_k$

separation-of-variables

$$\Phi = \Phi_1 \otimes I_2 \otimes \cdots \otimes I_d + I_1 \otimes \Phi_2 \otimes \cdots \otimes I_d + \cdots + I_1 \otimes I_2 \otimes \cdots \otimes \Phi_d$$

- transforms homogeneous linear system into eigenproblems:

$$\Phi(\mathbf{v}_1 \otimes \mathbf{v}_2 \otimes \cdots \otimes \mathbf{v}_d) = 0 \quad \longrightarrow \quad \begin{cases} \Phi_1(\mathbf{v}_1) = \lambda_1 \mathbf{v}_1 \\ \Phi_2(\mathbf{v}_2) = \lambda_2 \mathbf{v}_2 \\ \quad \vdots \\ \Phi_d(\mathbf{v}_d) = -(\lambda_1 + \cdots + \lambda_{d-1})\mathbf{v}_d \end{cases}$$

- Φ being linear, any sum, linear combination, integral of $\mathbf{v}_1 \otimes \mathbf{v}_2 \otimes \cdots \otimes \mathbf{v}_d$ is also a solution
- relies on just one simple fact: for any non-zero $\mathbf{v} \in \mathbb{V}$ and $\mathbf{w} \in \mathbb{W}$,

$$\mathbf{v} \otimes \mathbf{w} = \mathbf{v}' \otimes \mathbf{w}' \quad \Rightarrow \quad \mathbf{v} = \lambda \mathbf{v}', \ \mathbf{w} = \lambda^{-1} \mathbf{w}'$$

for some non-zero $\lambda \in \mathbb{R}$

proof

- $d = 3$ for illustration

$$(\Phi \otimes I \otimes I + I \otimes \Psi \otimes I + I \otimes I \otimes \Theta)(\mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w}) = 0$$

- equivalently

$$\Phi(\mathbf{u}) \otimes \mathbf{v} \otimes \mathbf{w} + \mathbf{u} \otimes \Psi(\mathbf{v}) \otimes \mathbf{w} + \mathbf{u} \otimes \mathbf{v} \otimes \Theta(\mathbf{w}) = 0$$

- since $\Phi(\mathbf{u}) \otimes (\mathbf{v} \otimes \mathbf{w}) = \mathbf{u} \otimes [-\Psi(\mathbf{v}) \otimes \mathbf{w} - \mathbf{v} \otimes \Theta(\mathbf{w})]$

$$\Phi(\mathbf{u}) = \lambda \mathbf{u}, \quad \mathbf{v} \otimes \mathbf{w} = -\lambda^{-1}[\Psi(\mathbf{v}) \otimes \mathbf{w} + \mathbf{v} \otimes \Theta(\mathbf{w})]$$

- rearranging second equation, $\Psi(\mathbf{v}) \otimes \mathbf{w} = \mathbf{v} \otimes [-\Theta(\mathbf{w}) - \lambda \mathbf{w}]$

$$\Psi(\mathbf{v}) = \mu \mathbf{v}, \quad \Theta(\mathbf{w}) = -(\mu + \lambda)\mathbf{w}$$

- transformed into three eigenproblems:

$$\begin{cases} \Phi(\mathbf{u}) = \lambda \mathbf{u}, \\ \Psi(\mathbf{v}) = \mu \mathbf{v}, \\ \Theta(\mathbf{w}) = (-\mu - \lambda)\mathbf{w} \end{cases}$$

example: partial differential equation

- wave equation:

$$\frac{\partial^2 f}{\partial x^2} - \frac{\partial^2 f}{\partial t^2} = 0.$$

- separation-of-variables

$$[\partial_x^2 \otimes I + I \otimes (-\partial_t^2)](\varphi \otimes \psi) = 0 \quad \longrightarrow \quad \begin{cases} \partial_x^2 \varphi = -\omega^2 \varphi \\ -\partial_t^2 \psi = \omega^2 \psi \end{cases}$$

- ODEs have solutions

$$\begin{cases} \varphi(x) = a_1 \, \mathrm{e}^{\omega x} + a_2 \, \mathrm{e}^{-\omega x} \\ \psi(x) = a_3 \, \mathrm{e}^{\omega t} + a_4 \, \mathrm{e}^{-\omega t} \end{cases} \qquad \begin{cases} \varphi(x) = a_1 + a_2 x \\ \psi(t) = a_3 + a_4 t \end{cases}$$

  for $\omega \neq 0$ and $\omega = 0$ respectively

- any finite linear combinations of $\varphi \otimes \psi$ give us general solutions

depends on coordinates

- change coordinates

$$\xi = x - t, \qquad \eta = x + t$$

- wave equation becomes

$$\frac{\partial^2 f}{\partial \xi \, \partial \eta} = 0$$

- operator here is $\partial_\xi \otimes \partial_\eta$ and is not separable
- solution easily seen to take the form

