

Resolution-Robust Machine Learning Heat Flux Closure for Inertial Confinement Fusion Plasmas

M. Luo^{1,*}, A. R. Bell^{1,2}, F. Miniati³, S. M. Vinko¹ and G. Gregori¹¹Department of Physics, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom²Central Laser Facility, STFC Rutherford Appleton Laboratory, Harwell, Oxfordshire OX11 0QX, United Kingdom³Mach42, Robert Robinson Avenue, Oxford Science Park, Oxford OX4 4GP, United Kingdom

(Received 21 March 2026; revised 14 June 2026; accepted 30 July 2026; published 3 September 2026)

Accurate modeling of heat flux in inertial confinement fusion plasmas requires closures that remain predictive far from local equilibrium and across disparate spatial and temporal resolutions. We develop a resolution-robust machine learning heat flux closure trained on particle-in-cell simulations using a Fourier neural operator. Two nonlocal electron thermal conduction models are trained and tested. When embedded self-consistently into the electron energy equation, the learned closure faithfully reproduces the temperature evolution and shows good temporal extrapolation and generalization capability. Remarkably, models trained on coarse-resolution data accurately predict heat flux when deployed in substantially finer-resolution implicit, iterative solvers of the energy equation, significantly enhancing the practicality of embedding data-driven closures into partial differential equation solvers. These results establish a data-driven closure that bridges kinetic and fluid descriptions and provides a viable pathway for treating machine learning as an iterative solver within the radiation-hydrodynamic simulations of inertial confinement fusion plasma.

DOI: 10.1103/914n-mnz6

Understanding heat transport in plasmas is essential for a broad range of applications, from inertial confinement fusion (ICF) [1] and high energy density physics to astrophysical systems [2,3] and laboratory astrophysics. Heat transport can be driven by temperature gradients [4] or currents [5], and in all these contexts, accurate closure of the heat flux in terms of lower-order fluid quantities such as density, temperature, and electromagnetic fields is essential for large-scale simulations, particularly where microscale kinetic instabilities may significantly impact macroscale dynamics and where fully kinetic simulations remain computationally prohibitive. In hot plasmas, the character of temperature-gradient-driven heat transport is governed by the Knudsen number, which is defined as the ratio of the electron mean free path λ_0 [6] to the temperature gradient scale length, $L_T = T_e/|\nabla T_e|$; here, T_e is the electron temperature, and the electron mean free path is $\lambda_0 = 4\pi\epsilon_0^2 m_e^2 v_{th}^4 / (Z n_e e^4 \ln \Lambda)$ [6], with ϵ_0 the vacuum permittivity, m_e the electron mass, $v_{th} = \sqrt{T_e/m_e}$ the electron thermal velocity, Z the ion charge state, n_e the electron number density, e the elementary charge, and $\ln \Lambda$ the Coulomb logarithm. When $\lambda_0/L_T \ll 0.002$, transport is well described by local theories, and classical models such as the Spitzer-Härm (SH) formulation [7] yield reliable predictions. As λ_0/L_T increases, kinetic and nonlocal effects become increasingly important, leading to the breakdown of local hydrodynamic closures [8].

Among existing nonlocal theories [9–11], the Schurtz-Nicolaï-Busquet (SNB) model [9] has been widely adopted

in radiation-hydrodynamic simulations of ICF and implemented in several ICF codes [12–15]. To further improve agreement with Vlasov-Fokker-Planck (VFP) [16] simulations, a number of SNB variants [17–19] have been proposed, primarily through modifications of the effective mean free path within the SNB framework. Despite these efforts, the discrepancies with kinetic simulations persist in strongly nonlocal regimes [16], and the traditional numerical solver of the SNB model introduces substantial computational overhead when coupled to a radiation-hydrodynamic code, and this extra computational expense motivates the exploration of data-driven approaches [20–26] as potential alternatives for nonlocal heat flux closure. However, conventional neural networks, such as multilayer perceptrons and convolutional networks, learn mappings tied to fixed discretizations, which limits their suitability for closures that must be embedded into partial differential equations and remain predictive across varying spatial and temporal resolutions.

In this article, we employ a neural operator framework [27] to learn the functional mapping from the electron temperature profile, $T_e(x)$, to the divergence of the heat flux, $\partial_x q(x)$. Specifically, we adopt the Fourier neural operator (FNO) [28], which exploits spectral representations to reduce the computational complexity to $\mathcal{O}(n \log n)$ (n is the number of discretized points), while naturally encoding the intrinsically nonlocal character of the heat transport through global interactions in Fourier space. Unlike the earlier data-driven approaches cited above, the present work makes three distinct advances: (1) operator learning via the FNO, which learns mappings between function spaces independent of discretization; (2) self-consistent embedding of the learned closure into an implicit iterative partial differential equation (PDE) solver; and (3) cross-resolution deployment, where models trained on coarse data are deployed at finer resolution without retraining, reducing data

*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.

requirements. Fully kinetic particle-in-cell (PIC) simulations were performed with the OSIRIS code [29] to generate the data. A comparison between published VFP results, PIC simulations, and the SNB model is provided in the Supplemental Material [30]. Our implementation of the SNB model follows that of Refs. [17,19], ensuring consistency with previous analyses. A total of 180 energy groups ($N_g = 180$) is employed, with the uniform energy grid ranging from 0 velocity up to 25 times the highest temperature value, although a convergence study confirms that $N_g = 15$ is sufficient; see also the Supplemental Material [30] for further details. The data used to train the model comprise two transport test cases: the relaxation of a *hot spot* [31] and the decay of a small-amplitude sinusoidal temperature perturbation, commonly referred to as the *Epperlein-Short (ES)* test [32], both considered at constant plasma density.

