ORIGINAL PAPER

Mallat Scattering Transformation based surrogate for Magnetohydrodynamics

Michael E. Glinsky[1] · Kathryn Maupin[2]

Abstract

A Machine and Deep Learning (MLDL) methodology is developed and applied to give a high fidelity, fast surrogate for 2D resistive MagnetoHydroDynamic (MHD) simulations of Magnetic Liner Inertial Fusion (MagLIF) implosions. The resistive MHD code GORGON is used to generate an ensemble of implosions with different liner aspect ratios, initial gas preheat temperatures (that is, different adiabats), and different liner perturbations. The liner density and magnetic field as functions of x, y, and z were generated. The Mallat Scattering Transformation (MST) is taken of the logarithm of both fields and a Principal Components Analysis (PCA) is done on the logarithm of the MST of both fields. The fields are projected onto the PCA vectors and a small number of these PCA vector components are kept. Singular Value Decompositions of the cross correlation of the input parameters to the output logarithm of the MST of the fields, and of the cross correlation of the SVD vector components to the PCA vector components are done. This allows the identification of the PCA vectors vis-a-vis the input parameters. Finally, a Multi Layer Perceptron (MLP) neural network with ReLU activation and a simple three layer encoder/decoder architecture is trained on this dataset to predict the PCA vector components of the fields as a function of time. Details of the implosion, stagnation, and the disassembly are well captured. Examination of the PCA vectors and a permutation importance analysis of the MLP show definitive evidence of an inverse turbulent cascade into a dipole emergent behavior. The orientation of the dipole is set by the initial liner perturbation. The analysis is repeated with a version of the MST which includes phase, called Wavelet Phase Harmonics (WPH). While WPH do not give the physical insight of the MST, they can and are inverted to give field configurations as a function of time, including field-to-field correlations.

Keywords Mallat Scattering Transform · Magnetohydrodynamics · Surrogate modeling · Machine learning

1 Introduction

The major challenge for physics-based machine learning is to replace expensive finite element and finite volume computer simulations with fast Machine and Deep Learning (MLDL) based surrogates that reproduce all the structure and emergent behaviors of the system. These surrogates can then be used for experimental design and analysis. Fundamentally, they can be used for hypothesis testing, theoretical model verification, and model extrapolation and scaling. The chal-

lenge has been to capture the rich texture of physical systems, then to reproduce them, not only to predict the texture of one field, but to predict the rich correlations between the fields. Attempts using both Gaussian process simulation [1] and convolutional network techniques combined with reduced order models [2,3] have had modest success. Inspired by the success of the Mallat Scattering Transformation (MST) [4,5] and the Wavelet Phase Harmonics (WPH) enhancement of the MST [6], which includes phase, in classifying and reproducing textures of physical systems [7], this paper presents a simple MLDL methodology based on the MST and WPH to give a high fidelity, fast surrogate of 2D resistive MagnetoHydroDynamics (MHD).

A major goal of this development is not only to reproduce the system evolution, but to do it in a way that can lead to physical insight. Much of the work to date in physics-based MLDL has approached the challenge using MLDL as a black-box of ingredients to be combined, with success judged by

✉ Kathryn Maupin
  kmaupin@sandia.gov

  Michael E. Glinsky
  glinsky@mac.com

[1] qiTech Consulting, Santa Fe, NM, USA

[2] Optimization and Uncertainty Quantification, Sandia National Laboratories, Albuquerque, NM, USA

final performance metrics. Attributes produced in the analysis are abstract vectors with little or no physical meaning. Displays of those vectors are rarely given. The methodology presented in this paper is inspired by fundamental understanding of the physics; has physical interpretations of the attributes, vectors, and MLDL structures; and has displays of the attributes, vectors, and MLDL performance that leads to important physical insights of the nonlinear and quantum dynamics of the system.

The system and geometry that is chosen to prototype this MLDL methodology is Magnetized Liner Inertial Fusion (MagLIF). MagLIF is a magneto-inertial fusion concept currently being explored at Sandia's Z Pulsed Power Facility [8–11]. MagLIF produces thermonuclear fusion conditions by driving mega-amps of current through a low-Z conducting liner. The subsequent implosion of the liner containing a preheated and pre-magnetized fuel of deuterium or deuterium-tritium compresses and heats the system, creating a plasma with fusion relevant conditions. Axial symmetry is assumed and the dynamics is simulated in the Cartesian, perpendicular plane.

It is well-known that the large accelerations of the liner, as it drives the implosion of the gas, cause the liner to go linearly unstable to the Magnetic Rayleigh Taylor (MRT) instability [12]. During the implosion, it has sufficient time to evolve well into the nonlinear regime. Increasing the AR will increase the linear growth rate by increasing the acceleration of the liner. Reducing the preheat temperature T_0 will put the implosion on a higher adiabat, allowing the implosion to reach higher compression ratios and a smaller radius at stagnation. This gives the instability more time to evolve, and puts a larger demand on the uniformity of the imploding surface. For laser-driven indirect ICF capsules, without a large applied magnetic field, the Rayleigh-Taylor instability exhibits a normal turbulent cascade that destroys the implosion well before it reaches the required compression, if allowed to grow. For this reason, the implosions need to have a larger AR, a larger T_0 (be put on a lower adiabat), and to have very small perturbations to the capsule surface. Unfortunately, this means that it has been very difficult to reach the needed conditions at stagnation for thermonuclear burn. There is indication that this is not the case for MagLIF. A double helical structure is observed by [9] with helical threads 30 μm in diameter, separated by 100 μm reaching very high compression ratios. Also, evidence of an inverse turbulent cascade in the liner structure has been seen by [13] on ultra-thin foils driven at less than 1 MA. This is the reason that AR and T_0, along with how the MRT is seeded, are the parameters that are varied. The scientific goal of this study is to characterize and understand the nonlinear evolution of the MRT.

The finite volume, parallel, resistive MHD code `GORGON` [14] is used to simulate the 2D MagLIF geometry. A single temperature, no radiation transport, and no thermonuclear burn is assumed. The focus of the MLDL is on reproducing and analyzing the emergent behavior of the liner dynamics. An ensemble of 539 simulations is done, with samples at different Liner Aspect Ratios (AR) and different preheat temperatures (T_0), so that the compression is done on different adiabats. Also, the ensemble has random liner perturbations in both amplitude and phase. The evolution was sampled every 2.5 ns over 200 ns, generating 87,318 images of liner density and magnetic field.

Based on this ensemble a MLDL workflow is developed to form a surrogate and to gain insight into the nonlinear physics. A Principle Components Analysis (PCA) is done on the logarithm of the MST of the logarithm of the liner density and the magnetic field. That is, the logarithm of the liner density is taken, and the MST is calculated. The logarithm of the MST is then calculated, and the resulting values are subjected to a PCA. It is found that most of the variation is in the first seven components. Singular Value Decompositions (SVDs) are done of the cross-correlations of the input parameters (AR, T_0, and liner perturbations) to the logarithm of the MST, and of the SVD vector components to the PCA vector components. This allows the PCA vectors to be identified with the input parameters. A Multi Layer Perceptron Neural Network (MLP/NN) [15] with a three-layer encoder/decoder structure is trained to predict the seven PCA vector components as a function of time, given the initial conditions. Excellent performance is found, as shown in Sect. 5. A permutation importance analysis is done on the inputs. It is determined that the quadrupole moment quickly decays and the energy inverse cascades into the dipole moment, where it remains through stagnation and the subsequent expansion. The evolution shows little to no dependence on the initial tripole or quadrupole moments but very strong dependence on time, AR, and the initial phase of the dipole moment, and modest dependence on the initial temperature and the size of the initial perturbation.

The analysis is repeated using the WPH in place of the MST. While the WPH cannot be interpreted physically, as it is currently implemented, it can be inverted due to the additional phase information. When this is done, the temporal evolution of the two fields is well-predicted, including the field-to-field correlations.

The MST and the WPH will be described in Sect. 2, along with the display and interpretation of the MST. How the training and testing dataset was constructed and generated appears in Sect. 3, and a description of the evolution is given. The details of the MLDL architecture is given in Sect. 4, followed by the results from applying the MLDL architecture in Sect. 5. A discussion of the results, conclusions that can be drawn, upcoming work, and possible improvements to the MLDL architecture is found in Sect. 6.

2 Mallat Scattering Transformation (MST) and Wavelet Phase Harmonics (WPH)

The Mallat Scattering Transformation (MST) can be viewed as a Convolutional Neural Network (CNN) [3,16,17] with predetermined weights. The filters are cleverly designed so that by a convolve-binate cycle, the CNN can span an exponentially large range in scale with a kernel of constant size. In other words, a very fast algorithm, analogous to a Fast Fourier Transform (FFT), can be constructed. For instance, on modern GPUs, the MST takes about 10 ms on a 512×512 image. A very useful way of defining the m-th order MST of a field $f(x)$ is

$$S_m[f(x)](p) \equiv \phi_{p_{\min}} \star \left(\prod_{k=1}^{m} \mathrm{mod}\, \psi_{p_k} \star \right) f(x), \qquad (1)$$

where $\phi(x)$ is the Father Wavelet, $\psi(x)$ is the Mother Wavelet, mod is the complex modulus,

$$\phi_p = p^2 \phi(px), \qquad (2)$$

$$\psi_p = p^2 \psi(px), \qquad (3)$$

$$\psi_p \star f \equiv \int \psi_p(x')\, f(x' - x)\, dx', \qquad (4)$$

$$p_k \in \{2^{-j}\}, \qquad (5)$$

$$j \in \{1, 2, \ldots, J\}, \qquad (6)$$

$$p_{k+1} < p_k, \qquad (7)$$

and

$$p = \sum_{k=1}^{\infty} p_k. \qquad (8)$$

Note that this is a path integral in p_k and a prescription needs to be chosen of how to go around the poles, see [18] for additional details. Equation (7) is the common choice of normal ordering. Also note that Eq. (4) is a Wigner-Weyl-like mapping from a cotangent bundle T^*M with coordinates (π, f) and symplectic metric $d\pi \wedge df$ (where f is the field and π is the canonically conjugate momentum) to the space of analytic functions on $\mathbb{C}$ with coordinate z.

This transform has been shown by [4] to be Lipschitz continuous to diffeomorphic deformation (unlike the Fourier Transformation), and by construction (through the final convolution with $\phi_{p_{\min}}$) to be translationally invariant. That is, it is stationary. The latter we view as an unfortunate step, because most physical processes are not stationary. This is not necessary and will be discussed at length in Sec. 6. The MST is rarely normalized in implementations despite the normalization being presented in [4] section on Dirac normalization of the MST.

The form of the MST in Eq. (1) can be identified as a CNN with a specified (fitting to data not necessary) and elegant structure. There is a multi layer (in m, the order of the MST) application of a bank of convolutional kernels (that is, ψ_p), a nonlinear activation (that is, mod), and a final pooling operation (that is, $\phi_{p_{\min}}$). We will then do a dimensional reduction using a PCA.

Reference [5] devised a way of visualizing the 2D MST that is shown in Fig. 1. The coefficients of the MST are plotted on radial plots, one for each order m, with the radial position being one of $|\bar{p}_m|$, $\ln |\bar{p}_m|$, $|1/\bar{p}_m|$, or $\ln(1/|\bar{p}_m|)$, where

$$\bar{p}_m \equiv \sum_{k=1}^{m} p_k. \qquad (9)$$

The "posting" or manifestation of the radial plot for the angular position is much more convoluted. For $m = 1$, it is simple: the angle is $\arg(\bar{p}_1)$. Because the basis used for the 2D transform is not orthogonal, the angle for $m = 2$ is calculated as $\arg(\bar{p}_1) + \arg(p_2)/L$, where L is the number of angular sectors calculated. This has the undesirable property that as the angular resolution is increased, the display is not simply a higher resolution version.

