Structure-preserving machine learning moment closure models for the radiative transfer equation

Juntao Huang

Texas Tech University

Joint work with Yingda Cheng (MSU), Andrew Christlieb (MSU), Luke Roberts (LANL), and Wen-An Yong (Tsinghua)

SITE Research Center, NYU Abu Dhabi
Nov 29, 2022

Outline

Outline

Background

Gradient-based hyperbolic moment closure
- symmetrizer-based hyperbolic closure
- eigenvalue-based hyperbolic closure

Numerical results

Conclusion

Kinetic equations

- Basic concept: statistical (distribution function) description of a large particle system
  - $f = f(x, v, t)$: particle densities with velocity $v \in \mathbb{R}^3$ (in velocity space) at position $x \in \mathbb{R}^3$ (in physical space) at time t
- Applications: rarefied gas dynamics, plasma physics, nuclear engineering, astrophysics...
- Numerical difficulty: multiscale, high dimension $(x, v, t) \in \mathbb{R}^7$ for $d = 3$

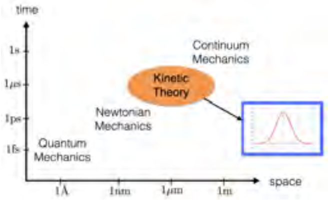

Figure: role of kinetic theory in multiscale modeling

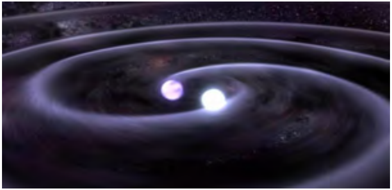

Figure: illustration of two neutron stars merging, along with the resulting gravitational waves, from [NASA/Goddard Space Flight Center]

Radiative transfer equation

- The radiative transfer equation (RTE) describe the propagation of radiation and interaction with a background medium:

$$\partial_t f + \Omega \cdot \nabla f = \sigma_s \langle f \rangle - \sigma_t f \tag{1}$$

  - $f = f(x, \Omega, t)$: specific intensity of radiation
  - $x \in \mathbb{R}^d$: location in physical space
  - $\Omega \in \mathbb{S}^{d-1}$: traveling direction in angular space
  - $\langle \cdot \rangle$ normalized integration in angular space: $\langle f \rangle := \frac{1}{|\mathbb{S}^{d-1}|} \int_{\mathbb{S}^{d-1}} f d\Omega$
  - $\sigma_s(x), \sigma_a(x) \geq 0$: scattering and absorption cross section
  - $\sigma_t(x) = \sigma_s(x) + \sigma_a(x)$: total cross section
- Moment closure:
  - introduced by Grad (1949) for Boltzmann equation
  - model/dimension reduction in angular space (3D + 2V $\rightarrow$ 3D if $d = 3$)
  - focus on evolution of moments of specific intensity, quantities of interest in physics
    - energy density: $\int f(x, \Omega, t) d\Omega$
    - radiation flux: $\int f(x, \Omega, t) \Omega d\Omega$
    - radiation pressure: $\int f(x, \Omega, t) \Omega \otimes \Omega d\Omega$

- The simplified RTE (1D and 1V): $f = f(x, v, t)$, $x \in \mathbb{R}$ and $v \in [-1, 1]$

$$\partial_t f + v\partial_x f = \sigma_s \left(\frac{1}{2} \int_{-1}^{1} f dv - f \right) - \sigma_a f. \tag{2}$$

- Define the k-th order moment by

$$m_k(x,t) = \frac{1}{2} \int_{-1}^{1} f(x,v,t) P_k(v) dv, \quad k \geq 0.$$

with $P_k(v)$ the k-th Legendre polynomial. We derive the unclosed moment equations

$$\begin{aligned} \partial_t m_0 + \partial_x m_1 &= -\sigma_a m_0, \\ \partial_t m_1 + \frac{1}{3}\partial_x m_0 + \frac{2}{3}\partial_x m_2 &= -(\sigma_s + \sigma_a) m_1, \\ &\cdots \\ \partial_t m_N + \frac{N}{2N+1}\partial_x m_{N-1} + \frac{N+1}{2N+1}\partial_x m_{N+1} &= -(\sigma_s + \sigma_a) m_N. \end{aligned} \tag{3}$$

- Traditional closures[1]: P_N closure, M_N closure, empirical assumptions

[1] Chandrasekhar (1944), Levermore (1996), Hauck and McClarren (2010), McClarren and Hauck (2010), Hauck (2011), Alldredge, Hauck and Tits (2012), Alldredge, Hauck, OLeary and Tits (2014), Laboure, McClarren and Hauck (2016)

