Proceedings of the ASME 2021
International Mechanical Engineering Congress Exposition
IMECE2021
November 1 – 5, 2021, Virtual, Online

IMECE2021-70463

DEFORMATION MANIFOLD LEARNING MODEL FOR DEFORMATION OF MULTI-WALLED CARBON NANO-TUBES: EXPLORING THE LATENT SPACE

Upendra Yadav
Shashank Pathrudkar
Mechanical Engineering - Engineering Mechanics
Michigan Technological University
Houghton, MI, USA

Susanta Ghosh*
Mechanical Engineering - Engineering Mechanics
and
The Center for Data Sciences
Michigan Technological University
Houghton, MI, USA

ABSTRACT

A novel machine learning model is presented in this work to obtain the complex high-dimensional deformation of Multi-Walled Carbon Nanotubes (MWCNTs) containing millions of atoms. To obtain the deformation of these high dimensional systems, existing models like Atomistic, Continuum or Atomistic-Continuum models are very accurate and reliable but are computationally prohibitive for these large systems. This high computational requirement slows down the exploration of physics of these materials. To alleviate this problem, we developed a machine learning model that contains a) a novel dimensionality reduction technique which is combined with b) deep neural network based learning in the reduced dimension. The proposed non-linear dimensionality reduction technique serves as an extension of functional principal component analysis. This extension ensures that the geometric constraints of deformation are satisfied exactly and hence we termed this extension as constrained functional principal component analysis. The novelty of this technique is its ability to design a function space where all the functions satisfy the constraints exactly, not approximately. The efficient dimensionality reduction along with the exact satisfaction of the constraint bolster the deep neural network to achieve remarkable accuracy. The proposed model predicts the deformation of MWCNTs very accurately when compared with the deformation obtained through atomistic-physics-based model. To simulate the complex high-dimensional deformation the atomistic-physics-based models takes weeks high performance computing facility, whereas the proposed machine learning model can predict the deformation in seconds. This technique also extracts the universally dominant pattern of deformation in an unsupervised manner. These patterns are comprehensible to us and provides us a better explanation on the working of the model. The comprehensibility of the dominant modes of deformation yields the interpretability of the model.

NOMENCLATURE

N_w	Number of walls
φ_0	Configuration map from parametric domain to the local reference configuration
φ	Configuration map from parametric domain to the deformed configuration
Φ	Deformation map
N_{cs}	Number of cross-sections
ξ_1,ξ_2,ξ_3	Representation of a point in parametric domain
m	m-th wall in a MWCNT
n	n-th cross-section along the length
r	Radial deformation
θ	Parameter defining the circumferential direction
N	Length of data set
g	Function representing the geometric constraint
μ	Mean of the data set
v	Covariance of the data set

*Address all correspondence to this author.

ψ_k k-th eigenfunction of the covariance function or k-th most dominant mode of variation

$\mathscr{B}$ Basis set out of which initial basis function must be chosen to perform functional principal component analysis while satisfying the constrain exactly

ϕ Chosen basis functions from the Hilbert space or $\mathscr{B}$

$\bar{c}_{ik}$ Coefficients of Functional principal components

INTRODUCTION

Carbon nanotubes have shown remarkable optical, electronic, and physical properties. Moreover, the deformation configuration that they achieve on loading influences their electronic properties, leading to a large number of applications including nanoelectromechanical systems (NEMS) [1, 2, 3]. In various experiments, a wavelike pattern has been seen when multi-walled carbon nanotubes undergo loading [4, 2]. These wavelike patterns are termed as rippling. The most important factor responsible for rippling pattern is the low bending modulus of MWCNTs as they act as crystalline membranes. Contrary to that, inter-wall interaction or the Van der Waals interaction between two consecutive walls in MWCNTs provide the stiffness against bending. On loading, MWCNTs obtain the configuration at which the combination of in-plane strain energy density and Van der Waals interaction energy reaches to a minimum, leading to a rippling pattern. These configurations are considered at equilibrium configuration.

Even though the deformation configurations of MWCNTs influences their chemical, and electronic properties, accurate and efficient tools to obtain these complex deformed configurations of MWCNTs are elusive. *Ab-initio* simulation are very accurate but they require huge amount of computational effort to simulate a sample containing millions of atoms. MWCNTs have also been modelled as continuum shell which makes these model very efficient but they ignore the underlying atomistic physics and hence lack in accuracy. To achieve the accuracy of atomistic models and efficiency of continuum models, Atomistic-Continuum (AC) models have been developed [5, 6, 7] by integrating atomistic and continuum models. AC models are accurate and efficient when compared to atomistic models but still require a significant amount of high-performance computing efforts for large MWCNTs. This slows down the process of exploration of the mechanics of these materials.

The advent of Machine Learning (ML) can help us in alleviating this bottleneck. Machine Learning (ML) methods such as Deep Neural Networks (DNNs) [8, 9] are intensely investigated for accelerating mechanics, and materials research [10, 11, 12], however, so far most of the applications are limited to the prediction of low-dimensional properties, such as material moduli. Whereas, the complex deformations that MWCNTs exhibit requires the prediction in high-dimensional space. While simulating MWCNTs using AC models, highest accuracy is achieved by discretizing the domain. This discretized data lies in a high-dimensional space [13, 7].

