*Computer Assisted Methods in Engineering and Science*
**32**(4): 385–408, 2025, https://doi.org/10.24423/cames.2025.1960

# Application of Neural Networks for Determining the Radiation Pressure in Two-Moment Radiation Hydrodynamics in Slab Geometry

Agnieszka PREGOWSKA*, Wiesław LARECKI,
Janusz SZCZEPANSKI

*Institute of Fundamental Technological Research, Polish Academy of Sciences,*
*Pawińskiego 5B, 02-106 Warsaw, Poland*

* *Corresponding Author e-mail: aprego@ippt.pan.pl*

In two-moment radiation transport, the closure is the constitutive relation that maps the energy and momentum to the radiation-pressure tensor. Among available closures, the maximum-entropy (ME) approach is the most reliable. However, it is associated with a high computational cost. In this paper, we propose a machine-learning approach for the rapid evaluation of the ME closure for bosonic, classical and fermionic radiation. We generate ME reference data using Gauss–Legendre angular quadrature combined with a robust bisection-based inversion of the moment constraints. Next, we train a small physics-constrained multilayer perceptron (MLP) with output restricted to the physically admissible range (between one third and one). Monotonicity in the reduced flux is enforced, and the derivative is matched to the ME reference. The neural network (NN)-based closure achieves a mean absolute error (MAE) of $9.0 \times 10^{-4}$ over the range $\phi \in [0, 0.98]$, which yields a latency reduction of about $\sim 10^3\times$ per closure evaluation. In the Marshak wave benchmark the full simulation runs about 247 × faster while the hyperbolicity indicator remains strictly positive (minimum $5.1 \cdot 10^{-2}$). Compared with the analytic Kershaw closure for bosonic and classical radiation, our model is substantially more accurate and faster. For practical adoption, we also provide a lightweight rational approximation (MAE $1.01\times \sim 10^{-3}$), and in the bosonic and classical cases we confirm positivity of the hyperbolicity indicator for degeneracy parameters between $-5$ and 5.

**Keywords:** radiation transport, Eddington factor, maximum entropy, neural surrogate, hyperbolicity, Marshak problem.

## 1. Introduction

Moment models of radiative transport are a standard tool for reducing the Boltzmann equations to systems of hyperbolic moment equations with a small number of macroscopic variables [1]. A key element of such models is the closure, a relation connecting higher-order moments with lower-order moments [2]. In the two-moment model, the closure is a constitutive relation expressed in terms of an Eddington factor, which uniquely determines the pressure tensor. This factor increases with the so-called reduced flux, ensuring the system remains hyperbolic [3–5]. Among the possible closures, a special place is played by the maximum-entropy (ME) closure [2]. It is obtained by maximizing the entropy under given constraints [4], which ensures realizability and the hyperbolicity of the system of the radiation hydrodynamics equations [2].

In a series of papers, several authors investigated the consistent derivations and properties of 1D ME closures for bosonic and fermionic radiation, including the two- and three-moment theories [6–8]. The proper choice of closure is important not only for accuracy but also for well-posed dynamics and the absence of nonphysical instabilities. However, a practical problem remains: the numerical cost of closing the ME [6, 7, 9]. In each cell and at every time step, one must solve an inverse problem to obtain the Lagrange multipliers and then compute the integrals. For comparison, the simple analytic Kershaw closure is employed [6, 10, 11]. It is fast but can significantly deviate from the ME closure and alter the wave properties of the system.

In recent years, a natural compromise between accuracy and computation cost has emerged: learned surrogates, which replace expensive input-output maps with a fast statistical model [12, 13]. In radiative transport, the simplest and most crucial map is the relation from the field directivity measure to the Eddington factor. In the fermionic variant, this map additionally depends on the degeneracy parameter. If the surrogate provides accuracy comparable to the ME closure, preserves basic physical properties (allowable range of values, monotonicity, and preservation of hyperbolicity [1, 13]) and enables speedups of hundreds to thousands of times, then in practice it can replace the ME closure in a solver without losing the quality of solutions. This approach is consistent with recent trends, ranging from learning moment closures with control over hyperbolicity and characteristic velocities [12, 13], to data-driven models of the variable tensor/Eddington factor in thermal radiative transfer [14], and to contemporary developments and tests of moment schemes in astrophysics [9, 15, 16].

In this paper, we propose an application of neural networks (NNs) for the fast calculation of the Eddington factor in the two-momentum transport description. We generate high-accuracy ME reference data and then train a compact neural surrogate on this dataset. Next, we train a small physics-constrained MLP with output restricted to monotonicity and derivative matching. Motivated by evidence that compact neural models can exhibit high information efficiency in other domains [17, 18], we employ a small architecture and investigate whether it can serve as a fast, physics-consistent approximation to the ME closure. For clarity and completeness of the paper, a brief characterization of two-moment radiation hydrodynamics is included.

As for the unit system, throughout this paper, we set $c = \hbar = k_B = 1$, where $c$ is the speed of light, $\hbar$ is the Planck constant divided by $2\pi$, and $k_B$ is the Boltzmann constant. The summation convention over repeated upper and lower indices is employed.

## 2. Preliminaries and theoretical background

The kinetic theory of radiation describes radiation as a gas of massless particles [19]. Each gas particle is equipped with a wave–vector $\mathbf{k} \in \mathbb{R}^3$, identified with a particle momentum, and with a particle energy $\omega = |\mathbf{k}|$, identified with its frequency. The state of a gas is represented by the one-point distribution function (phase density) $f(t, \mathbf{x}, \mathbf{k})$, where $t \in \mathbb{R}^+$ is the time and $\mathbf{x} = (x^1, x^2, x^3) \in \mathbb{R}^3$ is the position in the assumed laboratory frame.

The evolution of $f$ is governed by the kinetic equation:

$$\partial_t f + k^i \partial_i f = C[f], \tag{1}$$

where $\partial_t = \frac{\partial}{\partial t}$, $\partial_i = \frac{\partial}{\partial x^i}$, $i = 1, 2, 3$, and $C[f]$ is the collision term. The unit vector in the direction of $\mathbf{k}$ is denoted by $\mathbf{g} = \frac{\mathbf{k}}{|\mathbf{k}|} \in S^2$, $g^i = \frac{k^i}{|\mathbf{k}|}$.

There are two basic formulations of the macroscopic description of radiation. The spectral (frequency-dependent) formulation treats the particle frequency $k$ (equivalent to particle energy) as a fixed parameter. Hence, macroscopic quantities at $(t, \mathbf{x})$ are defined as weighted integrals (angular moments) of the angular distribution of particles of the same energy. In the gray (frequency-integrated) formulation, macroscopic quantities are weighted integrals (moments) of $f(t, \mathbf{x}, \mathbf{k})$ over $\mathbf{k} \in \mathbb{R}^3$. It is convenient to decompose this integration into radial integration over $k \in \mathbb{R}^+$ and angular integration over $\mathbf{g} \in S^2$ (spherical coordinates), i.e.,

$$\mathrm{d}^3 k = k^2 \,\mathrm{d}k \,\mathrm{d}^2 g, \tag{2}$$

and to define the spectral radiation intensity

$$\widetilde{I}(t, \mathbf{x}, k, \mathbf{g}) := k^3 f(t, \mathbf{x}, k\mathbf{g}), \tag{3}$$

and the angular radiation intensity

$$\widehat{I}(t, \mathbf{x}, \mathbf{g}) := \int_0^\infty k^3 f(t, \mathbf{x}, k\mathbf{g}) \,\mathrm{d}k = \int_0^\infty \widetilde{I}(t, \mathbf{x}, k, \mathbf{g}) \,\mathrm{d}k. \tag{4}$$