Machine learning moment closure

- Scientific machine learning (ML): multiscale modeling (molecular dynamics, turbulence, kinetic equations), solving PDEs in high dimensions
- ML closure:
  - Boltzmann BGK model: Han, Ma, Ma and E (2019), learn closure

$$m_{N+1} = \mathcal{N}(m_0, m_1, \cdots, m_N)$$

    with $\mathcal{N}$ a neural network (trained with data generated by kinetic model). Use auto-encoder to learn generalized moments
  - Williams-Boltzmann equation: Scoggins, Han and Massot (2021)
  - plasma physics:
    - Euler-Poisson system: nonlocal closure for heat flux, Bois, Franck, Navoret and Vigon (2020)
    - Hammett–Perkins Landau fluid closure: Ma, Zhu, Xu and Wang (2020), Wang, Xu, Zhu, Ma and Lei (2020), Maulik, Garland, Burby, Tang and Balaprakas (2020)
  - surrogate maximum entropy closure: convex splines and convex neural networks, Porteous, Laiu and Hauck (2021), Schotthöfer, Xiao, Frank and Hauck (2021)
  - preserve Galilean/reflecting/scaling invariance: Li, Dong and Wang (2021)
- Stability? only work for short time, blow up for long time

Hyperbolicity

- ML moment closure model:

$$\begin{aligned}\partial_t m_0 + \partial_x m_1 &= -\sigma_a m_0,\\ \partial_t m_1 + \frac{1}{3}\partial_x m_0 + \frac{2}{3}\partial_x m_2 &= -(\sigma_s + \sigma_a) m_1,\\ &\cdots\\ \partial_t m_N + \frac{N}{2N+1}\partial_x m_{N-1} + \frac{N+1}{2N+1}\partial_x \mathcal{N}(m_0, m_1, \cdots, m_N) &= -(\sigma_s + \sigma_a) m_N.\end{aligned} \tag{4}$$

  write into a system of first-order PDEs:

$$\partial_t \boldsymbol{m} + A(\boldsymbol{m})\partial_x \boldsymbol{m} = \boldsymbol{S}(\boldsymbol{m}) \tag{5}$$

  with $\boldsymbol{m} = (m_0, m_1, \cdots, m_N)^T$.

- Hyperbolicity
  - definition: The system (5) is hyperbolic if $A(\boldsymbol{m})$ is real diagonalizable
  - Hyperbolicity is crucial for long-time stability of the model!
  - difficulty: $A(\boldsymbol{m})$ depend on neural network, generally NOT real diagonalizable

Outline

Gradient-based closure

- Motivation: derive exact closure in free streaming limit with isotropic initial conditions

$$\begin{aligned} \partial_t f + v\partial_x f &= 0, \\ f(t=0) &= f_0(x). \end{aligned} \tag{6}$$

- Key idea: instead of learning a relation

$$m_{N+1} = \mathcal{N}(m_0, m_1, \cdots, m_N), \tag{7}$$

we propose to directly learn the gradient of the unclosed moment:

$$\partial_x m_{N+1} = \sum_{k=0}^{N} \mathcal{N}_k(m_0, m_1, \ldots, m_N)\partial_x m_k. \tag{8}$$

- Advantages:
  - accuracy: more accurate in optically thin regime (far away from equilibrium)
  - mathematical structure: more degrees of freedom to play with, enforce hyperbolicity and other properties

Outline

Approach 1: symmetrizer-based hyperbolic closure

- Gradient-based ML moment closure model:

$$\partial_t \boldsymbol{m} + A(\boldsymbol{m})\partial_x \boldsymbol{m} = \boldsymbol{S}(\boldsymbol{m}) \tag{9}$$

with $\boldsymbol{m} = (m_0, m_1, \cdots, m_N)$ and

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} & 0 & \dots & 0 \\ 0 & \frac{2}{5} & 0 & \frac{3}{5} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \frac{N-1}{2N-1} & 0 & \frac{N}{2N-1} \\ a_0 & a_1 & \dots & a_{N-2} & a_{N-1} & a_N \end{pmatrix}$$

with a_j related to neural networks:

$$a_j = \begin{cases} \frac{N+1}{2N+1}\mathcal{N}_j, & j \neq N-1 \\ \frac{N}{2N+1} + \frac{N+1}{2N+1}\mathcal{N}_j, & j = N-1 \end{cases}$$

