

Deep learning-based quasi-continuum theory for structure of confined fluids

Haiyi Wu¹ and N. R. Aluru^{1,2*}

1. Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin 78712, United States

2. Walker Department of Mechanical Engineering, The University of Texas at Austin, Austin 78712, United States

* Correspondence to aluru@utexas.edu

Abstract: Predicting the structural properties of water and simple fluids confined in nanometer scale pores and channels is essential in, for example, energy storage and biomolecular systems. Classical continuum theories fail to accurately capture the interfacial structure of fluids. In this work, we develop a deep learning-based quasi-continuum theory (DL-QT) to predict the concentration and potential profiles of a Lennard-Jones (LJ) fluid and water confined in a nano channel. The deep learning model is built based on a convolutional encoder-decoder network (CED) and is applied for high dimensional surrogate modeling to relate the fluid properties to the fluid-fluid potential. The CED model is then combined with the interatomic potential-based continuum theory to determine the concentration profiles of a confined LJ fluid and confined water. We show that the DL-QT model exhibits a robust predictive performance for a confined LJ fluid under various thermodynamic states and water confined in a nanochannel of different widths. The DL-QT model seamlessly connects the molecular physics at nanoscale with the continuum theory by using the deep learning model.

Keywords: convolutional encoder-decoder, fluid properties at nanoscale, deep learning-based quasi-continuum theory, interfacial structure, confinement, density, nanofluidics.

1. Introduction

Estimating the properties of confined water and Lennard-Jones (LJ) fluids in nanoscale channels and pores is fundamental to a wide range of engineering applications including, e.g., nanofiltration,¹⁻³ natural gas extraction from shale reservoirs,^{4, 5} and charging/discharging of Li-air batteries.⁶ Classical continuum theories fail to account for the atomic details at the solid-fluid interface. Hence, they are not accurate in predicting fluid properties in nanoscale confinement where the interfacial behavior can deviate significantly from bulk behavior. To overcome this difficulty, many theoretical and computational approaches have been developed.^{7, 8} For example, atomistic molecular dynamics (MD) simulations have been extensively used to investigate the structure, thermodynamic and dynamic properties of fluids confined at the nanoscale.⁹ The computational efficiency of MD simulations, however, depends on the critical length and time scales of the problem, i.e., MD simulations can be very expensive for many applications of practical interest, even with advanced computing resources. Therefore, it is of fundamental interest to develop multiscale models incorporating atomic physics into continuum theories.

Over the years, several multiscale models have been proposed to understand the fluid properties under confinement.¹⁰⁻¹⁴ For example, the classical density function theory (cDFT)¹⁵ has been used to study fluid properties under confinement by minimizing the grand potential. The cDFT suffers from the drawback that the precise relation between the Helmholtz free energy and the molecular density for many systems of practical interest is unclear.¹¹ An alternative multiscale approach to study fluids under confinement is the empirical potential based quasi-continuum theory (EQT) to predict the fluid density and potential in confinement.^{10, 12, 16} EQT was developed by incorporating the interatomic potential (or the coarse-grained potential) used in molecular dynamics into the continuum theory given by the Nernst-Planck equation. EQT has been used to predict the properties and potential of Lennard Jones (LJ) fluids, carbon dioxide and water in a nanochannel.^{10, 12, 17-19} EQT framework has also been extended to predict the structure of the electrical double-layer including charge inversion phenomena.^{16, 20} Though EQT has shown promise in estimating fluid density in confinement, it requires a precise mathematical expression for the wall-fluid and fluid-fluid potentials which can be complicated even for a simple LJ fluid.¹² In general, the exact formula of the fluid-fluid potential is unknown due to the complicated molecular interactions. Consequently, there is a need to develop new models and theories that not only can accurately approximate the fluid-fluid potential without knowing its exact potential formula but also can serve as a “bridge” that integrates the atomic details into the continuum theory.

Recently, deep learning-based methods have received enormous attention in scientific disciplines and have shown promise to be a powerful tool in a variety of fields.²¹⁻³² For instance,

deep learning methods have been used to build novel models for designing small-molecular organic structures²⁶ and for discovering new materials.^{22, 27-29} The deep learning-based models can reduce the time to deployment of new materials and lower the cost associated with the initial material discovery.^{27, 29} The deep learning assisted framework has also shown promise in solving the inverse problem of the liquid-state theory (DeepILST) to parameterize the standard 12-6 LJ potential.²⁵ In addition, deep learning-based surrogate models have been used to predict the flow transport and quantify the uncertainty of flow in stochastic media.^{21, 23, 24} However, direct combination of a deep learning-based surrogate model into the continuum theory faces many challenges. For example, the deep learning model combined with the Navier-Stokes (NS) equations can lead to a complex landscape of the loss function and make the training process more difficult.^{33, 34} In addition, new computational algorithms may be needed to bridge the different scales between deep learning model and continuum system.^{35, 36}

In this work, we build a deep learning-based quasi-continuum theory (DL-QT) to predict the concentration and potential profiles of a Lennard-Jones (LJ) fluid and water confined in a nano slit. Our first goal is to develop a deep learning-based surrogate model to extract important atomic features from MD data. Specifically, the surrogate model is trained to relate the fluid properties with the fluid-fluid potential. Because of the inhomogeneity imposed by the nanochannel, the input data (fluid properties) and output data (fluid potential) can be high dimensional. Hence, the deep learning model is built based on end-to-end regression by using the convolutional encoder-decoder network. The second goal is to integrate the atomic details extracted from the deep learning model into the continuum theory to compute accurate fluid properties in a nanochannel. The deep learning-based surrogate model serves as the connection between the molecular scale and continuum scale. We also show that the DL-QT model can predict the properties of fluids with more complex molecular structures like water.

The rest of the paper is structured as follows: In section 2, we describe the details of data generation using MD simulations, the convolutional encoder-decoder network and the computational framework and method for solving the DL-QT model. In section 3, the performance of the proposed computational framework is examined and discussed. We first quantify the performance of the convolutional encoder-decoder network model. Then, we investigate the performance of DL-QT model in predicting the concentration and potential profiles of a LJ fluid and water confined in nano channels. In section 4, we summarize the key findings of this study.

2. Computational Framework

In this section, we discuss the deep learning-based computational framework for predicting the fluid concentration and potential profiles in nano confinement. In the first part, we detail the data

generation method for the deep learning model by using the molecular dynamics simulation. In the second part, the architecture and parameterization of the convolutional encoder-decoder model is presented. Finally, the computational framework and algorithm of the deep learning-based quasi-continuum theory (DL-QT) are summarized.