Multiplying kinetic equation (1) by $k^3$, we obtain the spectral radiative equation of transfer for $\widetilde{I}(t,\mathbf{x},k,\mathbf{g})$. Similarly, multiplying Eq. (1) by $k^3$ and integratinh over $k \in \mathbb{R}^+$ results in the gray radiative equation of transfer for $\widehat{I}(t,\mathbf{x},\mathbf{g})$. These equations of transfer take the form:

$$\partial_t I + g^i\,\partial_i I = \overline{C}, \tag{5}$$

where $\overline{C} = k^3 C[k^{-3}\widetilde{I}]$ for $I = \widetilde{I}$, $\overline{C} = \int\limits_0^\infty k^3 C[\widehat{I}]\,\mathrm{d}k$ for $I = \widehat{I}$. The simplest two-moment formulation of radiation hydrodynamics involves the energy density $E$, the momentum $\mathbf{F}$, and the radiation pressure tensor $\mathbf{P}$. In view of Eqs. (2)–(5) these quantities are defined as the following spectral angular moments in the spectral formulation [3, 4, 7, 20]:

$$\widetilde{E}(t,\mathbf{x},k) := \int\limits_{S^2} kf(t,\mathbf{x},k\mathbf{g})\,k^2\,\mathrm{d}^2 g = \int\limits_{S^2} \widetilde{I}(t,\mathbf{x},k,\mathbf{g})\,\mathrm{d}^2 g, \tag{6}$$

$$\widetilde{F}^i(t,\mathbf{x},k) := \int\limits_{S^2} kg^i f(t,\mathbf{x},k\mathbf{g})\,k^2\,\mathrm{d}^2 g = \int\limits_{S^2} g^i \widetilde{I}(t,\mathbf{x},k,\mathbf{g})\,\mathrm{d}^2 g, \tag{7}$$

$$\widetilde{P}^{ij}(t,\mathbf{x},k) := \int\limits_{S^2} kg^i g^j f(t,\mathbf{x},k\mathbf{g})\,k^2\,\mathrm{d}^2 g = \int\limits_{S^2} q^i g^j \widetilde{I}(t,\mathbf{x},k,\mathbf{g})\,\mathrm{d}^2 g. \tag{8}$$

In the gray (frequency-integrated) formulation, they read [21, 22]:

$$\widehat{E}(t,\mathbf{x}) := \int\limits_{\mathbb{R}^3} kf(t,\mathbf{x},\mathbf{k})\,\mathrm{d}^3 k = \int\limits_{S^2} \widehat{I}(t,\mathbf{x},\mathbf{g})\,\mathrm{d}^2 g, \tag{9}$$

$$\widehat{F}^i(t,\mathbf{x}) := \int\limits_{\mathbb{R}^3} kg^i f(t,\mathbf{x},\mathbf{k})\,\mathrm{d}^3 k = \int\limits_{S^2} g^i \widehat{I}(t,\mathbf{x},\mathbf{g})\,\mathrm{d}^2 g, \tag{10}$$

$$\widehat{P}^{ij}(t,\mathbf{x}) := \int\limits_{\mathbb{R}^3} kg^i g^j f(t,\mathbf{x},\mathbf{k})\,\mathrm{d}^3 k = \int\limits_{S^2} g^i g^j \widehat{I}(t,\mathbf{x},\mathbf{g})\,\mathrm{d}^2 g. \tag{11}$$

Multiplying the radiative equation of transfer (5) by the weights $\{1, g^i\}$, performing the angular integration over $\mathbf{q} \in S^2$ and taking into account Eqs. (6)–(11), we arrive at the system of two-moment equations:

$$\partial_t E + \partial_i F^i = S, \tag{12}$$

$$\partial_t F^i + \partial_j P^{ij} = S^i. \tag{13}$$

For the spectral formulation, we put $(\widetilde{E}, \widetilde{F^i}, \widetilde{P^{ij}}, \widetilde{S}, \widetilde{S^i})$ in place of $(E, F^i, P^{ij}, S, S^i)$ in Eqs. (12) and (13). Replacing $(E, F^i, P^{ij}, S, S^i)$ with $(\widehat{E}, \widehat{F^i}, \widehat{P^{ij}}, \widehat{S}, \widehat{S^i})$ in Eqs. (12) and (13), we have the gray (frequency-integrated) system of moment equations. Here, $(\widetilde{S}, \widehat{S})$ and $(\widetilde{S^i}, \widehat{S^i})$ are the angular moments of $\overline{C}$ taken with weights 1 and $g^i$, respectively.

The moment equations (12) and (13), which may be considered to be either spectral (frequency-dependent) or gray (frequency-integrated), require a closure relation that relates $P^{ij}$ to $E$ and $F^i$ in order to obtain a closed system of equations for $E$ and $F^i$.

It follows from Eqs. (6), (8), (9) and (11) that

$$P^i{}_i = E, \tag{14}$$

since $\underline{g}$ is a unit vector. Assuming that the sought closure relation is a sufficiently smooth function $P^{ij}(E, F^k)$, and employing Eq. (5), along the representation of an isotropic tensor function, we can use the following general form of the closure relation:

$$P^{ij} = E\left[\frac{1 - f(E, |\mathbf{F}|)}{2}\delta^{ij} + \frac{3f(E, |\mathbf{F}|) - 1}{2}\frac{F^i F^j}{|\mathbf{F}|^2}\right]. \tag{15}$$

The scalar function $f(E, |\mathbf{F}|)$ is called the variable Eddington factor. Then, the term Eddington tensor is used for the normalized counterpart of $P^{ij}$ [3]:

$$p^{ij} := E^{-1}P^{ij} = \frac{1}{2}\left[(1 - f)\delta^{ij} + (3f - 1)\frac{F^i F^j}{|\mathbf{F}|^2}\right]. \tag{16}$$

In the case of two-moment radiation hydrodynamics, the so-called ME closure plays a special role. It involves the distribution $F(\omega)$ that maximizes the entropy functional under the constraints corresponding to a fixed value of energy $E$ and momentum $\mathbf{F}$.

Specifically, entropy functionals of the following form are considered:

$$S[f] := \int\limits_D \Big[\tau(1 + \tau f)\ln(1 + \tau f) - f(\tau(x) - 1 + \ln f)\Big]\,\mathrm{d}^m L, \tag{17}$$

where $\tau = 0, +1, -1$ corresponds to Maxwell–Boltzmann, Bose–Einstein and Fermi–Dirac statistics, respectively.

The maximization procedure involves Lagrange multipliers $\alpha$ and $a_i$ corresponding to the constraints $E$ and $F_i$, respectively, and concerns the functional

$$\Upsilon[f] := S[f] + \alpha\left(E - \int\limits_D \lambda\, f\,\mathrm{d}^m L\right) + a_i\left(F^i - \int\limits_D \lambda\, g^i\, f\,\mathrm{d}^m L\right). \tag{18}$$

Variation of $\Upsilon(f)$ leads to the formula

$$\delta\Upsilon[f] = \int\limits_D \left[\ln\!\left(\frac{1+\tau f}{f}\right) - \omega\right] \delta f \, \mathrm{d}^m L, \tag{19}$$

which gives the solution

$$f(\omega) = \frac{1}{e^{\omega} - \tau}, \tag{20}$$

where $\omega = (\alpha + a_i \, g^i)\lambda$.

These solutions reproduce the constraints and thus make the Lagrange multipliers $\alpha$ and $a_i$ implicit functions of $E$ and $F^i$. Resolving this relation and substituting $f(\omega)$ into the moment integral formula for $P^{ij}$ provides the closure relation. We set $D = \mathbb{S}^2$, $m = 2$, $L = \mathbf{g}$, $\lambda = 1$, $E = \widetilde{E}$ and $F^i = \widetilde{F}^{\,i}$ in Eqs. (17)–(20) for the spectral formulation, while we set $D = \mathbb{R}^3$, $m = 3$, $L = \mathbf{k}$, $\lambda = k$, $E = \widehat{E}$ and $F^i = \widehat{F}^{\,i}$ for the gray formulation.

In the bosonic case for the gray formulation, the maximum entropy closure is:

$$\begin{aligned} \widehat{P}^{\,ij}\left(\widehat{E}, F^k\right) := {}& \frac{1}{3}\left(\sqrt{4\widehat{E}^2 - 3|\widehat{\mathbf{F}}|^2 - \widehat{E}}\,\right)\delta^{ij} \\ &+ 3\left(\sqrt{4\widehat{E}^2 - 3|\widehat{\mathbf{F}}|^2 + 2\widehat{E}}\,\right)^{-1} \widehat{F}^{\,i}\widehat{F}^{\,j}. \end{aligned} \tag{21}$$

It was derived in [21], and it is proved in [22] that this formula also applies to the gray formulation of fermionic and classical radiations. The radiation pressure tensor (21) was also proposed in [5] motivated by the covariance of the strain-energy tensor and the isotropy of radiation in some inertial frame. In [3, 4] and [20], the ME closures in the spectral formulation were investigated. In contrast to the gray case, they are different for the different statistics. None of the Maxwell–Boltzmann, Bose–Einstein and Fermi–Dirac spectral maximum entropy closures is expressible in terms of elementary functions. For computational purposes, series expansions and approximate formulae were developed. The simplest exact closure relation for the Maxwell–Boltzmann statistics is

$$\widetilde{P}^{\,ij}(\widetilde{E}, \widetilde{\mathbf{F}}) = \frac{|\widetilde{\mathbf{F}}|}{L^{-1}\big(|\widetilde{\mathbf{F}}|/\widetilde{E}\big)}\,\delta^{ij} + \left[\widetilde{E} - 3\,\frac{|\widetilde{\mathbf{F}}|}{L^{-1}\big(|\widetilde{\mathbf{F}}|/\widetilde{E}\big)}\right]\frac{\widetilde{F}^{\,i}\widetilde{F}^{\,j}}{|\widetilde{\mathbf{F}}|^2}. \tag{22}$$

Here $L^{-1}(\,\cdot\,)$ denotes the inverse of the Langevin function,

$$L(x) := \coth x - x^{-1}, \tag{23}$$

and

$$L^{-1}(y) \approx \frac{15y}{(1-y)\left(5 + 5y + 2y^2 + 3y^3\right)} \tag{24}$$

is an empirical approximation proposed in [4]. In view of Eqs. (15) and (16), the closure relations of the bosonic and fermionic ME spectral hydrodynamics are entirely determined by the variable Eddington factor $f\Big(\widetilde{E},\, |\widetilde{\mathbf{F}}|\Big)$.