$$f(\xi, \eta) = \varphi(\xi) + \psi(\eta)$$

- so an ansatz of the form $f(\xi, \eta) = \varphi(\xi)\psi(\eta)$ will not work
- moral: separation-of-variables depends on coordinates

more generally

- wave equation

$$\Delta f - \frac{\partial^2 f}{\partial t^2} = 0,$$

- separation-of-variables

$$[\Delta \otimes I + I \otimes (-\partial_t^2)](\varphi \otimes \psi) = 0 \quad \longrightarrow \quad \begin{cases} \Delta\varphi = -\omega^2\varphi \\ -\partial_t^2\psi = \omega^2\psi \end{cases}$$

  with separation constant $-\omega^2$

- important for us later: Helmholtz equation

$$\Delta\varphi = -\omega^2\varphi$$

example: integro-differential equation

- heterogeneous heat transfer:

$$\frac{\partial f}{\partial t} = a\frac{\partial^2 f}{\partial x^2} + b\int_0^x f(y,t)\,\mathrm{d}y - f$$

- write

$$\Phi_x(f) := \frac{\partial^2 f}{\partial x^2} - f + b\int_0^x f(y,t)\mathrm{d}y, \quad \Psi_t(f) := -\frac{\partial f}{\partial t}$$

- separation-of-variables

$$[\Phi_x \otimes I + I \otimes \Psi_t](\varphi \otimes \psi) = 0 \quad \longrightarrow \quad \begin{cases} \Phi_x(\varphi) = \lambda\varphi \\ \Psi_t(\psi) = -\lambda\psi \end{cases}$$

- equivalently

$$a\frac{\mathrm{d}^2\varphi}{\mathrm{d}x^2} + (\lambda - 1)\varphi + b\int_0^x \varphi(y)\,\mathrm{d}y = 0, \quad \frac{\mathrm{d}\psi}{\mathrm{d}t} + \lambda\psi = 0$$

- solve to get $\varphi(x) = c_1\,\mathrm{e}^{r_1 x} + \mathrm{e}^{r_2 x}(c_2 \cos r_3 x + c_3 \sin r_3 x)$ and $\psi(t) = c_4\,\mathrm{e}^{-\lambda t}$ with c_i arbitrary constants

example: recurrence equations

- forward-time centred space discretization applied to heat equation

$$\begin{cases} u_{k,n+1} = ru_{k-1,n} + (1-2r)u_{k,n} + ru_{k+1,n} & k = 1, \ldots, m-1 \\ u_{0,n+1} = 0 = u_{m,n+1} & \\ u_{k,0} = f(k/m) & k = 0, 1, \ldots, m \end{cases}$$

with $n = 0, 1, 2, \ldots,$ and $r > 0$ some fixed constant

- write

$$\Phi_k(u_{k,n}) \coloneqq ru_{k-1,n} + (1-2r)u_{k,n} + ru_{k+1,n}, \quad \Psi_n(u_{k,n}) \coloneqq u_{k,n+1}$$

- separation-of-variables

$$[\Phi_k \otimes I + I \otimes (-\Psi_n)](a \otimes b) = 0 \quad \longrightarrow \quad \begin{cases} \Phi_k(a_k) = \lambda a_k \\ -\Psi_n(b_n) = -\lambda b_n \end{cases}$$

example: recurrence equations

- equivalently

$$ra_{k-1} + (1-2r)a_k + ra_{k+1} = \lambda a_k, \quad k = 1, \dots, m-1$$
$$b_{n+1} = \lambda b_n, \quad n = 0, 1, 2, \dots.$$

- second equation is easy: $b_n = \lambda^n b_0$
- first equation is tridiagonal eigenproblem with solution

$$\lambda_j = 1 - 4r \sin^2\left(\frac{j\pi}{2m}\right), \quad a_{jk} = \sin\left(\frac{jk\pi}{m}\right)$$

  where a_{jk} is kth coordinate of the jth eigenvector
- solution is

$$u_{k,n} = \sum_{j=1}^{m-1} c_j b_0 \left[1 - 4r \sin^2\left(\frac{j\pi}{2m}\right)\right]^n \sin\left(\frac{jk\pi}{m}\right)$$

general solutions

- summary: if operator Φ is separable

  $$\Phi = \Phi_1 \otimes I_2 \otimes \cdots \otimes I_d + I_1 \otimes \Phi_2 \otimes \cdots \otimes I_d + \cdots + I_1 \otimes I_2 \otimes \cdots \otimes \Phi_d$$

  then $\Phi(f) = 0$ has a multiplicatively separable solution

  $$f = \varphi \otimes \psi \otimes \cdots \otimes \theta$$

- linear combinations of such f's gives us more general solutions
- one way to write down such a linear combination

  $$f = \sum_{i=1}^{r} c_i \varphi_i \otimes \psi_i \otimes \cdots \otimes \theta_i$$

