Letter

Learning heat transport kernels using a nonlocal heat transport theory-informed neural network

Mufei Luo [1,*] Charles Heaton [1] Yizhen Wang [1] Daniel Plummer [1] Mila Fitzgerald [1] Francesco Miniati [2] Sam M. Vinko [1] and Gianluca Gregori [1]

[1]Department of Physics, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom
[2]Mach42, Robert Robinson Avenue, Oxford Science Park, Oxford OX4 4GP, United Kingdom

(Received 13 June 2025; accepted 24 September 2025; published 16 October 2025)

We present a data-driven framework for the modeling of nonlocal heat transport in plasmas using a nonlocal theory-informed neural network trained on kinetic particle-in-cell simulations that span both local and nonlocal regimes. The model learns spatio-temporal heat flux kernels directly from simulation data, capturing dynamic transport behaviors beyond the reach of classical formulations. Unlike time-independent kernel models such as Luciani-Mora-Virmont and Schurtz-Nicolaï-Busquet models, our approach yields physically grounded, time-evolving kernels that adapt to varying plasma conditions. The resulting predictions show strong agreement with kinetic benchmarks across regimes. This offers a promising direction for data-driven modeling of nonlocal heat transport and contributes to a deeper understanding of plasma dynamics.

DOI: 10.1103/m215-3f6c

Understanding heat transport [1,2] in plasmas is essential for a broad range of applications, from inertial confinement fusion [3] to astrophysical systems [4,5]. Heat transport in hot plasma is characterized by the ratio of the temperature gradient length $L_T = T_e/\nabla T_e$, where T_e is the electron temperature, to the electron mean free path $\lambda_{\text{free}} = 4\pi\varepsilon_0^2 m_e^2 v_{\text{th}}^4/(Z n_e e^4 \ln\Lambda)$ [6], where ε_0 is the vacuum permittivity, m_e is the electron mass, $v_{th} = \sqrt{T_e/m_e}$ is the electron thermal velocity, Z is the ion charge state, n_e is the electron number density, e is the elementary charge, and $\ln\Lambda$ is the Coulomb logarithm. In the local regime where $L_T/\lambda_{\text{free}} \gg 500$, classical models such as the Spitzer-Härm (SH) formulation [7] yield reliable predictions. However, as this ratio decreases, nonlocal effects emerge, invalidating local hydrodynamic closures [8] and necessitating more advanced descriptions. Existing transport models include the Luciani-Mora-Virmont (LMV) [9], Schurtz-Nicolaï-Busquet (SNB) [10–12], and flux-limited schemes, attempt to extend classical theory by incorporating empirical or semianalytic kernels. More recently, diffusion-type formulations have been revisited [13]. Yet these models generally assume temporal stationarity [14] and rely on simplified or heuristic approximations, limiting their validity in regimes where the electron collision time approaches hydrodynamic timescales.

At the core of the popular nonlocal transport theories lies the heat flux kernel, which mediates the spatial coupling between local SH flux and its nonlocal contributions. Classic examples include the LMV kernel [9], the kernel proposed by Epperlein and Short [15], as well as the variant derived in Ref. [16], which eliminates the adjustable parameter present in the LMV formulation while still achieving reasonable agreement with the LMV approximation. In current analytic models, this kernel is often time independent, thereby neglecting dynamic features that are critical for the accurate modeling in highly nonlocal regimes. More recently, Lamy *et al.* [17] constructed databases by running hydrodynamic simulations with the CHIC code based on the classical SNB kernel, thereby limiting their neural network (NN) to reproducing an existing analytic model rather than learning directly from kinetic data. This motivates the need for a more flexible framework—one capable of learning transport kernels directly from kinetic data, while capable of retaining interpretability and physical structure. In addition, in Lamy *et al.* [17], the neural network represents a correction to the heat flux divergence computed from the simulation data. This may be prone to unstable behavior because the noisiness of neural networks can predict a nonzero heat source even where the temperature profile is flat, and is grid dependent, i.e., unsuitable for use with different grids. This in contrast to the work in Ref. [18], where the grid dependence is factored out and sources of heat due to flux imbalance remain explicitly proportional to the gradient of the thermodynamic variable. Here, we follow a similar approach that relies on learning a transport kernel for convolving SH fluxes instead of the heat flux divergence itself, allowing for more stability and generalization without reliance on the details of the numerical grid.

Fully kinetic simulations, such as Vlasov-Fokker-Planck (VFP) and particle-in-cell (PIC) methods, offer high-fidelity descriptions of heat transport but vary greatly in computational cost. VFP solvers [19–21], while detailed, are often prohibitively expensive for large parameter studies. PIC simulations, when equipped with physically validated binary collision models [22,23], offer a computationally efficient yet accurate means to capture nonlocal heat transport [24–26], bridging the gap between fidelity and scalability. In this work, we applied a one-dimensional (1D) PIC code (OSIRIS [27]) to

*Contact author: mufei.luo@physics.ox.ac.uk

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

generate a large, structured dataset across local and nonlocal regimes, enabling data-driven construction of dynamic heat flux kernels via neural networks.