State-of-the-art deep learning models can predict the low dimensional (e.g. CNN [14, 15]) or high dimensional outputs (e.g. Encoder-Decoder [16, 17]). But in order to predict the high-dimensional outputs they require a high-dimensional input. State-of-the-art DNNs cannot accurately predict high-dimensional deformation of MWCNT from a few input features like geometry paramters and loading conditions. The objective of the present study is to create a ML model to accurately and efficiently predict high-dimensional discretized deformations of MWCNTs as output from low-dimensional inputs. This can be achieved by reducing the dimension of output labels. A simple dimensionality reduction technique can lead to inaccurate predictions as the deformed configuration is a highly non-linear manifold. Commonly used dimensionality reduction techniques [19] such as Principal Component Analysis (PCA) and classical Metric Multidimensional Scaling don't work for the deformation of MWCNTs as the deformation is highly non-linear and these dimensionality reduction techniques are based on linear models. There exists non-linear dimensionality reduction techniques such as Isomap, t-SNE, Locally-Linear Embedding etc. These techniques are most widely used to identify the low-dimensional non-linear manifold structure present in the data [20, 21]. An approximate low-dimensional neighborhood graph embedded in the high-dimension is obtained following the manifold structure of the data using the above mentioned techniques. To predict the high-dimensional deformation of MWCNTs an accurate and smooth map with functional representation is required. The map should also satisfy the intrinsic constraints of the deformation. Moreover, the functional representation of the mapping also ensures the reversibility of the map from low-dimensional to high-dimensional manifold. One such technique is Functional Principal Component Analysis (FPCA) [22, 23]. It provides a smooth functional representation of the data, which is analogous to Kosambi-Karhunen-Loève Expansion [24, 25]. FPCA is used to represent any stochastic process through a linear combination of an infinite number of orthogonal functions. These orthogonal functions are known as functional principal components and they are analogous to principal components in the standard Principal Component Analysis. However, it is not necessary that these orthogonal functions satisfy any constraint and hence are undesirable for the current work as MWCNTs has inherent periodicity constraint along the circumference due to its cylindrical structure. As a consequence FPCA yields discontinuous and erroneous predictions for MWCNTs.

In the present work, an extension to standard FPCA is proposed. The extension involves designing a function space where all included functions satisfies the constraint exactly. We named this novel technique as constrained-FPCA (c-FPCA). The proposed c-FPCA technique benefits us in two ways. *Firstly*, it alleviates the *curse of dimensionality* and provides a

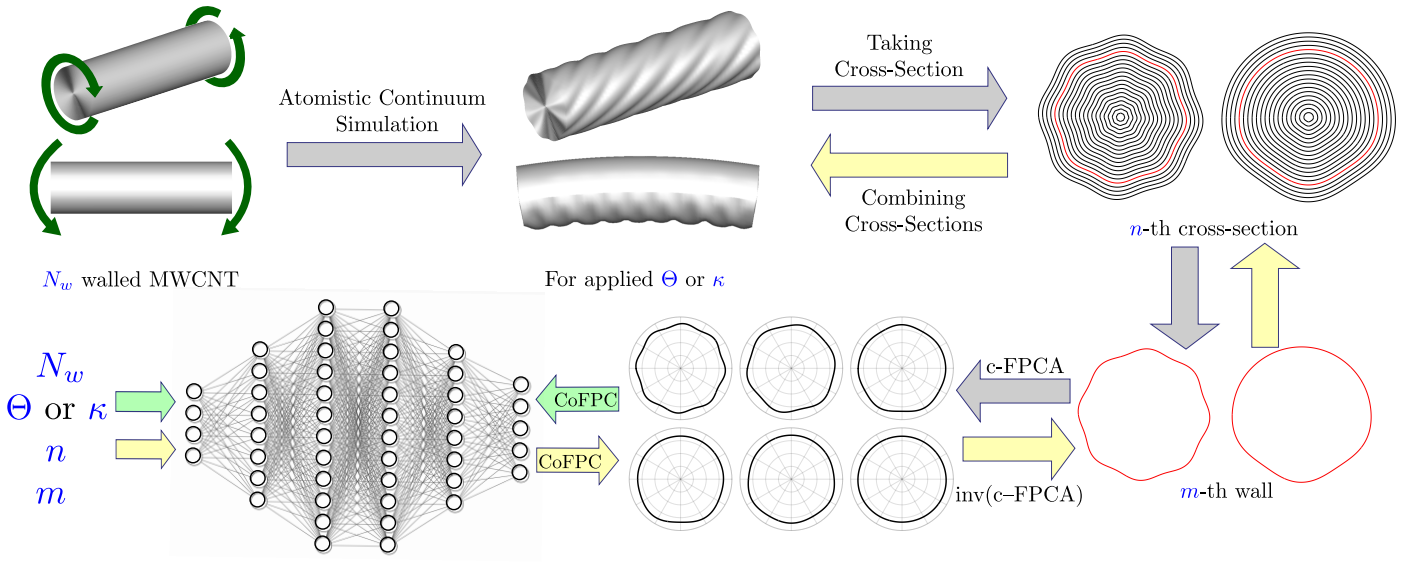

FIGURE 1. A DETAILED SCHEMATIC OF THE PRESENT FRAMEWORK INVOLVING THE DATA GENERATION VIA SIMULATION AND THE PROPOSED DEFORMATION MANIFOLD LEARNING (DML) MODEL. THE DML MODEL INCLUDES: THE DIMENSIONALITY REDUCTION PERFORMED THROUGH CONSTRAINED FUNCTIONAL PRINCIPAL COMPONENT ANALYSIS AND THE DNN-BASED LEARNING. INPUTS AND OUTPUTS OF DNN ARRE SHOWN WITH GREEN ARROWS. YELLOW ARROWS REPRESENT THE PREDICTION VIA THE DML MODEL TO OBTAIN THE DEFORMATION FOR A GIVEN INPUT [18].

low-dimensional representation in functional form. *Secondly*, it enforces the existing geometric constraint exactly hence ensuring a smooth deformation pattern in prediction. The proposed semi-supervised ML model consists of two steps (i) unsupervised dimensionality reduction via proposed c-FPCA to obtain the functional principal components present in the deformation and (ii) supervised learning (regression) via DNN in the reduced dimension. Henceforth, the proposed ML model will be referred to as the *Deformation Manifold Learning* (DML) model. The corresponding schematic is shown in Fig. 1. The proposed model takes the geometric information of MWCNTs and the applied loading condition as input and predicts the high-dimensional deformation.

The functional representation of the discrteized data also leads to interpretability. The interpretability arises as the *latent space* is comprehensible through the functions spanning it. The proposed interpretability also allows us to understand how the model predicts that makes it comprehensible.