- a better way: each factor given its own index

  $$f = \sum_{i=1}^{r_1} \sum_{i=1}^{r_2} \cdots \sum_{k=1}^{r_d} c_{ij\cdots k}\, \varphi_i \otimes \psi_j \otimes \cdots \otimes \theta_k$$

- both are sums of separable functions but latter gives us a way to impose structures on coefficients and indices

some background

Schrödinger equation

- time-dependent Schrödinger equation for d particles

$$\mathrm{i}\hbar\frac{\partial}{\partial t}f(\mathbf{x},t) = \left[-\frac{\hbar^2}{2m}\Delta + V(\mathbf{x})\right]f(\mathbf{x},t)$$

  - $\mathbf{x} = (\mathbf{x}_1,\ldots,\mathbf{x}_d) \in \mathbb{R}^{3n}$ positions of d particles
  - V real-valued function representing potential
  - $\Delta = \Delta_1 + \Delta_2 + \cdots + \Delta_d$ with $\Delta_i \colon L^2(\mathbb{R}^3) \to L^2(\mathbb{R}^3)$ Laplacian

- not necessarily Cartesian, may have say

$$\Delta_i = \frac{1}{r_i^2}\frac{\partial}{\partial r_i}\left(r_i^2\frac{\partial}{\partial r_i}\right) + \frac{1}{r_i^2\sin\theta_i}\frac{\partial}{\partial\theta_i}\left(\sin\theta_i\frac{\partial}{\partial\theta_i}\right) + \frac{1}{r_i^2\sin^2\theta_i}\frac{\partial^2}{\partial\phi_i^2}$$

  with $\mathbf{x}_i = (r_i,\theta_i,\phi_i)$

- enough for us: ignore constants, keep signs

$$(-\Delta + V)f - \mathrm{i}\partial_t f = 0$$

Schrödinger equation

- separation-of-variables

$$(-\Delta + V) \otimes I + I \otimes (-\mathrm{i}\partial_t) \quad \longrightarrow \quad \begin{cases} (-\Delta + V)\varphi = E\varphi \\ -\mathrm{i}\partial_t \psi = -E\psi \end{cases}$$

  where we write separation constant as $-E$
- second equation is easy $\psi(t) = \mathrm{e}^{-\mathrm{i}Et}$
- problem is first equation: time-independent Schrödinger equation
- need to solve for φ and E

toy example

- potential V additively separable

$$V(\mathbf{x}) = V_1(\mathbf{x}_1) + V_2(\mathbf{x}_2) + \cdots + V_d(\mathbf{x}_d)$$

- then Schrödinger equation has the form we need

$$\sum_{i=1}^{d}(-\Delta_i + V_i)\varphi - E\varphi = 0$$

- may apply separation-of-variables

$$\begin{aligned}
&[(-\Delta_1 + V_1) \otimes I \otimes \cdots \otimes I + I \otimes (-\Delta_2 + V_2) \otimes \cdots \otimes I \\
&\quad + I \otimes \cdots \otimes I \otimes (-\Delta_d + V_d - E)](\varphi_1 \otimes \varphi_2 \otimes \cdots \otimes \varphi_d) = 0 \\
&\quad \longrightarrow \quad \begin{cases}
(-\Delta_1 + V_1)\varphi_1 = E_1\varphi_1 \\
(-\Delta_2 + V_2)\varphi_2 = E_2\varphi_2 \\
\qquad\qquad \vdots \\
(-\Delta_d + V_d)\varphi_d = (E - E_1 - \cdots - E_{d-1})\varphi_d
\end{cases}
\end{aligned}$$

toy example

- write

$$E_d := E - E_1 - \cdots - E_{d-1}, \quad \varphi = \varphi_1 \otimes \cdots \otimes \varphi_d$$

- moral: if potential V additively separable

$$V(\mathbf{x}) = V_1(\mathbf{x}_1) + V_2(\mathbf{x}_2) + \cdots + V_d(\mathbf{x}_d)$$

- then eigenfunction φ is multiplicatively separable

$$\varphi(\mathbf{x}) = \varphi_1(\mathbf{x}_1)\varphi_2(\mathbf{x}_2)\cdots\varphi_d(\mathbf{x}_d)$$

- and the eigenvalue E is additively separable

$$E = E_1 + E_2 + \cdots + E_d$$

- separates d-particle Schrödinger into d one-particle Schrödinger

issue

- $V(\mathbf{x}) = V_1(\mathbf{x}_1) + \cdots + V_d(\mathbf{x}_d)$ unrealistic — says that the particles do not interact
- but even by including only pairwise interactions