For the bosonic case, it takes the following form:

$$f_{\mathrm{BE}} = \frac{1}{3} + \frac{2(1+\widetilde{E})(1+2\widetilde{E})}{3}\,\alpha_{\mathrm{BE}}\left(\widetilde{E},\, \frac{|\widetilde{F}|}{E}\right), \tag{25}$$

$$f_{\mathrm{FD}} = \frac{1}{3} + \frac{2(1-\widetilde{E})(1-2\widetilde{E})}{3}\,\alpha_{\mathrm{FD}}\left(\widetilde{E},\, \frac{|\widetilde{F}|}{(1-\widetilde{E})\,E}\right) \tag{26}$$

for the fermionic radiation. In [4], it is assumed that both $\alpha_{\mathrm{BE}}$ and $\alpha_{\mathrm{FD}}$ do not depend on $\widetilde{E}$ and can be expressed in terms of $L^{-1}(x)$. However, as shown in [20], this assumption is not valid. Nevertheless, the results of [4] may serve as a good approximation. Approximate formulae for $\alpha_{\mathrm{BE}}$ and $\alpha_{\mathrm{FD}}$ can also be found in [20].

The simplest setting for radiation hydrodynamics is slab geometry. We introduce a fixed Cartesian frame $(x^1, x^2, x^3)$ and assume that the radiation depends only on $x := x^1$ and is rotationally symmetric with respect to the $O_{x^1}$-axis.

It is convenient to employ spherical coordinates with the following basis: $\{g^1, g^2, g^3\} = \{\cos\Theta, \sin\Theta\sin\varphi, \sin\Theta\cos\varphi\}$, where $\Theta \in [0,\pi]$, $\varphi \in [0, 2\pi]$, and $\mu = \cos\Theta \in [-1, 1]$.

The differential solid angle element is $\mathrm{d}^2\underline{g} = \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi = -\,\mathrm{d}\,(\cos\theta)\,\mathrm{d}\varphi = -\,\mathrm{d}\mu\,\mathrm{d}\varphi$. The azimuthal angle $\Theta$ is the angle between a photon wave vector $\mathbf{k} = k\,\mathbf{g}$ (on the axis $O_{x^1}$). Rotational symmetry implies independence of the polar angle $\varphi$. In this case, the distribution function is $f(t, x, k, \mu)$ and, according to Eqs. (3) and (4), the radiation intensities assume the form:

$$\widetilde{I}(\mu) := \widetilde{I}(t, x, k, \mu) = k^3 f(t, x, k, \mu), \tag{27}$$

$$\widehat{I}(\mu) := \widehat{I}(t, x, \mu) := \int\limits_0^\infty k^3 f(t, x, k, \mu)\,\mathrm{d}k. \tag{28}$$

It follows from Eqs. (6)–(15) and (16) that the moments read:

$$E := 2\pi \int\limits_{-1}^{1} I(\mu)\,\mathrm{d}\mu, \tag{29}$$

$$F := F^1 = 2\pi \int\limits_{-1}^{1} \mu\, I(\mu)\,\mathrm{d}\mu, \quad F^2 = 0, \quad F^3 = 0, \tag{30}$$

$$P := P^{11} = 2\pi \int_{-1}^{1} \mu^2 \, I(\mu) \, \mathrm{d}\mu, \qquad P^{22} = P^{33} = \frac{1}{2}E - \frac{1}{2}P^{11}, \quad P^{ij} = 0 \quad \text{for} \quad i \neq j, \tag{31}$$

where $\widetilde{I}(\mu)$ and $\widehat{I}(\mu)$ stand for $I$ in the spectral and gray formulations, respectively.

The ME closures specified to the case of a slab geometry are obtained simply by substituting Eqs. (27) and (29) into the general three-dimensional formulae. For the gray formulation, formula (21) simplifies to the form:

$$\widehat{P}(\widehat{E}, \widehat{F}) = \widehat{E} \left( \frac{5}{3} - \frac{2}{3} \sqrt{4 - 3 \left( \frac{\widehat{F}}{\widehat{E}} \right)^2} \right), \tag{32}$$

and it holds for all radiation statistics. Also, the approximate formulae for the spectral ME closures significantly simplify when Eq. (29) is taken into account, since

$$\widetilde{P}(\widetilde{E}, \widetilde{F}) = f(\widetilde{E}, |\widetilde{F}|), \tag{33}$$

in view of Eq. (15). Substituting $E = [F, 0, 0]$ into the formulae for classical, bosonic, and fermionic approximate spectral Eddington factors yields the sought closures.

Besides the ME closures, the Kershaw closures ([10, 11] for the bosonic case, and [23] for the fermionic case) are of particular interest since they ensure moment realizability and imply the domain of admissible values for the primitive moments $E$ and $F$.

In the case of bosonic and classical radiation, $f(t, x, k, \mu) \geq 0$ and, according to Eqs. (26) and (27), $\widetilde{I}(\mu) \geq 0$ and $\widehat{I}(u) \geq 0$. Since $E$, $F$, and $P$ are the first three moments of $I(u)$, the solution of the truncated Hausdorff moment problem determines the lower and the upper bounds of $E$, $F$, and $P$ [10, 11]:

$$E > 0, \tag{34}$$

$$F_L(E) := -E < F < E =: F_U(E), \tag{35}$$

$$P_L(E, F) := E^{-1} F^2 < P < E =: P_U(E, F). \tag{36}$$

Then, the closure $P(E, F)$ is sought as a convex combination:

$$P(E, F) = \xi \, P_L(E, F) + (1 - \xi) \, P_U(E, F), \qquad \xi \in [0, 1]. \tag{37}$$

In order to determine the value of the parameter $\xi$, the condition that Eq. (36) recovers the equilibrium radiation pressure $P_{\mathrm{eq}} := \frac{1}{3}E = P(E,0)$ is imposed [10]. Then, the classical and bosonic Kershaw closure reads

$$P(E,F) = \frac{1}{3}\,E\left[1 + 2\left(\frac{F}{E}\right)^2\right], \tag{38}$$

which applies to both spectral and gray formulations.

It should be noted that inequalities (34) and (36) determine an open convex domain of admissible values for the primitive fields $E$ and $F$ in one-dimensional classical and bosonic radiation hydrodynamics.

The fermionic case is significantly different because the Pauli exclusion principle imposes bounds on the distribution function $f(\mu) := f(t,x,k,\mu)$,

$$0 \le f(\mu) \le 1. \tag{39}$$

As a consequence, the problem of moment realizability can be exploited only for the spectral formulation (fixed $k,\ 0 < k < +\infty$) of fermionic radiation hydrodynamics [7, 23]. In this case, the lower and the upper bounds on the moments follow from the solution of the truncated Markov moment problem [7, 23]:

$$0 < \widetilde{E} < 1, \tag{40}$$

$$F_L(\widetilde{E}) := -\widetilde{E}(1-\widetilde{E}) < \widetilde{F} < \widetilde{E}(1-\widetilde{E}) =: F_U(\widetilde{E}), \tag{41}$$

$$P_L(\widetilde{E},\widetilde{F}) := \frac{1}{3}\widetilde{E}^3 + \frac{\widetilde{F}^2}{\widetilde{E}} < \widetilde{P} < \widetilde{E}\left(1 - \widetilde{E} + \frac{\widetilde{E}^2}{3}\right) - \frac{\widetilde{F}^2}{1-\widetilde{E}}\,, \tag{42}$$

which determine the convex domain of admissible values of $(\widetilde{E},\widetilde{F})$. The spectral fermionic Kershaw closure is obtained similarly, as a convex combination of $\widetilde{P}_L(\widetilde{E},\widetilde{F})$ and $\widetilde{P}_U(\widetilde{E},\widetilde{F})$ under the same condition $\widetilde{P}(\widetilde{E},0) = \frac{1}{3}\widetilde{E}$ [23]:

$$\widetilde{P}(\widetilde{E},\widetilde{F}) = \frac{\widetilde{E}}{3} + \frac{2(1-2\widetilde{E})}{3(1-\widetilde{E})\widetilde{E}}\,\widetilde{F}^2. \tag{43}$$

It can be rearranged into a form similar to the spectral fermionic maximum-entropy closure:

$$\widetilde{P}^K(\widetilde{E},z) := \frac{\widetilde{E}}{3} + \frac{2}{3}\,\widetilde{E}(1-\widetilde{E})(1-2\widetilde{E})\,\Theta_K(z), \tag{44}$$

where

$$\Theta_K(z) = z^2, \qquad z := \frac{\widetilde{F}}{\widetilde{E}(1-\widetilde{E})}\,. \tag{45}$$

In summary, this section introduced the kinetic description of radiation, the two-moment system, and the main classes of closures (ME and Kershaw) that will be used as references in the subsequent analysis.