Spatial (x) and temporal (t) coordinates are normalized by the mean free path λ_0 and the corresponding collision time $1/\nu_0$, evaluated at a reference density n_{e0} and temperature T_{e0} (in the following, temperature profiles are also normalized by the reference temperature T_{e0} , and the heat flux is normalized by the free stream flux $n_{e0}T_{e0}v_{th0}$, $v_{th0} = \sqrt{T_{e0}/m}$). The Coulomb logarithm is taken to be constant, and the periodic boundary condition is applied. The simulations are performed in a one-dimensional domain of length $L = 100$ (100 times λ_0) over a time interval $\tau = 30$ (30 times $1/\nu_0$). The spatial step adopted in PIC simulation is $\Delta x = \lambda_{De,0}/2$, where $\lambda_{De,0} = k_{pe}^{-1}/50 = \lambda_0/8$ is the Debye length ($k_{pe} = \omega_{pe}/c$ is the plasma wave number and $\omega_{pe} = \sqrt{n_{e0}e^2/\varepsilon_0 m_e}$ denotes the electron plasma frequency), and the time step Δt is chosen to satisfy the Courant-Friedrichs-Lewy condition, with $\Delta t/\Delta x \approx 0.9$. A fully ionized hydrogen plasma with ion charge state $Z = 1$ is assumed. Each cell contains 6.4×10^4 electron and 6.4×10^4 ion macroparticles. Simulation data are recorded at spatial and temporal resolutions of $\mathbf{dx} = \Delta x \approx 0.06$ and $\mathbf{dt} = 9\Delta t = 0.01$, respectively.

For the hot-spot case, the initial temperature profile is prescribed as $T_e(x) = 1 + \exp[-x^2/(\alpha\ell)^2]$, where $\ell = 8.44$ and the parameter α is chosen from the set $\{1, 1.25, 1.5, 1.75, 2\}$. The Knudsen number reaches values of ~ 0.17 , at which strongly nonlocal behavior of heat transport occurs. In the ES case, the initial temperature profile is given by $T_e(x) = 1 + \delta T_0 \cos(\beta k_0 x)$, where $\delta T_0 = 4\%$, $k_0 = 2\pi/L$, and $\beta \in \{1, 4, 7, 10, 13, 16\}$. These choices correspond to $k\lambda_0 = \beta k_0 \lambda_0 \in \{0.063, 0.251, 0.440, 0.628, 0.817, 1.000\}$; here, $k = \beta k_0$ is the perturbation wave number. The model learns the operator $\partial_x q = \mathcal{F}_{\text{FNO}}(T_e)$ based on PIC simulation data (hot spot and ES) belonging to the time interval $t \in (0, 20]$. We expect the trained operator $\mathcal{F}_{\text{FNO}}(T_e)$ to capture the temperature evolution governed by

$$(3/2)\partial T_e/\partial t + \mathcal{F}_{\text{FNO}}(T_e) = S_{\text{ext}}, \quad (1)$$

which is solved implicitly using an iterative scheme [33]. The heating source or external energy deposition, S_{ext} , is temporally set to zero. The saved data in PIC are downsampled before being input into the same FNO architecture, to demonstrate learning rather than memorization. That dx and dt denote the spatial and temporal resolutions of the

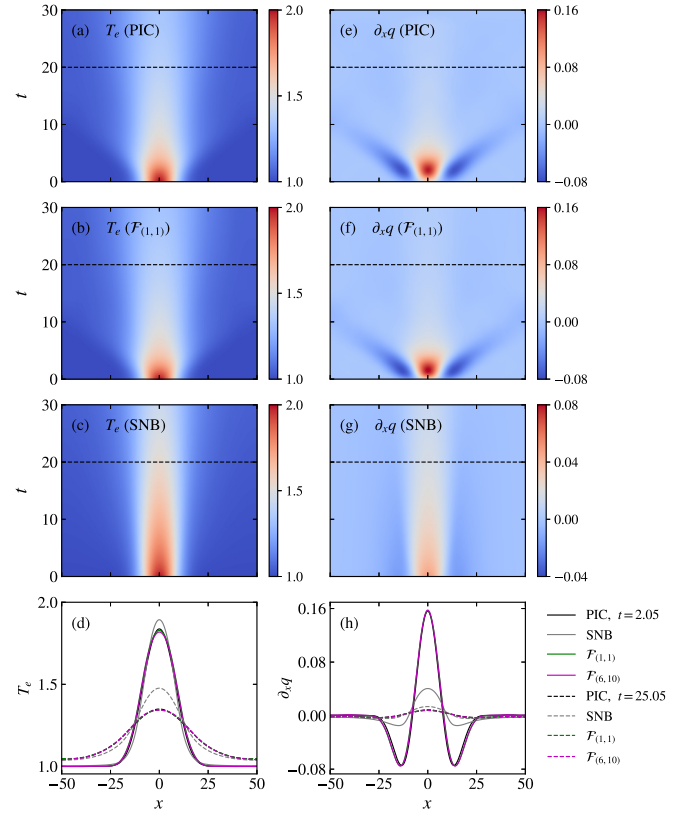


FIG. 1. Spatiotemporal evolution of the electron temperature T_e (left column) and the heat flux divergence $\partial_x q$ (right column) for $\alpha = 1.125$ of the hot-spot case. The first, second, and third rows correspond to results from PIC simulations, solutions of Eq. (1) using the FNO model $\mathcal{F}_{(1,1)}$, and solutions obtained using the SNB model, respectively. The horizontal dashed line at $t = 20$ separates the training interval $t \in (0, 20]$ from the temporal extrapolation interval $t \in (20, 30]$. The bottom row shows line plots of T_e and $\partial_x q$ at $t = 2.05$ and $t = 25.05$ of the PIC simulations, solutions of Eq. (1) using the FNO model $\mathcal{F}_{(1,1)}$, $\mathcal{F}_{(6,10)}$, and solutions obtained using the SNB model.

downsampled data. For convenience, the model trained at a specific resolution, $dx/\mathbf{dx} = n$ and $dt/\mathbf{dt} = m$, is denoted by $\mathcal{F}_{(n,m)}$. The training strategies and the model performance for $T_e \mapsto \partial_x q$ are detailed in the Supplemental Material [30].

By introducing the surrogate model into Eq. (1) via $\mathcal{F}_{\text{FNO}}(T_e) = \mathcal{F}_{(n,m)}(T_e)$, we assess its predictive capability as follows.