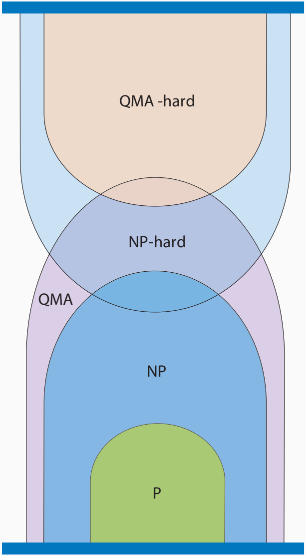
$$V(\mathbf{x}) = \sum_{i=1}^{d} V_i(\mathbf{x}_i) + \sum_{i<j} V_{ij}(\mathbf{x}_i, \mathbf{x}_j)$$

  i.e., no higher-order terms of the form $V_{ijk}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k)$
- and even by restricting to

$$V_{ij}(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{\|\mathbf{x}_i - \mathbf{x}_j\|}$$

- $(-\Delta + V)\varphi = E\varphi$ becomes computationally intractable in multiple ways

computational intractability



picture from [Aaronson, 2009]

- stunning results of [Schuch–Verstraete, 2009]
- Hartree–Fock is NP-hard
- density functional theory is QMA-hard
- see [Whitfield–Love–Aspuru-Guzik, 2013] for a survey, [Aaronson, 2009] for a summary

approximation schemes

- we know

$$V(\mathbf{x}) = V_1(\mathbf{x}_1) + \cdots + V_d(\mathbf{x}_d) \quad \Rightarrow \quad \begin{cases} \varphi(\mathbf{x}) = \varphi_1(\mathbf{x}_1) \cdots \varphi_d(\mathbf{x}_d) \\ E = E_1 + \cdots + E_d \end{cases}$$

- rougly, approximations based on the belief

$$V(\mathbf{x}) \approx V_1(\mathbf{x}_1) + \cdots + V_d(\mathbf{x}_d) \quad \Rightarrow \quad \begin{cases} \varphi(\mathbf{x}) \approx \varphi_1(\mathbf{x}_1) \cdots \varphi_d(\mathbf{x}_d) \\ E \approx E_1 + \cdots + E_d \end{cases}$$

- '$\approx$' interpreted differently and with different schemes
  - one-electron approximation: perturbation theory
  - Hartree–Fock approximation: calculus of variations

example: Hartree–Fock

- Rayleigh quotient

$$\mathcal{E}(\varphi) = \frac{\langle(-\Delta + V)\varphi, \varphi\rangle}{\|\varphi\|^2}$$

  is stationary, i.e., $\delta\mathcal{E} = 0$, if and only if $(-\Delta + V)\varphi = E\varphi$
- Hartree–Fock approximation seeks stationarity under multiplicative separability $\varphi = \varphi_1 \otimes \cdots \otimes \varphi_d$

$$\mathcal{L}(\varphi_1, \ldots, \varphi_d, \lambda_1, \ldots, \lambda_d) = \mathcal{E}(\varphi_1, \ldots, \varphi_d) - \lambda_1\|\varphi_1\|^2 - \cdots - \lambda_d\|\varphi_d\|^2$$

- $\delta\mathcal{L} = 0$ gives

$$\left[-\Delta_i + \sum_{j\neq i} \int_{\mathbb{R}^3} |\varphi_j(y)|^2 \, V(x, y) \, \mathrm{d}y\right]\varphi_i = \lambda_i\varphi_i$$

- makes physical sense:
  - particle i in a potential field due to the charge of particle j
  - charge spread over space with density $|\varphi_j|^2$
  - sum over potential fields created by all particles $j \neq i$

example: multiconfiguration Hartree–Fock

- as before but now with ansatz

$$f = a_1\varphi_1 \otimes \varphi_1 + a_2\varphi_2 \otimes \varphi_2$$

  where $\varphi_1, \varphi_2 \in L^2(\mathbb{R}^3)$ orthonormal and $(a_1, a_2) \in \mathbb{R}^2$ unit vector

- non-linear energy functional

$$\begin{aligned}\mathcal{E}(\varphi_1, \varphi_2, a_1, a_2) = a_1^2 &\langle(-\Delta + V)\varphi_1 \otimes \varphi_1, \varphi_1 \otimes \varphi_1\rangle \\ &+ 2a_1a_2\langle(-\Delta + V)\varphi_1 \otimes \varphi_1, \varphi_2 \otimes \varphi_2\rangle \\ &+ a_2^2\langle(-\Delta + V)\varphi_2 \otimes \varphi_2, \varphi_2 \otimes \varphi_2\rangle\end{aligned}$$

  with constraints $\|\varphi_1\|^2 = \|\varphi_2\|^2 = 1$, $\langle\varphi_1, \varphi_2\rangle = 0$, $\|a\|^2 = 1$

- Lagrangian is

$$\begin{aligned}&\mathcal{L}(\varphi_1, \varphi_2, a_1, a_2, \lambda_{11}, \lambda_{12}, \lambda_{22}, \lambda) \\ &\quad = \mathcal{E}(\varphi_1, \varphi_2, a_1, a_2) + \lambda_{11}\|\varphi_1\|^2 + \lambda_{12}\langle\varphi_1, \varphi_2\rangle + \lambda_{22}\|\varphi_2\|^2 - \lambda\|a\|^2\end{aligned}$$

example: multiconfiguration Hartree–Fock

- for $i, j \in \{1, 2\}$ write

$$b_{ij} = \langle(-\Delta+V)\varphi_i\otimes\varphi_i, \varphi_j\otimes\varphi_j\rangle, \quad c_{ij}(x) = \int_{\mathbb{R}^3} \varphi_i(x)\varphi_j(y)V(x,y)\,\mathrm{d}y$$