2.1 Generation of dataset

The dataset required for training, validation, and testing are generated via MD simulation. The MD systems are built with fluid sandwiched between two graphene layers. We study two types of fluids, namely, LJ fluid and water. For the graphene – LJ fluid system, we built 15012 distinct MD simulations with graphene channel widths of $L = 8 \text{ nm}, 10 \text{ nm}, 12 \text{ nm}, \dots, 28 \text{ nm}, 30 \text{ nm}$. Each channel width contains 3 different temperatures, $T = 303 \text{ K}, 323 \text{ K}, 353 \text{ K}$, and each temperature contains 416 different MD simulations with thermodynamic states sampled uniformly for bulk density ranging from $1.0 - 29.0 \text{ nm}^{-3}$. The lateral dimensions of the graphene layer are $8.0 \times 8.0 \text{ nm}^2$. For the graphene–water system, we built 240 distinct MD simulations with graphene channel widths of $L = 2.0, 2.025, 2.05, \dots, 7.95, 7.975 \text{ nm}$. The lateral dimensions of the graphene layer are $4.5 \times 4.5 \text{ nm}^2$. To ensure that the confined water in all channel widths corresponds to the standard thermodynamic state, i.e., $T = 300 \text{ K}$ and bulk density is $\rho_{\text{bulk}} = 1.0 \frac{\text{g}}{\text{cm}^3}$, we performed equilibrium simulations with the graphene channel system attached to the bulk water reservoir at 1 bar pressure and 300 K temperature.

In the MD simulation, the intermolecular interactions between an atom “i” and atom “j” are modeled using the Lennard-Jones (LJ) potential,

$$\varphi_{ij} = 4\varepsilon_{ij} \left(\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right) \quad (1)$$

where, ε_{ij} and σ_{ij} are the LJ parameters for the pair (i, j) and r_{ij} is the distance between the two atoms. Water is simulated using the extended simple point charge (SPC/E)³⁷ model. The LJ parameters for graphene – LJ fluid and graphene – water systems are summarized in Table 1. The short range interactions are calculated using a spherical cutoff of 1.4 nm, and water–water electrostatic interactions are calculated using the particle mesh Ewald (PME) method with a desired relative error of 1×10^{-5} in forces.³⁸

All the MD simulations are performed using the LAMMPS code³⁹ with a time step size of 1 fs. Periodic boundary conditions are applied in the x and y-directions. In the z-direction, slab correction is adopted for the PME method⁴⁰, where an extra vacuum space of three times the actual z dimension is inserted to effectively turn off the slab–slab interactions. During the simulation, the graphite wall is frozen; that is, their positions were fixed. The simulations are performed in the NVT (canonical) ensemble, and the temperature of the confined fluid is maintained using the Nosé-

Hoover thermostat⁴¹. Each MD simulation is performed for 2 ns and the fluid concentration and potential are calculated on the fly.

The generated MD dataset with the same channel and temperature are indexed based on their bulk density (index increases as bulk density increases). We then randomly divide these indexes into three groups: index for training data (80%), index for validation data (10%), and index for testing data (10%). We note that the testing data sets include various combinations of density, temperature and channel widths that are not seen during the training phase, but it does not include specific temperatures or channel width that are not part of the training sets. The channel width, confined fluid concentration, together with its corresponding thermodynamic state and fluid LJ parameters, are stored as a feature vector which serves as the input information to the machine learning model. The fluid-fluid potential of the confined fluid constitutes the output of the machine learning model.

2.2 Deep convolutional encoder-decoder network

Mathematically, the relation between fluid-fluid interactions and the confined fluid concentration can be expressed as:

$$U_{ff}(\mathbf{z}) = \mathcal{G}(\rho(\mathbf{z}), T, \rho_{bulk}, \varepsilon_{ff}, \sigma_{ff}, L) \quad (2)$$

where $U_{ff}(\mathbf{z})$ is the fluid-fluid potential, $\rho(\mathbf{z})$ is the fluid concentration, ε_{ff} and σ_{ff} are the LJ parameters for the fluid-fluid interaction, L is the channel width, T and ρ_{bulk} represent the temperature and bulk density, respectively, and $\mathcal{G}$ is the mathematical mapping. Due to the inhomogeneity imposed by the nanochannel and the many-body nature of fluid-fluid interactions, it is generally difficult and computationally expensive to develop an analytical mapping. To tackle this limitation, we adopt a deep learning model to approximate the above mapping, i.e.,

$$\tilde{U}_{ff}(\mathbf{z}) = \mathcal{H}(\rho(\mathbf{z}), T, \rho_{bulk}, \varepsilon_{ff}, \sigma_{ff}, L | \theta) \quad (3)$$

where θ represents the learning parameters of the model. The model $\mathcal{H}$ is trained via N sets of data $\mathcal{D} = \{U_{ff,i}(\mathbf{z}), (\rho_i(\mathbf{z}), T_i, \rho_{bulk,i}, \varepsilon_{ff}, \sigma_{ff})\}_{i=1}^N$ through minimization of the loss function,

$$\mathcal{H} = \arg \min_{\theta} (\text{loss}(\mathcal{H}(\rho(\mathbf{z}), T, \rho_{bulk}, \varepsilon_{ff}, \sigma_{ff}, L | \theta), U_{ff}(\mathbf{z}))) \quad (4)$$

In this work, we adopt the deep convolutional encoder-decoder network (CED) model to relate the concentration of the confined fluid with the fluid-fluid potential. The CED model has shown promise in dealing with the high dimensional to high dimensional regression problem.^{24, 42} In this work, because of the inhomogeneity imposed by the nanochannel, the input (fluid properties) and the output (fluid potential) of the deep learning model are both high dimensional data with size of 1×1005 and 1×1000 , respectively.

The CED model is constructed with three main parts, namely, an encoder network, a latent space, and a decoder network. Here the encoder part is constructed using a deep convolutional neural network (CNN). It takes the concentration $\rho(\mathbf{z})$ from MD data, the channel width L , and the LJ parameters ($\varepsilon_{ff}, \sigma_{ff}$) of the confined fluid, and its corresponding thermodynamic state (T, ρ_{bulk}) as input and extract the important correlations and features from the input. The extracted feature maps are stored in the latent space. The decoder part, constructed with transposed CNN layers, takes the information from the latent space and up-samples these feature maps to high-dimensional space to predict the fluid-fluid potential profile.

Before we discuss the CED model, we first briefly review the basic knowledge of CNN. Here we only outline the key ideas of these networks; a detailed introduction of the CNN model and the data workflow can be found in Ref⁴³. CNN model is designed based on shared-weight architecture. The output feature maps are calculated based on convolutional filters sliding along the input features. For example, the convolutional filters at the k^{th} layer are applied to compute the output (feature maps $F \cdot map_{conv}^k$) at this layer via convolution operation with the input x_i at the k^{th} layer, i.e.,

$$F \cdot map_{conv}^k = \mathcal{A}(w^k \oplus x_i + b^k \cdot J) \quad (5)$$

where $\mathcal{A}$ is the nonlinear activation function, J is a tensor with all elements equal to 1. w^k and b^k are the weights and bias of the filter at the k^{th} layer. $w^k \oplus x_i$ is the convolution between the filters' weight and the input. The exact formula relies on the padding (p) and strides (st) and can be found in Ref⁴⁴.