## 3. Modeling setup and applicability of the numerical procedure

All computations in this work are performed in the "gray" (frequency-integrated) formulation of the two-moment radiation-hydrodynamics model in 1D slab geometry. The applied closures, ME and Kershaw, are evaluated in this integrated setting. The ME pressure relation (Eq. (32)) retains the same form for all particle statistics (classical, bosonic, fermionic). The fermionic variant is analyzed only through the degeneracy parameter $a$ in $\Delta(\phi, a)$ tests, without retraining $f(\phi, a)$. Thus, the numerical procedure corresponds to the frequency-integrated ("gray") model, with all benchmarks and the Marshak wave test carried out in this configuration.

The fermionic variant considered here is intended primarily as an illustration of the flexibility of the proposed approach: the same numerical pipeline can be applied when the Eddington factor depends on both $\phi$ and the degeneracy parameter $a$. In this work we focus on verifying that the resulting fermionic closures preserve hyperbolicity over a wide range of $a$, while a full quantitative error analysis and dynamic benchmarks for the fermionic case are left for future investigations.

The measure of directivity (anisotropy) is the reduced flux:

$$\phi = \frac{|\mathbf{F}|}{E} \in [0, 1). \tag{46}$$

In this paper, we consider two closure maps: the classical $f(\phi)$ and the fermionic $f(\phi, a)$ with the degeneracy parameter $a$ [4, 20].

We construct the ME closure by solving the Lagrange multiplier problem for the angular intensity $w^*(\mu)$, which depends on the linear combination $\alpha+\beta\mu$ and additionally on the parameter $a$ in the fermionic variant [4, 20]. We determine the moments from the integrals in the direction $\mu \in [-1, 1]$:

$$E(\alpha, \beta[, a]) = \int\limits_{-1}^{1} w^*(\mu; \alpha, \beta[, a])\,\mathrm{d}\mu, \tag{47}$$

$$|\mathbf{F}|(\alpha,\beta[,a]) = \int\limits_{-1}^{1} \mu\, w^*(\mu;\alpha,\beta[,a])\,\mathrm{d}\mu, \tag{48}$$

$$\frac{\mathsf{P}}{E} \rightsquigarrow f(\alpha,\beta[,a]) = \frac{\int\limits_{-1}^{1} \mu^2 w^*(\mu;\alpha,\beta[,a])\,\mathrm{d}\mu}{\int\limits_{-1}^{1} w^*(\mu;\alpha,\beta[,a])\,\mathrm{d}\mu}. \tag{49}$$

For a given $\phi$ (and optionally $a$), we solve the monotonic mapping $\phi \mapsto \beta$ using the bisection method. We approximate the angular integrals by Gauss–Legendre quadrature using the number of nodes (quadrature points) in the Gauss–Legendre rule used to integrate over the angular cosine $\mu \in [-1, 1]$, i.e., $N_{\mathrm{GL}} = 512$ nodes, which stabilizes the accuracy near $\phi \to 1$. The derivative of the ME closures with respect to $\phi$, $\partial_\phi f_{\mathrm{ME}}$, is determined on a dense mesh using finite differences. We assign weights to samples closer to the radiation limit [24]:

$$w(\phi) = 1 + \beta_{\mathrm{edge}}\, \phi^{k_{\mathrm{edge}}} \quad (\beta_{\mathrm{edge}} = 5,\ k_{\mathrm{edge}} = 6), \tag{50}$$

which improves the approximation in the zone of the fastest changes of $f(\phi)$. The resulting dataset $\mathcal{D} = \{(\phi[;a],\ f_{\mathrm{ME}},\ \partial_\phi f_{\mathrm{ME}},\ w)\}$ is then partitioned into training and validation subsets using an 80/20 ratio. Without this edge weighting, preliminary experiments showed noticeably larger errors and oscillations in the vicinity of $\phi \approx 1$, especially in the derivative, which is critical for correct wave-speed predictions.

## 4. Materials and methods

The reference set was constructed according to the procedure for determining the maximum entropy proposed in [2, 3], as shown in Algorithm 1. For given moments $(E, \mathbf{F})$, after scaling, we reduce it to $\phi = |\mathbf{F}|/E)$. We then solve the monotonic inverse problem using the bisection method to determine the Lagrange multipliers. The higher-order moment is calculated using Gauss–Legendre quadrature with $N_{\mathrm{GL}} = 512$ nodes, which provides stable accuracy near $\phi \to 1$. The derivative with respect to $\phi$ is obtained from finite differences on a grid. Samples closer to the free-streaming ($\phi \to 1$) are weighted more strongly to better capture the sharp geometry of the function. Ablation runs performed without this weighting resulted in visibly larger errors near $\phi \to 1$, confirming that edge weighting is necessary to control the asymptotic behavior in the free-streaming regime. We use an analogous scheme in the fermionic variant, where the degeneracy parameter $a$ appears [4]. We divide the resulting dataset into training and validation sets using an 80/20 ratio.

**Algorithm 1.** Generate ME reference dataset.

```
Require: Grid $\{\phi_i\}_{i=1}^{M}$ on $[0, \phi_{\max}]$; optional $\{a_i\}$; $N_{\mathrm{GL}}$ nodes
Ensure: Dataset $\mathcal{D} = \{(\phi_i, a_i, f_{\mathrm{ME},i}, (\partial_\phi f)_{\mathrm{ME},i}, w_i)\}$
for $i = 1, \ldots, M$ do
    Invert $\phi_i \mapsto (\alpha, \beta)$ by robust bisection (monotone map)
    Integrate angular moments with GL ($N_{\mathrm{GL}}$ nodes): obtain $f_{\mathrm{ME},i}$
    Differentiate over $\phi$ on the grid $\rightarrow (\partial_\phi f)_{\mathrm{ME},i}$ (finite diff.)
    Edge weight $w_i \leftarrow 1 + \beta_{\mathrm{edge}} \phi_i^{k_{\mathrm{edge}}}$    ▷ higher weight near $\phi \rightarrow 1$
end for
Split $\mathcal{D} \rightarrow$ train/val
```

Compact neural models are known to achieve strong information efficiency across widely differing problem areas [17, 18]. On this basis, we employ a small MLP and assess whether it can act as a fast, physics-constrained approximation to the ME closure. We use a small MLP to approximate the map $\phi \mapsto f(\phi)$ (and $(\phi, a) \mapsto f(\phi, a)$ for fermions). To enforce the physical range, we parameterize the output with a smooth squashing function to $(1/3, 1)$, making $f_\theta \in (1/3, 1)$ independent of the weights. The objective function is the sum of three terms: (i) the error of $f$ with respect to the ME reference, (ii) the error of the derivative $\partial_\phi f$ with respect to the ME calculated by automatic differentiation, and (iii) a non-monotonicity penalty $\max(0, -\partial_\phi f)$, which softly enforces the growth of $f$ with $\phi$. Samples from the region $\phi \to 1$ are weighted higher. We optimize the network using Adam with the following parameters: a constant learning rate $10^{-3}$ and a mini-batch size 1024. After each epoch, we evaluate on a held-out validation split and report two metrics: the MAE and the maximum absolute error ($L^\infty$). We retain the checkpoint that attains the lowest validation MAE and employ early stopping with a patience of 50 epochs to avoid overfitting. All reported numbers (errors, timings) correspond to this best-validation checkpoint. Experiments use fixed random seeds and double precision. In the final phase of training, the validation MAE varies only slightly between successive checkpoints (typically by less than $10^{-4}$). Preliminary reruns with different random seeds produced final-validation MAE values within approximately 5% of each other, indicating low sensitivity of the surrogate to initialization. A full multi-seed variance analysis is left for future work.