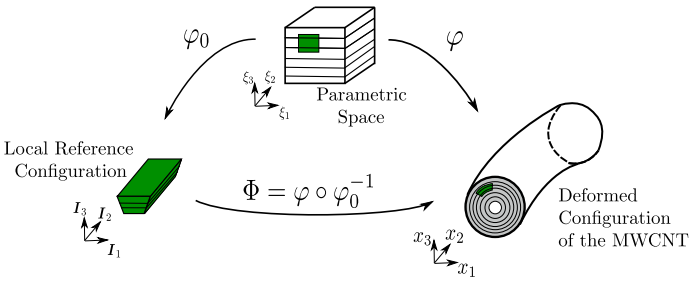

FIGURE 2. ILLUSTRATION OF THE KINEMATICS SHOWING THE PARAMETRIC DOMAIN, UNDEFORMED OR LOCAL REFERENCE CONFIGURATION AND DEFORMED CONFIGURATION ALONG WITH THE DEFORMATION MAPS, AND COORDINATE AXES [18].

DEFORMATION MANIFOLD LEARNING MODEL

As the electronic properties of MWCNTs are coupled with their mechanical deformation, it is important to capture the buckling pattern very accurately [26]. To generate the data for the DML model an atomistic-continuum model (*Foliation model* [7]) is used. By using *Foliation model* we simulated $(5,5),(10,10),\cdots,(5N_w,5N_w)$ MWCNTs walls, with $N_w = 10$ to $N_w = 40$ in the increment of 5, where N_w is the number of walls. These simulations represents the deformation pattern of thick MWCNTs that have been seen in the experiments. These 7 simulations ($N_w = 10,15,\cdots,40$) are used in training.

Kinematics of MWCNTs and Data Preparation

Following the iso-geometric analysis framework, a local reference configuration is defined through an injective map (φ_0) from the parametric space. Similarly, to define the deformed configuration another map (φ) from parametric space to Euclidean space, $\mathbb{R}^3$, is defined, as shown in Fig. 2. Hence the deformation map can be expressed as $\Phi = \varphi \circ \varphi_0^{-1}$. Having one simulation data for one MWCNT is not sufficient to learn the features, hence it is an usual practice to run multiple instances. In the proposed work to reduce the number of simulations we decomposed the domain obtained from the simulation into several cross-sections (N_{cs}) taken at regular intervals along its length. The intrinsic periodicity in the deformations of MWCNTs along the length helps this decomposition starategy to increase the size of data set. Following the decomposition technique, the total deformation of each point (ξ_1, ξ_2, ξ_3) for m-th wall at n-th cross-section of a MWCNT can be represented through two parts, (i) an in-plane radial deformation $r(\theta(\xi_1), \xi_2^n, \xi_3^m)$ in the undeformed cross-sectional plane, and (ii) axial deformation $\Phi_2((\theta(\xi_1), \xi_2^n, \xi_3^m))$.

Dimensionality Reduction through Proposed Constrained-FPCA

The dataset $\{r_i(\theta)\}_{i=1}^N$ represents the radial deformation of the deformed configuration of MWCNT cross-sections, where the cross-sections of the MWCNTs are given by the mapped (ξ_1, ξ_3) planes for different ξ_2. Energy of deformed MWCNT is a function of curvature of its walls. Therefore, to ensure smoothness of the deformation map, we assume that the radial deformations of each tube are sampled from a stochastic process $R(\theta)$, $\theta \in \mathscr{T} = (0, 2\pi)$, such that its second derivative is square-integrable. We suppose that $R(\theta)$ can take any of the values $r_i(\theta) \in \mathcal{H}^2(\mathscr{T}), i = 1, \cdots, N$, such that $g(r_i(\theta)) = 0$. Where $g(r_i(\theta)) = 0$, is a geometric constraint on the deformation of MWCNTs and $\mathcal{H}^2(\mathscr{T})$ is Hilbert space. We denote the $L^2(\mathscr{T})$ inner product of functions ϕ_i, $\phi_j \in \mathcal{H}^2(\mathscr{T})$ with $\langle \phi_i, \phi_j \rangle$: $\int_{\mathscr{T}} \phi_i(\theta)\, \phi_j(\theta)\, d\theta$.

Let the mean and the covariance functions of $R(\theta)$ be denoted by $\mu(\theta)$ and $v(\theta, \vartheta) = \mathrm{Cov}(R(\theta), R(\vartheta))$. Invoking the Kosambi-Karhunen-Loève Expansion theorem [24, 25], the centered process can be expressed as $R(\theta) - \mu(\theta) = \sum_{k=1}^{\infty} \bar{c}_k\, \psi_k(\theta)$. Here, $\bar{c}_k = \langle (R(\theta) - \mu(\theta)), \psi(\theta) \rangle$. Where $\psi_k(\theta), k = 1, 2, \cdots$, are the orthonormal eigenfunctions of the following eigenvalue problem $\int_{\mathscr{T}} v(\theta, \vartheta)\, \psi(\theta) d\vartheta = \lambda\, \psi(\theta)$.

These eigenfunctions, $\psi_k(\theta)$, are referred to as functional principal components (functional-PCs) hereafter. Assuming a finite set of eigenfunctions is sufficient to approximate the centered stochastic process, $R(\theta) - \mu(\theta)$, its i-th sample can be written as

$$r_i(\theta) - \mu(\theta) \approx \sum_{k=1}^{K} \bar{c}_{ik}\, \psi_k(\theta),\ i = 1, \cdots, N \tag{1}$$

Interpretation of eigenfunctions: The k-th eigenfunction ψ_k is the k-th most dominant mode of variation orthogonal to $\{\psi_i\}_{i=1}^{k-1}$. Solving an eigenvalue problem in $\mathcal{H}^2(\mathscr{T})$ while satisfying a constraint can be a difficult task, therefore, we choose a convenient finite-dimensional basis and look for solutions in terms of that predefined basis. In case of deformed MWCNTs, deformation configurations of MWCNTs have geometric constraints that need to be satisfied by the eigenfunctions and hence also needs to be satisfied by the basis. Therefore, any arbitrary basis will not work. FPCA with an arbitrary basis will not satisfy the constraint of MWCNT and would not work. We choose a basis $\mathscr{B} = \left\{\phi_k \in \mathcal{H}^2(\mathscr{T}), g(\phi_k) = 0, k = 1, \cdots, K\right\}$. We encode the geometric constraint (periodicity constraint) of the deformation of MWCNTs via the function $g(\phi_k) = 0$, which is crucially important and specializes the FPCA for the present system. We call this novel technique constrained-FPCA (c-FPCA).