Hot-spot test. In addition to the data used to train the model, namely, $\alpha \in \{1, 1.25, 1.5, 1.75, 2\}$, which define the width of the initial Gaussian temperature profile, the model is also evaluated on unseen intermediate values $\alpha \in \{1.125, 1.375, 1.625, 1.875\}$ to assess generalization. Figure 1 shows the spatiotemporal evolution of the temperature T_e (left column) and the divergence of the heat flux $\partial_x q$ (right column) for $\alpha = 1.125$, shown at the resolution $dx = \mathbf{dx}$ and $dt = \mathbf{dt}$ over $t \in (0, 30]$. That is, Eq. (1) is solved using $dx = \mathbf{dx}$ and $dt = \mathbf{dt}$. Note the model $\mathcal{F}_{(n,m)}$ is trained using data from $t \in (0, 20]$, whereas the interval $t \in (20, 30]$ represents the model's predicted future evolution, i.e., temporal extrapolation. The first, second, and third rows correspond to results

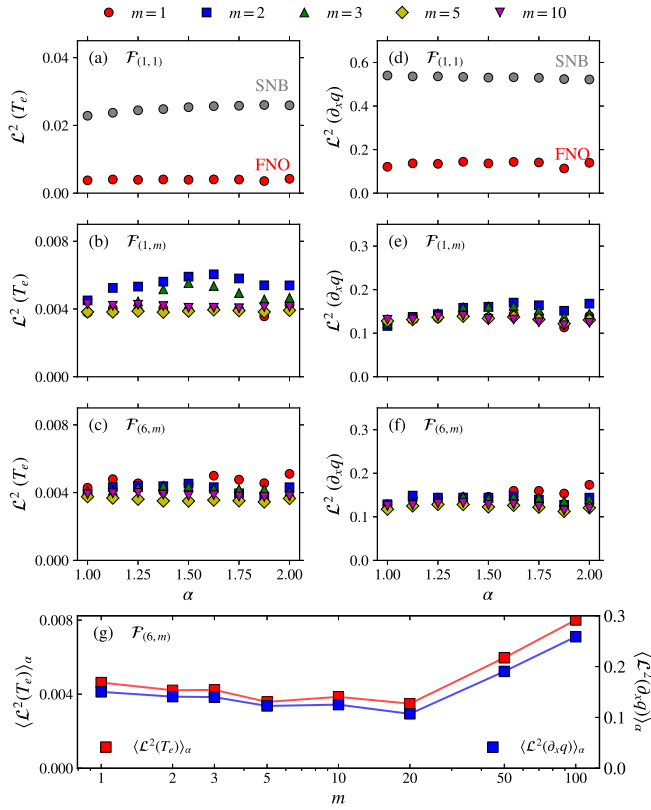


FIG. 2. Relative $\mathcal{L}^2$ errors for the hot-spot case obtained using different $\mathcal{F}_{(n,m)}$ models. The left and right columns show the relative $\mathcal{L}^2$ errors for T_e and $\partial_x q$, respectively. The top row shows results for the $\mathcal{F}_{(1,1)}$ -based (red circles) and SNB-based (gray circles) solutions. The second and third rows correspond to solutions obtained using $\mathcal{F}_{(1,m)}$ and $\mathcal{F}_{(6,m)}$, respectively, with different colors indicating $m = 1$ (red circles), 2 (blue squares), 3 (green up-triangles), 5 (yellow diamonds), and 10 (magenta down-triangles). The bottom row shows the averaged relative errors $\langle \mathcal{L}^2(T_e) \rangle_\alpha$ (left axis) and $\langle \mathcal{L}^2(\partial_x q) \rangle_\alpha$ (right axis) of the model $\mathcal{F}_{(6,m)}$ over all α as functions of the temporal resolution ($m = dt/dt$) in the training data.

from PIC simulations, from solving Eq. (1) using the FNO model $\mathcal{F}_{(1,1)}$, and the SNB model, respectively, and Eq. (1) is solved implicitly and iteratively, with two iteration cycles performed at each time step for both the $\mathcal{F}_{(1,1)}$ and SNB models. The solution of Eq. (1) using $\mathcal{F}_{(1,1)}$ agrees well with PIC results, whereas the SNB-based solution shows substantial discrepancies. In particular, the SNB model underestimates $\partial_x q$ compared to the PIC, leading to a reduced change in temperature. The bottom row in Fig. 1 shows line plots of T_e and $\partial_x q$ at $t = 2.05$ and $t = 25.05$ (temporal extrapolation). Besides the solution of Eq. (1) with $\mathcal{F}_{(1,1)}$, the solution of Eq. (1) with $\mathcal{F}_{(6,10)}$ is also plotted. We observe that the solution of Eq. (1) using $\mathcal{F}_{(6,10)}$ still provides accurate predictions despite being trained on coarse-resolution data; especially, these two time indices shown were never used during the training of $\mathcal{F}_{(6,10)}$.

The performance of all $\mathcal{F}_{(n,m)}$ models on hot-spot case is summarized in Fig. 2 and evaluated using the relative $\mathcal{L}^2$ error, defined as $\mathcal{L}^2(\mathcal{E}) = (\sum_{x,t} [E_g - \mathcal{F}_\mathcal{E}]^2)^{1/2} / (\sum_{x,t} E_g^2)^{1/2}$.