To enforce the physical range, the output is parameterized through a monotone squashing nonlinearity $\sigma : \mathbb{R} \to (0,1)$: we set $f_\theta = \frac{1}{3} + \frac{2}{3}\,\sigma(g_\theta(\cdot))$, which guarantees $f_\theta \in (1/3, 1)$ independently of the weights, see Algorithm 2. This construction is consistent with key properties of the ME closure, as it keeps the Eddington factor in the physically admissible interval $(1/3, 1)$ and promotes a globally increasing dependence on the reduced flux. At the same time, the ME curvature $\chi''(\varphi)$ and the full realizability structure are not enforced analytically but are instead reproduced numerically by training on ME reference data. The loss is the weighted sum of three terms: (i) the MAE of $f$ against the ME reference, (ii) the MAE of the derivative $\partial_\phi f$ against the ME computed by automatic differentiation (autograd), and (iii) a non-monotonicity penalty $\max(0, -\partial_\phi f)$, which softly enforces that $f$ increases with $\phi$. Samples near the free-streaming edge are emphasized by $w(\phi) = 1 + \beta_{\mathrm{edge}}\phi^{k_{\mathrm{edge}}}$. We optimize using Adam with the following parameters: a constant learning rate $10^{-3}$ and a mini-batch size 1024. The loss weights are $\lambda_{\mathrm{der}} = 5 \times 10^{-3}$ and $\gamma_{\mathrm{mono}} = 10^{-2}$. After each epoch we evaluate on a held-out validation split and report MAE

**Algorithm 2.** Train the physics-constrained neural Eddington closure.

```
Require: MLP $g_\theta : (\phi[, a]) \mapsto \mathbb{R}$ with two hidden layers (64,64), SiLU
Output clamp: $f_\theta \leftarrow \frac{1}{3} + \frac{2}{3}\,\sigma(g_\theta(\cdot))$        ▷ $\sigma$ maps to $(0,1)$
Hyperparams: Adam(lr = $10^{-3}$), batch=1024, $\lambda_{\mathrm{der}} = 5\times10^{-3}$, $\gamma_{\mathrm{mono}} = 10^{-2}$
while not early-stopped do
    for minibatch $B \subset \mathcal{D}_{\mathrm{train}}$ do
        $(\phi, a, f_{\mathrm{ME}}, (\partial_\phi f)_{\mathrm{ME}}, w) \leftarrow B$
        $f_\theta \leftarrow$ MLP output;    $\partial_\phi f_\theta \leftarrow$ autograd wrt $\phi$
        $L_{\mathrm{val}} \leftarrow \mathrm{MAE}(f_\theta, f_{\mathrm{ME}};\, w)$
        $L_{\mathrm{der}} \leftarrow \mathrm{MAE}(\partial_\phi f_\theta, (\partial_\phi f)_{\mathrm{ME}};\, w)$
        $L_{\mathrm{mono}} \leftarrow \mathrm{mean}(w \cdot \max(0, -\partial_\phi f_\theta))$
        $L \leftarrow L_{\mathrm{val}} + \lambda_{\mathrm{der}} L_{\mathrm{der}} + \gamma_{\mathrm{mono}} L_{\mathrm{mono}}$
        Backprop($L$);    Adam.step();    zero_grad()
    end for
    Evaluate on $\mathcal{D}_{\mathrm{val}}$ (MAE, $L^\infty$); update best model / patience
end while
```

and $L^\infty$. The derivative penalty promotes a faithful reconstruction of $\partial_\varphi f_\theta$ and improves the smoothness of the closure, while the monotonicity term suppresses small unphysical oscillations. The weights $\lambda_{\text{der}}$ and $\gamma_{\text{mono}}$ were chosen to improve the match to the ME derivative without noticeably degrading the MAE of $f_\theta$ itself; a more systematic optimization of these weights is left for future work. Then, we retain the checkpoint with the lowest validation MAE and apply early stopping with a patience of 50 epochs (i.e., training stops if the validation MAE fails to improve for 50 consecutive epochs by at least a small tolerance $\varepsilon$, and the best-performing checkpoint is restored). All reported numbers, including errors and timings, correspond to this best-validation checkpoint. Experiments use fixed random seeds and double precision. This physics-constrained approach complements classical entropy closure and two-moment models [2].

In the next step, we evaluate the accuracy of $f_\theta$ using the MAE and $L^\infty$ with respect to $f_{\text{ME}}$ on $\phi \in [0, 0.98]$ and, separately, in the boundary window $\phi \geq 0.95$, see Algorithm 3. We verify the hyperbolicity of the system using the discriminant $\Delta(\phi) = \big(\partial_\phi f\big)^2 + 4\big(f - \phi\,\partial_\phi f\big)$ and report $\min_\phi \Delta(\phi) > 0$. We also measure the per-call speedup of the closure invocation (the ratio of the ME evaluation time to the NN evaluation time for a set of queries), and examine the impact on the equations using the canonical Marshak wave test [25, 26]. Specifically, we compare $\|E(\cdot, T)\|_{L^1}$, $\|E(\cdot, T)\|_{L^\infty}$, and the end-to-end simulation time. As a benchmark, we include the analytical Kershaw closure [11, 27] and a lightweight rational fit to the ME data. In the fermionic variant, we scan $a \in [-5, 5]$ and plot the map $\min_\phi \Delta(\phi, a)$.

**Algorithm 3.** Physical checks and performance benchmarks.

```
Require: Best $\theta$; dense grid $\phi \in [0, \phi_{\max}]$
  Accuracy: report MAE, $L^\infty$ of $f_\theta$ vs. $f_{\text{ME}}$; edge-window stats for $\phi \geq 0.95$
  Hyperbolicity: compute $\Delta(\phi)$ from $f_\theta$ and $\partial_\phi f_\theta$; record $\min_\phi \Delta > 0$
  Per-call timing: compare $N$ batched calls of ME vs. NN (speedup)
  Marshak test (end-to-end): run solver with ME/NN/Kershaw; report $L^1/L^\infty$ vs. ME and wall-clock speedup
  Baselines: Kershaw closure; rational fit (report MAE vs. ME)
  Fermions: scan $a \in [-5, 5]$; map $\min_\phi \Delta(\phi, a)$
```

Table 1 summarizes the key hyperparameters and numerical settings used in the calculations, including the MLP architecture, optimization scheme, loss term weights, sample boundary weighting, precision, reduced flux range, ME reference parameters, and benchmark configuration.

TABLE 1. Hyperparameters and numerical settings applied in the calculations.

| Component | Value / Setting |
|---|---|
| Architecture | MLP, 2 hidden layers, 64 units each |
| Activation | SiLU (hidden), output clamped to $(1/3, 1)$ |
| Optimizer | Adam (learning rate $10^{-3}$) |
| Batch size | 1024 |
| Loss weights | $\lambda_{\text{der}} = 5 \times 10^{-3}$, $\gamma_{\text{mono}} = 10^{-2}$ |
| Edge weighting | $w(\phi) = 1 + \beta_{\text{edge}}\phi^{k_{\text{edge}}}$, $\beta_{\text{edge}} = 5$, $k_{\text{edge}} = 6$ |
| Early stopping | patience = 50 (best validation MAE) |
| Training / validation split | 80% / 20% |
| Precision | float64 |
| Random seed | fixed (reported in code) |
| Reduced flux range | $\phi \in [0,\ 0.98]$ |
| ME quadrature | Gauss–Legendre, $N_{\text{GL}} = 512$ nodes |
| ME inversion | monotone bisection (per point) |
| Baselines | Kershaw closure; rational fit $(a_2, a_4, b_2)$ |
| Benchmarks | Marshak wave (1D): $L^1/L^\infty$ vs. ME; end-to-end timing |

## 5. Results

We propose and evaluate a compact, physically consistent NN closure that replaces the ME closure in three-moment transport without compromising its essential properties while providing significant speedup. As a high-fidelity reference, we compute the ME Eddington factor on a dense grid using Gauss–Legendre angular quadrature and monotonic inversion. In the next step, we train a small MLP whose output is smoothly constrained to a physically feasible range, regularized to grow monotonically with anisotropy, and correspond to the derivative of the ME closure. We compare this NN-based closure with the ME and a standard analytical basis, i.e., the Kershaw closure [11, 27] in terms of local accuracy, hyperbolicity compliance, delay per evaluation, and performance on the Marshak wave benchmark [25, 26].