We rewrite the data set, $\{r_i(\theta)\}_{i=1}^N$, the eigenfunction $\psi(\vartheta)$, and the covariance function $v(\vartheta, \theta)$ in terms of the basis $\mathscr{B}$ and solve the aforementioned eigenvalue problem to obtain the functional-PCs, $\psi_k(\vartheta)$. The function $r_i(\theta)$ is represented in terms of functional-PCs using the Eq. 1 and their corresponding coefficients ($\bar{c}_{ik}$), that are referred here as coefficients of functional-PCs (CoFPCs). The dimension of the problem is significantly reduced by obtaining a K (number of functional-PCs) much smaller than the size of the discretized $r_i(\theta)$.

Learning in the Latent Dimension through Deep Neural Networks

Deep Neural Networks (DNNs) are used here to map the MWCNT geometry parameters and Boundary conditions to its deformation in the latent dimension (CoFPCs). The 4 Inputs for the proposed DNN are: Geometry parameters (i) total number of walls in the MWCNT (N_w), (ii) the wall number $(m, m = 1, \cdots, N_w)$, and (iii) the length coordinate $(\Phi_2(\xi_2^n), n = 1, \cdots, N_{\mathrm{cs}})$; (iv) Boundary Conditions: Angle of twist (Θ) or Curvature (κ), per unit length. The dimension of the output layer is the number of CoFPCs(K), which is decided based on the percentage of variation captured by Functional-PCs.

DNN($\mathscr{N}$) [8] is a function of weights $\bar{\mathbf{w}}$ and biases β that maps features x_i to labels y_i. The objective is to update the weights $\mathbf{w}$ and biases β in order to minimize the dissimilarity between true labels y and predicted labels $\hat{y}$, defined by a cost function $\mathscr{J}(y, \hat{y})$:

$$\min_{\bar{\mathbf{w}}} \quad \mathscr{J}(y, \mathscr{N}(x_i, \bar{\mathbf{w}})) \tag{2}$$

Updating the weights $\bar{\mathbf{w}}$ and biases β is done iteratively by using Stochastic Gradient Descent (SGD) and Backpropagation algorithm [27]. Once trained, the network can be used to predict outputs $y_i = \mathscr{N}(x_i, \bar{\mathbf{w}})$ for given new inputs x_i by using forward propagation technique.

If not properly trained DNNs tend to work accurately for training data but fail to predict for testing data. This condition is termed as overfitting and hinders generalization (accurate prediction for new inputs). To alleviate the overfitting we have used multiple regularization [28,29] and normalization [30] strategies. Inputs are normalized using z-score normalization. Along with that, we use early stopping to cease training when validation loss starts to increase. We used elastic net regularization which is a combination of L_1 and L_2 regularization. Details of these techniques are as follows,

Early Stopping: Number of epochs used for training purposes was set at a maximum limit of 2000, however, early stopping was used with a "patience" of 100 epochs to avoid overfitting of the network [28, 31]. Use of early stopping ensured generalization performance of our model but to further improve performance for test data we used naive elastic net regularization [29].

Elastic Net Regularization: A combination of $\mathscr{L}_1$ and $\mathscr{L}_2$ regularization called elastic net regularization [29] is used, which overcomes the individual drawbacks of $\mathscr{L}_1$ and $\mathscr{L}_2$ regularization. The cost function including $\mathscr{L}_1$ and $\mathscr{L}_2$ regularization can be written as

$$\tilde{\mathscr{J}}(y,\mathscr{N}(x_i,\bar{\mathbf{w}})) = \mathscr{J}(y,\mathscr{N}(x_i,\bar{\mathbf{w}})) + \alpha\Omega(\bar{\mathbf{w}})\ \alpha \in \mathbb{R}$$

Where $\Omega(\bar{\mathbf{w}}) = ||\bar{\mathbf{w}}||_{\mathscr{L}_1}$ and $\Omega(\bar{\mathbf{w}}) = ||\bar{\mathbf{w}}||_{\mathscr{L}_2}$ for $\mathscr{L}_1$ and $\mathscr{L}_2$ regularization respectively. The cost function for elastic net regularization can be written as

$$\begin{aligned}\tilde{\mathscr{J}}(y,\mathscr{N}(x_i,\bar{\mathbf{w}})) = \mathscr{J}(y,\mathscr{N}(x_i,\bar{\mathbf{w}})) &+ \alpha_1\Omega(||\bar{\mathbf{w}}||_1) \\ &+ \alpha_2\Omega(||\bar{\mathbf{w}}||_2) \\ &\alpha_1,\alpha_2 \in \mathbb{R}\end{aligned}$$

We also performed a resampling procedure of k-fold cross-validation on the data set to validate the performance of the model on independent fractions of the data set [32].

The DNN architectures used in our work consist of approximately 40 thousand learning parameters, 6 hidden layers of varying sizes between 64 to 128 neurons each. The activation function used was ReLU. The optimizer used was Adam optimizer. We trained three DNNs for predicting the following deformations of MWCNTs: (i) In-plane deformation under torsion, (ii) in-plane, and (ii) out-of-plane (axial) deformation under bending. Unlike torsion, in bending the axial deformation is not negligible, hence we have used two DNNs for in-plane and axial deformations.

RESULTS

Dimensionality Reduction

The proposed c-FPCA reduces the dimension of the deformation, moreover it also extracts the most dominant modes of deformation present in the data set. Based on the accuracy required the number of functional-PCs can be chosen. In order to achieve the reduced dimension we started with 64 basis functions to represent data set where each data set size was up to several hundred. From those 64 basis function, only 16 and 6 functional-PCs are required to capture 99.9% variability for torsion and bending respectively and 14 and 4 functional-PCs are required to capture 99% variability for torsion and bending respectively, as shown in Fig. 3. This demonstrates up to two orders of magnitude dimensionality reduction via the current method. This dimensionality reduction with high-accuracy of c-FPCA improves the accuracy of DNNs as it need to learn in significantly reduced dimensions.