One deficiency of the MST is that it discards the phase when it takes the modulus. This does not matter when it is used for image classification or regression, but it is a serious problem when the MST, predicted by the regression, needs to be inverted to get the predicted image. It is analogous to inverting a Fourier transform with only the modulus. To address this situation, Wavelet Phase Harmonics (WPH) were developed by [6] and [7]. This theory has conceptually replaced the modulus with a phase harmonic expansion

$$\mathrm{mod}(z) \rightarrow \sum_{k=0}^{\infty} \mathrm{mod}(z)\, \mathrm{e}^{\mathrm{i}\,k\,\arg(z)}. \qquad (10)$$

Unfortunately, there is no obvious way to plot the transformation, as it is not orthogonal and is significantly over-determined. Because of this, a fast inverse cannot be constructed and a conventional gradient descent optimization must be done, where the objective is the L_2 norm of the difference in the WPHs. It is extremely slow. Where the forward WPH transform takes less than a second, the inverse transform may take more than an hour. Even with these deficiencies, this transformation has been used with great success, most notably to analyze cosmological simulations [19,20].

Fig. 1 Coefficients produced by applying MST to 2D images in this work will be displayed on radial plots as shown. Bins are created according to scale (radial positioning, $|\bar{p}_m|$ as defined in Eq. (9), and rotations $(\arg(\bar{p}_1), \arg(\bar{p}_1) + \arg(p_2)/L)$ for first and second order, with magnitude (color scale, not shown) representing the size of the coefficient at that scale and rotation

$\arg(\bar{p}_1) + \arg(p_2)/L$

$|\bar{p}_1|$

$|\bar{p}_2|$

$\arg(\bar{p}_1)$

$(m = 1)$

$(m = 2)$
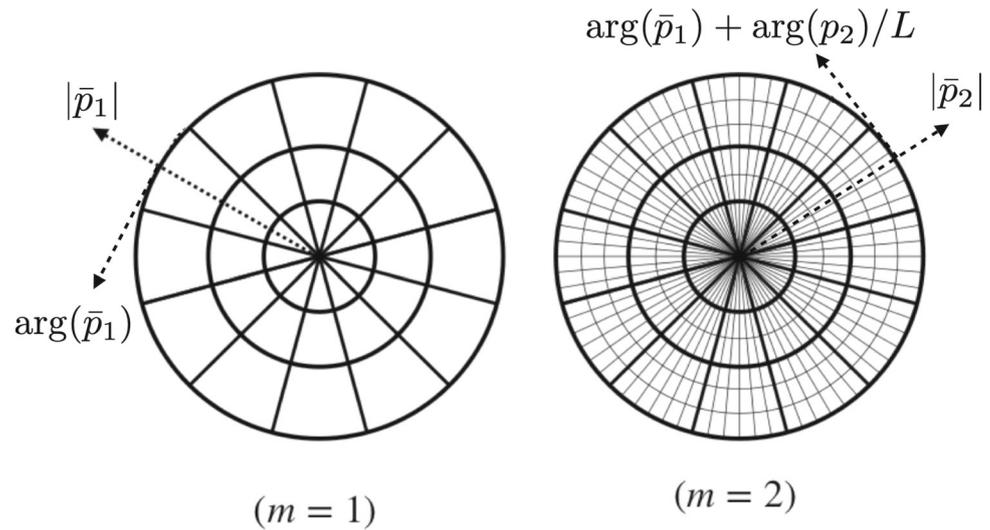
3 Generation of the dataset

An ensemble of 2D resistive MHD simulations were done using the finite volume GORGON code. Axial symmetry was assumed and the simulation was done in the 2D (x, y)-plane as shown in Fig. 2. A geometry relevant to MagLIF was used with a cylindrical beryllium liner with inner and outer radii

$$R_{\text{outer}} \equiv \left(1 + \frac{w}{2 R_0}\right) R_0,$$
$$R_{\text{inner}} \equiv \left(1 - \frac{w}{2 R_0}\right) R_0, \tag{11}$$

where R_0 is the mean radius and w is the liner thickness. We define the liner aspect ratio as $\text{AR} = R_0/w$. Note that the larger the AR, the larger the liner acceleration, and, therefore, the larger the linear growth rate of the MRT. The liner inner and outer surface is perturbed with dipole (m = 2) to quadrupole (m = 4) azimuthal perturbations to seed the MRT of form

$$\delta_o = \sum_{m=2}^{4} \delta_{o0m} \cos(2\pi m \varphi + \varphi_{om}),$$
$$\delta_i = \sum_{m=2}^{4} \delta_{i0m} \cos(2\pi m \varphi + \varphi_{im}). \tag{12}$$

Inside the liner is a Deuterium (D_2) gas with initial density n_0, preheated to a temperature T_0. Note that a larger value of T_0 will put the implosion on a lower adiabat, causing the implosion to reach lesser compression ratios and a larger radius at stagnation. This gives the MRT less time to grow. A uniform axial magnetic field B_{z0} is initially established, then a large axial current with a sinusoidal profile is driven

through the liner, with a peak current of I_{z0} at 100 ns and a total duration of 200 ns.

An ensemble of 539 simulations were done with the liner density $n_l(x, y; t)$ and magnetic field strength $B(x, y; t)$ sampled every 2.5 ns with a 10 μm grid spacing in x and y. This took over 200 k core×hrs on a high performance cluster at LLNL, and generated 87, 318 256×256 images. To reduce the number of parameters and simplify the analysis, the liner perturbations were limited to $w \Delta = \delta_{i0m} = -\delta_{o0m}$ and $\varphi_{im} = \varphi_{om}$ for $m = 2, 3, 4$. The parameters $R_0 = 2.4$ mm, $B_{z0} = 10$ T, $n_0 = 1$ mg/cc, and $I_{z0} = 20$ MA were held constant. That left 6 parameters to be sampled, including the initial conditions and the stochastic parameters. The two initial conditions included AR = [3, 9] and $\log_{10} T_0 = [1, 2.8]$ (that is, $T_0 = [10 \text{ eV}, 630 \text{ eV}]$). The four stochastic parameters included $\log_{10} \Delta = [-2, -1]$ (that is, $\Delta = [1\%, 10\%]$) and $\varphi_{im} = [0, 2\pi]$ for each of $m = 2, 3, 4$. The smallest Δ corresponds to 1% of the thinnest liner, which is about 5 μm, about half a grid cell. Latin Hypercube Sampling (LHS) was used to generate 27 samples of (AR, T_0, Δ), with the φ_{im} being chosen from uniform distributions. Another 512 parameter vectors were randomly sampled from their uniform distributions.

An example of one of the evolutions of the liner density n_l and the gas density n_g is shown in Fig. 3. This simulation was done with an AR = 3, $T_0 = 631$ eV, and $\Delta = 1\%$. Note how gas stagnates into a dipole pattern in a sausage-like mode, then that pattern is imprinted on the liner as it expands post stagnation.

4 MLDL architecture

The following MLDL workflow was constructed and used to analyze the data. The logic behind the construction will be

Fig. 2 Geometry of the simulations. (Left) Beryllium cylinder in which fuel is held. (Right, top) Illustration of aspect ratio quantities. (Right, bottom) Demonstration of azimuthal perturbations
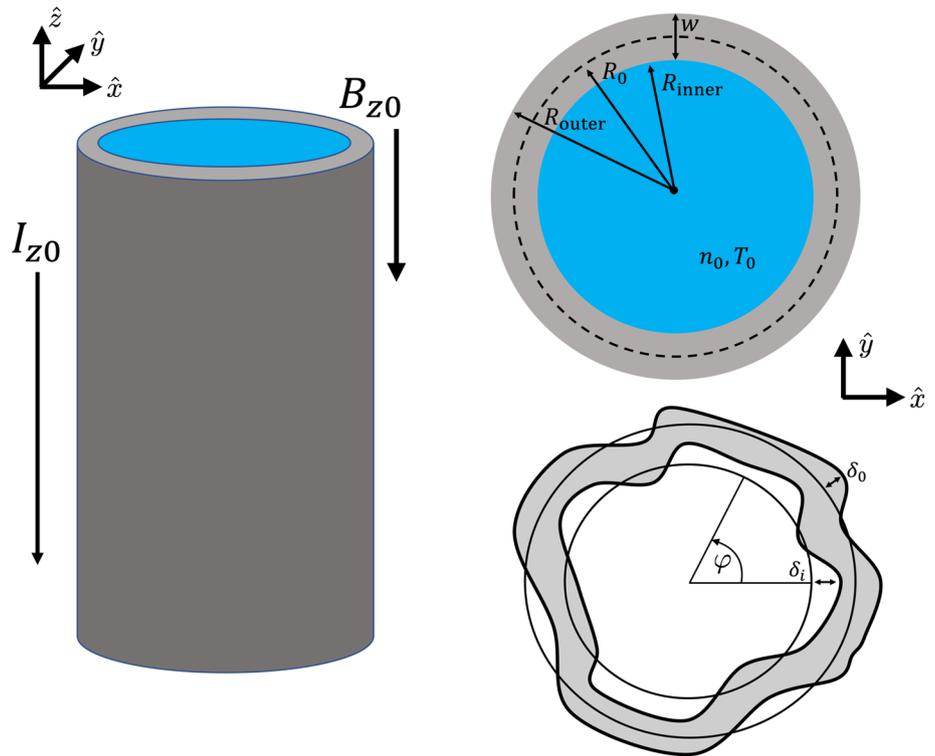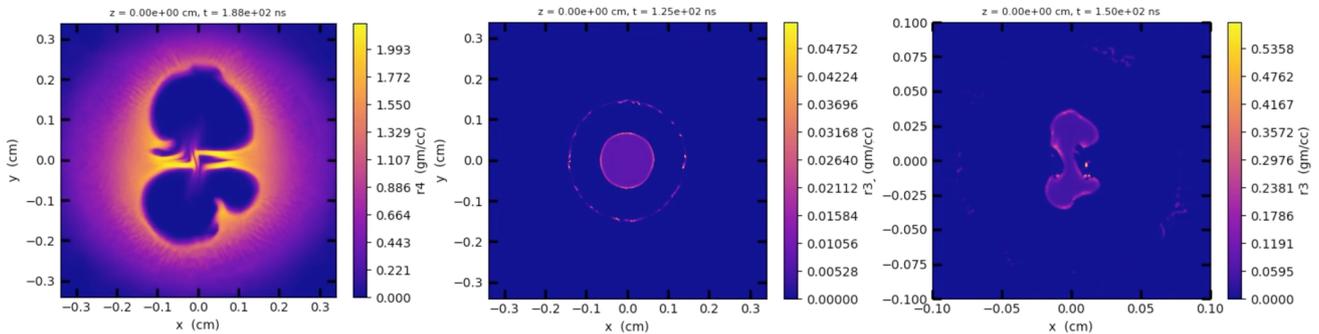
Fig. 3 Evolution of the liner density $n_l(x, y; t)$ and the gas density $n_g(x, y; t)$ as simulated by GORGON. The picture on the right is a zoom in on the stagnation of the gas. The animation of this figure can be found at this link to a https://youtu.be/GmIr3O5GLR0

discussed in Sect. 6. The pipeline is shown in Fig. 4. First, the logarithm of each pixel in the images was calculated, and the MST and WPH were subsequently found for the logarithmic image. It took 16 GPU×hrs (on Nvidia GTX 3090) to take the MSTs of the images and 27 GPU×hrs to take the WPH of the images. Values of $J = 8$ and $L = 16$ were used to take the MST, and values of $J = L = 8$, $\Delta_j = 5$, $\Delta_k = 0$, $\Delta_l = 8$, and $\Delta_n = 2$ to take the WPH. Though these key specifications are not discussed in this paper, they are included here for completeness and reproducibility; details can be found in [4] and [6] for MST and WPH, respectively. The logarithm (base 10) was then taken of the MST and the mean was subtracted on a coefficient by coefficient basis. Note that subtracting the mean from the $\log_{10}$ is equivalent to applying a multiplicative

scaling to the original transform. It is known that this implementation of the MST is not properly Dirac-normalized. This leads to an imprinted pattern that distracts from the natural variation in the image. Subtracting the mean removes this imprinted pattern. It should be noted that this does not affect the subsequent PCA. The analysis of the MST is followed by a PCA. The data is projected onto the top seven PCA vectors for subsequent analysis. On the other side of the pipeline, the input vector of the three control variables (AR, T_0, t) and the four stochastic variables (Δ, φ_{i2}, φ_{i3}, φ_{i4}) are Z-normalized.