Here, $\mathcal{E}$ denotes either T_e or $\partial_x q$, the subscript g indicates the ground truth from PIC results, and $\mathcal{F}_\mathcal{E}$ represents the solution of Eq. (1) obtained using the corresponding $\mathcal{F}_{(n,m)}$. $\mathcal{L}^2(T_e)$ and $\mathcal{L}^2(\partial_x q)$ are shown in the left and right columns of Fig. 2, respectively. The top row presents $\mathcal{L}^2(\mathcal{E})$ of the $\mathcal{F}_{(1,1)}$ -based (red circles) and SNB-based (gray circles) solutions of Eq. (1), showing the superior accuracy of $\mathcal{F}_{(1,1)}$ -based solution, as already observed in Fig. 1. Although improving the SNB accuracy is possible by fine-tuning the effective mean free path definition, such tuning is inherently problem dependent, as evidenced by the several SNB variants in the literature. $\mathcal{L}^2(\mathcal{E})$ of the solutions obtained using $\mathcal{F}_{(1,m)}$ and $\mathcal{F}_{(6,m)}$ are shown in the second and third rows, respectively, with different colors indicating $m = 1$ (red circles), 2 (blue squares), 3 (green up-triangles), 5 (yellow diamonds), and 10 (magenta down-triangles). $\mathcal{L}^2(\mathcal{E})$ remains stable across α and is largely insensitive to the choice of $\mathcal{F}_{(n,m)}$. Notably, models trained on coarse-resolution data (e.g., $\mathcal{F}_{(6,10)}$) still achieve high accuracy when deployed within the fine resolution solver of Eq. (1). Figure 2(g) shows the averaged relative errors $\langle \mathcal{L}^2(T_e) \rangle_\alpha$ (left axis) and $\langle \mathcal{L}^2(\partial_x q) \rangle_\alpha$ (right axis) of the model $\mathcal{F}_{(6,m)}$ over all α as functions of the temporal resolution in the training data. We find even for a coarser resolution of $dt/dt = 20$, the model $\mathcal{F}_{(6,20)}$ retains a comparable $\mathcal{L}^2$ error around 0.4% for the temperature evolution, which is practically important for embedding the learned closure into energy equation solvers that may operate at different spatial and temporal resolutions. In addition, solving Eq. (1) using $\mathcal{F}_{(6,10)}$ with larger time steps, $dt = 10dt$ and $dt = 100dt$, still yields quantitative agreement with the PIC results (shown in the Supplemental Material [30]), although more iterations are required to advance each time step.

To further assess the model's temporal extrapolation capability, Fig. 3(a) shows the temporal evolution of the temperature obtained from PIC simulations (black) and from solutions of Eq. (1) using the FNO models $\mathcal{F}_{(1,1)}$ (green) and $\mathcal{F}_{(6,10)}$ (magenta). Results are shown at $t = 20$ (solid), $t = 25$ (dashed), and $t = 30$ (dotted) for an initial Gaussian temporal profile with $\alpha = 1.5$. Excellent agreement between the PIC results and the FNO predictions is observed. The temporal relative $\mathcal{L}^2$ error of the temperature, defined as $\mathcal{L}_t^2(T_e) = (\sum_x [T_e(t)_g - \mathcal{F}_{T_e(t)}]^2)^{1/2} / (\sum_x T_e(t)_g^2)^{1/2}$, is shown in Fig. 3(b) with $\alpha \in \{1.125, 1.375, 1.5, 1.625, 1.875\}$, and models $\mathcal{F}_{(1,1)}$ (solid, trained using the highest resolution data) and $\mathcal{F}_{(6,10)}$ (dashed, trained using the lowest resolution data) are examined. Over nearly the entire time domain, the relative error $\mathcal{L}_t^2(T_e)$ remains below 1%, including during the temporal extrapolation interval $t \in (20, 30]$, demonstrating the high predictive accuracy of the learned model. Furthermore, the error curves from $\mathcal{F}_{(1,1)}$ (solid) and $\mathcal{F}_{(6,10)}$ (dashed) for the same α show comparable behavior, except in the far-tail region, further indicating the resolution robustness of the training and the performance.

Figures 3(c) and 3(d) show the performance of the model $\mathcal{F}_{(1,1)}$ for different initial temperature profiles (black curves): a sub-Gaussian profile $T_e = 1 + \exp[-|x/(\alpha\ell)|^1]$ in panel (c) and a super-Gaussian profile $T_e = 1 + \exp[-|x/(\alpha\ell)|^3]$ in panel (d). Here, $\alpha = 1.5$, and the reference Gaussian profile is also shown in light blue. The temperature profiles at two time instances, $t = 5$ (strong nonlocality) and $t = 25$ (less non-

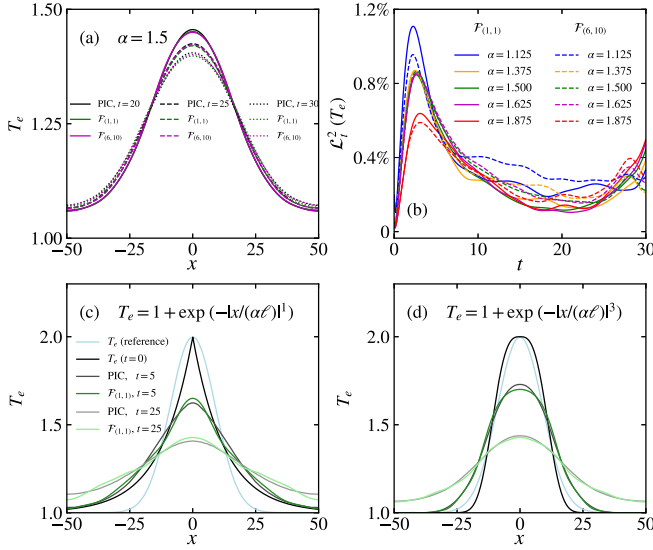


FIG. 3. Assessment of temporal extrapolation and generalization of the FNO model. (a) Temperature evolution obtained from PIC simulations (black) and from solutions of Eq. (1) using the FNO models $\mathcal{F}(1, 1)$ (green) and $\mathcal{F}(6, 10)$ (magenta) for an initial Gaussian profile with $\alpha = 1.5$. Profiles at $t = 20$ (solid), $t = 25$ (dashed), and $t = 30$ (dotted) are shown. (b) Temporal relative $\mathcal{L}^2$ error of the temperature, $\mathcal{L}_t^2(T_e)$, for $\alpha \in \{1.125, 1.375, 1.5, 1.625, 1.875\}$ using $\mathcal{F}(1, 1)$ (solid, trained with the highest-resolution data) and $\mathcal{F}(6, 10)$ (dashed, trained with the lowest-resolution data). Panels (c) and (d) show model performance for different initial temperature profiles (black): sub-Gaussian $T_e = 1 + \exp[-|x/(\alpha\ell)|^1]$ in panel (c), super-Gaussian $T_e = 1 + \exp[-|x/(\alpha\ell)|^3]$ in panel (d), respectively, with $\alpha = 1.5$. The reference Gaussian profile is shown in light blue. Temperature profiles are shown at $t = 5$ (strong nonlocality) and $t = 25$ (less nonlocality).