- stationarity conditions $\nabla_a \mathcal{L} = 0$ and $\delta\mathcal{L} = 0$ give

$$\begin{bmatrix} b_{11}-\lambda & b_{12} \\ b_{12} & b_{22}-\lambda \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = 0$$

  and

$$\begin{bmatrix} -\Delta_1\varphi_1 \\ -\Delta_2\varphi_2 \end{bmatrix} = \begin{bmatrix} c_{11}(x)-\lambda_{11} & (a_2/a_1)(c_{12}(x)-\lambda_{12}) \\ (a_1/a_2)(c_{12}(x)-\lambda_{12}) & c_{22}(x)-\lambda_{22} \end{bmatrix} \begin{bmatrix} \varphi_1 \\ \varphi_2 \end{bmatrix}$$

- solved numerically with a combination of finite-difference and quadrature [Fischer, 1977]

tensor networks

assumptions

- instead of restricting to $L^2(\mathbb{R}^3)$ assume arbitrary separable Hilbert spaces $\mathbb{H}_1, \dots, \mathbb{H}_d$ to allow for spin, i.e., $L^2(\mathbb{R}^3) \otimes \mathbb{C}^{2s-1}$
- seek solution $f \in \mathbb{H}_1 \otimes \cdots \otimes \mathbb{H}_d$ to d-particle Schrödinger equation
- by definition of $\otimes$, f finite rank even if $\mathbb{H}_1, \dots, \mathbb{H}_d$ infinite-dimensional
- so there is a decomposition

$$f = \sum_{i=1}^{r_1} \sum_{j=1}^{r_2} \cdots \sum_{k=1}^{r_d} c_{ij\cdots k}\, \varphi_i \otimes \psi_j \otimes \cdots \otimes \theta_k$$

- may assume orthogonal factors

$$\langle \varphi_i, \varphi_j \rangle = \langle \psi_i, \psi_j \rangle = \cdots = \langle \theta_i, \theta_j \rangle = \begin{cases} 0 & i \neq j, \\ 1 & i = j \end{cases}$$

motivation

- issue with

$$f = \sum_{i=1}^{r_1} \sum_{j=1}^{r_2} \cdots \sum_{k=1}^{r_d} c_{ij\cdots k}\, \varphi_i \otimes \psi_j \otimes \cdots \otimes \theta_k$$

  is exponential number of rank-one terms as d increases
- if $r_1 = \cdots = r_d = r$, then there are r^d summands
- good ansatz supposed to capture small region of the space where solution likely lies
- goal of tensor networks is to provide such an ansatz by limiting the coefficients $[c_{ij\cdots k}] \in \mathbb{R}^{r_1 \times \cdots \times r_d}$ to a much smaller set

example: matrix product states

- impose on the coefficients the structure

$$c_{ij\cdots k} = \operatorname{tr}(A_i B_j \cdots C_k)$$

  with

$$A_i \in \mathbb{R}^{n_1 \times n_2}, B_j \in \mathbb{R}^{n_2 \times n_3}, \dots, C_k \in \mathbb{R}^{n_d \times n_1}$$

- due to [Anderson, 1959], [Affleck, Kennedy, Lieb and Tasaki, 1987], [White, 1992], [White–Huse, 1993]
- MPS is ansatz of the form

$$f = \sum_{i=1}^{r_1} \sum_{j=1}^{r_2} \cdots \sum_{k=1}^{r_d} \operatorname{tr}(A_i B_j \cdots C_k)\, \varphi_i \otimes \psi_j \otimes \cdots \otimes \theta_k$$

- coefficients parametrized by $r_1 + \cdots + r_d$ matrices of various sizes
- if $r_1 = \cdots = r_d = r$ and $n_1 = \cdots = n_d = n$, then MPS has rdn^2 as opposed to r^d degrees of freedom

open and periodic MPS

- when $n_1 = 1$, first and last matrices are a row and a column vector respectively

$$A_i = \mathbf{a}_i^\top, \quad C_k = \mathbf{c}_k$$

  with $\mathbf{a}_i \in \mathbb{R}^{n_2}$ and $\mathbf{c}_k \in \mathbb{R}^{n_d}$

- trace of a 1×1 matrix is itself and may drop the 'tr'

$$f = \sum_{i=1}^{r_1} \sum_{j=1}^{r_2} \cdots \sum_{k=1}^{r_d} \mathbf{a}_i^\top B_j \cdots \mathbf{c}_k \, \varphi_i \otimes \psi_j \otimes \cdots \otimes \theta_k$$