In order to identify the PCA vectors, an SVD is done on the cross-correlation of the inputs to the output, before the projection onto the PCA vectors. This allows for a labeling of the SVD vectors according to the inputs. A second SVD

Fig. 4 The MLDL workflow
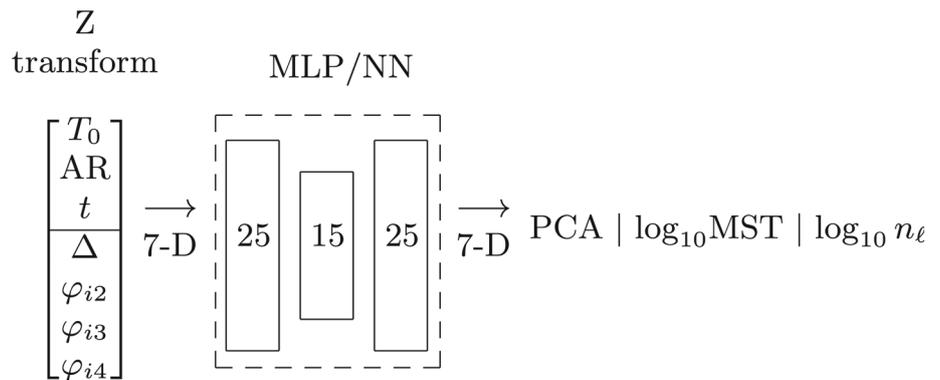
is taken of the cross-correlation between the projection onto the PCA vectors and the projection onto the SVD vectors. This allows the labeled SVD vectors to be associated with the correct PCA vector, thereby labeling the PCA vectors with the input.

The last step is to train a MLP neural network with ReLU activation to predict the seven PCA vector components given the inputs. The structure of the MLP/NN was optimized, as well as the regularization parameter, via a grid search. There was a bias in this choice toward more regularization (without significantly compromising performance) to prevent over-training. The optimal structure was one with three hidden layers (with 25-15-25 nodes) with an encoder/decoder structure. A permutation importance analysis was also done to determine the importance of the input parameters in the estimation. A five-fold cross validation was used in all analyses. As will be shown in the next section, the MLP/NN had very good performance, and a very interpretable result. A shallow Support Vector Regression (SVR) with a radial basis was attempted. The results were very low frequency and were unable to capture the details of the stagnation. A higher frequency result could be obtained by decreasing the regularization, but the result did not cross-validate.

For this application, the primary downside to the MST is that it throws out the phase of the transformation. While this does not matter for classification or some of the physical interpretation, it prevents the transform from having a useful inversion. It is akin to throwing out the phase of a Fourier Transform, then inverting with a random phase. For this reason we repeated our analysis using the WPH (the MST with phase). While this transformation does have reasonable inversions, it has no physically interpretable display. It is a black-box vector. Care also needed to be taken with the treatment of the complex-valued transform. The complex analytic $\ln(z)$ function was used, yielding $\ln | z | + i \arg(z)$, and a circular correlation [21] was used in the PCA with respect to the cyclic $\mathrm{Im}(\ln(z))$. Finally, due to the unreasonable translational invariance built into the WPH, there is an arbitrary x and y translation that must be removed. A fiducial at the

vertical and horizontal edges of the images was added to aid in this task.

The PCA only took 1 core×sec, and the MLP/NN took 20 core×sec to train. The resulting forward model surrogate takes 0.1 core×sec to evaluate, compared to the 360 core×hrs required for the GORGON simulation – a factor of 10^7 acceleration.

The implementation of the MST used was the Kymatio package [22], available on https://github.com/kymatio/kymatio. The version of the WPH used is based on the work of [7], also available on https://github.com/bregaldo/pywph as the pyWPH package.

5 Results

An overview of the results are shown in Fig. 5. Images corresponding to the MLDL pipeline are shown beneath the schematic of Fig. 4. On the right-hand side is the $\log_{10}$ image of the liner density. To the left of it are the first and second order MST of that image. Continuing to the left of that are the seven principal vectors of the MST, which explain 94% of the variance. The identification of the PCA vectors with the inputs along with the five-fold cross validated score for each PCA vector, as predicted by the MLP/NN, are included. The total score was 81%. A typical example of the PCA vector evolution with respect to time is shown, as predicted by the MLP/NN. Finally, the permutation feature importance result is shown on the left of the image. More details of these results follow, and they will be discussed in detail in the following section.

The analysis starts by taking the MST of all the images. Shown in Fig. 6 is the MST of a simulation at three different times: at time zero, at 12.5 ns before bang time (maximum compression), and at 25 ns after bang time. There is also a https://youtu.be/b-p09GZigNA to an animation showing all of the realizations. It is followed by Fig. 7, which shows the MST at time zero of three simulations with the highest adiabat of $T_0 = 631$ eV and the smallest liner perturbation of $\Delta = 1\%$, for three different values of AR $= 3, 6, 9$. There is
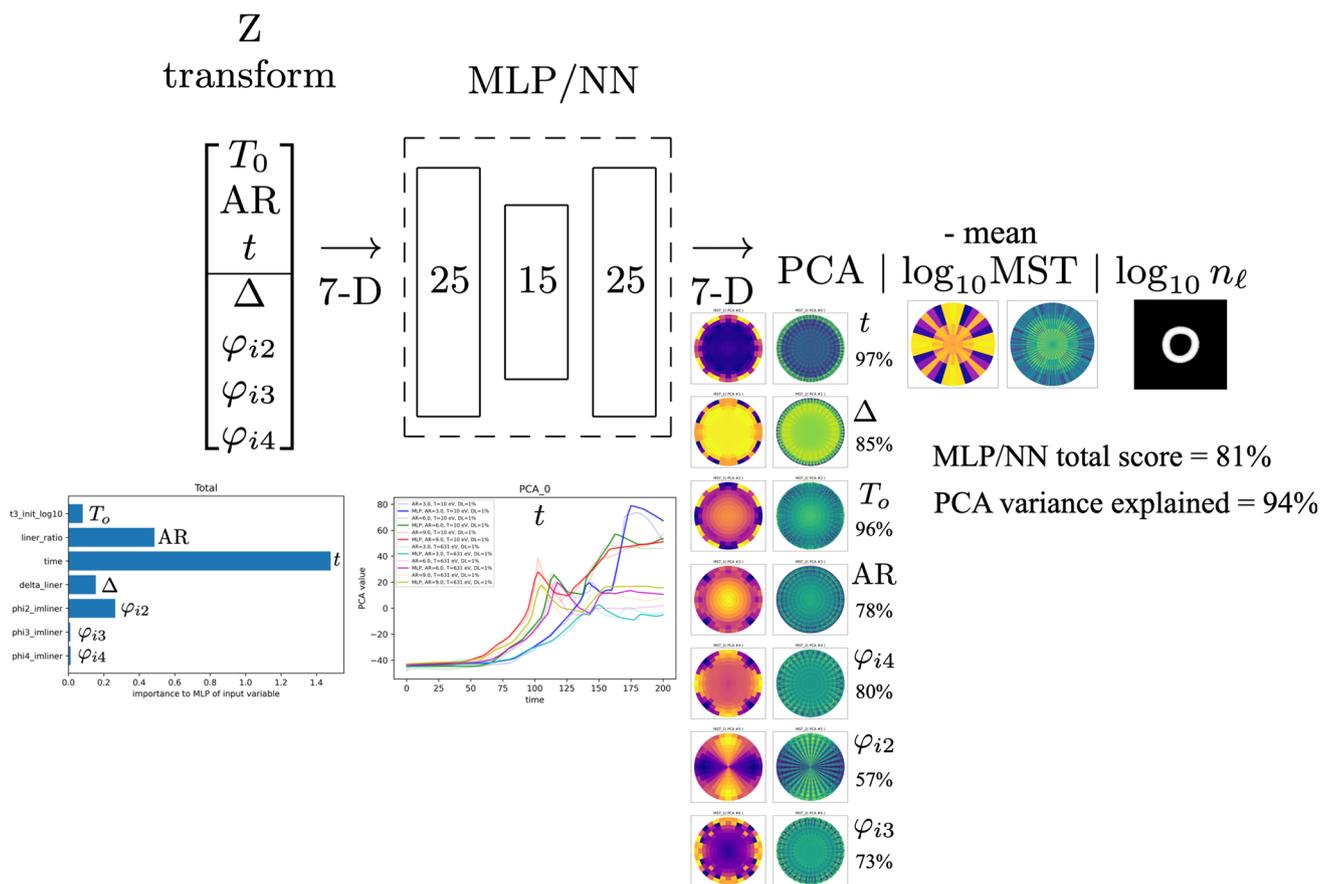
Fig. 5 The complete MLDL workflow with key results. Underneath a diagram of the MLDL pipeline, going right to left, are: (1) grey scale image of n_l, (2) first and second order MST of the grey scale image, (3) the MST PCA vectors, (4) some predictions of the first PC vector component with respect to time, and (5) the permutation feature importance analysis

a https://youtu.be/b-p09GZigNA in the figure caption to an animation of the full time evolution of each case.

The mean value for the MST (shown in Fig. 8) is then subtracted from the values, and a PCA is done, as well as an SVD analysis of the cross correlation of the inputs to the outputs. The PCA vectors are shown in Fig. 9, and the SVD vectors are shown in Fig. 10. The eigenvalues for both decompositions are shown in Fig. 11. Note that the PCA is capturing more of the variation in fewer components. The first seven PCA vectors have captured 94% of the total variation. The SVD has a strong correlation of each component with one input parameter. When the PCA projection of the output is cross-correlated with the SVD projection of the output, an interpretation key is generated for the PCA vectors by the display of the singular vectors, as shown in Fig. 12.

A MLP/NN is then trained to predict the PCA vector components of the fields $n_l(x, y; t)$ and $B(x, y; t)$, given the six initial condition parameters. The performance is shown in Fig. 13. The results of a grid search for network structure and the regularization parameter α are shown in the upper left-hand corner. Note that the value of α is increased from the

best parameter to one with almost the same performance but more regularization in order to prevent overfitting. The overall performance is 81% and the performance on individual PCA vector components range from a high of 97% to a low of 57%. The performance is quantified by the linear correlation of the predicted to the actual values. All of the correlations look very linear with no pathologies. The performance of the MLP/NN on predicting the PCA vector components is quite remarkable and is shown in Fig. 14. Shown are the predicted versus actual for six simulations that span AR = 3, 6, 9 and $T_0 = 10$ eV, 631 eV. Note that the MLP/NN decided to use few points where the function was linear and captured the stagnation behavior well where the function is singular. The MLP/NN found the stagnation points and put an interpolation point at the cusp. The permutation feature importance shows interesting results that will be discussed in the next section. Do note that AR, t, and φ_{i2} are the most important, and φ_{i3} and φ_{i4} are of little importance.