To investigate heat transport across a wide range of plasma conditions, we impose an initial electron temperature profile of the form

$$T_e(x) = \frac{2T_{e0}(R-1)/(R+1)}{1+\exp(x/L)} + \frac{2T_{e0}}{R+1}, \tag{1}$$

where $T_{e0} = 1$ keV sets the central temperature, representative of conduction-zone plasmas in laser-fusion targets, where temperatures typically span ∼50 eV–2 keV [28]. $R = T_h/T_c$ is the hot-to-cold temperature ratio, and the steepness length L determines the gradient scale length via $L_T = 2L(R+1)/(R-1)$. The temperature profile described by Eq. (1) features a hot region on the left-hand side and a cold region on the right-hand side, with the gradient localized at the center over a width of approximately L. The corresponding density profile is chosen as $n_e(x) \propto 1/T_e(x)$, consistent with isobaric heat conduction commonly observed in the conduction zone of laser-driven fusion plasmas [28]. To control the degree of nonlocality, we vary the background electron density, thereby tuning the mean free path λ_{free} relative to L_T. This allows systematic exploration of transport behavior from the classical to the strongly nonlocal regime.

The initial condition is parametrized by the tuple $(R,\ L,\ n_{e0})$, where R is the temperature ratio, L determining the gradient scale length, and n_{e0} the central electron density. Note that R and L are not entirely independent. We will discuss this further below. To ensure physically meaningful regimes and nondegenerate conditions, we consider electron densities ranging from 10^{19} to 10^{25} cm^{-3}, corresponding to Fermi temperatures up to approximately 0.17 keV. We assume a fully ionized, high-Z plasma ($Z = 16$) to enhance collisionality and reduce relaxation times. The Coulomb logarithm $\ln\Lambda$, calculated at the box center, is held constant for each simulation.

All simulations are performed in 1D geometry with temporal and spatial units normalized to the plasma frequency $\omega_{\text{pe}} = (n_{e0}e^2/\epsilon_0 m_e)^{1/2}$ and wavenumber $k_{\text{pe}} = \omega_{\text{pe}}/c$, with c the light velocity in vacuum, then the normalized mean free path is given as $\lambda_{\text{free}} = 4\pi\varepsilon_0^{3/2} m_e^{3/2} v_{\text{th}}^4/(Zn_e^{1/2}e^3 c\ln\Lambda)$. The box length is fixed at $L_{\text{box}} \approx 1500$, with spatial resolution $\Delta x = 0.04$ and time step $\Delta t = 0.035$, satisfying Courant-Friedrichs-Lewy condition ($\Delta t/\Delta x \approx 0.9$) and resolving the Debye length ($\lambda_{\text{De},0} = 1.2\Delta x$). Each cell contains 25 000 electron and 1600 ion macroparticles. A total of 829 data points are sampled from the parameter space, covering a range of $L_T/\lambda_{\text{free}}$ from 10 to 5000, spanning regimes from strongly nonlocal to classical heat transport. All PIC simulations are evolved until the heat flux relaxation saturates, ensuring that the resulting heat flux can be directly benchmarked against the steady-state LMV model [9].

The heat flux is calculated by $Q_e = m_e \int v^3 f(v) dv$ as no current is driven, where $f(v)$ is the electron velocity distribution. After noise filtering, the maximum electron heat flux $Q_{\text{e,m}}$, normalized by the free-streaming flux ($Q_{\text{free}} = n_{e0}T_{e0}v_{\text{th0}}$, where $v_{\text{th0}} = \sqrt{T_{e0}/m_e}$ is the electron thermal velocity at the center of the simulation box), is shown in Fig. 1. The x axis represents the initial value of $L_T/\lambda_{\text{free}}$, evaluated

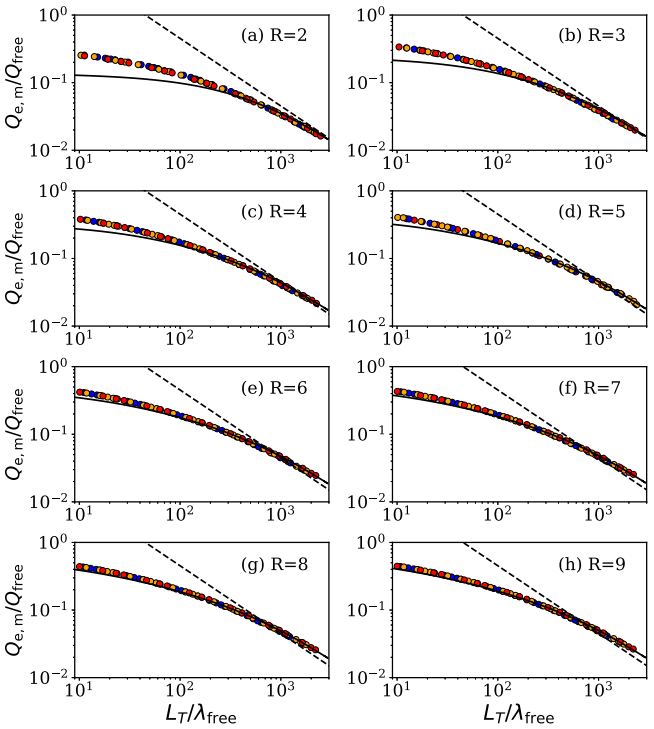


FIG. 1. Maximum normalized electron heat flux $Q_{\text{e,m}}/Q_{\text{free}}$ as a function of the initial local transport ratio $L_T/\lambda_{\text{free}}$ evaluated at the center of the domain ($x = 0$). Each subplot corresponds to a different temperature ratio $R = T_h/T_c$, increasing from (a) $R = 2$ to (h) $R = 9$. In panel (a), the blue, orange, and red points correspond to steepness lengths of $L \approx 16$, 24, and 32, respectively. In panels (b)–(h), the steepness length L associated with each color [as defined in panel (a)] is adjusted for each temperature ratio R using the relation $L_T = 2L(R+1)/(R-1)$, so that the ratio $L_T/\lambda_{\text{free}}$ at the domain center remains the same for one given central density n_{e0} across all panels. This design enables systematic exploration of nonlocal transport behavior across a range of physical regimes. The black dashed and solid lines represent the predictions of the SH and LMV models, respectively.