The architecture of our CED model is shown in Fig. 1a. The encoder network (blue color) is constructed with six encoder blocks. Each encoder block consists of four parts as presented in Fig. 1b: (1) a 1×3 convolutional layer with same padding and $st = 1$; (2) the convolutional layer is followed by a tanh nonlinear activation function; (3) the output is then down-sampled via a 1×2 maxpooling layer with $st = 2$; (4) finally a dropout layer is applied before feeding the data to the next encoder block. The decoder network (green color) consists of six decoder blocks, with each block structured in five parts (Fig. 1c): (1) a 1×3 transposed convolution layer with the same padding and $st = 2$; (2) followed by a tanh activation function; (3) and a dropout layer; (4) the output from the dropout layer is fed to a regular convolutional layer; and (5) followed by another tanh activation function. The encoder and decoder parts are connected using a latent space z , consisting of one convolutional layer. The output of the decoder network will be fed to a fully connected neural network layer to generate the final prediction (fluid-fluid potential).

The CED model is trained using the training dataset generated in section 2.1 by minimizing the loss function:

$$loss = \frac{1}{N} \sum_{i=1}^N |U_{ff,i} - \tilde{U}_{ff,i}| \quad (6)$$

where N is the number of training data, and $U_{ff,i}$ is the ground-truth fluid-fluid potential of sample i , computed from the MD simulation. $\tilde{U}_{ff,i}$ is the prediction by the CED model. The CED model is implemented using Tensorflow⁴⁵ and is trained on an NVIDIA V100 (“Volta”) GPU with 32 GB memory. We initialize the weights of all the convolutional filters with a truncated normal distribution with a standard deviation of 0.1. The initial value of all biases is set to 0.1. We adopt Adam Optimizer⁴⁶ with a decay learning rate (start with 8×10^{-4} and decrease by 20% every 600 epochs) to minimize the corresponding loss function in Eq. (6). The model is trained using a batch size of 32 and the training process stops at 15000 steps when the validation error is minimized. Once the model is well trained, the optimized learning parameters at the final step are saved and will then be restored for evaluating the testing dataset without retraining the model.

2.3 Deep learning-based quasi-continuum theory (DL – QT)

The equilibrium concentration profile $\rho(z)$, of a confined fluid in a nanochannel with channel width L , can be determined by solving the 1-D steady state Nernst-Planck equation,

$$\frac{d}{dz} \left(\frac{d\rho(z)}{dz} + \frac{\rho(z)}{RT} \frac{dU}{dz} \right) = 0 \quad (7)$$

with boundary conditions:

$$\rho(z = 0) = 0 \quad (8a)$$

$$\rho(z = L) = 0 \quad (8b)$$

$$\frac{1}{L} \int_0^L \rho(z) dz = \rho_{ave} \quad (8c)$$

where T is the fluid temperature, R is the ideal gas constant, U is the total potential of mean force (PMF), and ρ_{ave} is the average concentration of fluid inside the channel. If we know the total PMF, $U(z)$, the fluid concentration can be solved with Eq. (7) together with boundary conditions in Eq. (8a), (8b) and (8c).

A. Model for total PMF $U(z)$

For a simple fluid in nanoconfinement, the total potential is due to the wall-fluid and fluid-fluid interactions. Thus, $U(z)$ can be calculated by the summation of the wall-fluid and fluid-fluid potential, i.e.,

$$U(z) = U_{wf}(z) + U_{ff}(z) \quad (9)$$

The wall-fluid potential. Using a continuum approximation, the wall-fluid potential can be computed by integrating the contribution of wall atoms with a continuous distribution⁴⁷

$$U_{wf}(z) = \int \varphi_{wf}(z - r) c_{wall} dV \quad (10)$$

where φ_{wf} is the wall-fluid pair interaction, c_{wall} is the wall atom density in a continuum approximation. For a graphene layer, $c_{wall} = 38.287 \text{ nm}^{-2}$ and the wall-fluid potential can be calculated analytically¹²

$$U_{wf}(z) = \frac{4}{5}\pi c_{wall}\varepsilon_{wf}\sigma_{wf}^{12}\left(\frac{1}{z^{10}} + \frac{1}{(L-z)^{10}}\right) - 2\pi c_{wall}\varepsilon_{wf}\sigma_{wf}^6\left(\frac{1}{z^4} + \frac{1}{(L-z)^4}\right) \quad (11)$$

where ε_{wf} and σ_{wf} are the pair interaction parameters for wall-fluid atoms.

The fluid-fluid potential. The exact formula for the fluid-fluid potential, however, can be extremely complicated. Here we adopt the convolutional encoder-decoder network to approximate the relation between fluid-fluid potential with the fluid concentration,

$$U_{ff}(z) \approx \tilde{U}_{ff}(z) = \mathcal{H}(\rho(z), T, \rho_{bulk}, \varepsilon_{ff}, \sigma_{ff}, L | \theta) \quad (12)$$

where $\tilde{U}_{ff}(z)$ is the fluid-fluid potential predicted by the well trained deep learning model.

In summary, we developed a deep learning-based quasi-continuum theory to determine the concentration of fluid confined in a nanochannel as defined in Eq. (7) with boundary conditions given by Eq. (8). The key idea is to obtain the total PMF which has contributions from wall-fluid interactions in Eq. (11) and the fluid-fluid interactions in Eq. (12). Notice that the fluid-fluid potential $U_{ff}(z)$ in Eq. (12) is a function of concentration and Eq. (7) is nonlinear. We use an iterative scheme to find the self-consistent concentration and potential.

B. Implementation

The computational framework and implementation of the DL-QT method are summarized in figure 2. We first compute the wall-fluid potential $U_{wf}(z)$ in Eq. (11), by taking the wall-fluid LJ parameters, ε_{wf} and σ_{wf} , channel width, L , the concentration of wall atoms, c_{wall} as input. Since the wall-fluid potential does not depend on the fluid concentration, we only need to compute $U_{wf}(z)$ once. We then calculate the initial fluid concentration,

$$\rho^{n=0}(z) = \rho_{bulk} \exp\left(-\frac{U_{wf}(z)}{RT}\right) \quad (13)$$

The fluid-fluid potential $U_{ff}(z)$ in Eq. (12) is computed using the initial fluid concentration as input to the well trained CED model. With the wall-fluid potential and fluid-fluid potential, the total potential can be determined using Eq. 9. Considering the nonlinearity of Equation 7, we apply a relaxation factor when calculating the new concentration in the next iteration. In this paper, we use $\alpha = 1.3$. The iterative process is repeated until the solver converges to a self-consistent solution.