A Support Vector Regression (SVR) with a radial basis was attempted. The results were disappointing. The overall performance was 46%, with a high of 90% and a low of 10%.
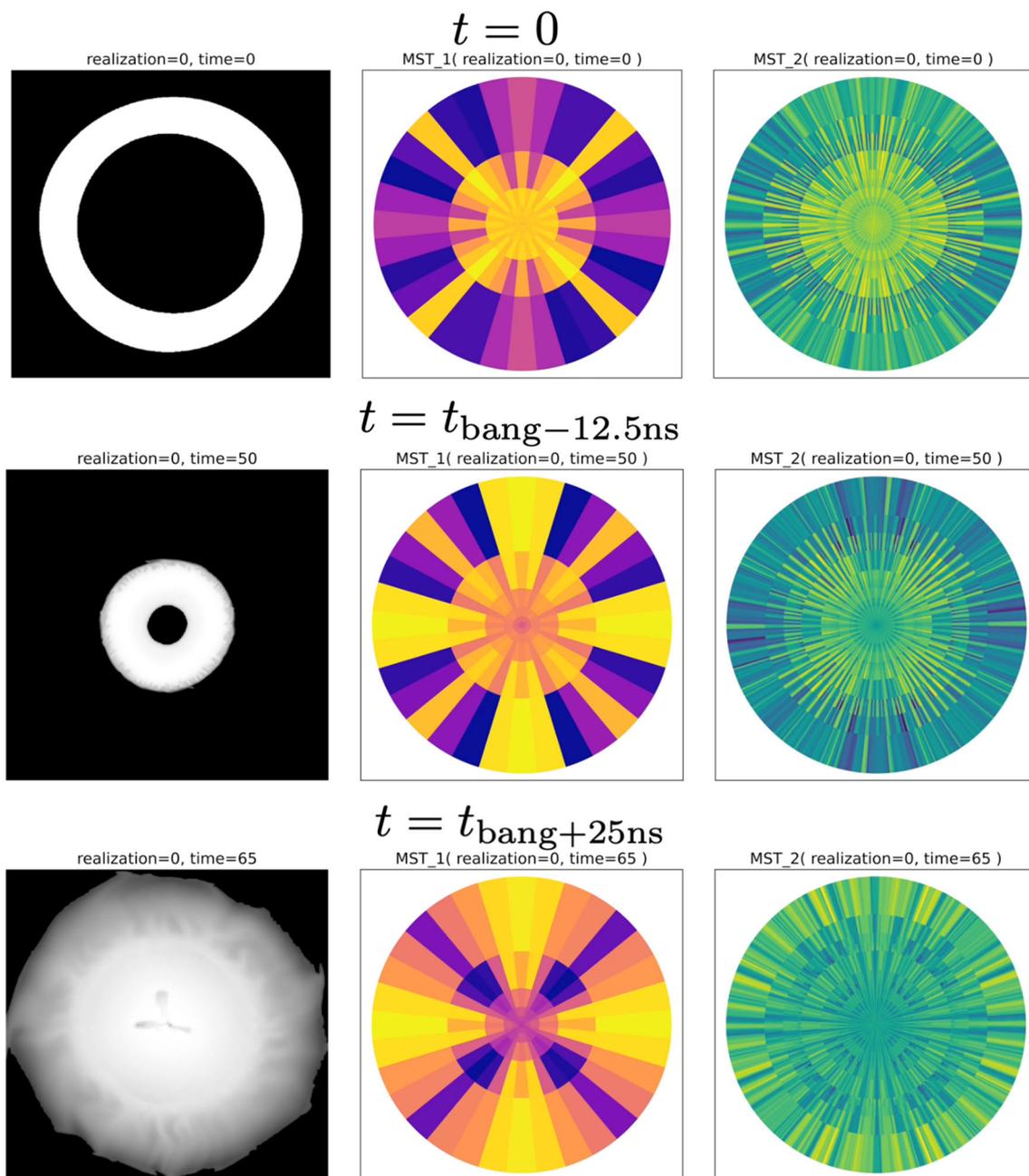
Fig. 6 Ensemble of realizations at three select times. Shown are the $\log_{10} n_l(x, y; t)$, then the first and second order MST of the liner density. From top to bottom: (1) $t = 0$, (2) $t = t_{\text{bang}-12.5\,\text{ns}}$, and (3) $t = t_{\text{bang}+25\,\text{ns}}$. The radial axis is linear in scale. The animation of this figure can be found at this link to a https://youtu.be/b-p09GZigNA

The predictions of the PCA vector components are shown in Fig. 15. Note how (overly) smooth the regressions are, and how the stagnation (singularity) is missed completely. When the regularization was reduced to reduce the smoothing, the performance went down significantly, evidenced by the result not cross validating.

Finally, the mapping back to the fields was done using the WPH. The analysis previously described was repeated using the WPH in place of the MST. The evolution was predicted by the MLP/NN, and the results were inverted to give the evolution at four key times for both the liner density and the magnetic field demonstrating the correlation. The results are shown in Fig. 16.
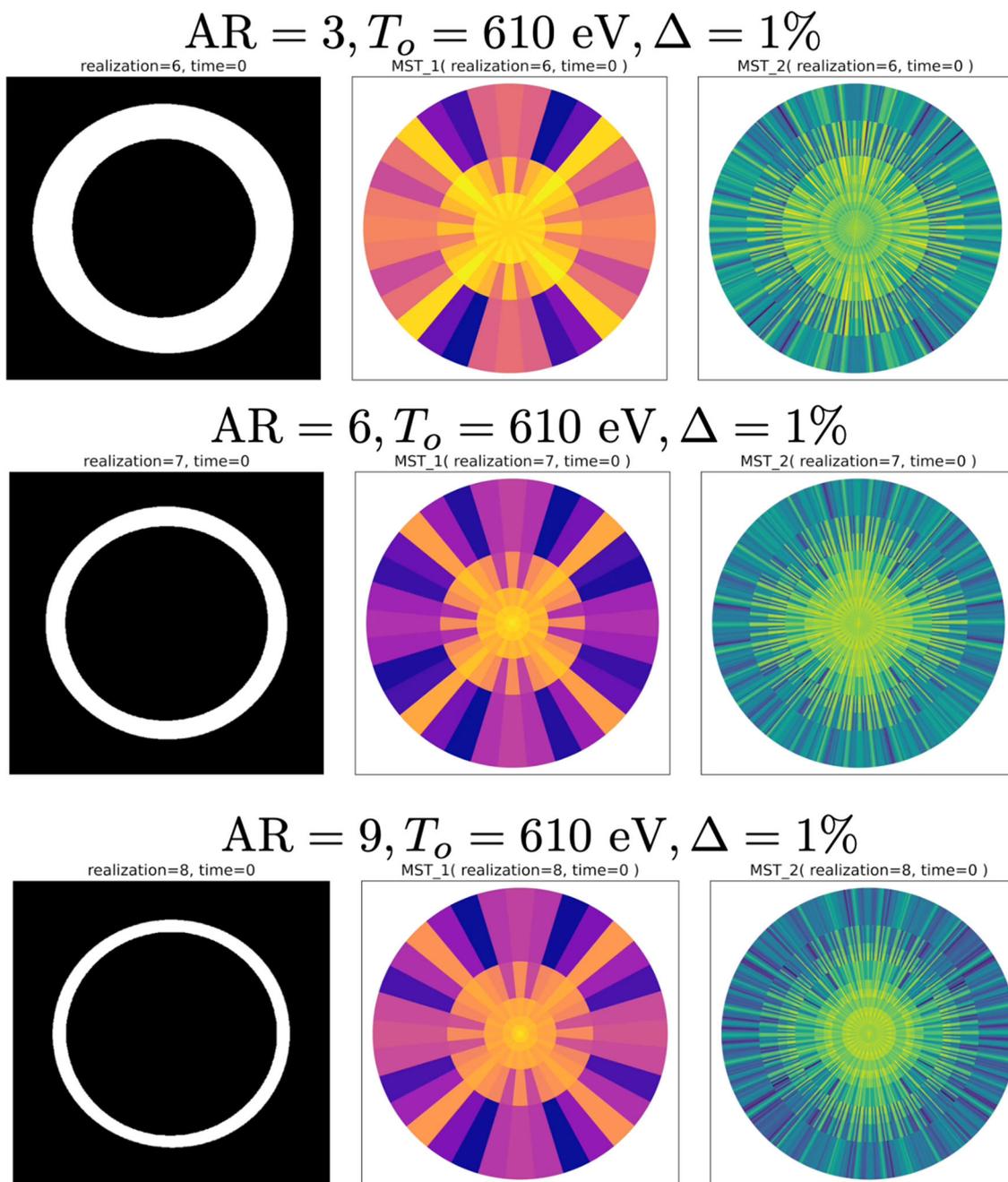
Fig. 7 Three time evolutions of the MST of liner density. Shown for $T_0 = 0$ and $\Delta = 1\%$. From top to bottom: (1) AR = 3, (2) AR = 6, and (3) AR = 9. The radial axis is linear in scale. The animation of this figure can be found at this link to a https://youtu.be/b-p09GZigNA

6 Conclusions and discussion

The results of the previous section give insight into the emergent behavior of this nonlinear system. This behavior is exposed in the structure of two critical PCA vectors (φ_{i4}, the quadrupole moment with $m = 4$, and φ_{i2}, the dipole moment with $m = 2$). These PCA vectors are featured in Fig. 17, where the scale (radial) axis is plotted on a logarithmic scale, and exposed in the permutation feature importance, as illus-

trated in Fig. 18. Note that the φ_{i4} PCA vector starts at the largest radii with a very clear quadrupole pattern in both the first order and the second order MST. Due to the way that the second order MST is displayed, there will be two beats in every one of the 32 sectors. This pattern disappears at the smaller radii as the plasma nears stagnation. The opposite is true of the φ_{i2} PCA vector. The dipole pattern persists to the smallest radii. Note that there is only one beat per sector for a dipole pattern. There is even a strong overall dipole pat-