locality), are shown. The sub-Gaussian profile has a sharper peak at the center and longer tails compared to the reference Gaussian, while the super-Gaussian profile has a flatter top but sharper wings, concentrating the temperature gradient at the profile boundaries. Despite these structural differences from the Gaussian profiles used during training, the solutions of Eq. (1) obtained with $\mathcal{F}(1,1)$ faithfully reproduce the PIC simulation results, demonstrating the model's generalization capability. Further generalization tests are shown in the Supplemental Material [30].

ES test. We next examine the performance of all $\mathcal{F}(n,m)$ models for the ES case. The temperature decay ($\delta T = T_e - T_{e0}$) for the case $\beta = 7$, leading to $k\lambda_0 \approx 0.44$, is shown in Fig. 4. The left and right panels of the top row show results from PIC simulations and the solutions of Eq. (1) obtained using the FNO model $\mathcal{F}(1, 1)$, respectively. Again, the PIC ground truth is shown at the reference resolution $dx = \mathbf{d}\mathbf{x}$ and $dt = \mathbf{d}\mathbf{t}$, and Eq. (1) is solved at the same resolution. The solution of Eq. (1) obtained using $\mathcal{F}(1,1)$ exhibits high accuracy. The bottom row of Fig. 4 shows line plots of δT at $t = 1.05$ and $t = 5.05$ of the PIC simulations, solutions of Eq. (1) using the FNO model $\mathcal{F}(1,1)$, $\mathcal{F}(6,10)$, and solutions obtained using the SNB model. Despite being trained on coarser data, $\mathcal{F}(6,10)$ produces solutions that remain quantitatively ac-

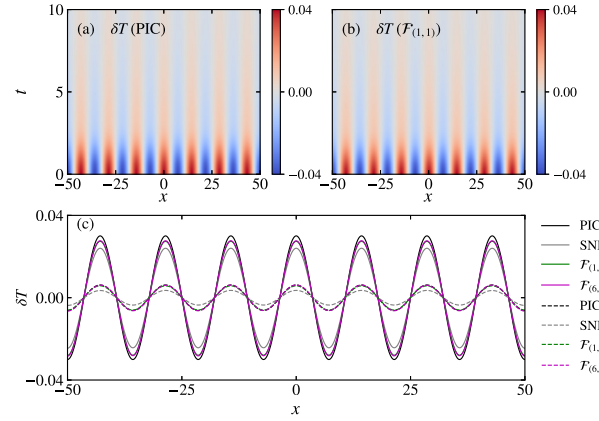


FIG. 4. Spatiotemporal evolution of the temperature perturbation $\delta T = T_e - T_{e0}$ for the ES case with $\beta = 7$ or $k\lambda_0 \approx 0.44$. The left and right panels of the top row show results from PIC simulations and the solutions of Eq. (1) obtained using the FNO model $\mathcal{F}(1, 1)$, respectively. The bottom row shows line plots of δT at $t = 1.05$ and $t = 5.05$ of the PIC simulations, solutions of Eq. (1) using the FNO model $\mathcal{F}(1,1)$, $\mathcal{F}(6,10)$, and solutions obtained using the SNB model.

curate, whereas the solution obtained with the SNB model shows substantial deviations from the PIC results. We again emphasize that these time instances were not included in the training of $\mathcal{F}(6,10)$.

A key quantity in the ES case is the decay rate γ of the temperature perturbation $\delta T(t)$, defined through $\delta T(t) \propto \exp(-\gamma t)$, and the decay rate can be given as $\gamma = 2k^2\kappa/3n_{e0}$, where k is the perturbation wave number and κ is the thermal conductivity. Consequently, the effective thermal conductivity normalized to the SH thermal conductivity can be inferred from the ratio $\kappa/\kappa_{\text{SH}} = \gamma/\gamma_{\text{SH}}$. When the SH heat flux closure is employed, the corresponding decay rate satisfies $\gamma_{\text{SH}}/\nu_0 = (2/3)(k\lambda_0)^2(128/\sqrt{2\pi})(Z + 0.24)/(Z + 4.2)$. Figure 5(a) shows two representative examples of temperature decay corresponding to $\beta = 4$ and 7 , i.e., $k\lambda_0 \approx 0.25$ and 0.44 , respectively. The black dashed curves indicate the PIC results, while the gray dotted curves correspond to solutions of Eq. (1) obtained using the SNB model. In both cases, the temperature continues to decay, as the heat flux remains driven by a persistent temperature gradient throughout the time window shown. However, the PIC results exhibit a slower decay at very early times, a more rapid decay at intermediate times, and a slower late-time evolution compared with the SNB model. Regarding the $\mathcal{F}(n,m)$ -based solutions of Eq. (1), indicated by different colors (red for $\mathcal{F}(1,m)$, blue for $\mathcal{F}(3,m)$, and green for $\mathcal{F}(6,m)$), all consistently reproduce the PIC temperature decay, irrespective of the resolution of the training data.