3. Results and Discussion

In this section, we examine the performance of our deep learning-based quasi-continuum theory in determining the concentration and potential of a fluid confined in a nanochannel. The DL-QT model is applied to solve two systems, namely, LJ fluid confined in a graphene channel and water confined in a graphene channel. Below, we first show that our deep learning model can be applied as a surrogate model to relate the fluid-fluid interactions with the fluid concentration and its corresponding thermodynamic state without resolving the complex mathematical formula. We show that the deep learning-based surrogate model performs well in predicting the fluid-fluid potential for both confined LJ fluid and water. Next, we demonstrate that the well-trained deep learning model can be seamlessly combined with the quasi-continuum theory to determine the concentration of fluid in a nanochannel. The DL-QT model has a robust predictive performance for LJ fluid confined in different channel widths and under various thermodynamic states. Finally, we demonstrate that the DL-QT model can be extended to solve the concentration of water which has a more complex molecular structure.

3.1 Performance of CED model in predicting the fluid-fluid potential

The CED model is trained using the dataset generated in section 2.1 by following the protocol described in section 2.2. The evolution of the loss function in Eq. (6) over the training process is shown in Fig. 3. We observe that both the training loss and validation loss decrease sharply in the first ~ 1000 epochs and the training loss decreases faster than the validation loss. The training process stops at 15000 epochs when the training and validation loss are no longer decreasing. The optimized learning parameters at the final epoch are saved and will be restored for evaluating the testing dataset in the next step.

We then test the well-trained CED model using the testing dataset. We note that the testing data sets include various combinations of density, temperature and channel widths that are not seen during the training phase, but it does not include specific temperatures or channel width that are not part of the training sets. For the LJ fluid sandwiched between two graphene walls, the MD system in the testing datasets have distinct thermodynamic states, i.e., the bulk density in the testing data is different from those in the training dataset. For the graphene-water system, the testing dataset and the training dataset are under the same thermodynamic state, but the test dataset has distinct channel widths compared to the training dataset.

CED model for confined LJ fluid. Fig. 4a shows a one-to-one comparison of the CED predicted and ground-truth (MD results) fluid-fluid potential for the LJ fluid confined in channel widths of $L = 8$ nm, 16 nm and 28 nm, and for different bulk density. We find that the fluid-fluid interactions predicted by our CED model agree well with the true values. Because the concentration and potential profiles are symmetric about the channel center, we only present the

profiles for half the channel. Notice that U_{ff} has significant layering only near the interface and is close to zero far away (> 2 nm) from the wall surface, where the fluid concentration is nearly homogeneous. Considering the important role of interfaces, here we examine the predictive performance of CED model near the interface ($z < 2$ nm). The results in the inset of Fig. 4a indicate that the CED model can accurately resolve the complex fluid-fluid interactions near the interface by capturing the interfacial behavior accurately.

To quantify the performance of our CED model in the interface region over the entire testing dataset (~ 1500 data samples), we define the average relative error of the i^{th} data sample by averaging the error over the interfacial region ($z < 2$ nm), i.e., $error_i = \frac{1}{N_i} \sum_{j=1}^{N_i} \frac{|U_{ff,i}(z_j) - \tilde{U}_{ff,i}(z_j)|}{\max U_{ff,i}(z_j)}$.

The histogram of the prediction error for the testing dataset is updated in Fig. 4b. We can see that $> 90\%$ of the testing data have relative errors less than 1%. The distribution of the prediction error for different channel widths is presented in Fig. 4c, and most of these relative error points are grouped below 1% for all the channel widths in this work. In addition, we found that the testing error of the CED model does not follow a monotonic trend with respect to the channel width. This is because the CED model is trained via a stochastic optimization procedure and there is no theoretical guarantee that the well-trained deep learning model will converge to the global minima. Even if the CED model converges to a local minima state, the results in figure 4c indicate that it does offer an accurate prediction of the fluid potential.

CED model for confined water. While the CED model shows a good performance for the confined LJ fluid, its ability to determine the fluid-fluid interactions for confined water is unknown. Water has a more complex molecular structure compared to the LJ fluid and is able to form directional hydrogen bonds between water molecules. Furthermore, the structural properties, such as the concentration, RDF, tetrahedral structure, and other properties of water near the wall can be different from the structural properties of water in bulk. Hence, resolving the fluid-fluid interactions of confined water is a formidable challenge.

In this part, we investigate whether the CED model can tackle the above challenge. The CED model is trained using MD data from the graphene-water system. The performance of the CED model for predicting the fluid-fluid interactions of confined water is shown in Fig. 5a. We note that the CED model performs well for water confined in different channel widths at a standard thermodynamic state ($\rho_{bulk} = 1 \text{ g/cm}^3, T = 300 \text{ K}$). The histogram of the relative prediction error in Fig. 5b indicates that the overall performance of the CED model for the testing dataset is again very good with all the average prediction relative errors less than 0.01.

To this end, it is worth mentioning that the CED model implicitly treats one water molecule as one coarse-grained bead when calculating the fluid-fluid interactions of confined water. While a simpler molecular structure is considered, the CED model can still accurately resolve the water concentration near the wall without introducing a coarse-grained potential.

3.2 Performance of DL-QT model for confined LJ fluid

The results above show that the CED model performs well in predicting the fluid-fluid interactions for confined LJ fluid and confined water. Next, we assess the performance of DL-QT model which integrates the well trained CED model with the quasi-continuum theory. The detailed solution of DL-QT model is discussed in section 2.3. The concentration profiles given by the DL-QT model are compared with those from MD simulations in Fig. 6. Here we randomly select 9 samples from the large testing dataset. When the channel width is $L = 8$ nm, we choose three samples with different thermodynamic states: ($\rho_{bulk1} = 11.8 \text{ nm}^{-3}, T_1 = 303 \text{ K}$), ($\rho_{bulk2} = 24.2 \text{ nm}^{-3}, T_2 = 303 \text{ K}$), ($\rho_{bulk3} = 26.1 \text{ nm}^{-3}, T_3 = 303 \text{ K}$); when $L = 16$ nm, we have three samples with different thermodynamic states: ($\rho_{bulk4} = 15.3 \text{ nm}^{-3}, T_4 = 323 \text{ K}$), ($\rho_{bulk5} = 22.5 \text{ nm}^{-3}, T_5 = 303 \text{ K}$), ($\rho_{bulk6} = 27.3 \text{ nm}^{-3}, T_6 = 323 \text{ K}$); when $L = 28$ nm, we choose three samples with different thermodynamic states: ($\rho_{bulk7} = 10.7 \text{ nm}^{-3}, T_7 = 303 \text{ K}$), ($\rho_{bulk8} = 22.4 \text{ nm}^{-3}, T_8 = 323 \text{ K}$), ($\rho_{bulk9} = 28.2 \text{ nm}^{-3}, T_9 = 353 \text{ K}$). We observe that the concentration profiles from both the DL-QT model and MD simulation have a “bulk-like” region in the channel center and an interfacial structure near the wall. The DL-QT concentrations are in good agreement with the MD results for both the “bulk-like” and interfacial regions. The accuracy of the results indicates that atomic physics is integrated into the continuum theory.