at the center of the domain ($x = 0$). In Figs. 1(a)–1(h), the temperature ratio $R = T_h/T_c$—defined as the ratio of the hot-side to cold-side boundary temperatures—is varied from 2 to 9. In Fig. 1(a) with $R = 2$, the blue, orange, and red points correspond to steepness lengths of $L \approx 16$, $L \approx 24$, and $L \approx 32$, respectively. We avoid choosing excessively large values of L to prevent heat flux from being driven near the boundaries. From Figs. 1(b) to 1(h), as the temperature ratio R increases, the steepness length L associated with each color [as defined in Fig. 1(a)] is adjusted according to $L_T = 2L(R+1)/(R-1)$. For the given central density n_{e0}—and hence a given mean free path—this ensures that the ratio $L_T/\lambda_{\text{free}}$ remains constant at the center of the domain. It should be noted that all simulations are terminated at $t\omega_{\text{pe}} = 2500$ for their respective background densities. This duration is generally sufficient to saturate the transport process in all samples.

The black dashed lines in each subplot represent the classical local heat transport prediction of SH flux, given by $Q_{\text{SH}}/Q_{\text{free}} = a\,\lambda_{\text{SH}}/L_T$, where $a = 128(Z+0.24)/[3\pi(Z+4.2)]$ and $\lambda_{\text{SH}} = 3\lambda_{\text{free}}/\sqrt{\pi/2}$ is the classical Spitzer-Härm

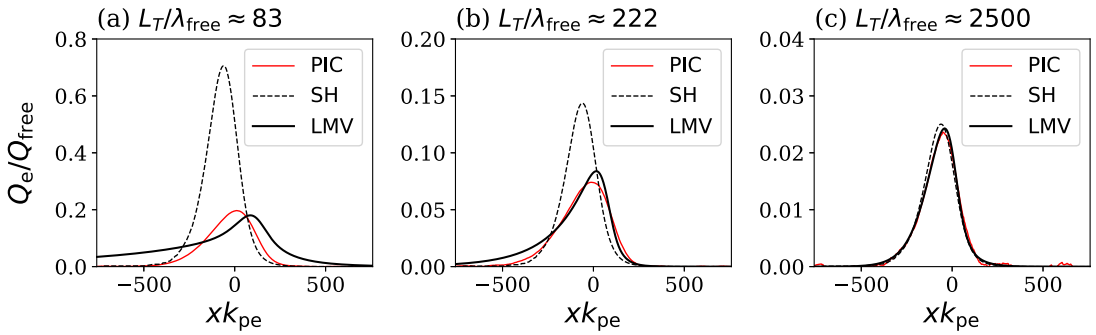


FIG. 2. Comparison of the spatial profiles of the saturated heat flux obtained from PIC simulations (red solid lines), SH theory (black dashed lines), and the LMV model (black solid lines) for different plasma densities. The temperature ratio is fixed at $R = 5$, and the gradient scale length is set to $L_T \sim 195$. From panels (a) to (c), the background electron density increases from 10^{21} to $10^{25}\,\mathrm{cm}^{-3}$, corresponding to $L_T/\lambda_{\rm free} \approx 83$, 222, and 2500, respectively.

mean free path. To validate the accuracy of our data in the nonlocal regime, we also include the 1D nonlocal theoretical prediction as black solid lines. This LMV model follows the original formulation by Luciani and Mora [9], in which the heat flux is expressed as

$$Q_e(x) = \int Q_{\rm SH}(x')\,\mathcal{W}(x,x')\,dx', \tag{2}$$

where the kernel $\mathcal{W}(x,x')$ describes the nonlocal contribution of the SH flux from all spatial positions x' to the location x, and is defined as

$$\mathcal{W}(x,x') = \frac{1}{2\lambda(x')}\exp\left[-\frac{X}{\lambda(x')}\right], \tag{3}$$

where the density-weighted distance measure X is given by $X = |\int_x^{x'} n_e(x'')\,dx''|/n_e(x')$ and $\lambda(x')$ represents the effective propagation range of electrons originating from position x' and is given by $\lambda(x') \simeq a\sqrt{Z+1}\,\lambda_{\rm free}(x')$. Here, the empirical parameter is taken as $a = 30$, following the original calibration of the LMV model [9], where $a \approx 32$ was found to reproduce Fokker–Planck results. This choice provides good agreement with our PIC benchmarks in the moderately nonlocal regime.