Figure 5(b) shows the conductivity ratio $\kappa/\kappa_{\text{SH}}$ obtained from different models. The red diamonds and triangles denote published results from the VFP codes OSHUN [18] and KIPP [19], respectively, while the solid and dashed pink lines correspond to the fitting function in Eq. (25) and the SNB model reported in Ref. [19]. The solution of Eq. (1) obtained using the SNB model applied here, shown by gray crosses, coincides with the dashed pink line. Importantly, both $\kappa_{\text{PIC}}/\kappa_{\text{SH}}$ (black squares) and $\kappa_{\mathcal{F}(n,m)}/\kappa_{\text{SH}}$ (green diamonds), extracted

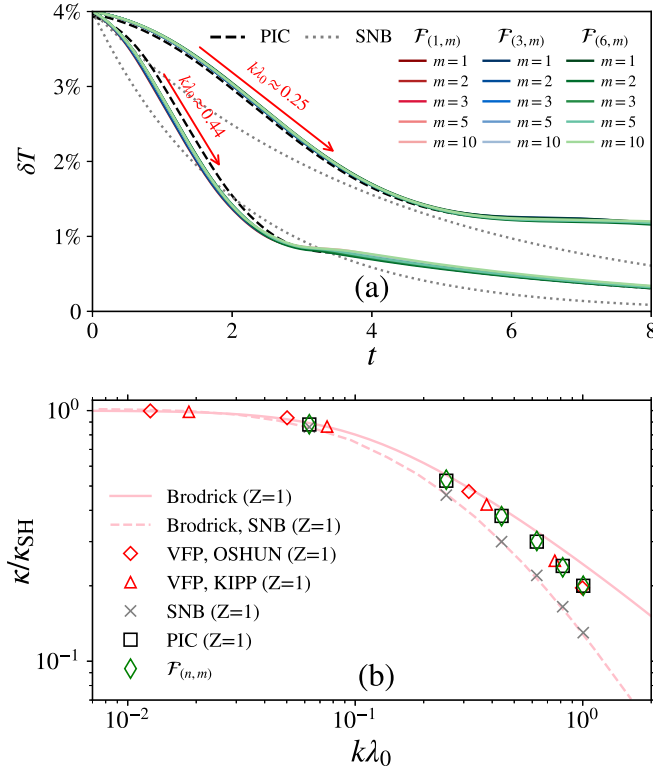


FIG. 5. (a) Temporal decay of the temperature perturbation $\delta T(t)$ for the ES cases with $\beta = 4$ and 7 , corresponding to $k\lambda_0 \approx 0.25$ and 0.44 , respectively. Black dashed curves show PIC results, gray dotted curves show solutions of Eq. (1) obtained using the SNB model, and colored curves correspond to solutions obtained using the $\mathcal{F}(n, m)$ models trained at different resolutions (red for $\mathcal{F}_{(1, m)}$, blue for $\mathcal{F}_{(3, m)}$, and green for $\mathcal{F}_{(6, m)}$). (b) Effective thermal conductivity ratio κ/κ_{SH} as a function of $k\lambda_0$. The red diamonds and red triangles denote published results from the VFP codes OSHUN [18] and KIPP [19], respectively. The solid and dashed pink lines correspond to the fitting function in Eq. (25) and the SNB model reported in Ref. [19]. The gray crosses indicate results obtained by solving Eq. (1) using the SNB model applied here. And the black squares and green diamonds show κ_{PIC}/κ_{SH} and $\kappa_{\mathcal{F}(n, m)}/\kappa_{SH}$, respectively, extracted from the early time decay rates of the temperatures in panel (a).

from the intermediate-time decay rates shown in Fig. 5(a), exhibit trends consistent with the VFP results. Besides, another generalization test showing the competition between spatially localized energy deposition and thermal relaxation is discussed in the Appendix.

In summary, this article demonstrates a pathway to replace the conventional SNB heat flux closure with a machine learning operator trained directly on kinetic simulations. Although the hot-spot and ES problems differ significantly in amplitude and modulation, the learned operator captures the essential kinetic features of nonlocal heat transport of both problems while enabling fast and stable iterative solutions (replacing the $N_g = 15$ SNB closure with the learned operator $\mathcal{F}_{(n, m)}$ is approximately $3\times$ faster. This speedup is expected to be substantially enhanced in larger-scale and multidimensional configurations). At the same time, the learned model shows

good temporal extrapolation and generalization capability. Importantly, models trained on coarse-resolution data can be seamlessly deployed within electron energy solvers at arbitrary resolution, including finer grids, significantly enhancing the practicality of embedding data-driven closures into partial differential equation solvers. Meanwhile, training on coarse data further improves the model training efficiency and reduces data requirements without compromising the accuracy of electron temperature evolution. In practice, solving Eq. (1) at the time resolution $dt = \mathbf{d}t$ requires 6000 iterations, corresponding to 3000 time steps with two iterations per step, while only 2.3% (1.1%) of these profiles are used to train the model $\mathcal{F}_{(n, 10)}$ ($\mathcal{F}_{(n, 20)}$). Even using only 0.2% of these profiles ($\mathcal{F}_{(n, 100)}$), the error just increases by a factor of 2 while reducing the data requirement by a factor of 5 compared to $\mathcal{F}_{(n, 20)}$, as shown in Fig. 2(g). However, the model's predictive capability remains limited when the initial conditions deviate substantially from those represented in the training set. For example, a model trained on hot spot cannot reliably predict the ES case, as such extrapolation leads to large generalization errors for conditions outside the training distribution. More broadly, robust generalization across widely separated initial conditions remains challenging. Nevertheless, our study represents an important step toward replacing complex closures in hydrodynamic simulations and highlights the potential of treating machine learning as an iterative solver rather than a black-box model.

In future work, to address the robust generalization challenge, we aim to learn the underlying first-order spherical harmonic components and to reconstruct the nonlocal heat flux indirectly through the corresponding corrections to the SH flux within the SNB framework, with the long-term goal of extending the approach to explore the multidimensional configurations. Particularly, nonuniform laser intensity creates temperature modulations in the transverse direction due to laser imprint, along with steep temperature gradients in the longitudinal direction near the critical density surface. Nonlocal heat transport in this multidimensional setting may significantly influence subsequent fluid dynamics, including hydrodynamic instabilities such as the Rayleigh-Taylor instability, making the extension of the present framework to multidimensional configurations a compelling and practically important direction. In addition, we will further investigate how externally applied and self-generated magnetic fields influence the internal representations of machine learning models within the magnetized SNB framework [34]. These directions critically rely on access to high-fidelity VFP data.