To quantitatively investigate the performance of DL-QT model over the entire testing data, we define an average relative prediction error, $relative\ error_i = \frac{1}{N_i} \sum_{j=1}^{N_i} \frac{|\rho_{MD,i}(z_j) - \rho_{DL-QT,i}(z_j)|}{\rho_{MD,i}(z_j)}$, where $\rho_{MD,i}(z_j)$ and $\rho_{DL-QT,i}(z_j)$ represent, respectively, the MD and DL-QT density at position, z_j . Again, we focus on the relative error within the near-wall region with $z < 2$ nm and the histogram of the average relative error for the overall testing data is presented in Fig. 7a. We observe that the relative prediction error is less than 10% for almost all the testing data and about 95% of the testing data have an average relative error less than 6%. Knowing that the concentration depends on the temperature and the channel width, we divide the testing data into three groups based on the temperature in the MD simulation: $T=303\text{K}$, 323K , and 353K . The distribution of the average relative error with different temperatures is shown in Fig. 7b. We can see that the performance of the DL-QT model is about the same under different temperatures, with roughly 75% of the testing data having an average relative error less than 2%. To assess the ability of the DL-QT model for the LJ fluid confined in different channel widths, we divide our testing data into 12 groups, and the distribution of the average relative error is shown in Fig. 7c. Although the average relative error is not the same for different channel widths, we still find $\sim 75\%$ of the testing data have an average relative error less than 5% for all the channel widths considered in this work. This result confirms that the DL-QT model has a robust performance for the LJ fluid confined in different channel widths.

In summary, the DL-QT model performs well in determining the concentration of the LJ fluid confined in different channel widths and under various thermodynamic states. These results

suggest the possibility of connecting atomic physics with continuum theories by using the deep learning model. In DL-QT, we adopt the well trained CED model as the “bridge” to build connections between the continuum theory and molecular physics at the nanoscale. The CED model is trained as a surrogate model to approximate the relation between the fluid-fluid potential and the LJ fluid concentration using a large MD dataset. The trained CED model can predict the fluid-fluid interactions accurately without resolving the complex molecular physics and can be directly combined into the quasi-continuum theory.

3.3 Performance of DL-QT model for water in nanoconfinement

The DL-QT model in section 3.2 shows a good performance in determining the concentration of the LJ fluid confined in a nanochannel. To investigate the performance of DL-QT for confined water, we consider confined water in different channel widths at a standard thermodynamic state, i.e., $T = 300 \text{ K}$ and $\rho_{\text{bulk}} = 1.0 \frac{\text{g}}{\text{cm}^3}$. The one-to-one comparison between the concentration given by MD simulation (ground-truth) and DL-QT is presented in Fig. 8a. We can see that the water concentrations from DL-QT are in good agreement with the MD results. Fig. 8b shows the histogram of the average relative error over the whole testing data, and we observe that DL-QT can predict the confined water concentration accurately, with 90% of the testing data having an average relative error less than 6%.

4. Conclusions

In summary, we developed a deep learning-based quasi-continuum theory to solve the concentration and fluid-fluid potential profiles of confined LJ fluid and confined water in a nano channel. The deep learning model is built based on convolutional encoder-decoder network, and we train the model using the dataset generated by MD simulations. In the first part, we examine the performance of the CED model. We show that the CED model can predict the fluid-fluid interactions accurately without resolving its complex mathematical formula. The model performs well with a prediction error < 0.05 for LJ fluid confined in a nanochannel with different widths and various thermodynamic states. We demonstrate that the CED model can also be used to predict the complex fluid-fluid interactions for confined water accurately by implicitly treating one water molecule as one coarse-grained bead without introducing a coarse-grained potential.

In the second part, we developed the DL-QT model that integrates the well trained CED model into the quasi-continuum theory. The DL-QT model has a robust predictive performance for fluid confined in a nanochannel with different widths and under various thermodynamic states. We find that about 95% of the overall testing data for the LJ fluid and 90% of the testing data for confined water have a relative prediction error less than 6%. The good performance of DL-QT suggests the possibility of connecting molecular physics at the nanoscale with the continuum theories by using the deep learning model. In DL-QT, the CED model serves as the “bridge” to connect the

molecular physics at the nanoscale to the continuum theory. The molecular physics in the CED model is learnt by training the model using a large MD dataset. Conceivably, models built here can be extended to determine the properties of more complex structures, e.g., ionic liquids, organic solvents, and so on.

The DL-QT methodology is not limited to any channel width for the prediction of structure of a confined fluid. While the method is rigorous, deviations from MD simulations can be attributed to inadequate training of the model. The current DL-QT focuses only on the structure of the fluid, but the deep learning-based continuum model can be combined with the Navier-Stokes theory to understand confined fluid transport.^{33, 34}

Acknowledgements

The work on water was supported by the Center for Enhanced Nanofluidic Transport (CENT), an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences (Award # DE-SC0019112). All other aspects of this work were supported by the National Science Foundation under Grants 2140225 and 2137157. The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing access to Lonestar6 resource that has contributed to the research results reported within this paper. We also acknowledge the use of the Extreme Science and Engineering Discovery Environment (XSEDE) Stampede2 at the Texas Advanced Computing Center through allocation TG-CDA100010.

Competing interests: The authors declare no competing interests.