From Fig. 1, we observe that for a given temperature ratio R, simulations with different values of n_{e0} and L—but with the product $n_{e0}L\ln\Lambda$ held constant to preserve the same value of $L_T/\lambda_{\rm free}$ for the fixed central temperature—yield the same maximum normalized heat flux $Q_{\rm e,m}/Q_{\rm free}$. Note that in our simulation setup, $L_T/\lambda_{\rm free} \sim n_{e0}L\ln\Lambda(R+1)/(R-1)$. Moreover, all such cases exhibit consistent trends. Specifically, when $L_T/\lambda_{\rm free} > 500$, corresponding to the local transport regime, the normalized heat flux approaches the classical SH limit (The SH limit is exceeded within the classical regime at higher R values, attributable to the presence of hotter electrons originating from the hot side). In this regime, the LMV nonlocal theory also agrees well with the simulation results because the kernel $\mathcal{W}(x,x')$ effectively reduces to a Dirac delta function, and therefore recovers local heat flux.

As the system transitions into the nonlocal regime ($L_T/\lambda_{\rm free} < 500$), a clear mitigation of the heat flux is observed with respect to the SH model, i.e., $Q_{\rm e} < Q_{\rm SH}$. Notably, the simulation results capture the overall magnitude of the heat flux, consistent with the predictions of the LMV model in moderately nonlocal conditions. However, as the degree of nonlocality increases, particularly in cases with lower temperature ratios R, the LMV model begins to break down and significantly underestimates the heat flux simulated by the PIC.

In Fig. 2, we show that the LMV model fails to accurately reproduce the spatial profile of the heat flux even under moderately nonlocal conditions. In Figs. 2(a)–2(c), the background electron density is varied from 10^{21} to $10^{25}\,\mathrm{cm}^{-3}$, while the temperature ratio is held constant at $R = 5$, and the characteristic gradient scale length is fixed at $L_T \sim 195$. As a result, the transport parameter varies across the cases: $L_T/\lambda_{\rm free} \approx 83$ in Fig. 2(a), $L_T/\lambda_{\rm free} \approx 222$ in Fig. 2(b), and $L_T/\lambda_{\rm free} \approx 2500$ in Fig. 2(c). In Fig. 2(c), the simulation results show excellent agreement with both the SH and LMV models, indicating a return to local transport behavior in the high-collisionality regime. However, in Figs. 2(a) and 2(b), significant deviations are observed in the spatial shape of the heat flux when compared to both models.

The complex and often unpredictable behavior of heat transport in this regime underscores the limitations of traditional theoretical models. This motivates the adoption of data-driven approaches that can capture the underlying physics more accurately. In particular, modern neural networks provide a powerful and flexible framework for learning the transport dynamics directly from high-fidelity simulation data [18,29–34], without the need for empirical closures or restrictive analytical approximations.

Motivated by the structure of nonlocal transport theory as expressed in Eq. (2) and SNB model, where the kernel plays a central role in convolving the local SH heat flux across space, this work proposes a data-driven framework for learning such kernels using an NN architecture. Drawing inspiration from the physical definitions of X and λ [9,15,16], the NN is trained to learn the mapping from (X,λ) to the kernel function $\mathcal{W}$. To facilitate a dimensionless and numerically stable learning process, we express the input λ in units of the NN-grid spacing dx, so that λ becomes the dimensionless input. At the same time, the explicit dx appearing in the discrete sum over X is removed, exactly as done in the LMV discretization [Eq. (2)]. This treatment ensures that the NN inputs are scale invariant and that the kernel representation remains consistent with the LMV formalism. Furthermore, recognizing that heat transport

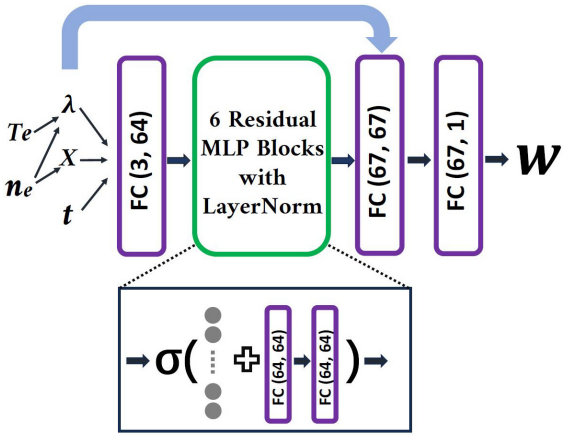


FIG. 3. Schematic of *LINN* used to learn the kernel function $\mathcal{W}$. The input features (λ, X) are derived from the density and temperature profiles, along with the time coordinate t are first projected into a 64-dimensional latent space via an FC layer followed by LayerNorm and GELU. The core of the network consists of six residual MLP blocks (shown enlarged below), each composed of two FC layers with LayerNorm, GELU, and a skip connection. A high-level skip connection (light-blue arrow) reintroduces the original input vector via concatenation, enabling the model to preserve and leverage key physical features. The resulting 67-dimensional representation is passed through additional FC layers to produce a positive scalar output via Softplus, yielding the kernel weight $\mathcal{W}$.

is inherently dynamic, the temporal coordinate (i.e., the time label, which is normalized by the saturation time corresponding to different samples) is introduced as a third input variable, enabling the model to capture the dynamic transport kernel conditioned on the given X and λ. Since the LMV model knowledge is embedded into the neural network, we refer to it as the *LMV-informed neural network (LINN)* hereafter. Then, the heat flux can be calculated by

$$\mathbf{Q}_e(x,t) = \sum_{x'} \mathcal{W}\left(\begin{matrix}\lambda(x')\\ X(x,x')\\ t\end{matrix}\right) \cdot \mathbf{Q}_{\mathrm{SH}}(x'). \quad (4)$$