### 5.1. The quality of the closure

Apart from pointwise errors $f(\phi)$, we quantify the local geometry by comparing $\partial_\phi f$ to the ME reference, see Table 2. The NN-based closure matches the ME slope to within a few percent most of the range and remains competitive even in the rapid-variation region near $\phi \to 1$, where the absolute error peaks yet stays localized. This derivative control is important for preserving the eigen-speeds of the flux Jacobian and, consequently, hyperbolicity. The NN-based closure

TABLE 2. Accuracy of the closures against the ME reference for $f(\phi)$, on the grid $\phi \in [0, 0.98]$.

| Closure | MAE | Max error | MAE (edge $\phi \geq 0.95$) |
|---|---|---|---|
| Kershaw (analytic) | $4.346 \times 10^{-2}$ | $8.464 \times 10^{-2}$ | $2.160 \times 10^{-2}$ |
| NN | $8.917 \times 10^{-4}$ | $2.608 \times 10^{-2}$ | $1.217 \times 10^{-2}$ |
| Rational fit $(a_2, a_4, b_2)$ | $1.012 \times 10^{-3}$ | – | – |

achieves MAE $= 8.917 \times 10^{-4}$ on the entire grid $\phi \in [0, 0.98]$, which is $\sim 50\times$ more accurate than the analytical Kershaw closure ($4.346 \times 10^{-2}$). In the difficult boundary $\phi \geq 0.95$, the error increases due to the rapid variation of $f$. However, the NN remains significantly better than Kershaw. An analytic rational baseline

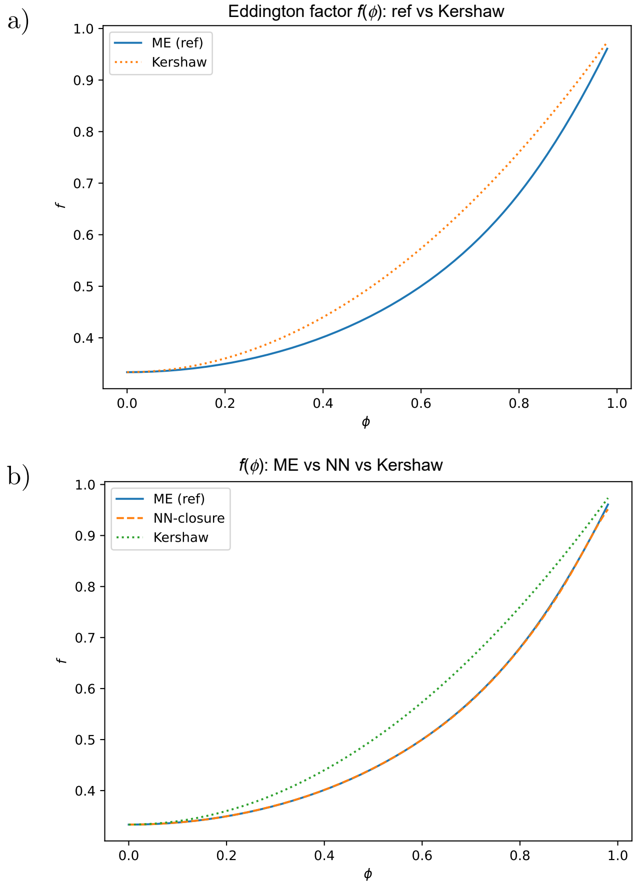

FIG. 1. Eddington factor comparison: a) Kershaw closure departs visibly from the ME reference, especially near extreme $\phi$, b) NN-based closure overlaps the ME curve across almost the entire admissible range, while Kershaw shows systematic bias; color/line coding: blue solid – ME, orange dotted – Kershaw; green dashed – NN-based.

gives MAE $= 1.012 \times 10^{-3}$, which is a useful compromise between simplicity and accuracy. Beyond scalar errors, the surrogate provides a smoother and more accurate approximation of $\partial_\phi f$ near $\phi \to 1$ than the Kershaw closure, which tends to underpredict the steepness of the ME Eddington factor in the free-streaming regime. This improved control of the slope is important for obtaining physically consistent characteristic speeds in the moment system.

Figure 3a summarizes the trade-off between accuracy and computational cost. The NN-based closure maintains near-ME accuracy while delivering substantial speedups, outperforming the Kershaw closure in both quality and efficiency. Learning dynamics are presented in Fig. 3b. The error drops rapidly during the first $\sim 50$ epochs (from $\sim 2 \times 10^{-1}$ to $\sim 10^{-2}$), then decays slowly and stabilizes at $\sim 9 \times 10^{-4}$. Validation closely tracks training, and the monotonicity term decays to $\sim$ zero, consistent with $f(\phi)$ being globally increasing over the discretized training domain.

Figure 1 compares Eddington factors across closures. Figure 1a shows that the Kershaw closure departs from the ME reference, with the largest deviations near the edges of the admissible domain. Figure 1b includes the NN-based closure. This curve essentially overlaps the ME reference over almost the entire range of $\phi$. By contrast, Kershaw exhibits a persistent systematic bias. This indicates that the NN-based closure has learned the structure required to reproduce the ME mapping $f(\phi)$ with correct asymptotics near the boundaries.

### 5.2. Hyperbolicity

As shown in Fig. 2a, the discriminant remains non-negative across the admissible domain, in agreement with ME, thereby ensuring a physically consistent hyperbolic formulation. This satisfies a key physical-consistency requirement that purely heuristic closures may violate.

In Table 3, we report the minimum discriminant for the NN-based closure $\min_\phi \Delta(\phi)$ for NN: $5.1 \cdot 10^{-2} > 0$, implying no hyperbolicity violations over the tested range. The fermionic variant also exhibits zero violations across $a \in [-5, 5]$ on the $(\phi, a)$ lattice. A positive margin of $\Delta$ is critical for stable hyperbolic dynamics.

To quantify the hyperbolicity margin across thermodynamic regimes, we report $\min_\phi \Delta$ as a function of the degeneracy $a$, see Fig. 2b. Across $a \in [-5, 5]$, the minimum discriminant remains strictly positive ($\approx 0.5$–$1.3$), indicating hyperbolicity throughout all regimes considered. The above configuration was selected after preliminary experiments, indicating qualitatively similar learning dynamics and final errors for nearby choices of learning rate and batch size, suggesting limited sensitivity of the method to these hyperparameters. Although the present analysis is based on discrete sampling and numerically evaluated derivatives,

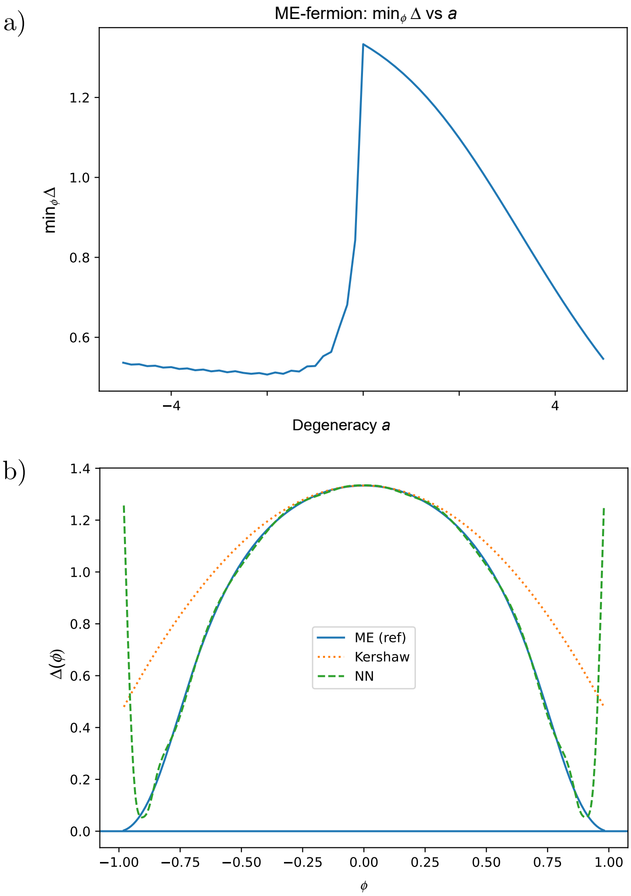

FIG. 2. a) Minimum discriminant $\min_\phi \Delta$ versus degeneracy $a$. The horizontal line at $y = 0$ would mark the hyperbolicity boundary, b) hyperbolicity check across the admissible range of $\phi$. The horizontal line at $y = 0$ marks zero error relative to ME.

TABLE 3. Hyperbolicity checks.

| Case | $\min_\phi \Delta(\phi)$ | Fermions: $\min_\phi \Delta(\phi, a)$ for $a \in [-5, 5]$ |
|---|---|---|
| NN-based closure | $5.1 \cdot 10^{-2} > 0$ | $> 0$ (no violations) |

the positive discriminant margin ($\min_\phi \Delta \approx 5 \times 10^{-2}$) suggests robustness with respect to small perturbations. A systematic perturbation study is deferred to future work.