Fig. 8 The mean MST of liner density. Shown is the first order MST on the left and the second order MST on the right. The radial axis is linear in scale
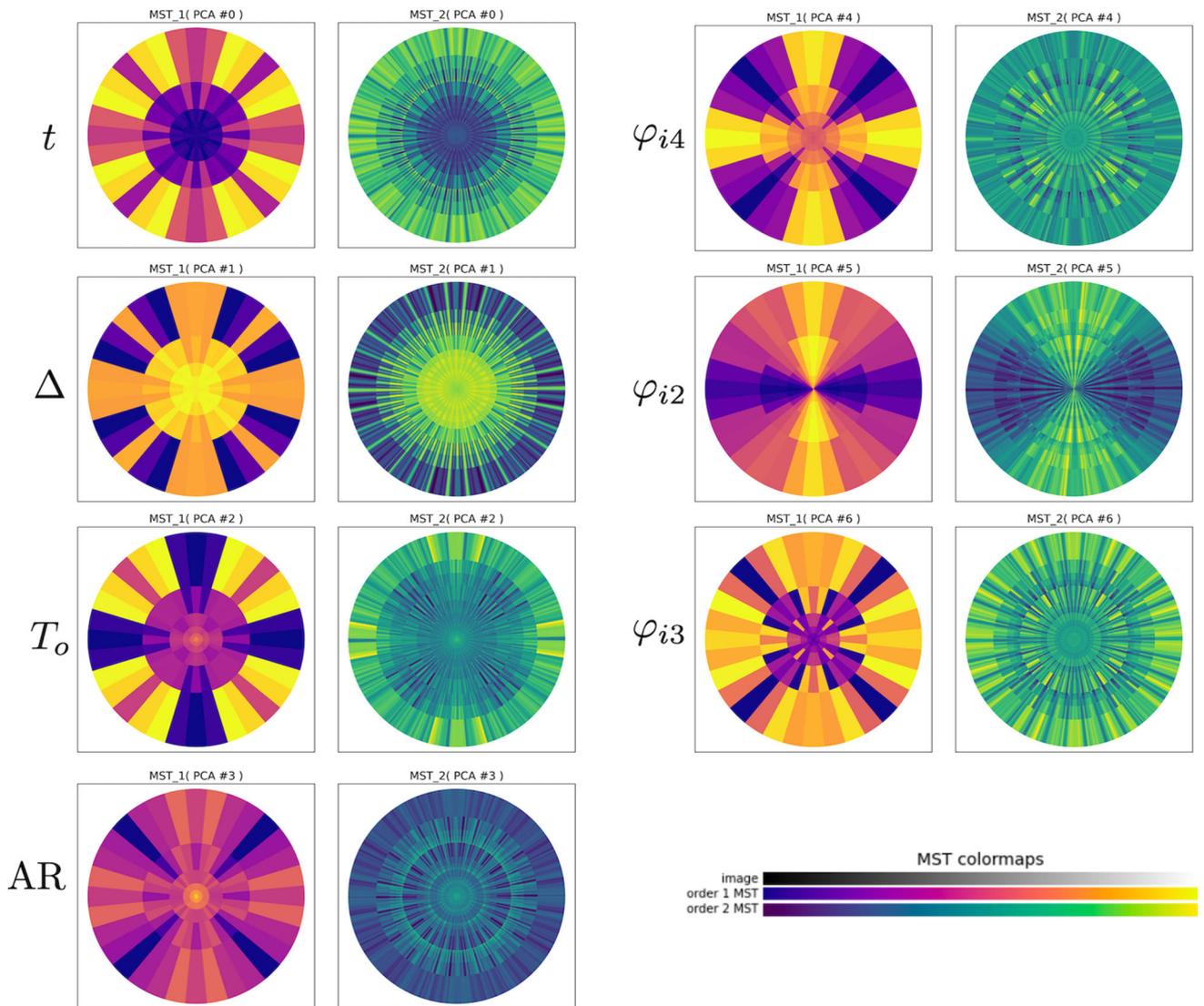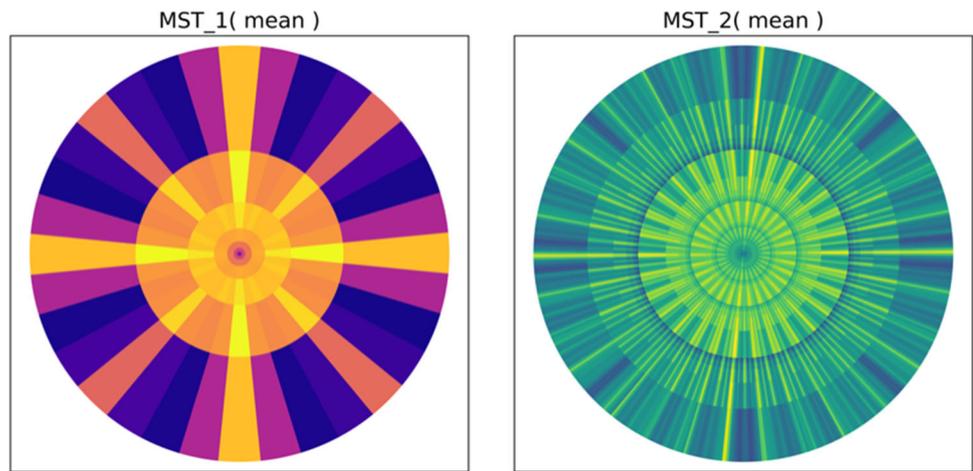
Fig. 9 PCA vectors. Shown are the PCA vectors of liner density in MST space. The radial axis is linear in scale. The color bar for all MST displays is in the lower right hand corner. The identification of each PCA vector with an input is displayed to the left of the PCA vector. See Fig. 12 for this interpretation key
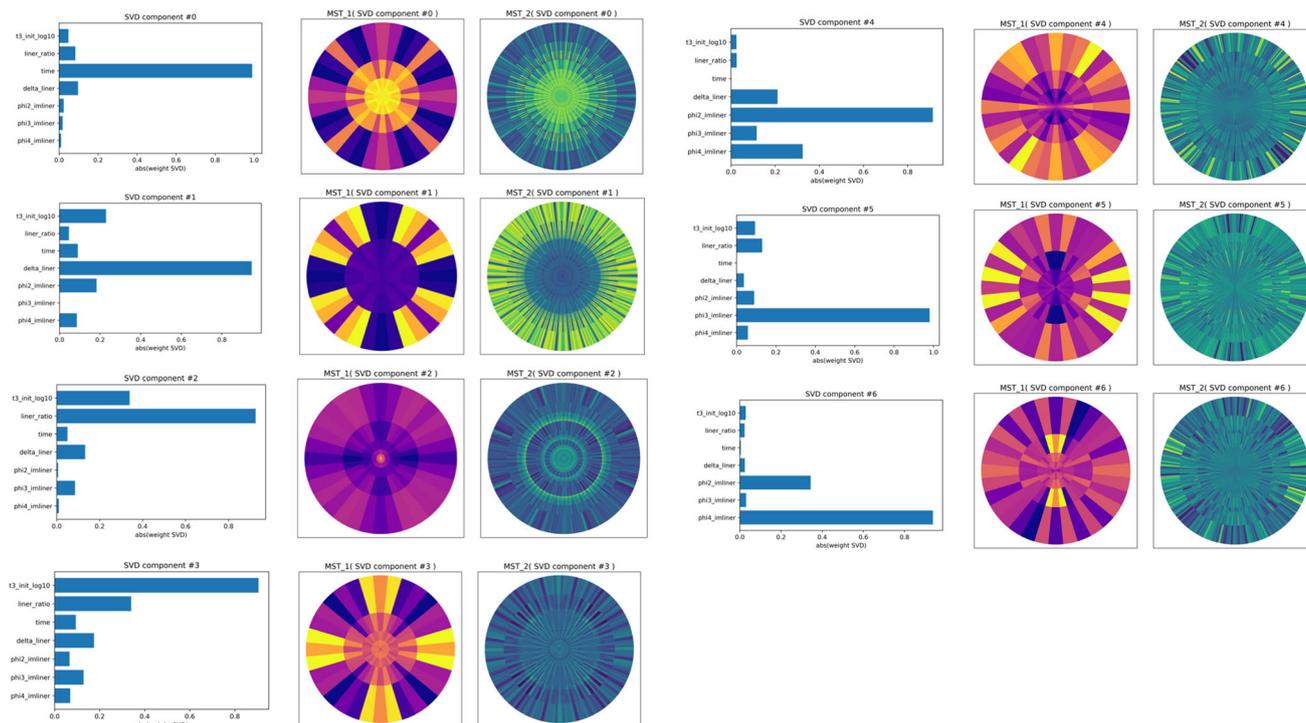
Fig. 10 SVD vectors. Shown are the SVD vectors in both input space and liner density MST space. Note the clear identification of each SVD component with one input. The radial axis is linear in scale
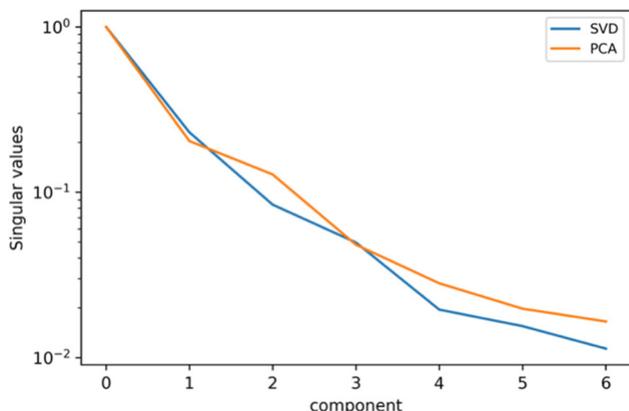
Fig. 11 The eigenvalues of the PCA vectors compared to those of the SVD. Note that the PCA is more efficient at capturing the variance in fewer components than the SVD

tern on the second order MST. The pattern seems to appear stronger as the radius gets smaller. You can see this by focusing your attention on the outer three rings of the first order MST. Note that as the φ_{i4} PCA vector loses its structure, the structure of the φ_{i2} PCA vector increases. This is happening because there is an inverse cascade from the φ_{i4} PCA vector to the φ_{i2} PCA vector, forming a self-organized dipole state. This state persists as the plasma expands, post stagnation, to a larger radius. This is further highlighted by the permutation feature importance, shown in Fig. 18. It should be no

surprise that time is the most important feature, followed by AR. The liner aspect ratio changes the acceleration, which effects all facets of the evolution. An interesting result is that the phase of the $m = 2$ mode is the next most important feature. On reflection, this is not surprising. The initial phase of the $m = 2$ mode, although a stochastic variable, is very quickly reinforced by the inverse cascade into this mode, and sets the phase of the resulting dipole mode. The size of the perturbation does not matter as much, as demonstrated by Δ being less important than φ_{i2}, and φ_{i3} and φ_{i4} having little to no importance. The preheat temperature T_0 only effects the evolution near stagnation, so its modest importance is expected.

Let us now focus our attention on the detail of the evolution being predicted by the MLP/NN by looking closely at the prediction of the first PCA vector component in Fig. 19. Remember that a MLP/NN with ReLU activation layers is a piecewise linear universal function approximator. There are few tie points where the function is linear, but the number increases near the singularities in the mapping, where the slope is changing rapidly. A couple of tie points are always put near stagnation. The liner with the largest AR has the largest acceleration and should converge first. Indeed, the two simulations with AR = 9, shown in red and yellow, do converge first. This time should be roughly independent of the preheat temperature T_0 (which it is), and the stagnation with the smaller $T_0 = 10\,\text{eV}$ should have more compression
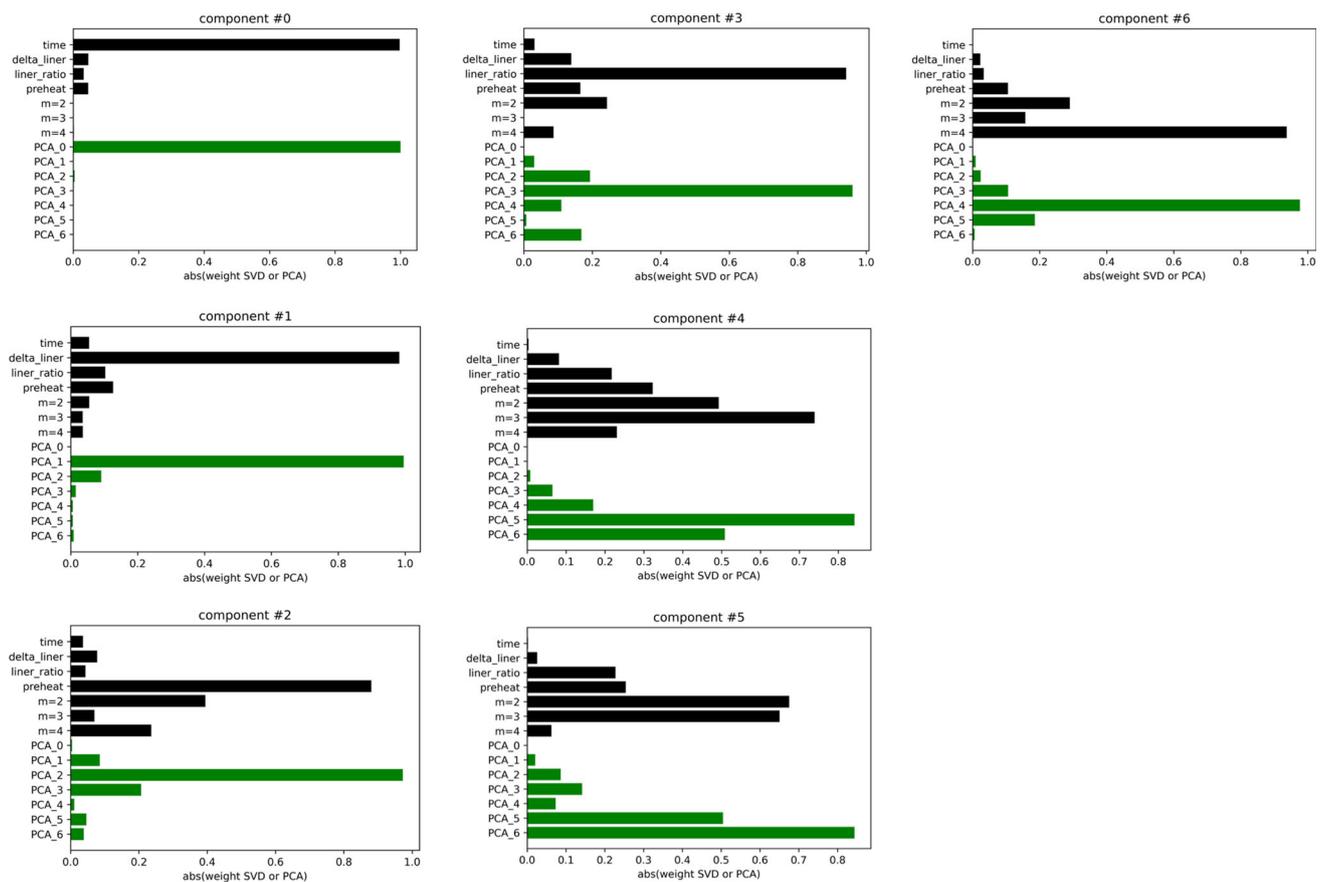
Fig. 12 The singular vectors of the PCA-to-SVD cross variance analysis. This analysis allows the identification of the PCA vectors with a unique input

than the one with $T_0 = 631$ eV (which it does). This feature is also present with the AR $= 6$ stagnations (green and magenta), and the AR $= 3$ stagnations (blue and cyan). The MLP/NN captures this feature well.