The architecture of our machine learning model, including its inputs and outputs, is illustrated in Fig. 3. The PIC output (38 400 cells) is downsampled to 64 points to retain the heat flux profile while reducing noise and NN complexity. The inputs—λ and X, computed from the given density and temperature profiles, along with the time coordinate t—are first passed through a fully connected (FC) layer that projects the input into a 64-dimensional latent space, followed by LayerNorm and GELU activation (see the definitions in the Supplemental Material [35]). The core of the network comprises six residual blocks, each consisting of two FC layers with interleaved LayerNorm and GELU activations, wrapped in a residual skip connection (an example block is shown at the bottom of the figure). To preserve and emphasize the original physical input features, a skip connection (light-blue arrow) introduces the raw input vector directly into the postresidual output via concatenation. This combined representation (67 dimensional) is further processed through another FC layer with LayerNorm and GELU, and finally mapped to a non-negative scalar weight using a fully connected output layer with Softplus activation. This architecture balances expressive feature transformation with the retention of physically meaningful input structure, making it well suited for learning structured kernel functions $\mathcal{W}$ in nonlocal models (see the Supplemental Material [35] for details on the NN architecture and model performance).

To further enlarge the dataset, we apply a mirror transformation to the PIC data, since the transport physics is invariant under reflection, allowing both temperature and heat flux profiles to be mirrored. All 829×2 samples are partitioned into training, validation, and test sets using a 70:15:15 ratio. Upon constructing the model, we compare the kernels and heat flux predictions of *LINN* with those of the SH and LMV models (as shown in Fig. 4). The kernel contours $\mathcal{W}(x, x')$ are plotted in terms of their natural logarithm values. In Fig. 4(a), the central background density is $n_{e0} = 2.5 \times 10^{20}$ cm^{-3}, with a temperature ratio $R = 5$, and a central value of $L_T/\lambda_{\mathrm{free}} \approx 33$. Figure 4(b) presents results for $n_{e0} = 7.5 \times 10^{21}$ cm^{-3}, again with $R = 5$ and $L_T/\lambda_{\mathrm{free}} \approx 333$. In Fig. 4(c), the conditions are $n_{e0} = 10^{25}$ cm^{-3}, $R = 2$, and $L_T/\lambda_{\mathrm{free}} \approx 5000$. These three scenarios represent a transition from strongly nonlocal to moderately nonlocal and finally to the local transport regime. It is worth noting that none of these cases were included in the training set. The first (1) and second (2) rows correspond to the NN-predicted kernels at the early stage ($t = t_{\mathrm{sa}}/4$) and the saturated stage ($t = t_{\mathrm{sa}}$), respectively, where t_{sa} denotes the saturation time specific to each case. In comparison, the third (3) row shows the theoretical LMV kernels independent of time. While in the fourth (4) row, the SH and LMV heat flux are presented by the black dashed and solid lines. Because of the correct prediction of the kernels, the calculated heat flux (blue dashed lines: $t = t_{\mathrm{sa}}/4$; red dashed lines: $t = t_{\mathrm{sa}}$) via Eq. (4) obtained good agreement with the ground truth data (blue solid lines: $t = t_{\mathrm{sa}}/4$; red solid lines: $t = t_{\mathrm{sa}}$).

Physically, in the local transport regime, the heat flux kernel is expected to approach a diagonal matrix, reflecting purely local interactions. However, due to temperature asymmetry, the kernel near the hot side should be slightly broader than on the cold side. This is accurately captured by the NN predictions [as shown in Figs. 4(c.1) and 4(c.2)]. As the system becomes increasingly nonlocal, the kernel broadens more prominently on the hot side, while remaining relatively sharp on the cold side. This trend is clearly observed in Figs. 4(a.1), 4(a.2) and 4(b.1), 4(b.2), where the NN predictions remain more confined compared to the broader LMV kernels. We note that the time-independent kernels derived by Epperlein and Short [15] are even broader in form than the LMV kernels, yet the resulting heat flux is more localized than that of LMV, particularly in the strongly nonlocal regime. Notably, the influence of hot particles on the downstream region should decay as the distance from the heat source increases. However, in Figs. 4(a.3) and 4(b.3), the LMV model exhibits a concave feature (highlighted by the black boxes). In contrast, the NN-predicted kernels in Figs. 4(a.2) and 4(b.2) significantly suppress this nonphysical effect, adhering more faithfully to the expected physical behavior. The less smooth structures observed in Figs. 4(a.1) and 4(a.2) result from the fact that within the strongly nonlocal regime, the initial temperature profile would evolve significantly during the transient [36].