The authors thank 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.

The data are not publicly available. The data are available from the authors upon reasonable request.

Appendix. Different from the above study, a heating source term was incorporated into both the PIC simulation and

Eq. (1) to drive the system away from its initial homogeneous temperature profile. Under the action of this source, the initially homogeneous temperature profile evolves through a sustained competition between spatially localized energy deposition and thermal relaxation. The heating term appearing in Eq. (1) takes the form $S_{\text{ext}} = \sigma S(x)S(t)$, expressed in units of $n_{e0}m_e c^2 \omega_{pe}$. Within the OSIRIS framework, in which density, spatial coordinates, and time are normalized by the background electron density n_{e0} , the plasma wave number k_{pe} , and the plasma frequency ω_{pe} , respectively, the heating is applied as a stochastic momentum kick to each macroparticle at every simulation time step Δt according to $(u_i/c)|_{t+\Delta t} = (u_i/c)|_t + \sqrt{2S_{\text{ext}}/3}\xi_i$, where $\xi_i \sim \mathcal{N}(0, 1)$ is a Gaussian random variate with zero mean and unit variance, and the index $i \in \{x, y, z\}$ denotes the three Cartesian velocity components. For spatially uniform heating, $S(x) = S(t) = 1$, the electron temperature evolves analytically as $T_e(t) = T_{e0} + 2\sigma m_e c^2 \omega_{pe} t / 3$. Figure 6(a) presents the temporal evolution of the electron temperature under uniform heating at $\sigma = 2 \times 10^{-6}$ (corresponding to a power density of $1.7 \times 10^{28} \text{ W m}^{-3}$ for $n_{e0} = 1.5 \times 10^{30} \text{ m}^{-3}$), confirming quantitative agreement with the analytical prediction and thereby validating the implementation of the heating source within the code.

In the following, a spatial profile is incorporated into the heating source prescribed as $S(x) = \exp[-x^2/(1.5\ell)^2]$, and two distinct temporal modulations are considered: a smooth turn-on/turn-off profile, $S_1(t) = [1 - e^{-t/4}][1 - e^{-(t-30)/4}]$, and a periodic sinusoidal modulation, $S_2(t) = \sin^2(2\pi t/15)$, the latter comprising four complete heating cycles. Both temporal profiles are shown in Fig. 6(b). Each heating scenario, $\sigma S(x)S_1(t)$ and $\sigma S(x)S_2(t)$, is implemented in both the PIC simulation and Eq. (1) using the $\mathcal{F}_{(1,1)}$ closure. The results are presented in Fig. 6(c) with S_1 and Fig. 6(d) with S_2 , respectively. In Fig. 6(c), spatial temperature profiles are compared at two time instants, $t = 5$ and $t = 25$.

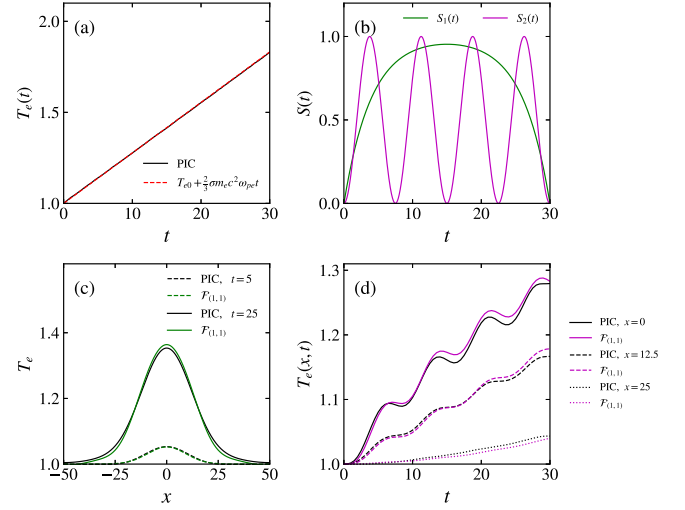


FIG. 6. (a) Temporal evolution of the temperature under spatially uniform heating, $S(x) = S(t) = 1$, with $\sigma = 2 \times 10^{-6}$. The PIC simulation results, shown in black, are compared with the analytical prediction $T_e(t) = T_{e0} + 2\sigma m_e c^2 \omega_{pe} t / 3$, shown in red. (b) Temporal profiles of the two heating modulations: the smooth turn-on/turn-off function $S_1(t) = [1 - e^{-t/4}][1 - e^{-(t-30)/4}]$ and the periodic sinusoidal modulation $S_2(t) = \sin^2(2\pi t/15)$. Panels (c) and (d) compare the temperatures obtained from the PIC simulations with the solutions of Eq. (1) using the $\mathcal{F}_{(1,1)}$ closure. The heating sources are $\sigma S(x)S_1(t)$ in panel (c) and $\sigma S(x)S_2(t)$ in panel (d), where $S(x) = \exp[-x^2/(1.5\ell)^2]$. Panel (c) presents the spatial temperature profiles at $t = 5$ and $t = 25$, while panel (d) shows the temporal evolution of the local temperature at $x = 0, 12.5$, and 25 .

Figure 6(d) tracks the temporal evolution of the electron temperature at three different positions, $x = 0, 12.5$, and 25 . In both cases, the $\mathcal{F}_{(1,1)}$ closure reproduces the temperature dynamics observed in the PIC simulations.