There are several astonishing things that have happened in this MLDL workflow, that lead to the following questions. Why did the dynamics reduce down to the evolution of a few PCA vector components? Another way of asking this is: why are the dynamics constrained to a very low dimensional linear subspace of this complex (in the case of the WPH) Hilbert space? What is the physical interpretation of the basis vectors that span this low-dimensional linear subspace? Why are both the SVDs (inputs to outputs, and PCA projection to SVD projection) nearly diagonal? Why were the $\log_{10}$ operations needed before and after the functional convolutional transformation (the MST and WPH)? Why was the best MLP/NN architecture an encoder/decoder architecture, and what is the physical interpretation of the coordinates of the middle layer? Why was it so easy for the MLP/NN to approximate the function with so few neurons? The answer to these questions will be the subject of an upcoming paper, now in draft form. Because the design of this workflow was

not accidental or the product of a large amount of experimentation (in fact, all the best design choices were the first thing that we tried, including the structure of the MLP/NN), we will present the answers to these questions, which you should view as hypotheses at this time. The efficacy of this MLDL is tantalizing evidence supporting these hypotheses.

Could the MST, if properly formulated, be a transformation to a complex "renormalization" space where the basis vectors are the solution to the Renormalization Group Equations (RGE)? (Remember that Renormalization is the study of how the physics changes as a function of scale, in our case p. The solution to the RGEs, a coupled set of ODEs, gives the scaling exponents as a function of scale, where the fields and coupling constants scale as $\sim f_i^{\beta_i(p)}$, where the $\beta_i(p)$ are the solutions to the RGEs for field f_i.) Is a natural logarithm needed to flatten this space? Is there then a simple mapping to decoupled action-angle coordinates on the low-dimensional linear subspace, where the action and the angle are uniform circular functions of time? After all, the RGEs are coupled ODEs of the logarithmic derivative. The complex $\ln(z)$ function, as a conformal mapping, takes polar coordinates about the origin to the cylinder where
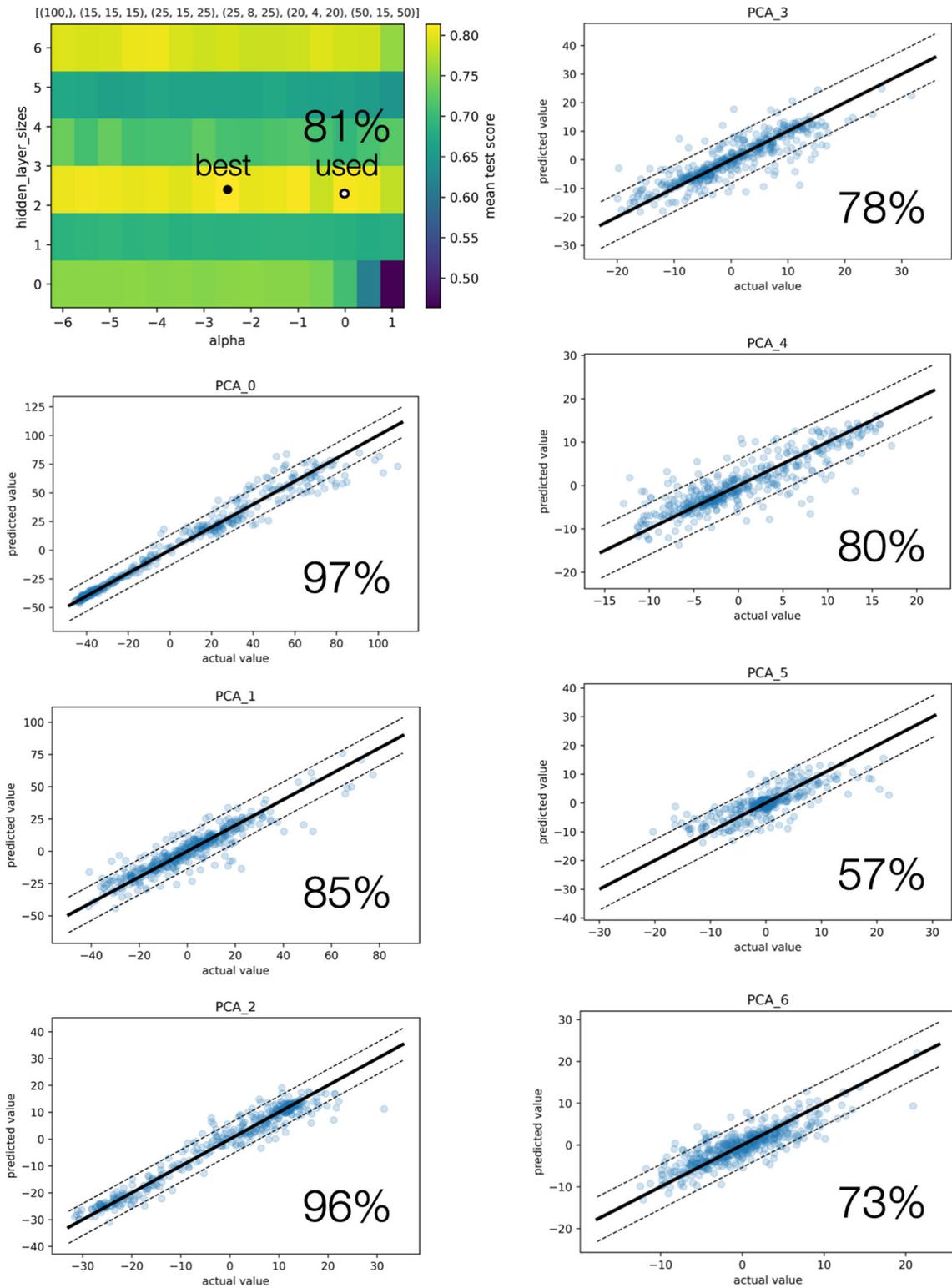
Fig. 13 Cross plots of predicted versus actual values for each PCA vector component, by the MLP/NN. The score for each PCA vector component is shown on the respective plot. The grid search for the MLP/NN structure and regularization parameter α is shown in the upper left corner. A structure of hidden layers of 25-15-25, and a value of $\alpha = 0$ is used
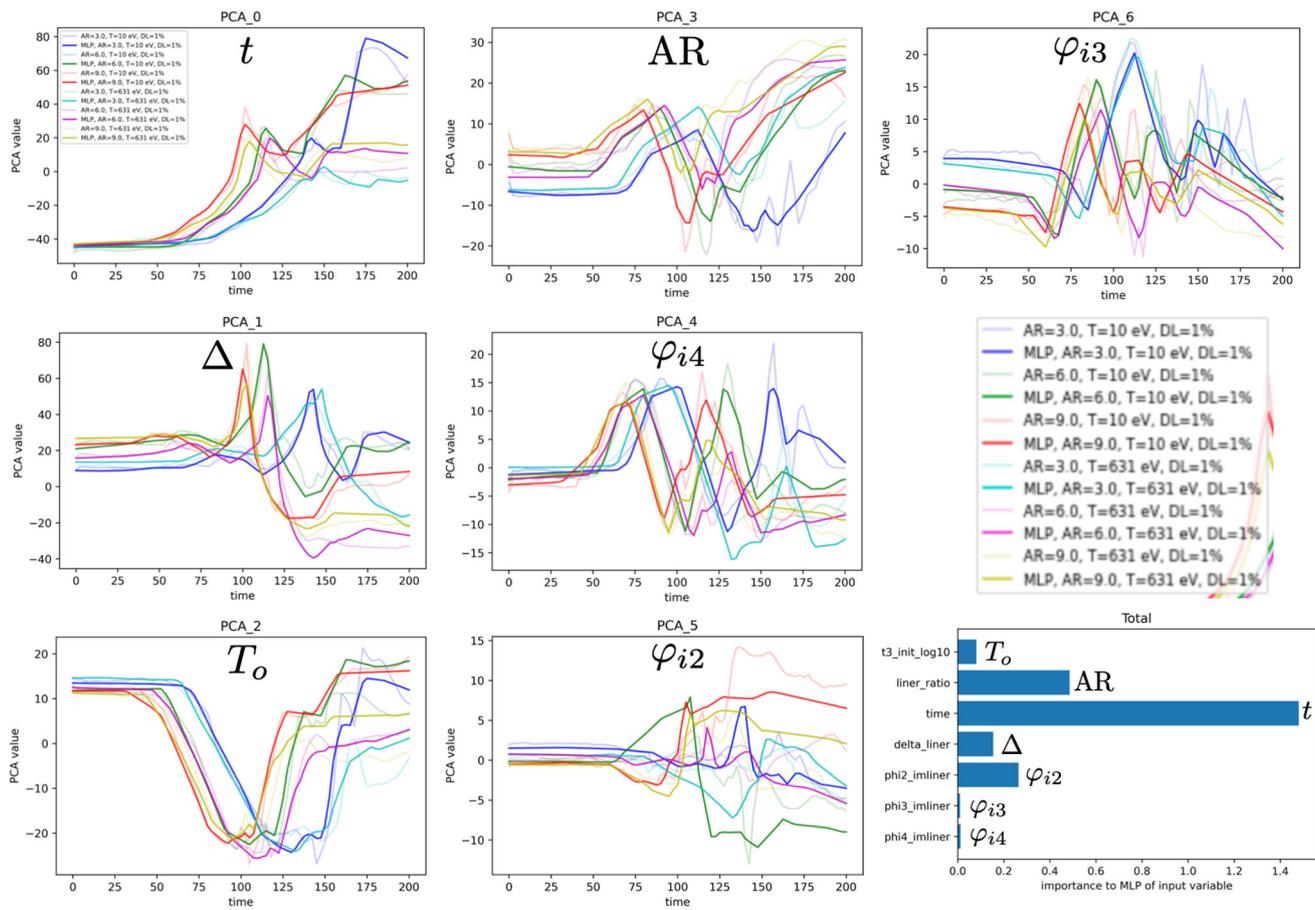
Fig. 14 The evolution of the PCA vector components as predicted by the MLP/NN compared to the actual values. The predicted values are the bold lines and the actual values are the light lines of the same color. The permutation feature importance is shown in the lower right hand corner.

The identification of each PCA vector component with the unique input parameter is indicated on each plot. Shown are six simulations that span $AR = 3, 6, 9$ and $T_0 = 10$ eV, 631 eV

$$\ln(z) = \ln | z | + i \ \arg(z), \qquad (13)$$

which flattens the space. A PCA then would identify this linear subspace. An SVD analysis of the fields and coupling constants would diagonally correlate to the basis vectors of this subspace. The PCA vectors would be the solutions to the RGEs. The number of important PCA vectors would therefore be equal to the number of fields and coupling constants. The number of actions plus angles would be twice the number of PCA vectors. Given that there are 6 fields (density, charge density, and 4 E&M fields) and 4 coupling constants (charge of the electron and ion, and mass of the electron and ion, which were held constant) in MHD, the use of 7 PCA vectors and 15 nodes in the middle hidden layer is not surprising. There are six fields and another adjoint basis vector for the resistivity, giving a total of seven. The encoding in the middle layer needs to have a node for the action and the angle for each field and another for time, giving a total of 15. The motion on this low dimensional linear subspace

would be geodesic motion determined by an analytic function that the MLP/NN is approximating. It is interesting to note that knowing the topology of this low-dimensional complex space is equivalent to knowing the analytic function (it is the solution to Laplace's equation). For a simple topology, this would be knowing the location and order of the poles and zeros. It would be very easy for a MLP/NN to approximate this function since it is a solution to Laplace's equation. It would need few tie points away from the poles and zeros because the space would be flat.