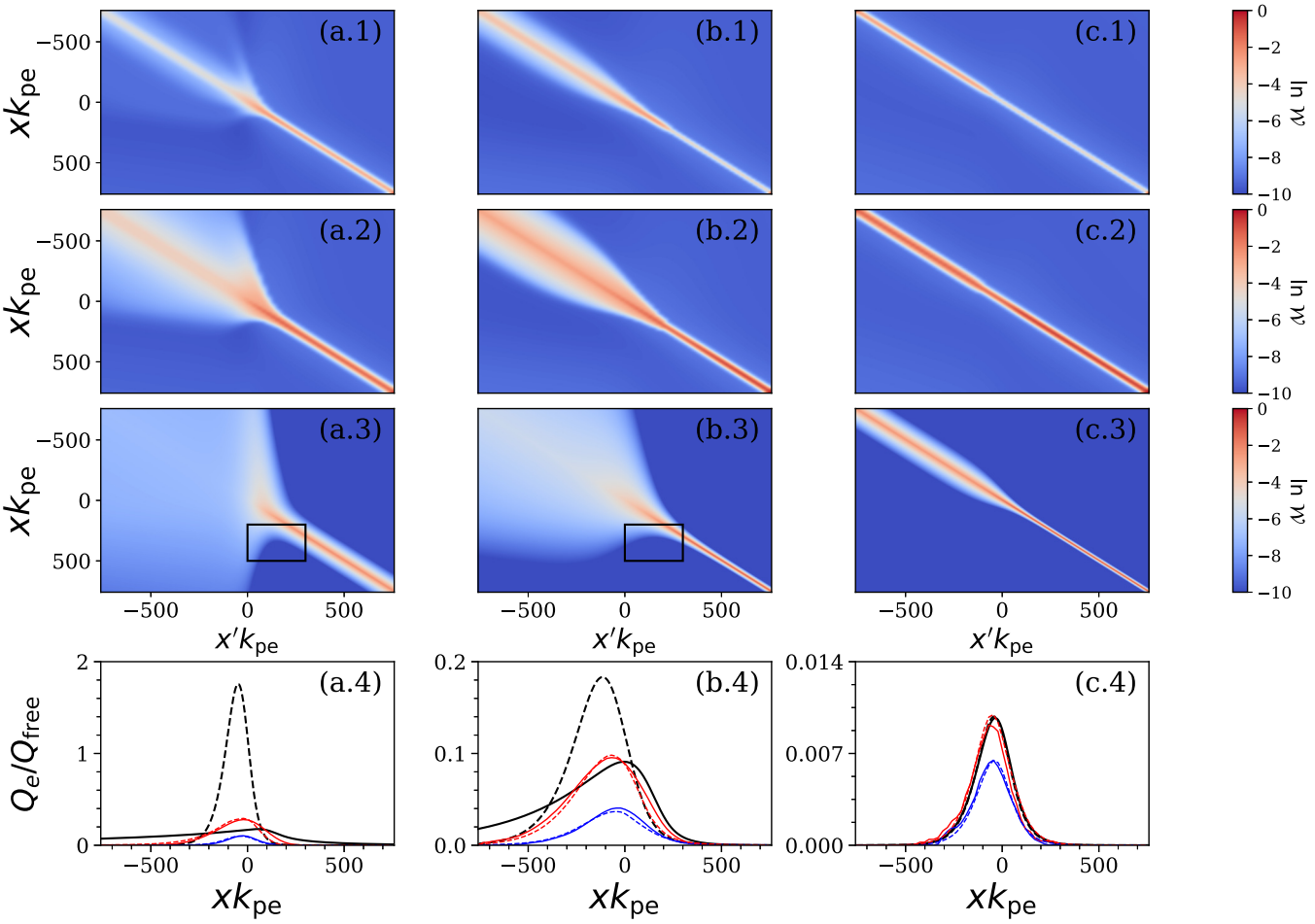


FIG. 4. Comparison of heat flux kernels and resulting heat fluxes for three representative test cases not seen during training, and the kernel contours $\mathcal{W}(x, x')$ are plotted in terms of their natural logarithm values. Panels (a)–(c) correspond to cases with increasing collisionality: (a) $n_{e0} = 2.5 \times 10^{20}$ cm^{-3}, $R = 5$, $L_T/\lambda_{free} \approx 33$; (b) $n_{e0} = 7.5 \times 10^{21}$ cm^{-3}, $R = 5$, $L_T/\lambda_{free} \approx 333$; and (c) $n_{e0} = 10^{25}$ cm^{-3}, $R = 2$, $L_T/\lambda_{free} \approx 5000$, representing the transition from strongly nonlocal to local transport. For each case, rows (1) and (2) show NN-predicted kernels at the early ($t = t_{sa}/4$) and saturated ($t = t_{sa}$) stages, where t_{sa} denotes the saturation time. Row (3) displays time-independent LMV theoretical kernels. Row (4) presents heat flux comparisons: SH (black dashed), LMV (black solid), and NN predictions (dashed red: $t = t_{sa}$, dashed blue: $t = t_{sa}/4$) against ground truth (solid red: $t = t_{sa}$, solid blue: $t = t_{sa}/4$).

In such cases, convolving the local SH flux obtained from the initial temperature profile across the whole domain is not correct. This suggests that the kernel-based approach breaks down in strongly nonlocal regimes [9]. An alternative model that avoids the use of kernels will be explored in future work.

In summary, we present a nonlocal LMV model-informed NN that learns nonlocal heat transport kernels directly from structured physical inputs. By incorporating key physical priors—including normalized mean free path, spatial separation, and time—the model captures essential dynamical features while maintaining interpretability, though the interpretability of the NN-predicted kernel diminishes in regimes of strong nonlinearity, a limitation that originates from the LMV model itself. The key outcome is that well-defined nonlocal kernels can be constructed directly from kinetic data, avoiding the complex empirical assumptions underlying analytic models such as LMV and SNB. This provides a systematic pathway for refining nonlocal transport theory itself. Furthermore, unlike classical approaches that assume static, time-independent kernels, our framework yields space- and time-dependent kernels that evolve dynamically during relaxation, thereby capturing transient transport physics inaccessible to stationary models. While classical kernels such as SH, LMV, and SNB are re-evaluated at each hydrodynamic time step with updated plasma conditions, their functional form remains stationary, implicitly assuming instantaneous flux saturation. This assumption breaks down in regimes where the transport relaxation time is comparable to the hydrodynamic timescale, so that the heat flux can no longer be regarded as instantaneous. While the present study is restricted to 1D configurations for computational feasibility—consistent with both the LMV model and studies in Ref. [16]—the proposed framework is readily generalizable to multidimensional settings once suitable multidimensional kinetic datasets become available. The learned kernels act as flexible, data-driven operators that respond to different plasma conditions while embedding core physical behavior, offering a natural path toward generalization across regimes. Beyond predictive accuracy, the model provides insight into