### 5.3. Efficiency

Table 4 shows per-evaluation times for a batch of $N = 20\,000$ samples/ points. The ME procedure is a baseline. The Kershaw closure is ∼64× faster,

TABLE 4. Per-evaluation timing of the closure $f(\phi)$ (batch $N = 20\,000$).

| Method | Runtime | Speedup vs. ME |
|---|---|---|
| ME (reference) | 16.51 s | 1× |
| Kershaw (analytic) | 0.2572 s | 64.2× |
| NN-based closure | 0.01573 s | 1049.6× |

while the NN-based model is $\sim 1050\times$, due to its constant cost of $O(1)$ (matrix multiplications) instead of integration and map inversion. These speedups translate into real-time savings in the solvers, see Fig. 3a.

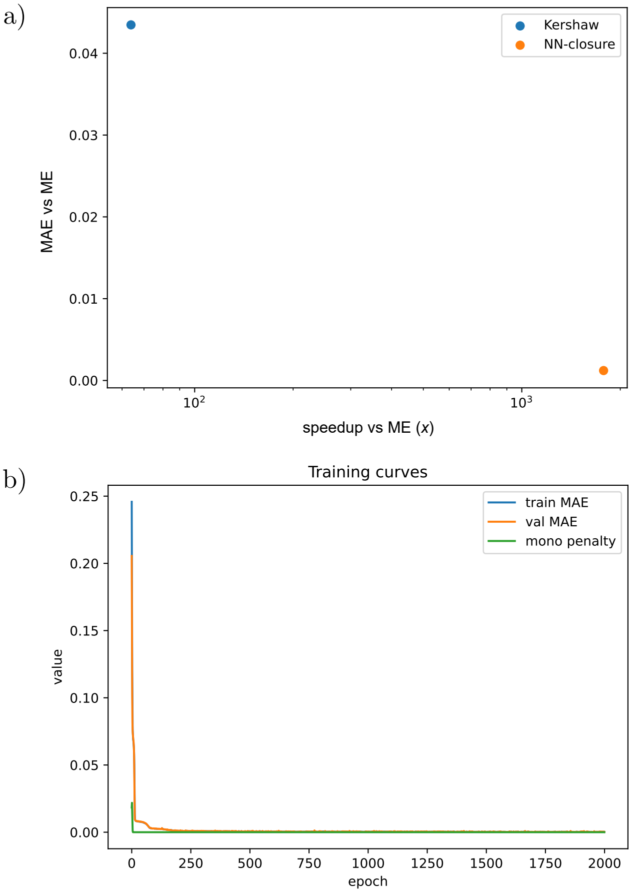

FIG. 3. Eddington factor: a) accuracy–efficiency trade-off, b) learning dynamics for the NN-based Eddington closure.

The learning dynamics are presented in Fig. 3b. Both training and validation MAE drop sharply in the first ≈50 epochs and then flatten near $10^{-3}$.

## 5.4. Marshak wave test

In the Marshak wave test, we compare the profile of $E(x,T)$ in the $L^1$ and $L^\infty$ norms and the total computation time, see Table 5. The NN-based closure remains very close to the ME ($L^1 = 8.928 \times 10^{-4}$) and speeds up the entire run

TABLE 5. Marshak benchmark.

| Closure in solver | $L^1$ against ME | $L^\infty$ vs. ME | Speedup (end-to-end) |
|---|---|---|---|
| ME (reference) | – | – | 1× |
| Kershaw (analytic) | $1.490 \times 10^{-2}$ | $7.419 \times 10^{-2}$ | 48.2× |
| NN-based closure | $8.928 \times 10^{-4}$ | $1.473 \times 10^{-2}$ | 247.4× |

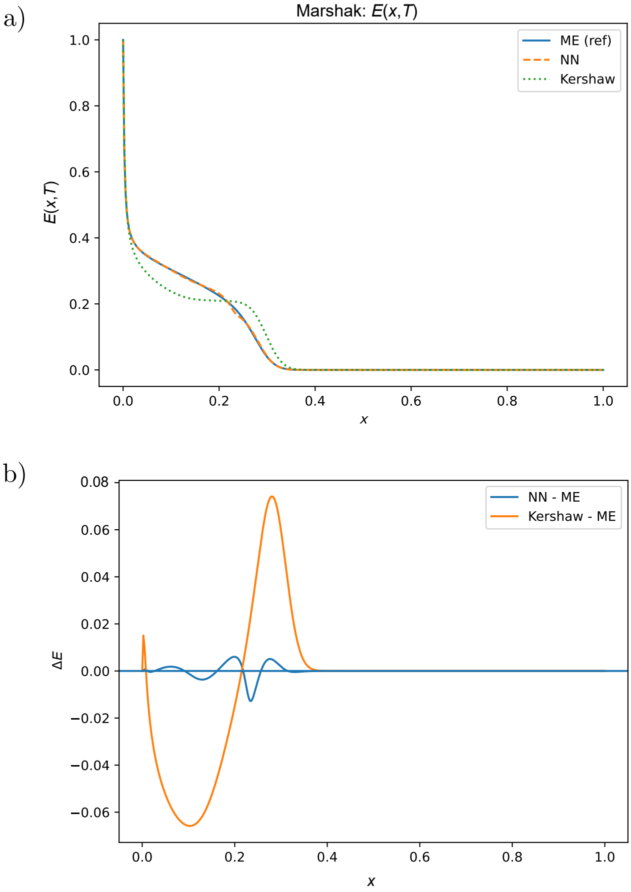

FIG. 4. Marshak benchmark: a) NN closure reproduces $E(x,T)$ from the ME reference, while Kershaw deviates more strongly, b) pointwise differences show near-zero error for NN across most of the slab and larger, structured errors for Kershaw.

by $\sim 247\times$. The Kershaw closure is faster than the ME closure, but generates larger deviations ($L^1 = 1.49 \times 10^{-2}$, $L^\infty = 7.419 \times 10^{-2}$) and a smaller time gain of $\sim 48\times$. Figures 4a and 4b assess solution quality for the Marshak problem. The NN-based closure reproduces the reference $E(x, T)$ profile (Fig. 4a) and yields pointwise errors close to zero across most of the slab (Fig. 4b), while Kershaw exhibits larger, structured discrepancies. These results confirm that closure-level enhancements propagate to higher-quality solutions for the canonical Marshak benchmark. In particular, the NN-based closure reproduces both the position and the shape of the Marshak wave front at the final time, with only small deviations from the ME profile. The Kershaw closure, while still capturing the overall behavior, yields a more diffuse and slightly displaced front, consistent with its larger $L^1$ and $L^\infty$ errors. The trajectory of the Marshak front, estimated as the position where $E(x, T)$ drops to half of its peak value, matches the ME reference within the grid resolution, whereas the Kershaw closure exhibits a slightly delayed and more diffuse front.

## 5.5. Ablation study

We quantify the contribution of each component in Table 6. FULL denotes the complete training objective (DERIV + MONO + EDGE), no DERIV removes the derivative regularization term, no MONO removes the monotonicity penalty, and EDGE removes the edge-weighted error term near domain boundaries. We consider the following metrics: MAE, MAE_edge: MAE computed on edge regions only, dMAE: MAE of the first derivative (shape fidelity), min $\boldsymbol{\Delta}$: minimum discriminant over the scan (hyperbolicity margin), neg $\boldsymbol{\Delta}$: count (or fraction) of sampled points with $\Delta < 0$ (hyperbolicity violations), and speedup: computational speedup relative to the ME reference. Disabling regularization, including the monotonicity penalty, increases the MAE. Architectural improvements that increase expressivity reduce error while maintaining stability. Ablation runs performed without this edge-weighting resulted in visibly larger errors near $\phi \to 1$, confirming that edge-weighting is necessary to control the asymptotic behavior in the free-streaming regime.

TABLE 6. Ablation study.

| Name | min $\boldsymbol{\Delta}$ | neg $\boldsymbol{\Delta}$ | MAE | MAE_edge | dMAE | Speedup |
|---|---|---|---|---|---|---|
| No DERIV | 0.119 | 0 | 0.002 | 0.040 | 0.076 | 748 |
| No MONO | 0.108 | 0 | 0.003 | 0.039 | 0.082 | 922 |
| FULL | 0.166 | 0 | 0.004 | 0.043 | 0.088 | 1178 |
| No EDGE | 0.265 | 0 | 0.005 | 0.069 | 0.106 | 2299 |

## 6. Limitations

The results presented here concern an energetically integrated, one-dimensional test case and the range $\phi \leq 0.98$, where the boundary $\phi \to 1$ is particularly challenging numerically. Training is performed on ME data generated for a fixed class of problems; changes in medium properties would require regenerating the reference data set and/or performing a short fine-tuning stage. Although no violations of hyperbolicity were observed, this guarantee is currently numerical, based on verifying $\min_\phi \Delta(\phi) > 0$ on a dense grid, rather than an analytical proof for the continuous NN model.