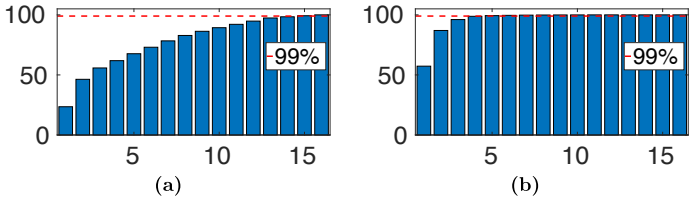

FIGURE 3. CUMULATIVE % VARIANCE CAPTURED BY PRINCIPAL COMPONENTS FOR MWCNTs UNDER (a) TORSION AND (b) BENDING [18].

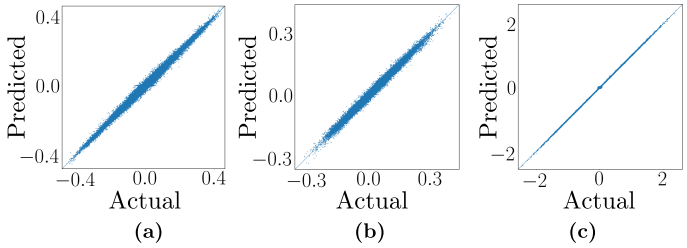

FIGURE 4. CORRELATION PLOTS FOR TEST SET CoFPCS OF (a) IN-PLANE DEFORMATION IN TORSION, (b) IN-PLANE AND (c) OUT-OF-PLANE DEFORMATION IN BENDING. THE CORRESPONDING CORRELATION VALUES ARE 0.9943, 0.9931 AND 0.9991 RESPECTIVELY [18].

Accuracy

Once the dimensionality reduction is performed on the data set, the prediction by the DNN becomes very accurate for a given

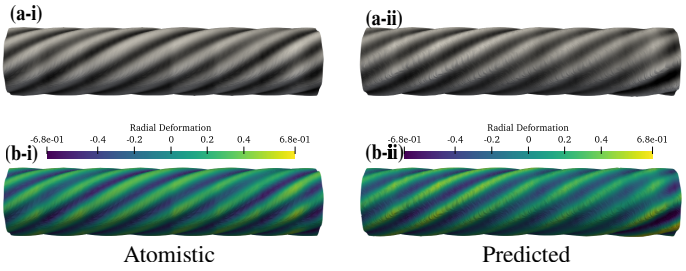

FIGURE 5. (a-i, a-ii) DEFORMATION PATTERN OBTAINED FOR 30 WALLED CNT UNDER TWISTING TWISTED VIA AC (LEFT) AND DML (RIGHT) MODEL. (b-i, b-ii) COLORMAP REPRESENTING THE RADIAL DEFORMATION. (c-i, c-ii) REPRESENTS THE CROSS-SECTION COMPARISON ALONG THE LENGTH AND DEMONSTRATES THAT THE DML MODEL IS ABLE TO PREDICT THE WAVE NUMBER ALONG THE LENGTH VERY ACCURATELY.

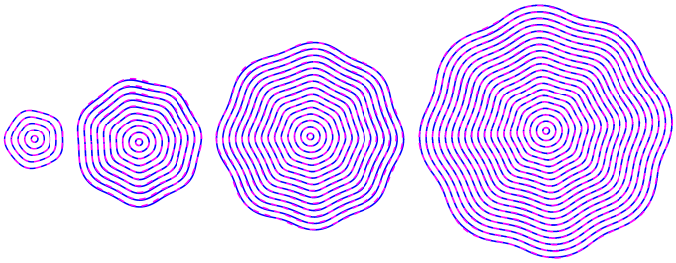

FIGURE 6. CROSS SECTION OF TWISTED 10, 20, 30 AND 40 WALLED TUBE OBTAINED USING ATOMISTIC-CONTINUUM SIMULATION(——) AND DML MODEL (——). ONLY ALTERNATE WALLS ARE SHOWN FOR CLARITY.

MWCNT and loading. The DNN predicts the CoFPCS which lie in the latent space. Through CoFPCS the cross-sections of individual walls are obtained through inverse of c-FPCA. Similarly the cross-section is obtained by adding all the walls. These cross-sections are then concatenated along the length to obtain the 3D deformation pattern of MWCNTs. The accuracy in prediction of CoFPCS via DNN is of utmost important. To make the DNNs predict accurately we have adopted the following strategies: (i) regularization techniques, and (ii) features-normalization.

The high accuracy of DNNs is demonstrated through excellent correlation coefficient ($R > 0.993$) for the test data as shown in Fig. 4 and very low relative-mean squared error (order of 10^{-4}) for the validation data.

We compared the deformation pattern obtained from DML model with the atomistic-continuum model for two types of systems: (i) known MWCNTs (10, 15, 20, 25, 30, 35, 40 walled) but unknown loading, (ii) unknown MWCNTs and unknown loading. Comparison of the deformation patterns under torsion and bending obtained through AC and DML models are provided for

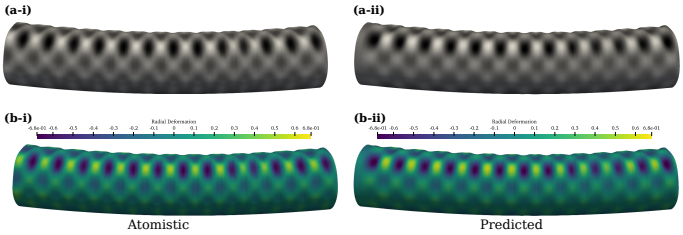

FIGURE 7. DEFORMATION PATTERN FOR 40 WALLED MWCNT UNDER BENDING OBTAINED VIA AC AND DML MODEL. THE CORRESPONDING COLORMAP REPRESENTING THE RADIAL DEFORMATION OBTAINED VIA AC AND DML MODEL.