- Key idea: seek a symmetric positive definite matrix A_0 (also called a symmetrizer) such that $A_0 A$ is symmetric $\Rightarrow$ symmetrizable hyperbolic
- Diffusion limit: ML closure model $\rightarrow$ diffusion limit of RTE (as $Kn \rightarrow 0$)

Theorem (symmetrizable hyperbolic)

Consider matrix $A \in \mathbb{R}^{(N+1)\times(N+1)}$ with $N \geq 3$ and $a_i = 0$ for $i = 0, 1, \cdots, N-4$. If the coefficients a_i for $i = N-3, N-2, N-1, N$ satisfy the following constraints:

$$a_{N-3} > -\frac{(N-1)(N-2)}{N(2N-3)}, \quad a_{N-1} > \frac{g(a_{N-3}, a_{N-2}, a_N; N)}{(N-2)(a_{N-3}(2N-3)N + (N-1)(N-2))^2} \tag{10}$$

where $g = g(a_{N-3}, a_{N-2}, a_N; N)$ is a function given by

$$g = a_{N-3}^3(N-1)N^2(3-2N)^2 + a_{N-2}(2N-1)(N-2)^3(a_{N-2}N - a_N(N-1)) + a_{N-3}(N-2)^2(a_N(4N^2-8N+3)(a_{N-2}N - a_N(N-1)) + (N-1)^3) + 2a_{N-3}^2(N-1)^2N(2N-3)(N-2),$$

then there exist a SPD matrix $A_0 = diag(D, B) \in \mathbb{R}^{(N+1)\times(N+1)}$ such that $A_0 A$ is symmetric. Here, $D = diag(1, 3, 5, \cdots, 2N-5) \in \mathbb{R}^{(N-2)\times(N-2)}$ and $B \in \mathbb{R}^{3\times3}$ is a SPD matrix.

Outline

Approach 2: eigenvalue-based hyperbolic closure

Key idea: using the algebraic structure of A (lower Hessenberg matrix), we relate the real diagonalizablity to the roots of its associated polynomial.

Theorem (hyperbolicity and physical characteristic speeds)

For the coefficient matrix A, the associated polynomial sequence defined in Elouafi and Hadj (2009) satisfies:

$$q_{N+1}(x) = \frac{N+1}{2N+1} P_{N+1}(x) + \frac{N}{2N+1} P_{N-1}(x) - \sum_{k=0}^{N} a_k P_k(x), \tag{11}$$

where $P_n(x)$ denotes the Legendre polynomial of degree n.

1. *If all the roots of $q_{N+1}(x)$ are simple, then the characteristic polynomial of A is:*

$$\det(xI_{N+1} - A) = \frac{N!}{(2N-1)!!} \left(\frac{N+1}{2N+1} P_{N+1}(x) + \frac{N}{2N+1} P_{N-1}(x) - \sum_{k=0}^{N} a_k P_k(x) \right)$$

2. *If further assuming all the roots of $q_{N+1}(x)$ are simple, real and lie in the interval $[-1, 1]$, then the moment closure system is strictly hyperbolic with physical characteristic speeds.*

- Two new neural network architectures:
  - Step 1: represent eigenvalues with fully connected neural networks
  - Step 2: mapping from eigenvalues to coefficient matrix

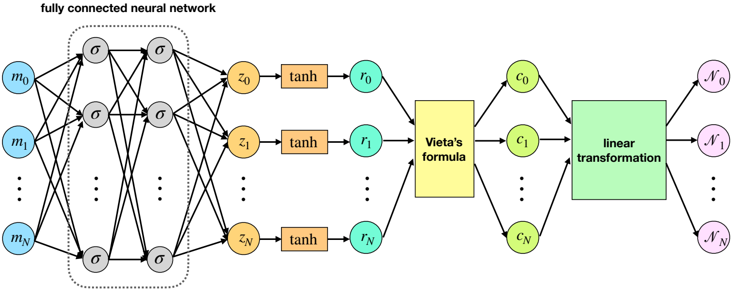

Figure: weakly hyperbolic with physical characteristic speeds

Outline

Training data

We numerically solve the RTE to generate training data:

- unit interval $[0, 1]$ in the physical domain with periodic boundary conditions
- initial conditions, truncated Fourier series[2]:

$$f_0(x, v) = a_0 + \sum_{k=1}^{k_{\max}} a_k \sin(2k\pi x + \phi_k), \tag{12}$$

  $k_{\max} = 10$, $a_k \sim U(-\frac{1}{k}, \frac{1}{k})$ for $k \geq 1$, $\phi_k \sim U[0, 2\pi]$, $a_0 = c + \sum_{k=1}^{k_{\max}} \frac{1}{k}$ with $c \sim U[0, 1]$
- $\sigma_s \sim U[0.1, 100]$ and $\sigma_a \sim U[0, 10]$, constants over the domain
- take 100 different initial data
- space time DG method[3]: $N_x = 512$, $\Delta t = 8\Delta x$, $T = 1$