- special case called MPS with open boundary conditions [Anderson, 1959]
- general case called MPS with periodic conditions

computing MPS

- another difference between Hartree–Fock and tensor network is that the factors $\varphi_i, \psi_j, \dots, \theta_k$ are often fixed in advance as some standard bases of $\mathbb{H}_1, \dots, \mathbb{H}_d$, called a local basis
- computational effort then reduces to determining coefficients $c_{ij\cdots k}$
- for MPS this can be done via several SVDs [Orús, 2014]
- in fact coefficients of MPS ansatz sometimes represented as

$$\operatorname{tr}(Q_1 \Sigma_1 Q_2 \Sigma_2 \dots \Sigma_d Q_{d+1}), \quad Q_i \in \mathrm{U}(n_i),\ \Sigma_i \in \mathbb{R}^{n_i \times n_{i+1}}$$

- follows from singular value decomposing $A_i = U_i \Sigma_i V_i^{\mathsf{T}}$ and setting

$$Q_{i+1} = V_i^{\mathsf{T}} U_{i+1}, \quad i = 1, \dots, d-1,$$

with $Q_1 = U_1,\ Q_{d+1} = V_d^{\mathsf{T}}$

graph structure

- take $d = 3$ and denote

$$A_i = [a^{(i)}_{\alpha\beta}], \quad B_j = [b^{(j)}_{\beta\gamma}], \quad C_k = [c^{(k)}_{\gamma\alpha}]$$

- then MPS is

$$\begin{aligned} f &= \sum_{i,j,k=1}^{r_1,r_2,r_3} \operatorname{tr}(A_i B_j C_k)\, \varphi_i \otimes \psi_j \otimes \theta_k \\ &= \sum_{i,j,k=1}^{r_1,r_2,r_3} \Big[\sum_{\alpha,\beta,\gamma=1}^{n_1,n_2,n_3} a^{(i)}_{\alpha\beta} b^{(j)}_{\beta\gamma} c^{(k)}_{\gamma\alpha}\, \varphi_i \otimes \psi_j \otimes \theta_k\Big] \\ &= \sum_{\alpha,\beta,\gamma=1}^{n_1,n_2,n_3} \Big[\sum_{i=1}^{r_1} a^{(i)}_{\alpha\beta}\, \varphi_i\Big] \otimes \Big[\sum_{j=1}^{r_2} b^{(j)}_{\beta\gamma}\, \psi_j\Big] \otimes \Big[\sum_{k=1}^{r_3} c^{(k)}_{\gamma\alpha}\, \theta_k\Big] \\ &= \sum_{\alpha,\beta,\gamma=1}^{n_1,n_2,n_3} \varphi_{\alpha\beta} \otimes \psi_{\beta\gamma} \otimes \theta_{\gamma\alpha} \end{aligned}$$

- where

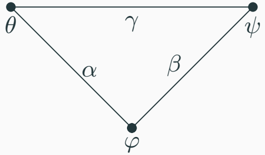
$$\varphi_{\alpha\beta} := \sum_{i=1}^{r_1} a^{(i)}_{\alpha\beta}\, \varphi_i, \quad \psi_{\beta\gamma} := \sum_{j=1}^{r_2} b^{(j)}_{\beta\gamma}\, \psi_j, \quad \theta_{\gamma\alpha} := \sum_{k=1}^{r_3} c^{(k)}_{\gamma\alpha}\, \theta_k$$

graph structure

- so MPS may alternatively be written in the form

$$f = \sum_{\alpha,\beta,\gamma=1}^{n_1,n_2,n_3} \varphi_{\alpha\beta} \otimes \psi_{\beta\gamma} \otimes \theta_{\gamma\alpha}$$

- indices have the incidence structure of an undirected graph, in this case a triangle



- bottom line: any tensor network state is a sum of separable functions indexed by a graph [Landsberg–Qi–Ye, 2012]

common examples

- periodic matrix product states

$$f(\mathbf{x},\mathbf{y},\mathbf{z}) = \sum_{i,j,k=1}^{n_1,n_2,n_3} \varphi_{ij}(\mathbf{x})\psi_{jk}(\mathbf{y})\theta_{ki}(\mathbf{z})$$

- tree tensor network states [Shi–Duan–Vidal, 2006]

$$f(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{w}) = \sum_{i,j,k=1}^{n_1,n_2,n_3} \varphi_{ijk}(\mathbf{x})\psi_{i}(\mathbf{y})\theta_{j}(\mathbf{z})\pi_k(\mathbf{w})$$