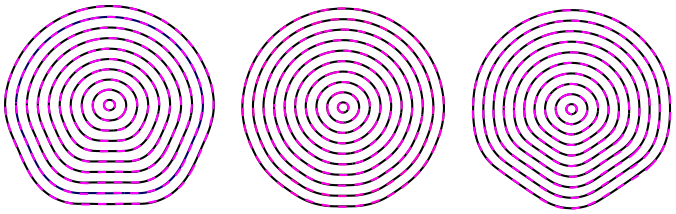

FIGURE 8. CROSS SECTIONs OF BENT 20 WALLED TUBE OBTAINED USING ATOMISTIC-CONTINUUM SIMULATION(——) AND DML MODEL (——). ONLY ALTERNATE WALLS ARE SHOWN FOR CLARITY.

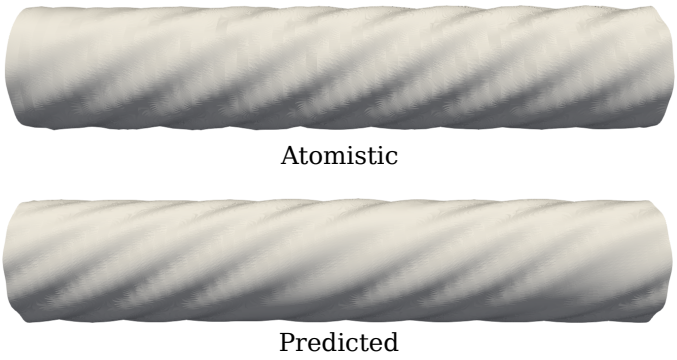

FIGURE 9. COMPARISON OF DEFORMATION CONFIGURATIONS OBTAINED VIA AC (TOP) AND DML (BOTTOM) MODELS FOR 27 WALLED CNT UNDER TORSION. THIS COMPARISON DONE FOR UNKNOWN MWCNT AND UNKNOWN LOADING CONDITION AS THE DML MODEL HAS NO INFORMATION ABOUT 27 WALLED CNT [18].

the known and unknown MWCNTs in Fig.5, Fig.7 and in Fig. 9, Fig. 10 respectively. The deformation obtained through the proposed DML model matches remarkably well with the AC model for unknown loading. The match of deformation is very accurate even for the unknown MWCNT (32 walled) and the loading (as

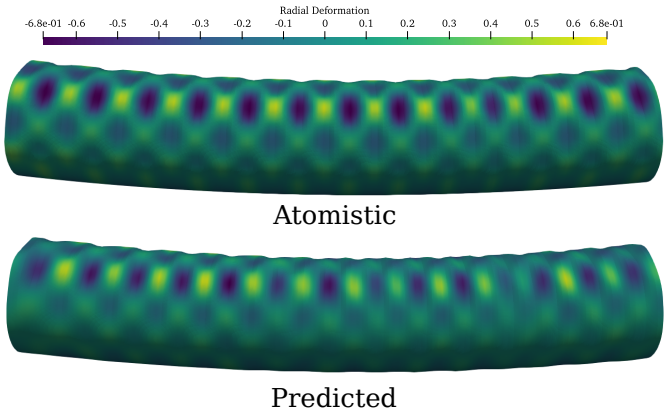

FIGURE 10. COMPARISON OF DEFORMATION CONFIGURATIONS OBTAINED VIA AC (TOP) AND DML (BOTTOM) MODELS FOR 31 WALLED CNT UNDER BENDING. THIS COMPARISON DONE FOR UNKNOWN MWCNT AND UNKNOWN LOADING CONDITION AS THE DML MODEL HAS NO INFORMATION ABOUT 31 WALLED CNT [18].

long as the unknown MWCNT is within the range of the training data). This obviates the need for AC simulations for such MWCNTs, yielding huge computational savings. However, if an MWCNT lies way outside the regime of the training data (10, 15, 20, 25, 30, 35, 40 walled CNTs) its accuracy might go down since the MWCNT might exhibit a deformation pattern that the DNN is not aware of as it doesn't occur in the training.

The maximum relative L_2 error is found to be $\approx 1\%$ for the 32-walled CNT (an unknown MWCNT under unknown loading).

The proposed model is significantly more computationally efficient than the AC model. Inference via the proposed model (upon training), requires only about ten seconds for an unknown MWCNT at all loading, whereas AC simulations for the same systems requires tens or hundreds of *total CPU* hours in parallel processing.

Discontinuity in Longitudinal Direction

As mentioned in the 'Kinematics of MWCNTs and Data Preparation' section we discretize the high dimensional MWCNT by taking several cross sections. Nonlinear dimensionality reduction technique, c-FPCA is used on the cross sections to obtain coFPCs that are then used as labels for the deep neural network. Once trained, we predict the coFPCs for new inputs and map them to cross sections (predicted) using inverse c-FPCA. We collocate these predicted cross sections along the longitudinal axis of tube to obtain deformed tube as shown in figure 11.

The smoothness of the predicted tube depends on accurate prediction of each and every cross section. Inaccurate prediction of any cross section will result in discontinuity along the length of the tube as shown in Fig. 12. Accurate prediction of cross

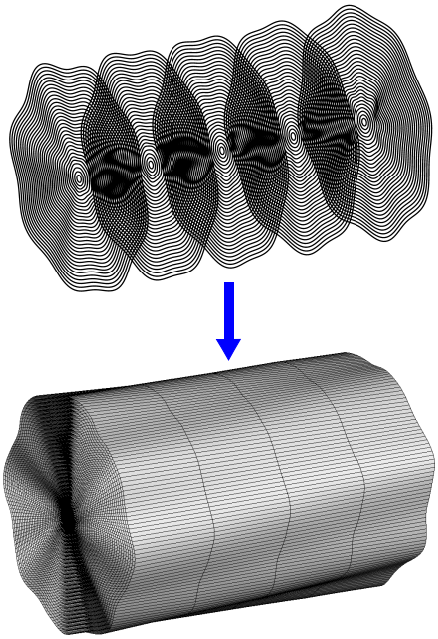

FIGURE 11. PREDICTED INDIVIDUAL CROSS SECTIONS *(TOP)* AND OBTAINED TUBE FROM THE PREDICTED CROSS SECTIONS *(BOTTOM)* . ONLY 5 CROSS SECTIONS ARE SHOWN HERE FOR SAKE OF SIMPLICITY. AXIS IS Y DIRECTION IS SCALED BY A FACTOR OF 5.

section has 2 attributes, (i) accurate crests and troughs prediction and (ii) accurate orientation w.r.t to succeeding and previous cross section. These requirements on the prediction of cross sections can limit the smoothness of predicted tube obtained using the proposed model. Addressing this challenge and eliminate these stringent requirements on cross section prediction is the future scope of this work.