These hypotheses, if true, lead to some interesting applications of this MLDL workflow. For example, the predictions will extrapolate well, as long as the extrapolation is made going away from the poles and zeros of the topology. This can be experimentally tested. First, make a prediction using the MLDL predictor of scaling into into a new regime, then do the experiment. If the prediction is better than expected, then the extrapolation is away from the poles and zeros. If the prediction is worse than expected, then the extrapolation
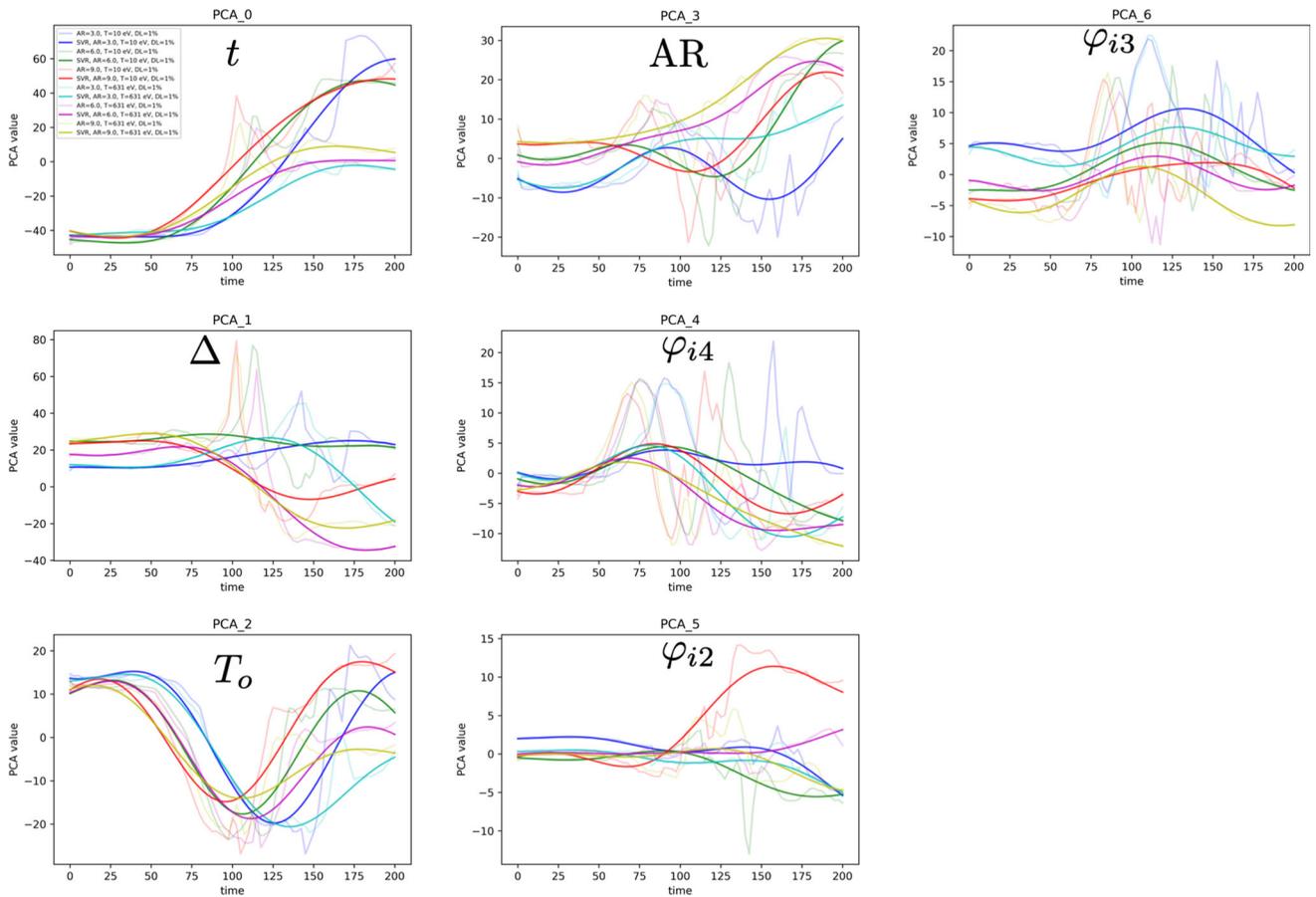
Fig. 15 The evolution of the PCA vectors as predicted by the SVM compared to the actual values. The predicted values are the bold lines and the actual values are the light lines of the same color. The identifi-cation of each PCA vector with the unique input parameter is indicated on each plot. Shown are six simulations that span AR = 3, 6, 9 and $T_0 = 10$ eV, 631 eV

is towards new poles and zeros. The theoretical model needs to have additional physics added to it. In fact, the MLDL workflow, if augmented with the experimental points and the deletion of the theoretical (computer simulation) in this extrapolation regime, will determine what topology needs to be added. This process is one of topological discovery or causal analysis.

Another way of looking at the MLDL workflow presented and implemented in this paper is a redefinition of the MST/WPH to

$$S_m[f(x)](p, x) \equiv \phi_{px} \star \left(\prod_{k=1}^{m} i \ln R_0 \, \psi_{p_k} \star \right) i \ln R_0 \, f(x),$$

$$(14)$$

where

$$R_0(z) \equiv z + e^{i \arg(z)}.$$

$$(15)$$

It should be noted that this transformation is no longer stationary since the Father Wavelet only averages over as large

of a patch as it has to do. Nothing prevents this partition of unity from being summed over a larger domain in x, if the process is stationary over that domain. This transformation is complex from the beginning to the end. The real part is the $\ln(\text{mod})$ and the imaginary part is the arg. Not only does $\ln(R_0(z))$ conformally flatten the space onto the cylinder, it now (with the addition of R_0) exponentially (for large deviation), then logarithmically (for small deviations) converges to the origin $z = 0$. The connection to the MST/WPH work can be seen by examining the limiting behavior of the $\ln(R_0(z))$ mapping

$$\ln(R_0(z)) \xrightarrow[|z| \to 0]{} z$$

$$\xrightarrow[|z| \to \infty]{} \ln |z| \quad \text{(this work)}$$

$$\xrightarrow[|z| \to \infty]{} \ln |z| \, e^{i \arg(z)/\ln|z|}$$

$$\sim \sum_{k=0}^{n} |z| \, e^{ik \arg(z)} \quad \text{(that is, WPH)}$$

$$\xrightarrow[n=0]{} |z| \quad \text{(that is, MST)}.$$

$$(16)$$
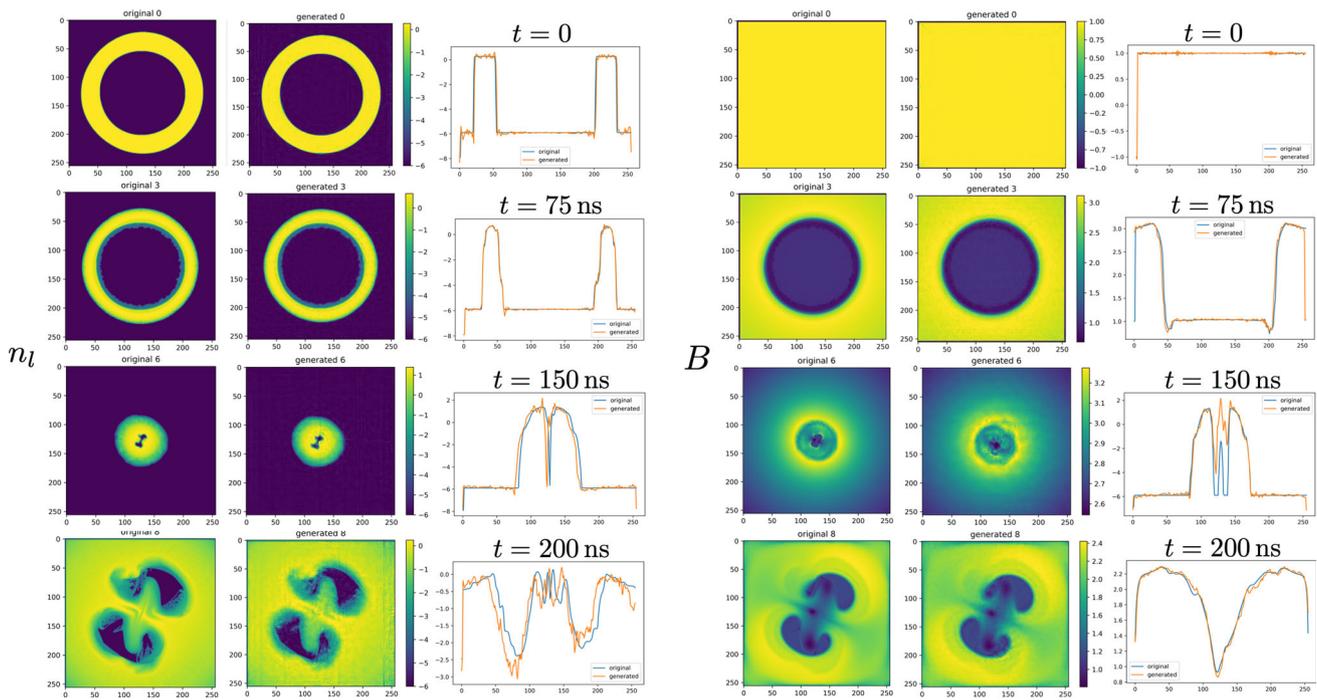
Fig. 16 The evolution of the liner density (left) and the magnetic field (right), and the reconstruction from the inverse WPH (iWPH) and horizontal lineouts through both profiles. The fields are displayed at four key times (top to bottom): $t = 0$, $t = 75$ ns, $t = 150$ ns, and $t = 200$ ns. The initial conditions are AR $= 3$, $T_0 = 631$ eV, and $\Delta = 1\%$