the spatiotemporal structure of heat transport, demonstrating the potential of machine learning embedded with physics models to uncover complex plasma physics in ways that extend beyond traditional analytical frameworks. Looking ahead, at the implementation level, embedding our NN-based kernel into radiation–hydrodynamics codes would follow the same strategy already used for LMV or SH models. We clarify that in our PIC approach, the initial temperature profile used for evaluating fluxes corresponds to the instantaneous profile at a hydrodynamic time step, entirely consistent with hydrodynamic practice. But implementation remains a major challenge, since our present training datasets are restricted to short kinetic scales, and direct PIC simulations at hydrodynamic scales would be prohibitively expensive. Therefore, as a forward-looking step, we are exploring strategies to bridge the scale gap by dividing hydrodynamic domains into smaller independent grids that can be treated with PIC simulations, thereby enabling the training of neural networks that could be coupled back into hydrodynamic codes in a multiscale framework. Nevertheless, the data driven and temporal form of the kernels demonstrated here may serve as a guide for the further development of nonlocal transport theory, especially in constructing physically consistent kernels that go beyond the stationary approximations used in most classical models.

The authors thank S. Nathaniel, S. Hüller, T. Fülöp, and C. Riconda for their useful discussions. They are also grateful for the computing resources provided by the STFC Scientific Computing Department's SCARF cluster. This work was supported by EPSRC and First Light Fusion under the AMPLIFI prosperity partnership, Grant No. EP/X025373/1.

Data availability. The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

[1] G. Gregori, S. H. Glenzer, J. Knight, C. Niemann, D. Price, D. H. Froula, M. J. Edwards, R. P. J. Town, A. Brantov, W. Rozmus, and V. Y. Bychenkov, Effect of nonlocal transport on heat-wave propagation, Phys. Rev. Lett. **92**, 205006 (2004).

[2] C. Arran, M. Holec, M. Sherlock, M. Marinak, and C. P. Ridgers, The accuracy of multi-group models for nonlocal electron transport in magnetized plasmas, Phys. Plasmas **32**, 043908 (2025).

[3] O. A. Hurricane, P. K. Patel, R. Betti, D. H. Froula, S. P. Regan, S. A. Slutz, M. R. Gomez, and M. A. Sweeney, Physics principles of inertial confinement fusion and U.S. Program overview, Rev. Mod. Phys. **95**, 025005 (2023).

[4] E. L. Yerger, M. W. Kunz, A. F. Bott, and A. Spitkovsky, Collisionless conduction in a high-beta plasma: A collision operator for whistler turbulence, J. Plasma Phys. **91**, E20 (2025).

[5] S. Komarov, A. A. Schekochihin, E. Churazov, and A. Spitkovsky, Self-inhibiting thermal conduction in a high-β, whistler-unstable plasma, J. Plasma Phys. **84**, 905840305 (2018).

[6] J. D. Huba, *NRL Plasma Formulary* (Naval Research Laboratory, Washington, DC, 2023).

[7] L. Spitzer and R. Härm, Transport phenomena in a completely ionized gas, Phys. Rev. **89**, 977 (1953).

[8] A. V. Brantov, V. Y. Bychenkov, V. T. Tikhonchuk, and W. Rozmus, Nonlocal electron transport in laser heated plasmas, Phys. Plasmas **5**, 2742 (1998).

[9] J. F. Luciani, P. Mora, and J. Virmont, Nonlocal heat transport due to steep temperature gradients, Phys. Rev. Lett. **51**, 1664 (1983).

[10] G. P. Schurtz, P. D. Nicolaï, and M. Busquet, A nonlocal electron conduction model for multidimensional radiation hydrodynamics codes, Phys. Plasmas **7**, 4238 (2000).

[11] M. Sherlock, J. P. Brodrick, and C. P. Ridgers, A comparison of non-local electron transport models for laser-plasmas relevant to inertial confinement fusion, Phys. Plasmas **24**, 082706 (2017).

[12] W. Q. Yuan, Z. H. Zhao, S. P. Zhu, X. T. He, and B. Qiao, Space-time dependent non-local thermal transport effects on laser ablation dynamics in inertial confinement fusion, Plasma Phys. Controlled Fusion **66**, 045023 (2024).

[13] O. Michel, R. Duclous, P.-E. Masson-Laborde, C. Enaux, and P. Lafitte, A nonlocal electron transport model in the diffusion scaling of hydrodynamics, Phys. Plasmas **30**, 022712 (2023).

[14] A. Chrisment, P. Loiseau, J.-L. Feugeas, P.-E. Masson-Laborde, J. Mathiaud, V. Tikhonchuk, and P. Nicolaï, Analysis of a kinetic model for electron heat transport in inertial confinement fusion plasmas, Phys. Plasmas **29**, 062301 (2022).