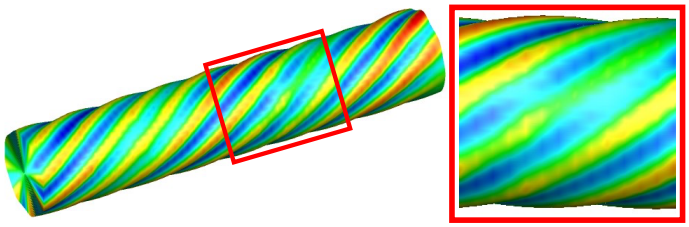

FIGURE 12. PRELIMINARY DML MODEL YIELDS UNDESIRED DISCONTINUITY ALONG THE LENGTH OF A TWISTED MWCNT. COMPARE AGAINST SMOOTH TWISTED MWCNT BY AC SIMULATION IN FIG. 5 , USED IN TRAINING. COLOR INDICATES RADIAL DEFORMATION, RED IS FOR OUTWARD BLUE IS FOR INWARD.

Latent Space Exploration

Herein, we explore the model-based-interpretability (as defined in [33]) of the proposed model by examining the principal modes and their coefficients present in the latent space. The proposed DML model is able to extract the dominant (principal) modes of deformation and their relative contribution in an unsupervised manner. Fig. 13 shows few principal components. The top row shows the principal components obtained for torsion. These principal components resemble with the rippling deformation that occur in MWCNT during torsional loading. Similarly, the bottom row in Fig. 13 represents the principal components for bending which resemble to the diamond buckling pattern. The similar pattern can be seen from the atomistic results too. These key features are obtained through constrained-Functional Principal Component Analysis. So far these key deformed patterns were approximately-identified manually for individual MWCNTs, whereas in the present work these principal components of deformation are automatically identified. The automatically identified most dominant modes of deformation shows qualitative similarity with those identified manually in [26, 6, 34]. These functional-PCs are universal since they are obtained from the entire data set. This fact enhances the model's predictive capability on unseen MWCNTs and hence explains the generalizability (performance for unseen MWCNTs) of the model. Once the dimensionality of the output is reduced, the DNNs also learn the reduced dimension and predicts the deformation very accurately. The proposed c-FPCA not only enables us to comprehend the principal modes of deformation but also increases the accuracy of DNN which leads to better and smooth prediction of high-dimensional deformation of MWCNTs.

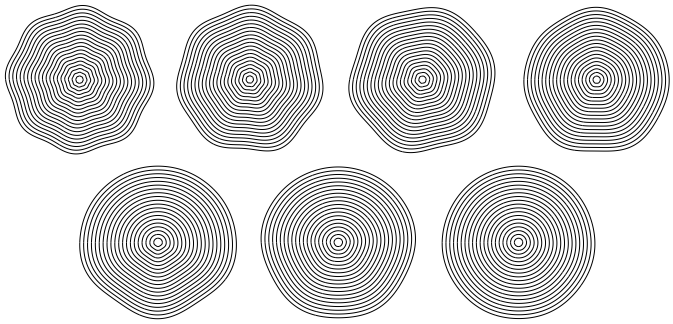

FIGURE 13. FUNCTIONAL PRINCIPAL COMPONENTS OF MWCNTS UNDER TORSION (TOP) AND BENDING (BOTTOM) OBTAINED FROM CONSTRAINED FUNCTIONAL PRINCIPAL COMPONENT ANALYSIS [18].

Computational Details

DNNs used in DML model are built and trained in Tensorflow 1.1.14 on Python 3.6. We used the Bridges facility available at Pittsburgh Supercomputing Center for training purposes. Bridges' GPU-AI partition consists of NVIDIA Tesla V100 GPU, which was used for training purposes. The time required per iteration was $\approx 12s$ and $\approx 8s$ for DNNs used for MWCNTs under torsion and bending respectively.

MWCNT simulations were performed using an in-house FORTRAN code which is based on foliation model [7]. We used XSEDE-Bridges supercomputer for MWCNT simulations which is equipped with Intel Haswell (E5-2695 v3) CPUs with 14 cores per CPU running at 2.3-3.3 GHz with 128GB RAM and 35MB cache. Postprocessing of the simulation data was performed on the same system. We used an in-house desktop system: 2 processors, 18-core Intel Xeon Gold 5220 CPU running at 2.20 GHz with 64 GB RAM, 24.65MB cache and 512 GB SATA SSD for i) Discretization of the data and other data preparation tasks ii) Non-Linear dimensionality reduction iii) Inference from the Deep Neural Network iv) Plotting the high dimensional deformation data. Discretization of the data was performed in MATLAB. Dimensionality reduction was done using an in-house non-linear dimensionality reduction code.

CONCLUSION AND DISCUSSIONS

In this study, a novel interpretable machine learning model is proposed, which predicts high-dimensional deformed configurations of MWCNTs accurately and efficiently using only 4 inputs. This work introduces a novel dimensionality reduction technique that extends FPCA to respect the constraints of MWCNT deformation exactly. The dimensionality reduction obtained through this not only improves accuracy in low-dimensional representation of deformation but also enables accurate prediction of high-dimensional deformation of MWCNTs. The dimensionalilty reduction technique enables us to represent the data in functional representation and hence allows the generability. This generability leads to remarkably accuracy for unknown MWCNTs and unknown loading. The proposed DML model completely eliminates the need of expensive AC simulations for MWCNTs (within the trainind data regime), yielding a massive gain in computational efficiency. The functional representation of the most dominant modes of deformation or the principal components are comprehensible and thus help to elucidate how the model predicts high-dimensional deformation through learning the space of functional-PCs, yielding interpretability [18].

ACKNOWLEDGMENT

The work was supported by the NSF (CMMI MoMS) under Grant No. 1937983. We acknowledge Superior, a high-performance computing facility at MTU. This work used the

Extreme Science and Engineering Discovery Envi- ronment (XSEDE), which is supported by NSF Grant No. ACI-1548562. This work used the XSEDE Bridges at the Pittsburgh Supercomputing Center through allocation No. MSS200004.