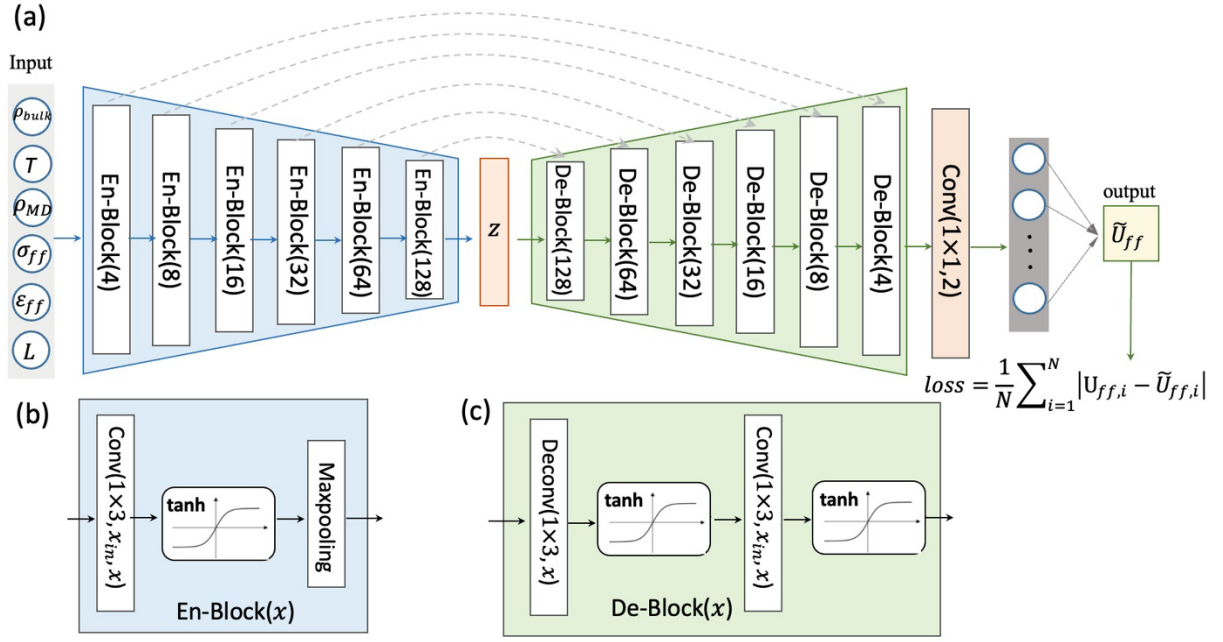


Fig. 1. (a) The convolutional encoder-decoder (CED) model for predicting the fluid-fluid interactions. Blue trapezoid box represents the encoder network, and the green trapezoid box is the decoder network. The encoder and decoder networks are connected via latent space z (orange box). Both encoder and decoder parts contain six repeated blocks. The number of output features is marked in each block. (b) The detailed architecture of a single encoder block. (c) The detailed architecture of a single decoder block.

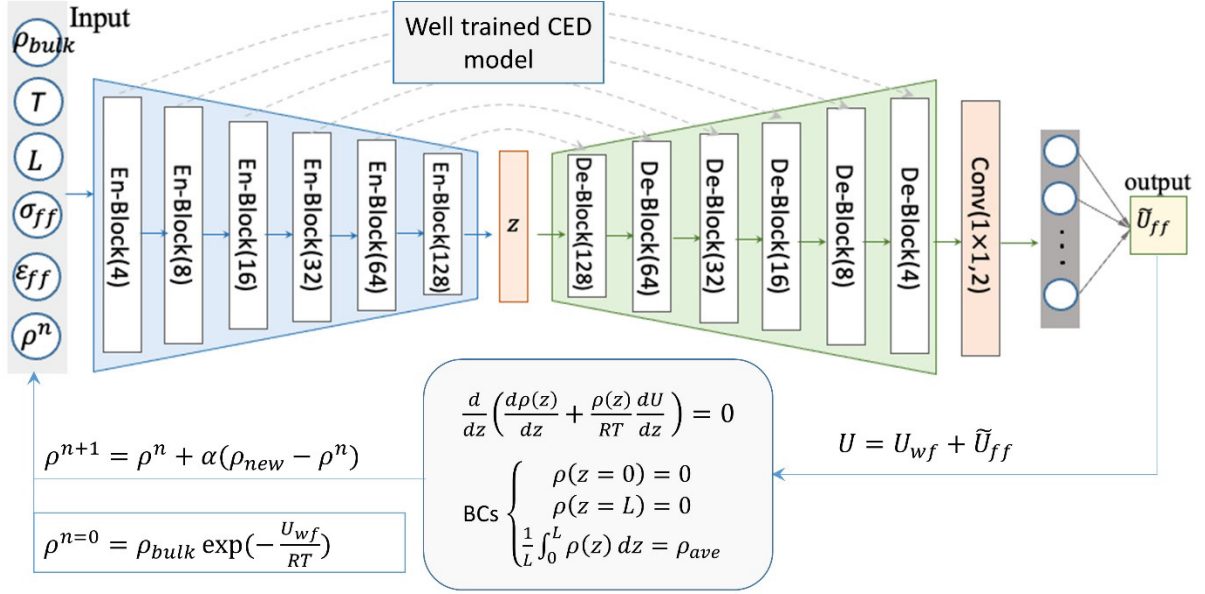


Fig. 2. The schematic of the computational framework of the deep learning-based quasi-continuum theory (DL-QT) model for determining fluid structure in a nanochannel. The model integrates a well-trained CED network with a quasi-continuum theory.

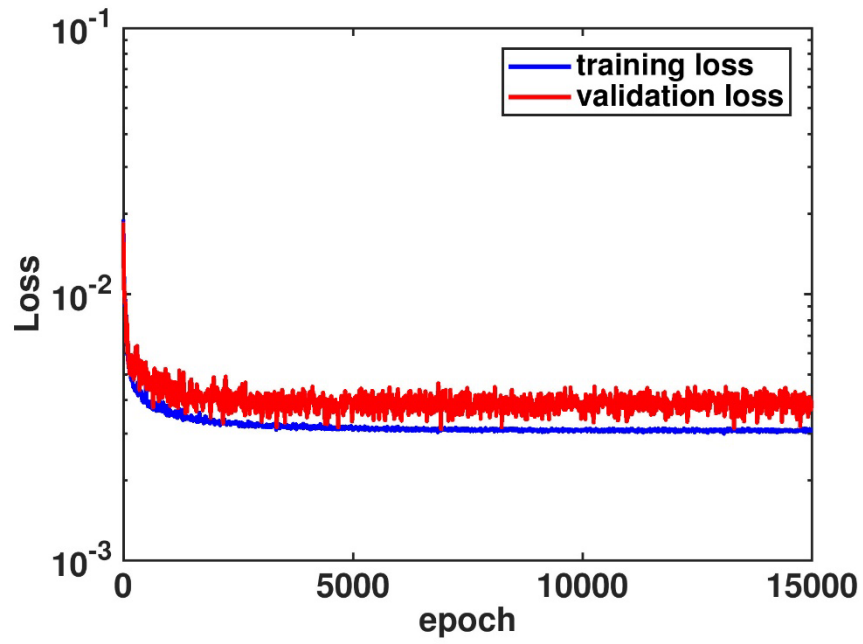


Fig. 3. The evolution of the training loss (Eq. 6) and validation loss by the CED model in the training process.