## 7. Conclusions

The presented NN-based Eddington closure achieved a MAE of $9.0 \times 10^{-4}$ on the reduced flux $\phi \in [0, 0.98]$ and provided a $\sim 10^3\times$ speedup for a single closure invocation. In the Marshak wave test, i.e., a canonical 1D transient test that combined closure with wave speeds and front shape, it speeds up the entire simulation by $\sim 247\times$ while maintaining hyperbolicity, with $\min_\phi \Delta(\phi) = 5.1 \times 10^{-2} > 0$. Compared to the analytic Kershaw closure $f(\phi) = \frac{1}{3} + \frac{2}{3}\phi^2$, our model is clearly more accurate (MAE $4.35 \times 10^{-2}$ for Kershaw) and faster end-to-end. For simplicity of implementation, we also provided a lightweight rational fit (MAE $1.01 \times 10^{-3}$) and showed that in the fermionic variant $\min_\phi \Delta(\phi, a) > 0$ for $a \in [-5, 5]$. Thus, a small, physically constrained NN trained on ME data can practically replace the ME closure in the three-moment model, preserving hyperbolicity and accuracy of the order of MAE $\sim \times 10^{-3}$, while speeding up computations from $\sim \times 10^2$ to $\sim \times 10^3$ times, as confirmed by the Marshak wave test. The model can be directly used in three-moment solvers in place of the ME procedure, achieving orders-of-magnitude savings in computational time without losing the quality of the solutions. It should be emphasized that the same numerical approach can be employed for radiation hydrodynamics in the spectral formulation, where different ME closures correspond to different statistics (Maxwell–Boltzmann, Bose–Einstein, and Fermi–Dirac). Moreover, the method naturally extends to three-moment radiation hydrodynamics. This extension is of special interest, since there are no closed analytic formulae for ME closures of the three-moment theories [6–8].

## References

1. J.A. Pons, J.Ma. Ibáñez, J.A. Miralles, Hyperbolic character of the angular moment equations of radiative transfer and numerical methods, *Monthly Notices of the Royal Astronomical Society*, **317**(3): 550–562, 2000, https://doi.org/10.1046/j.1365-8711.2000.03679.x.

2. C.D. Levermore, Moment closure hierarchies for kinetic theories, *Journal of Statistical Physics*, **83**: 1021–1065, 1996, https://doi.org/10.1007/BF02179552.
3. G.N. Minerbo, Maximum entropy Eddington factors, *Journal of Quantitative Spectroscopy and Radiative Transfer*, **20**(6): 541–545, 1978, https://doi.org/10.1016/0022-4073(78)90024-9.
4. J. Cernohorsky, S. Bludman, Maximum entropy distribution and closure for Bose-Einstein and Fermi-Dirac radiation transport, *The Astrophysical Journal*, **433**: 250, 1994, https://doi.org/10.1086/174640.
5. C.D. Levermore, Relating Eddington factors to flux limiters, *Journal of Quantitative Spectroscopy and Radiative Transfer*, **31**(2): 149–160, 1984, https://doi.org/10.1016/0022-4073(84)90112-2.
6. Z. Banach, W. Larecki, Entropy-based mixed three-moment description of fermionic radiation transport in slab and spherical geometries, *Kinetic and Related Models*, **10**(4): 879–900, 2017, https://doi.org/10.3934/krm.2017035.
7. Z. Banach, W. Larecki, Spectral maximum entropy hydrodynamics of fermionic radiation: a three-moment system for one-dimensional flows, *Nonlinearity*, **26**(6): 1667–1701, 2013, https://doi.org/10.1088/0951-7715/26/6/1667.
8. Z. Banach, W. Larecki, One-dimensional maximum entropy radiation hydrodynamics: Three-moment theory, *Journal of Physics A: Mathematical and Theoretical*, **45**(38): 385501, 2012, https://doi.org/10.1088/1751-8113/45/38/385501.
9. R. Chu, E. Endeve, C.D. Hauck, A. Mezzacappa, Realizability-preserving DG-IMEX method for the two-moment model of fermion transport, *Journal of Computational Physics*, **389**(C): 62–93, 2019, https://doi.org/10.1016/j.jcp.2019.03.037.
10. F. Schneider, Kershaw closures for linear transport equations in slab geometry I: Model derivation, *Journal of Computational Physics*, **322**: 905–919, 2015, https://doi.org/10.1016/j.jcp.2016.02.080.
11. D.S. Kershaw, *Flux limiting nature's own way – A new method for numerical solution of the transport equation*, Technical Report, Lawrence Livermore Laboratory, CA, 1976, https://doi.org/10.2172/104974.
12. J. Huang, Y. Cheng, A.J. Christlieb, L.F. Roberts, Machine learning moment closure models for the radiative transfer equation I: Directly learning a gradient based closure, *Journal of Computational Physics*, **453**: 110941, 2022, https://doi.org/10.1016/j.jcp.2022.110941.
13. J. Huang, Y. Cheng, A.J. Christlieb, L.F. Roberts, Machine learning moment closure models for the radiative transfer equation III: Enforcing hyperbolicity and physical characteristic speeds, *Journal of Scientific Computing*, **94**: 7, 2023, https://doi.org/ 10.1007/s10915-022-02056-7.
14. J.M. Coale, D.Y. Anistratov, A variable Eddington factor model for thermal radiative transfer with closure based on data-driven shape function, *Journal of Computational and Theoretical Transport*, **53**(2): 153–172, 2024, https://doi.org/10.1080/23324309.2024.2327992.
15. C. Musolino, L. Rezzolla, A practical guide to a moment approach for neutrino transport in numerical relativity, *Monthly Notices of the Royal Astronomical Society*, **528**(4): 5952–5971, 2024, https://doi.org/10.1093/mnras/stae224.

16. H. Andresen, E.P. O'Connor, O.E. Andersen, S.M. Couch, Gray two-moment neutrino transport: Comprehensive tests and improvements for supernova simulations, *Astronomy & Astrophysics*, **687**: A55, 2024, https://doi.org/10.1051/0004-6361/202449776.

17. B. Paprocki, A. Pregowska, J. Szczepanski, Does adding of neurons to the network layer lead to increased transmission efficiency?, *IEEE Access*, **12**: 42701–42709, 2024, https://doi.org/10.1109/ACCESS.2024.3379324.

18. B. Paprocki, A. Pregowska, J. Szczepanski, Optimizing information processing in brain-inspired neural networks, *Bulletin of the Polish Academy of Sciences Technical Sciences*, **68**(2): 225–233, 2020, https://doi.org/10.24425/bpasts.2020.131844.

19. G.C. Pomraning, *The Equations of Radiation Hydrodynamics*, in series: International Series of Monographs in Natural Philosophy, Vol. 54, Dover Publications, Mineola, NY, 1973.

20. W. Larecki, Z. Banach, Entropic derivation of the spectral Eddington factors, *Journal of Quantitative Spectroscopy and Radiative Transfer*, **112**(15): 2486–2506, 2011, https://doi.org/10.1016/j.jqsrt.2011.06.011.

21. W. Larecki, Symmetric conservative form of low-temperature phonon gas hydrodynamics. II. – Equations of heat transport and thermal waves in the case of linear isotropic approximation of phonon frequency spectrum, *Il Nuovo Cimento D*, **14**: 141–176, 1992, https://doi.org/10.1007/BF02457349.

22. W. Larecki, Z. Banach, Two-field radiation hydrodynamics in $n$ spatial dimensions, *Journal of Physics A: Mathematical and Theoretical*, **49**(12): 125501, 2016, https://doi.org/10.1088/1751-8113/49/12/125501.

23. Z. Banach, W. Larecki, Kershaw-type transport equations for fermionic radiation, *Zeitschrift für angewandte Mathematik und Physik*, **68**(4): 100-1–24, 2017, https://doi.org/10.1007/s00033-017-0847-z.

24. W.H. Press, S.A. Teukolsky, W.L. Vetterling, B.P. Flannery, *Numerical Recipes in FORTRAN. The Art of Scientific Computing*, 3rd ed., Cambridge University Press, Cambridge–New York, NY, 2007.

25. R.E. Marshak, Effect of radiation on shock wave behavior, *Physics of Fluids*, **1**(1): 24–29, 1958, https://doi.org/10.1063/1.1724332.

26. T. Shussman, S.I. Heizler, Full self-similar solutions of the subsonic radiative heat equations, *Physics of Plasmas*, **22**(8): 082109, 2015, https://doi.org/10.1063/1.4927756.

27. E.M. Murchikova, E. Abdikamalov, T. Urbatsch, Analytic closures for M1 neutrino transport, *Monthly Notices of the Royal Astronomical Society*, **469**(2): 1725–1737, 2017.

*Received October 22, 2025; revised version November 23, 2025;*
*accepted November 24, 2025; published online December 4, 2025.*