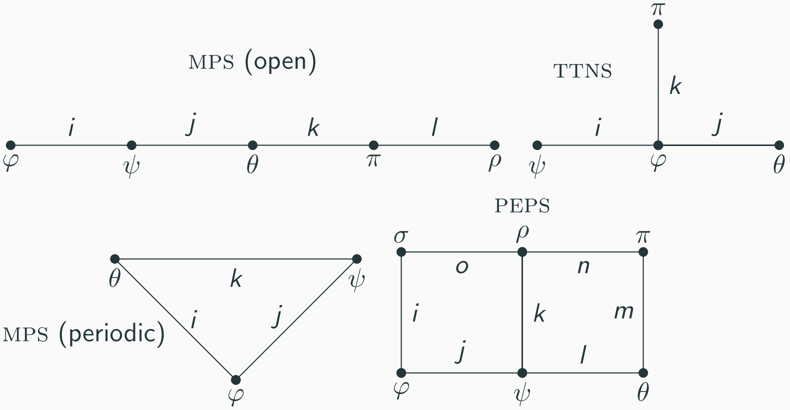
- open matrix product states

$$f(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{v}) = \sum_{i,j,k,l=1}^{n_1,n_2,n_3,n_4} \varphi_{i}(\mathbf{x})\psi_{ij}(\mathbf{y})\theta_{jk}(\mathbf{z})\pi_{kl}(\mathbf{u})\rho_l(\mathbf{v})$$

- projected entangled pair states [Verstrate–Cirac, 2004]

$$f(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{v},\mathbf{w}) = \sum_{i,j,k,l,m,n,o=1}^{n_1,n_2,n_3,n_4,n_5,n_6,n_7} \varphi_{ij}(\mathbf{x})\psi_{jkl}(\mathbf{y})\theta_{lm}(\mathbf{z})\pi_{mn}(\mathbf{u})\rho_{nko}(\mathbf{v})\sigma_{oi}(\mathbf{w})$$

associated graphs



- all tensor network ansätze are sums of separable functions
- differ only in how their factors are indexed

deeper look

Helmholtz equation

- recall Helmholtz equation

$$\Delta f + \omega^2 f = 0,$$

- $n = 2$ in Cartesian and polar coordinates:

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \omega^2 f = 0, \quad \frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} + \omega^2 f = 0$$

- separation-of-variables works for both but entirely different solutions
  - Cartesian:

$$f_k(x,y) = a_1 \,\mathrm{e}^{\mathrm{i}[kx+(\omega^2-k^2)^{1/2}y]} + a_2 \,\mathrm{e}^{\mathrm{i}[-kx+(\omega^2-k^2)^{1/2}y]} + a_3 \,\mathrm{e}^{\mathrm{i}[kx-(\omega^2-k^2)^{1/2}y]} + a_4 \,\mathrm{e}^{\mathrm{i}[-kx-(\omega^2-k^2)^{1/2}y]}$$

  - polar:

$$f_k(r,\theta) = a_1 \,\mathrm{e}^{\mathrm{i}k\theta} J_k(\omega r) + a_2 \,\mathrm{e}^{-\mathrm{i}k\theta} J_k(\omega r) + a_3 \,\mathrm{e}^{\mathrm{i}k\theta} J_{-k}(\omega r) + a_4 \,\mathrm{e}^{-\mathrm{i}k\theta} J_{-k}(\omega r)$$

- for $n = 2$, there are exactly four systems of separable coordinates: Cartesian, polar, parabolic and elliptic

separable coordinates

- for $n = 3$, there are exactly eleven
  (i) Cartesian
  (ii) cylindrical
  (iii) spherical
  (iv) parabolic
  (v) paraboloidal
  (vi) ellipsoidal
  (vii) conical
  (viii) prolate spheroidal
  (ix) oblate spheroidal
  (x) elliptic cylindrical
  (xi) parabolic cylindrical
- how do I know this?

Stäckel condition

- n-dimensional Helmholtz equation in coordinates $x_1, \ldots, x_n$ can be solved using the separation-of-variables technique if and only if
  ① the Euclidean metric tensor g is a diagonal matrix in this coordinate system
  ② if $g = \operatorname{diag}(g_{11}, \ldots, g_{nn})$, then there exists an invertible matrix of the form

$$S = \begin{bmatrix} s_{11}(x_1) & s_{12}(x_1) & \cdots & s_{1n}(x_1) \\ s_{21}(x_2) & s_{22}(x_2) & \cdots & s_{2n}(x_2) \\ \vdots & \vdots & & \vdots \\ s_{n1}(x_n) & s_{n2}(x_n) & \cdots & s_{nn}(x_n) \end{bmatrix}$$

  with

$$g_{jj}^{-1} = (S^{-1})_{1j}, \quad j = 1, \ldots, n$$

- S is called a Stäckel matrix for g in coordinates $x_1, \ldots, x_n$
- note that the ith row of S depends only on the ith coordinate

Stäckel condition for $n = 3$

Euclidean metric in Cartesian, cylindrical, spherical, parabolic coordinates

$$g(x, y, z) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad g(r, \theta, z) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$g(r, \theta, \phi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2 \theta \end{bmatrix}, \quad g(\sigma, \tau, \phi) = \begin{bmatrix} \sigma^2 + \tau^2 & 0 & 0 \\ 0 & \sigma^2 + \tau^2 & 0 \\ 0 & 0 & \sigma^2 \tau^2 \end{bmatrix}$$