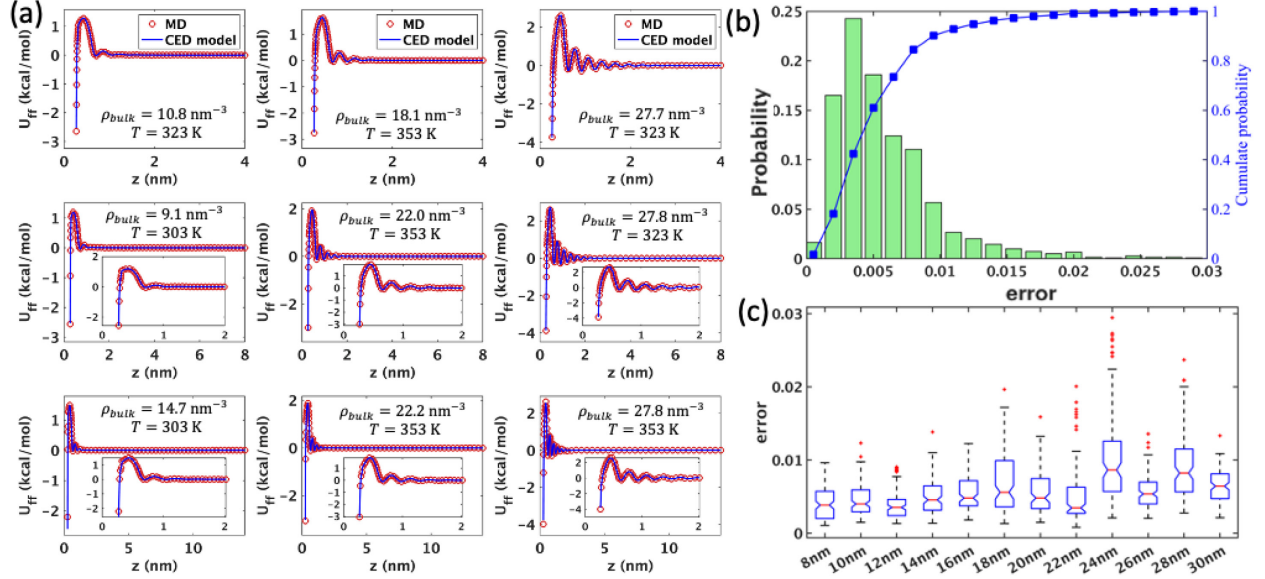


Fig. 4. The predictive performance of the CED model for a confined LJ fluid. (a) One-to-one comparison of the CED predicted and ground-truth (MD results) fluid-fluid interactions for a LJ fluid confined in a channel of width, $L = 8 \text{ nm}$, 16 nm and 28 nm , and under different thermodynamic states. (b) The histogram of the average relative prediction error between the CED predicted and the ground-truth fluid-fluid potential in the near wall region ($z < 2 \text{ nm}$) over all the LJ fluid testing dataset. (c) The distribution of the average prediction error for the LJ fluid confined in a nanochannel of different widths.

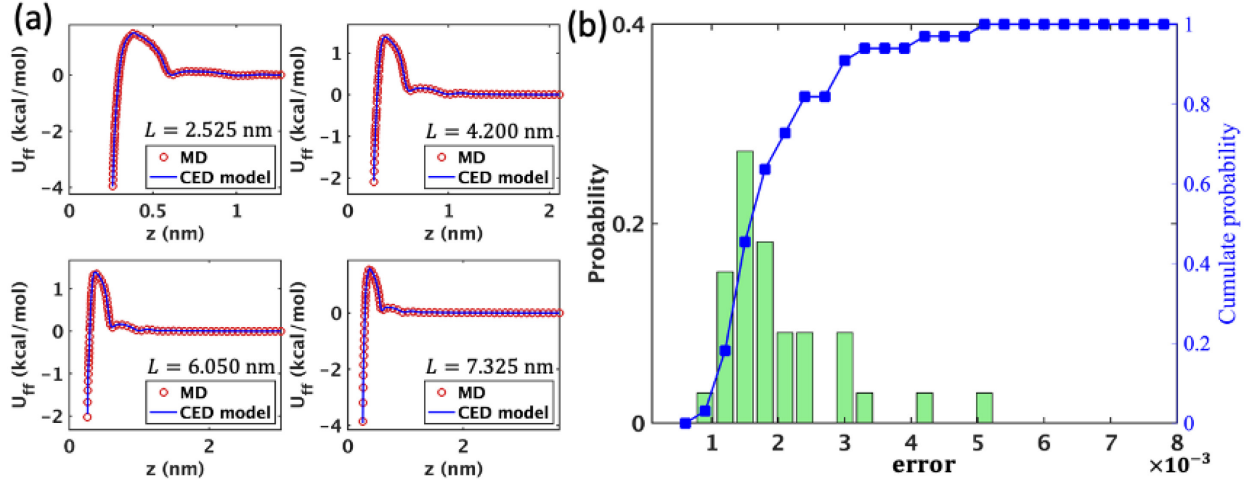


Fig. 5. The predictive performance of the CED model for confined water. (a) One-to-one comparison of the CED predicted and ground-truth (MD results) fluid-fluid interactions for water confined in a nanochannel of different widths. (b) The histogram of the average relative prediction error between CED predicted and ground-truth fluid-fluid potential in the near wall region ($z < 2$ nm) over the entire testing dataset for a confined water system.