[1] J. Lindl, O. Landen, J. Edwards, E. Moses, and N. Team, Review of the national ignition campaign 2009-2012, *Phys. Plasmas* **21**, 020501 (2014).
[2] S. Silva, J. Santos, J. Büchner, and M. Alves, Nonlocal heat flux effects on temperature evolution of the solar atmosphere, *Astron. Astrophys.* **615**, A32 (2018).
[3] B. Somov, *Plasma Astrophysics, Part I: Fundamentals and Practice, Astrophysics and Space Science Library* (Springer, New York, 2012).
[4] 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).
[5] N. Mitchell, D. Chapman, and G. Kagan, Nonlocal current-driven heat flow in ideal plasmas, *Phys. Rev. E* **112**, L053202 (2025).
[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] S. Atzeni and J. Meyer-ter Vehn, *The Physics of Inertial Fusion: Beam-Plasma Interaction, Hydrodynamics, Hot Dense Matter* (Oxford University Press, Oxford, 2004), Vol. 125.
[9] G. P. Schurtz, P. D. Nicolai, and M. Busquet, A nonlocal electron conduction model for multidimensional radiation hydrodynamics codes, *Phys. Plasmas* **7**, 4238 (2000).
[10] A. Chrisment, P. Loiseau, J.-L. Feugeas, P.-E. Masson-Laborde, J. Mathiaud, V. Tikhonchuk, and P. Nicolai, Analysis of a kinetic model for electron heat transport in inertial confinement fusion plasmas, *Phys. Plasmas* **29**, 062301 (2022).
[11] O. Michel, R. Ducloux, 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).
[12] M. M. Marinak, G. B. Zimmerman, T. Chapman, G. D. Kerbel, M. V. Patel, J. M. Koning, S. M. Sepke, B. Chang, C. R. Schroeder, J. A. Harte, *et al.*, How numerical simulations helped to achieve breakeven on the NIF, *Phys. Plasmas* **31**, 070501 (2024).
[13] J. Delettrez, R. Epstein, M. C. Richardson, P. A. Jaanimagi, and B. L. Henke, Effect of laser illumination nonuniformity on the

- analysis of time-resolved x-ray measurements in UV spherical transport experiments, *Phys. Rev. A* **36**, 3926 (1987).
- [14] E. Buresi, J. Coutant, R. Dautray, M. Decroisette, B. Duborgel, P. Guillauneux, J. Launsbach, P. Nelson, C. Patou, J. M. Reisse, *et al.*, Laser program development at CEL-V: Overview of recent experimental results, *Laser Part. Beams* **4**, 531 (1986).
- [15] J. Breil, S. Galera, and P.-H. Maire, Multi-material ALE computation in inertial confinement fusion code CHIC, *Comput. Fluids* **46**, 161 (2011).
- [16] A. R. Bell, R. G. Evans, and D. J. Nicholas, Electron energy transport in steep temperature gradients in laser-produced plasmas, *Phys. Rev. Lett.* **46**, 243 (1981).
- [17] 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).
- [18] A. Marocchino, M. Tzoufras, S. Atzeni, A. Schiavi, P. D. Nicolaï, J. Mallet, V. Tikhonchuk, and J.-L. Feugeas, Comparison for non-local hydrodynamic thermal conduction models, *Phys. Plasmas* **20**, 022702 (2013).
- [19] J. P. Brodrick, R. J. Kingham, M. M. Marinak, M. V. Patel, A. V. Chankin, J. T. Omotani, M. V. Umansky, D. Del Sorbo, B. Dudson, J. T. Parker, G. D. Kerbel, M. Sherlock, and C. P. Ridgers, Testing nonlocal models of electron thermal conduction for magnetic and inertial confinement fusion applications, *Phys. Plasmas* **24**, 092309 (2017).
- [20] 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).
- [21] M. Luo, C. Heaton, Y. Wang, D. Plummer, M. Fitzgerald, F. Miniati, S. M. Vinko, and G. Gregori, Learning heat transport kernels using a nonlocal heat transport theory-informed neural network, *Phys. Rev. Res.* **7**, L042017 (2025).
- [22] M. Luo, C. Heaton, Y. Wang, D. Plummer, M. Fitzgerald, F. Miniati, S. M. Vinko, and G. Gregori, Time-embedded convolutional neural networks for modeling plasma heat transport, *Phys. Rev. E* **113**, 035303 (2026).
- [23] F. Miniati and G. Gregori, Learning transport processes with machine intelligence, *Sci. Rep.* **12**, 11709 (2022).
- [24] S. Zhang, M. Mallon, M. Luo, J. Thiyaalingam, P. Tzeferacos, R. Bingham, and G. Gregori, Data-driven modeling of shock physics by physics-informed MeshGraphNets, [arXiv:2602.14918](https://arxiv.org/abs/2602.14918).
- [25] S. Schotthöfer, M. P. Laiu, M. Frank, and C. D. Hauck, Structure-preserving neural networks for the regularized entropy-based closure of a linear, kinetic, radiative transport equation, *J. Comput. Phys.* **533**, 113967 (2025).
- [26] J. Huang, Y. Cheng, A. J. Christlieb, and L. F. Roberts, Machine learning moment closure models for the radiative transfer equation III: Enforcing hyperbolicity and physical characteristic speeds, *J. Sci. Comput.* **94**, 7 (2023).
- [27] J. Kossaifi, N. Kovachki, Z. Li, D. Pitt, M. Liu-Schiaffini, V. Duruisseaux, R. J. George, B. Bonev, K. Azizzadenesheli, J. Berner, and A. Anandkumar, A library for learning neural operators, [arXiv:2412.10354](https://arxiv.org/abs/2412.10354).
- [28] Z. Li, N. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, and A. Anandkumar, Fourier neural operator for parametric partial differential equations, [arXiv:2010.08895](https://arxiv.org/abs/2010.08895).
- [29] 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.
- [30] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/914n-mnz6> for a comparison of published VFP results, PIC simulations, the SNB model, the SNB model implementation, machine learning model training strategies, and performance.
- [31] 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).
- [32] E. M. Epperlein and R. W. Short, A practical nonlocal model for electron heat transport in laser plasmas, *Phys. Fluids B* **3**, 3092 (1991).
- [33] D. Cao, G. Moses, and J. Delettrez, Improved non-local electron thermal transport model for two-dimensional radiation hydrodynamics simulations, *Phys. Plasmas* **22**, 082308 (2015).
- [34] P. D. Nicolaï, J.-L. A. Feugeas, and G. P. Schurtz, A practical nonlocal model for heat transport in magnetized laser plasmas, *Phys. Plasmas* **13**, 032701 (2006).