Stäckel condition for $n = 3$

Stäckel condition is satisfied with following matrices

Cartesian $$S = \begin{bmatrix} 0 & -1 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \qquad S^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ -1 & -1 & 0 \end{bmatrix}$$

cylindrical $$S = \begin{bmatrix} 0 & -\frac{1}{r^2} & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \qquad S^{-1} = \begin{bmatrix} 1 & \frac{1}{r^2} & 1 \\ 0 & 1 & 0 \\ -1 & -\frac{1}{r^2} & 0 \end{bmatrix}$$

spherical $$S = \begin{bmatrix} 1 & -\frac{1}{r^2} & 0 \\ 0 & 1 & -\frac{1}{\sin^2\theta} \\ 0 & 0 & 1 \end{bmatrix}, \quad S^{-1} = \begin{bmatrix} 1 & \frac{1}{r^2} & \frac{1}{(r^2\sin^2\theta)} \\ 0 & 1 & \frac{1}{\sin^2\theta} \\ 0 & 0 & 1 \end{bmatrix}$$

parabolic $$S = \begin{bmatrix} \sigma^2 & -1 & -\frac{1}{\sigma^2} \\ \tau^2 & 1 & -\frac{1}{\tau^2} \\ 0 & 0 & 1 \end{bmatrix}, \qquad S^{-1} = \begin{bmatrix} \frac{1}{(\sigma^2+\tau^2)} & \frac{1}{(\sigma^2+\tau^2)} & \frac{1}{(\sigma^2\tau^2)} \\ -\frac{\tau^2}{(\sigma^2+\tau^2)} & \frac{\sigma^2}{(\sigma^2+\tau^2)} & \frac{1}{\tau^2-1/\sigma^2} \\ 0 & 0 & 1 \end{bmatrix}$$

Stäckel condition for $n = 3$

- when metric tensor is diagonal, $g = \text{diag}(g_{11}, \ldots, g_{nn})$, Laplacian is

$$\Delta = \sum_{i=1}^{n} \frac{1}{\sqrt{\det(g)}} \frac{\partial}{\partial x_i} \frac{\sqrt{\det(g)}}{g_{ii}} \frac{\partial f}{\partial x_i}$$

- Helmholtz equation in cylindrical, spherical, parabolic coordinates:

$$\frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} + \frac{\partial^2 f}{\partial z^2} + \omega^2 f = 0$$

$$\frac{\partial^2 f}{\partial r^2} + \frac{2}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2 \sin^2\phi}\frac{\partial^2 f}{\partial \theta^2} + \frac{\cos\phi}{r^2 \sin^2\phi}\frac{\partial f}{\partial \phi} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \phi^2} + \omega^2 f = 0$$

$$\frac{1}{\sigma^2 + \tau^2}\left[\frac{\partial^2 f}{\partial \sigma^2} + \frac{1}{\sigma}\frac{\partial f}{\partial \sigma} + \frac{\partial^2 f}{\partial \tau^2} + \frac{1}{\tau}\frac{\partial f}{\partial \tau}\right] + \frac{1}{\sigma^2\tau^2}\frac{\partial^2 f}{\partial \phi^2} + \omega^2 f = 0$$

- not so obvious that these are amenable to separation-of-variables, speaking to the power of the Stäckel condition

general Stäckel condition

- more generally, Stäckel condition can be extended to
  - any higher-order semilinear PDE [Koornwinder, 1980]
  - any Riemannian manifold M [Eisenhart, 1934]
- a system of local coordinates on M is separable if and only if
  ① Riemannian metric tensor g is a diagonal matrix in this coordinate system
  ② Ricci curvature tensor $\bar{R}$ is a diagonal matrix in this coordinate system
  ③ g satisfies the Stäckel condition
- ancient results that have largely been forgotten

thanks again

- Simons Foundation Award no. 663281 granted to the Institute of Mathematics of the Polish Academy of Sciences for 2021–2023
- Arieh for the opportunity to write book chapter
  - L.-H. Lim, "Tensors in computations," *Acta Numer.*, **30** (2021), pp. 555–764
- Arieh and Karolina for the opportunity to give this series of lectures

more thanks

- helpful discussions: K. Conrad, Z. Dai, S. Friedland, E. Gnang, S. Hu, R. Kondor, J. M. Landsberg, B. Nelson, J. Nie, P. McCullagh, Y. Qi, E. Riehl, K. Ye
- image credits: A. Krizhevsky (CIFAR-10), J. Jumper (AlphaFold 2), J. Richardson (protein backbone), University of Göttingen (Voigt), K. Ye (tensor networks), Sanpaz (Cauchy stress), Waveguide2 (Ohm's law), svjo (Hooke's law)
- funding: Defense Advanced Research Projects Agency HR00112190040, National Science Foundation DMS-11854831