REFERENCES

[1] Williams, P., Papadakis, S., Patel, A., Falvo, M., Washburn, S., and Superfine, R.“, 2002.”. *Phys. Rev. Lett.,* ***89***, p. 255502.

[2] Papadakis, S. J., Hall, A. R., Williams, P. A., Vicci, L., Falvo, M. R., Superfine, R., and Washburn, S.“, 2004.”. *Phys. Rev. Lett.,* ***93***(14), p. 146101.

[3] De Volder, M. F. L., Tawfick, S. H., Baughman, R. H., and Hart, A. J.“, 2013.”. *Science,* ***339***(6119), pp. 535–539.

[4] Poncharal, P., Wang, Z. L., Ugarte, D., and de Heer, W. A.“, 1999.”. *Science,* ***283***, pp. 1513–1516.

[5] Arroyo, M., and Belytschko, T.“, 2002.”. *J. Mech. Phys. Solids,* ***50***(9), pp. 1941–1977.

[6] Arroyo, M., and Belytschko, T.“, 2004.”. *International Journal for Numerical Methods in Engineering,* ***59***(3), pp. 419–456.

[7] Ghosh, S., and Arroyo, M.“, 2013.”. *J. Mech. Phys. Solids,* ***61***(1), pp. 235–253.

[8] Lecun, Y., Bengio, Y., and Hinton, G.“, 2015.”. *Nature,* ***521***(7553), pp. 436–444.

[9] Hornik, K., Stinchcombe, M., and White, H.“, 1989.”. *Neural Networks,* ***2***(5), pp. 359–366.

[10] Xie, T., and Grossman, J. C.“, 2018.”. *Phys. Rev. Lett.,* ***120***, p. 145301.

[11] Iten, R., Metger, T., Wilming, H., del Rio, L., and Renner, R.“, 2020.”. *Phys. Rev. Lett.,* ***124***, p. 010508.

[12] Lu, L., Dao, M., Kumar, P., Ramamurty, U., Karniadakis, G. E., and Suresh, S.“, 2020.”. *Proc. Natl. Acad. Sci. U.S.A.,* ***117***(13), pp. 7052–7062.

[13] Arias, I., and Arroyo, M.“, 2008.”. *Phys. Rev. Lett.,* ***100***(085503).

[14] Hanakata, P. Z., Cubuk, E. D., Campbell, D. K., and Park, H. S.“, 2018.”. *Phys. Rev. Lett.,* ***121***(25), p. 255304.

[15] Hanakata, P. Z., Cubuk, E. D., Campbell, D. K., and Park, H. S.“, 2020.”. *arXiv preprint arXiv:2008.05298*.

[16] Gómez-Bombarelli, R., Wei, J. N., Duvenaud, D., Hernández-Lobato, J. M., Sánchez-Lengeling, B., Sheberla, D., Aguilera-Iparraguirre, J., Hirzel, T. D., Adams, R. P., and Aspuru-Guzik, A.“, 2018.”. *ACS central science,* ***4***(2), pp. 268–276.

[17] Balu, A., Nallagonda, S., Xu, F., Krishnamurthy, A., Hsu, M.-C., and Sarkar, S.“, 2019.”. *Scientific reports,* ***9***(1), pp. 1–12.

[18] Yadav, U., Pathrudkar, S., and Ghosh, S., 2021. “Interpretable machine learning model for the deformation of multiwalled carbon nanotubes”. *Physical Review B,* ***103***(3), pp. 1–7.

[19] Van Der Maaten, L., Postma, E., and Van den Herik, J., 2009. “Dimensionality reduction: a comparative review”. *J Mach Learn Res,* ***10***, pp. 66–71.

[20] Lee, J. A., and Verleysen, M., 2007. *Nonlinear Dimensionality Reduction*, 1st ed. Springer.

[21] McInnes, L., Healy, J., and Melville, J.“, 2018.”. *arXiv preprint arXiv:1802.03426*.

[22] Yao, F., Müller, H.-G., and Wang, J.-L.“, 2005.”. *Annals of Statistics,* ***33***(6), 12, pp. 2873–2903.

[23] Ramsay, J., and Silverman, B., 2005. *Functional Data Analysis*. Springer.

[24] Stark, H., and Woods, J. W., eds., 1986. *Probability, Random Processes, and Estimation Theory for Engineers*. Prentice-Hall, Inc., USA.

[25] Jorgensen, P., and Song, M.“, 2007.”. *J. of Mathematical Physics,* ***48***(10), p. 103503.

[26] Arroyo, M., and Belytschko, T.“, 2003.”. *Phys. Rev. Lett.,* ***91***(21), p. 215505.

[27] LeCun, Y., Touresky, D., Hinton, G., and Sejnowski, T., 1988. “A theoretical framework for back-propagation”. In Proceedings of the 1988 connectionist models summer school, Vol. 1, CMU, Pittsburgh, Pa: Morgan Kaufmann, pp. 21–28.

[28] Prechelt, L., 2012. *Early Stopping — But When?* Springer, pp. 53–67.

[29] Zou, H., and Hastie, T.“, 2005.”. *Journal of the royal statistical society: series B (statistical methodology),* ***67***(2), pp. 301–320.

[30] Jayalakshmi, T., and Santhakumaran, A.“, 2011.”. *Int. J. of Computer Theory and Engineering,* ***3***(1), pp. 1793–8201.

[31] Sarle, W. S.“, 1996.”. *Computing sci. and stats.*, pp. 352–360.

[32] Bengio, Y., and Grandvalet, Y.“, 2004.”. *J. machine learning research,* ***5***(Sep), pp. 1089–1105.

[33] Murdoch, J., Singh, C., Kumbier, K., Abbasi-Asl, R., and Yu, B.“, 2019.”. *Proc. Natl. Acad. Sci. U.S.A.,* ***116***(44), pp. 22071–22080.

[34] Zou, J., Huang, X., Arroyo, M., and Zhang, S.“, 2009.”. *Journal of applied physics,* ***105***(3), p. 033516.