Fig. 17 A closeup of two critical PCA vectors that illuminate the inverse cascade. The radial axes are logarithmic in scale. Shown are the first and second order MST for the φ_{i4} (top) and φ_{i2} (bottom) components
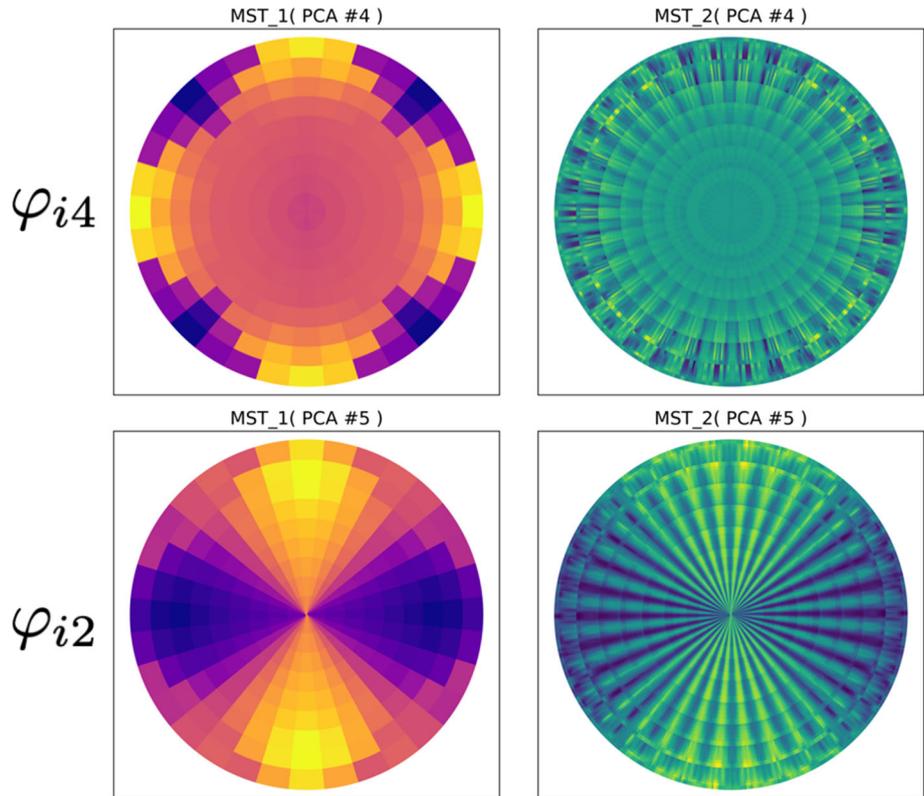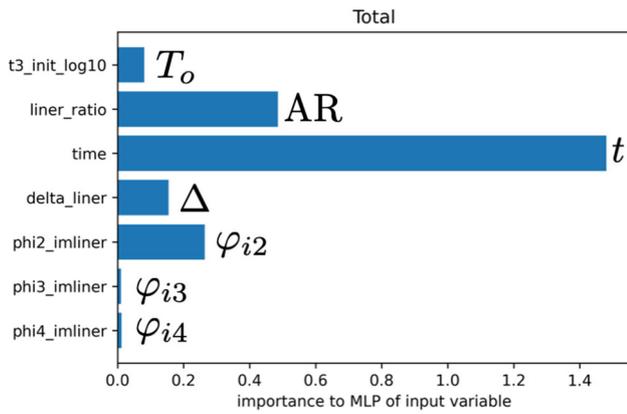
Fig. 18 MLP/NN permutation feature importance
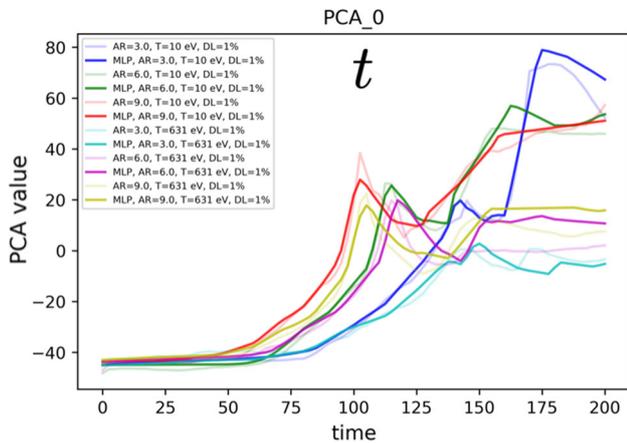
Fig. 19 The time evolution of the first PCA vector component with respect to time. The MLP/NN estimated values are the bold lines, the GORGON simulated values are the light lines in corresponding colors. Shown are six cases for $T_0 = 10\,\text{eV}$, $631\,\text{eV}$ and $AR = 3, 6, 9$

Fig. 20 Three different variations of the MLDL workflows: (top) a composite workflow, (middle) a decomposed approach, and (bottom) the hybrid case presented in this paper

The characteristic of the fixed point at the origin is the statement of the first limit. The second limit is effectively what has been used in this paper. In the third limit (which keeps the small arg imaginary term), the ln chirps the pulse, generating the harmonics that are explicitly generated in the fourth line by the WPH. The conventional MST is just using the first term in that series.

This MLDL architecture can be modified to make it a sequential, modular approach as shown in Fig. 20. In general, one starts with initial conditions $\tilde{A}$ and coupling constants C, does a computer simulation of the process, followed by a computer simulation of the diagnostics, to predict the diagnostic response $\tilde{D}$. Here the tilde signifies the PCA(ln(MST/WPH(ln))) of the field quantity. The composite approach can be taken, where there is one MLP/NN, $\tilde{D}(\tilde{A}, C)$. There is a more flexible decomposed approach where the approximation is separated into three approximators. The first is a mapping of the initial conditions to the initial fields $\tilde{z}_i(\tilde{A})$. This is followed by a general dynamic

mapper of the initial fields, coupling constants and time to the fields $\tilde{z}(\tilde{z}_i, C; t)$. Finally, there is a mapping from the fields to the diagnostics, $\tilde{D}(\tilde{z})$. The current work is a restricted hybrid of initial conditions and time to the fields $\tilde{z}(\tilde{A}; t)$.

This work needs to be extended into 3D to see if the inverse cascade persists and what the characteristics of the 3D self-organized state are. It is well-known that the inverse cascade in 2D fluid flow is caused by the topological invariant of vorticity. For the case of the magnetized plasmas encountered in MagLIF, there is still a topological invariant: helicity [23–25]. There is experimental evidence of helical stagnations showing the dipole structure in the plane perpendicular to the axis [9]. In addition, there is also an axial sausage mode. The result is a double helical structure that looks like a DNA atom.

There are obvious improvements that also need to be made to the MST/WPH. It needs to be made non-stationary, such that it is not translationally invariant. The complex natural logarithm function needs to be made an integral part of it.

The orthogonal local wavelet basis needs to be a carefully constructed partition of unity, so that a fast inverse transformation can be constructed. Given this basis, a resolution independent, physical display needs to be constructed.

We end with a summary of what this research has shown or given indications may be true. It has demonstrated a fast, high fidelity surrogate for resistive MHD. This surrogate is 10^7 times faster than conventional computational prediction. It is based on a simple, fast to train, physics-based machine learning. It gives field to field correlation, physically interpretable results, and meaningful graphical displays. It has the potential to give fundamental insights into nonlinear dynamics, physical kinetics, quantization and second quantization, renormalization, and the topology of dynamics. This surrogate will either extrapolate well or give insights into additional causality. Finally, from the practical MagLIF physics perspective, it has shown an emergent behavior of 2D MagLIF implosions into a self organized dipole state.

References

1. Rasmussen CE (2006) Gaussian processes in machine learning. MIT Press, Cambridge
2. Goodfellow I, Bengio Y, Courville A (2016) Deep learning. MIT press, Cambridge
3. le Cun Y (1989) Generalization and network design strategies. Technical Report CRG-TR-89-4. Tech. rep., University of Toronto, Department of Computer Science
4. Mallat S (2012) Group invariant scattering. Commun Pure Appl Math 65(10):1331–1398
5. Bruna J, Mallat S (2013) Invariant scattering convolution networks. IEEE Trans Pattern Anal Mach Intell 35(8):1872–1886. https://doi.org/10.1109/TPAMI.2012.230. arXiv:1203.1513
6. Mallat S, Zhang S, Rochette G (2020) Phase harmonic correlations and convolutional neural networks. Inform Inference J IMA 9(3):721–747
7. Zhang S, Mallat S (2021) Maximum entropy models from phase harmonic covariances. Appl Comput Harmon Anal 53:199–230
8. Slutz SA, Herrmann MC, Vesey RA, Sefkow AB, Sinars DB, Rovang DC, Peterson KJ, Cuneo ME (2010) Pulsed-power-driven cylindrical liner implosions of laser preheated fuel magnetized with an axial field. Phys Plasmas 17(5):056303. https://doi.org/10.1063/1.3333505
9. Awe TJ, McBride RD, Jennings CA, Lamppa DC, Martin MR, Rovang DC, Slutz SA, Cuneo ME, Owen AC, Sinars DB, Tomlinson K, Gomez MR, Hansen SB, Herrmann MC, McKenney JL, Nakhleh C, Robertson GK, Rochau GA, Savage ME, Schroen DG, Stygar WA (2013) Observations of modified three-dimensional instability structure for imploding z-pinch liners that are premagnetized with an axial field. Phys Rev Lett 111(23):1–5. https://doi.org/10.1103/PhysRevLett.111.235005
10. Gomez MR, Slutz SA, Sefkow AB, Sinars DB, Hahn KD, Hansen SB, Harding EC, Knapp PF, Schmit PF, Jennings CA, Awe TJ, Geissel M, Rovang DC, Chandler GA, Cooper GW, Cuneo ME, Harvey-Thompson AJ, Herrmann MC, Hess MH, Johns O, Lamppa DC, Martin MR, McBride RD, Peterson KJ, Porter JL, Robertson GK, Rochau GA, Ruiz CL, Savage ME, Smith IC, Stygar WA, Vesey RA (2014) Experimental demonstration of fusion-relevant conditions in magnetized liner inertial fusion. Phys Rev Lett 113(15):1–5. https://doi.org/10.1103/PhysRevLett.113.155003
11. McBride RD, Slutz SA, Jennings CA, Sinars DB, Cuneo ME, Herrmann MC, Lemke RW, Martin MR, Vesey RA, Peterson KJ, Sefkow AB, Nakhleh C, Blue BE, Killebrew K, Schroen D, Rogers TJ, Laspe A, Lopez MR, Smith IC, Atherton BW, Savage M, Stygar WA, Porter JL (2012) Penetrating radiography of imploding and stagnating beryllium liners on the *z* accelerator. Phys Rev Lett 109(135):004. https://doi.org/10.1103/PhysRevLett.109.135004
12. Seyler CE, Martin MR, Hamlin ND (2018) Helical instability in maglif due to axial flux compression by low-density plasma. Phys Plasmas 25(6):062,711. https://doi.org/10.1063/1.5028365
13. Yager-Elorriaga DA, Lau Y, Zhang P, Campbell PC, Steiner AM, Jordan NM, McBride RD, Gilgenbach RM (2018) Evolution of sausage and helical modes in magnetized thin-foil cylindrical liners driven by a z-pinch. Phys Plasmas 25(5):056,307
14. Chittenden JP, Lebedev SV, Jennings CA, Bland SN, Ciardi A (2004) X-ray generation mechanisms in three-dimensional simulations of wire array Z-pinches. Plasma Phys Controll Fus 46(12B):B457. https://doi.org/10.1088/0741-3335/46/12B/039

15. Haykin S, Haykin S (1994) Neural networks: a comprehensive foundation (Macmillan). https://books.google.com/books?id=PSAPAQAAMAAJ

16. Ning F, Delhomme D, le Cun Y, Piano F, Bottou L, Barbano P (2005) Toward automatic phenotyping of developing embryos from videos. IEEE Trans Image Process 14(9):1360–1371

17. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2014) Advances in neural information processing systems. Curran Assoc 27:2672–2680

18. Zinn-Justin J (2010) Path integrals in quantum mechanics. Oxford Graduate Texts (OUP Oxford). https://books.google.com/books?id=MWQBAQAAQBAJ

19. Allys E, Marchand T, Cardoso JF, Villaescusa-Navarro F, Ho S, Mallat S (2020) New interpretable statistics for large-scale structure analysis and generation. Phys Rev D 102(10):103506

20. Régaldo-Saint Blancard B, Allys E, Auclair C, Boulanger F, Eickenberg M, Levrier F, Vacher L, Zhang S (2022) Generative models of multi-channel data from a single example–application to dust emission. arXiv e-prints pp. arXiv–2208

21. Jammalamadaka SR, SenGupta A (2001) Topics in circular statistics. World Scientific, Singapore

22. Andreux M, Angles T, Exarchakis G, Leonarduzzi R, Rochette G, Thiry L, Zarka J, Mallat S, Andén J, Belilovsky E, Bruna J, Lostanlen V, Chaudhary M, Hirn MJ, Oyallon E, Zhang S, Cella C, Eickenberg M (2020) Kymatio: scattering transforms in python. J Mach Learn Res 21(60):1–6

23. Perez JC, Boldyrev S (2009) Role of cross-helicity in magnetohydrodynamic turbulence. Phys Rev Lett 102(2):025,003

24. Glinsky ME, Hjorth PG (2019) Helicity in hamiltonian dynamical systems. Tech. Rep. SAND2019-14731, Sandia National Laboratories arXiv:1912.04895

25. Taylor J (1986) Relaxation and magnetic reconnection in plasmas. Rev Mod Phys 58(3):741