[2] Ma, Zhu, Xu and Wang (2020), Bois, Franck, Navoret and Vigon (2020)

[3] Crockatt, Christlieb, Garrett and Hauck (2017)

Comparison with non-hyperbolic ML closure and P_N closure

- Two-material problem: $N = 6$ at $t = 0.5$ and $t = 1$. Gray part: optically thin regime; other part: intermediate regime.

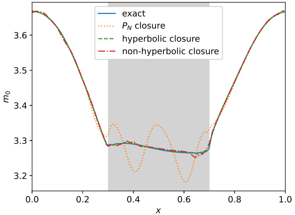

Figure: m_0 *at* $t = 0.5$

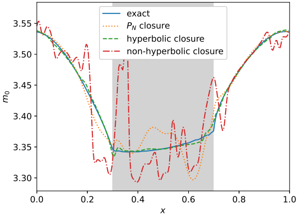

Figure: m_0 *at* $t = 1$

Comparison:

- P_N closure: stable, but not accurate in optically thin regime
- Non-hyperbolic ML closure: accurate for short time, but blow up for long time
- Hyperbolic ML closure: stable and accurate for long time

Higher order moments (ML, P_N and filtered P_N closure)

- Two-material problem: $N = 9$ at $t = 0.5$.

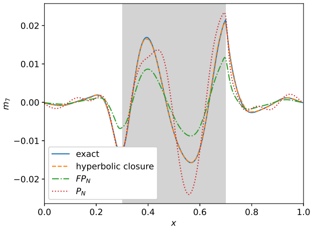

Figure: m_7 **at** $t = 0.5$

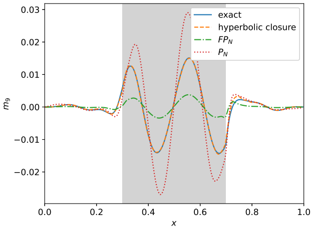

Figure: m_9 **at** $t = 0.5$

Hyperbolic ML closure in different regimes

- Relative L^2 error vs. scattering coefficient σ_s
  - hyperbolic ML closure is stable and more accurate than P_N closure

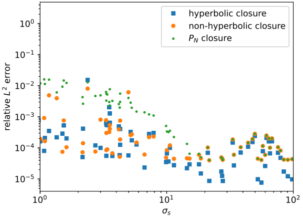

Figure: error of m_0 at $t = 0.5$

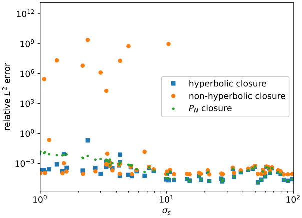

Figure: error of m_0 at $t = 1$

Computational cost with traditional closure

	ML ($N = 6$)	P_6	P_9	P_{15}	P_{18}
relative L^2 error	4.69e-4	9.66e-3	3.78e-3	6.87e-4	4.54e-4
computational time (sec)	1.94e-2	1.61e-2	2.13e-2	3.43e-2	6.40e-2

Table: two-material problem. Comparison between the computational time per time step and the relative L^2 error of m_0 for the hyperbolic ML closure with $N = 6$ and the P_N closure with $N = 6, 9, 12, 15, 18$ at $t = 0.5$.

Outline

- Our contributions
  - learning gradient moment closure
  - approach 1: symmetrizer-based hyperbolic closure, diffusion limit
  - approach 2: eigenvalue-based hyperbolic closure, physical characteristic speeds

- Future work
  - Generalization to multi-dimension
  - Generalization to other models: Boltzmann equations, Vlasov Maxwell/Poisson equations
  - Preserve other properties (realizability, rotational symmetry)
  - Boundary conditions
  - Applications in astrophysics and plasma simulations...

References:

- J. Huang, Y. Cheng, A. J. Christlieb, and L. F. Roberts. Journal of Computational Physics. 453:110941, 2022.
- J. Huang, Y. Cheng, A. J. Christlieb, L. F. Roberts, and W.-A. Yong. arXiv:2105.14410, accepted by Multiscale Modeling and Simulation, 2021.
- J. Huang, Y. Cheng, A. J. Christlieb, and L. F. Roberts. arXiv:2109.00700, accepted by Journal of Scientific Computing, 2021.

The END!

Thank You!