[15] E. M. Epperlein and R. W. Short, A practical nonlocal model for electron heat transport in laser plasmas, Phys. Fluids B **3**, 3092 (1991).

[16] O. V. Batishchev, V. Y. Bychenkov, F. Detering, W. Rozmus, R. Sydora, C. E. Capjack, and V. N. Novikov, Heat transport and electron distribution function in laser produced plasmas with hot spots, Phys. Plasmas **9**, 2302 (2002).

[17] C. Lamy, B. Dubroca, P. Nicolaï, V. Tikhonchuk, and J.-L. Feugeas, Modeling of electron nonlocal transport in plasmas using artificial neural networks, Phys. Rev. E **105**, 055201 (2022).

[18] F. Miniati and G. Gregori, Learning transport processes with machine intelligence, Sci. Rep. **12**, 11709 (2022).

[19] A. R. Bell, R. G. Evans, and D. J. Nicholas, Elecron energy transport in steep temperature gradients in laser-produced plasmas, Phys. Rev. Lett. **46**, 243 (1981).

[20] H. Zhao, S. Weng, Z. Sheng, S. Jin, and J. Zhang, A Vlasov-Fokker-Planck-Landau code for the simulation of colliding supersonic dense plasma flows, J. Comput. Phys. **503**, 112843 (2024).

[21] A. R. Bell and M. Sherlock, The fast VFP code for solution of the Vlasov–Fokker–Planck equation, Plasma Phys. Controlled Fusion **66**, 035014 (2024).

[22] Y. Sentoku, K. Mima, Y. Kishimoto, and M. Honda, Effects of relativistic binary collisions on PIC simulation of laser plasmas, J. Phys. Soc. Jpn. **67**, 4084 (1998).

[23] Y. Sentoku and A. Kemp, Numerical methods for particle simulations at extreme densities and temperatures: Weighted particles, relativistic collisions and reduced currents, J. Comput. Phys. **227**, 6846 (2008).

[24] A. J. Kemp, Y. Sentoku, V. Sotnikov, and S. C. Wilks, Collisional relaxation of superthermal electrons generated by relativistic laser pulses in dense plasma, Phys. Rev. Lett. **97**, 235001 (2006).

[25] M. Sherlock, E. G. Hill, R. G. Evans, S. J. Rose, and W. Rozmus, In-depth plasma-wave heating of dense plasma irradiated by short laser pulses, Phys. Rev. Lett. **113**, 255001 (2014).

[26] T. D. Arber, K. Bennett, C. S. Brady, A. Lawrence-Douglas, M. G. Ramsay, N. J. Sircombe, P. Gillies, R. G. Evans, H. Schmitz, A. R. Bell, and C. P. Ridgers, Contemporary particle-in-cell approach to laser-plasma modelling, Plasma Phys. Controlled Fusion **57**, 113001 (2015).

[27] R. A. Fonseca, L. O. Silva, F. S. Tsung, V. K. Decyk, W. Lu, C. Ren, W. B. Mori, S. Deng, S. Lee, T. Katsouleas, and J. C. Adam, Osiris: A three-dimensional, fully relativistic particle in cell code for modeling plasma based accelerators, in *Computational Science—ICCS 2002*, edited by P. M. A. Sloot, A. G. Hoekstra, C. J. K. Tan, and J. J. Dongarra (Springer, Berlin, 2002), pp. 342–351.

[28] S. X. Hu, K. A. Nichols, N. R. Shaffer, B. Arnold, A. J. White, L. A. Collins, V. V. Karasiev, S. Zhang, V. N. Goncharov, R. C. Shah, D. I. Mihaylov, S. Jiang, and Y. Ping, A review on charged-particle transport modeling for laser direct-drive fusion, Phys. Plasmas **31**, 040501 (2024).

[29] Z. Huang, C. Dong, and L. Wang, Machine-learning heat flux closure for multi-moment fluid modeling of nonlinear Landau damping, Proc. Natl. Acad. Sci. USA **122**, e2419073122 (2025).

[30] E. R. Ingelsten, M. C. McGrae-Menge, E. P. Alves, and I. Pusztai, Data-driven discovery of a heat flux closure for electrostatic plasma phenomena, J. Plasma Phys. **91**, E64 (2025).

[31] E. P. Alves and F. Fiuza, Data-driven discovery of reduced plasma physics models from fully kinetic simulations, Phys. Rev. Res. **4**, 033192 (2022).

[32] A. Döpp, C. Eberle, S. Howard, F. Irshad, J. Lin, and M. Streeter, Data-driven science and machine learning methods in laser–plasma physics, High Power Laser Sci. Eng **11**, e55 (2023).

[33] C. J. McDevitt and X. Tang, A physics-informed deep learning description of Knudsen layer reactivity reduction, Phys. Plasmas **31**, 062701 (2024).

[34] D. D. Carvalho, D. R. Ferreira, and L. O. Silva, Learning the dynamics of a one-dimensional plasma model with graph neural networks, Mach. Learn.: Sci. Technol. **5**, 025048 (2024).

[35] See Supplemental Material at http://link.aps.org/supplemental/10.1103/m215-3f6c for more details on the network configuration and model performance.

[36] J. P. Matte and J. Virmont, Electron heat transport down steep temperature gradients, Phys. Rev. Lett. **49**, 1936 (1982).