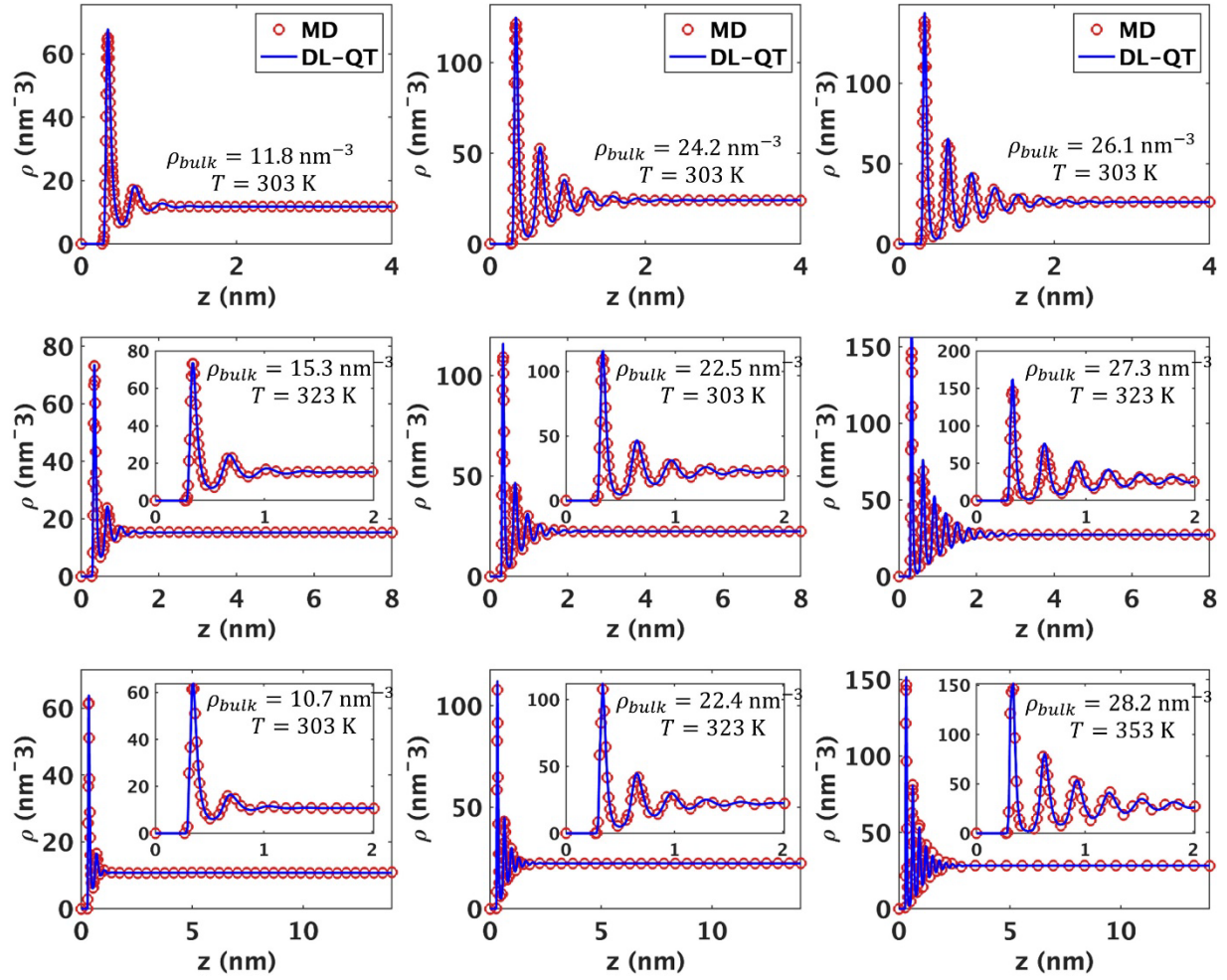


Fig. 6. Comparison of the concentration profiles predicted by the DL-QT model and MD simulations with different channel widths, $L = 8 \text{ nm}$, 16 nm , and 28 nm and for different thermodynamic states.

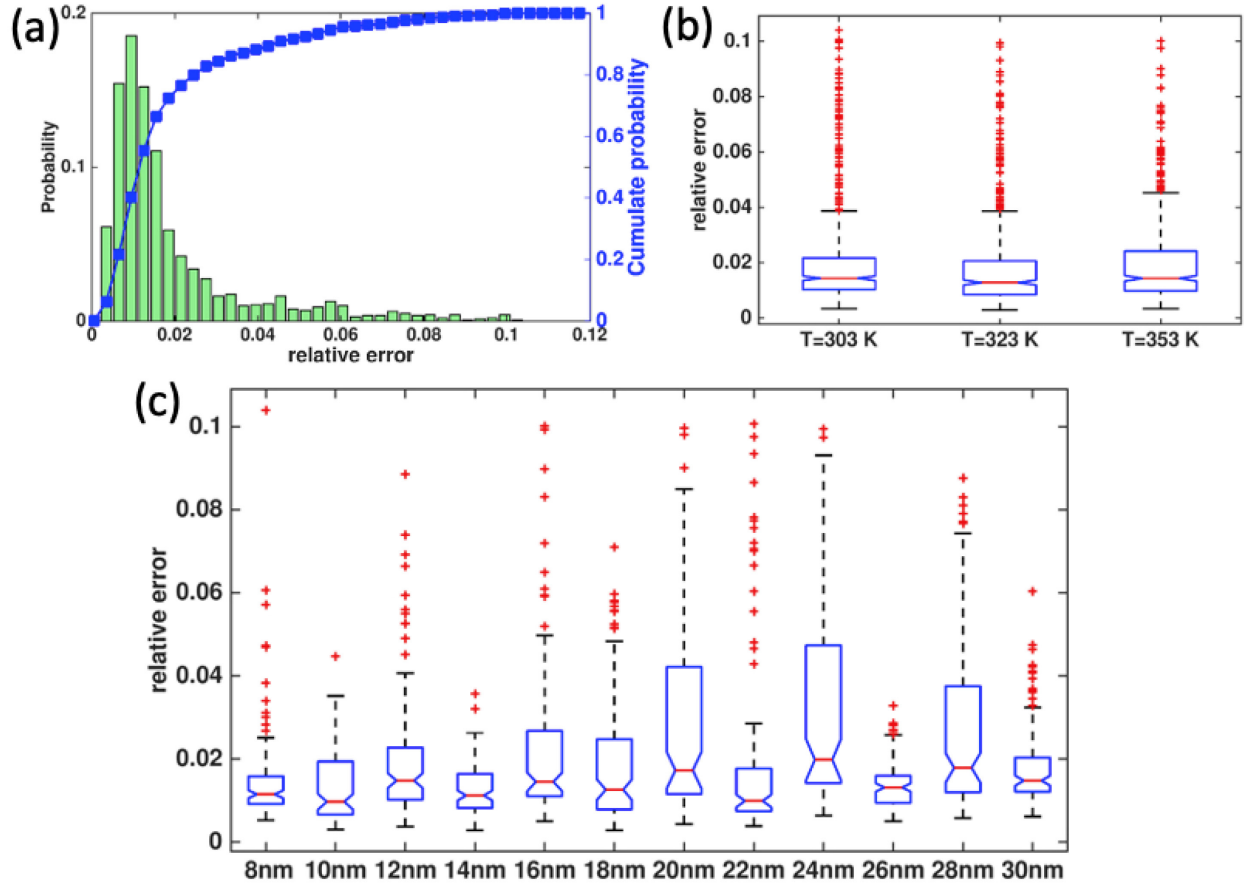


Fig. 7. The predictive performance of the DL-QT model in predicting the concentration of confined LJ fluid. (a) The histogram of the average relative error between the DL-QT solution and the ground-truth in the near wall region ($z < 2$ nm) over the entire testing dataset for a confined LJ fluid. (b) The distribution of the average relative error for a confined LJ fluid for different temperatures. (c) The distribution of the relative error for a LJ fluid confined in a nanochannel of different channel widths.

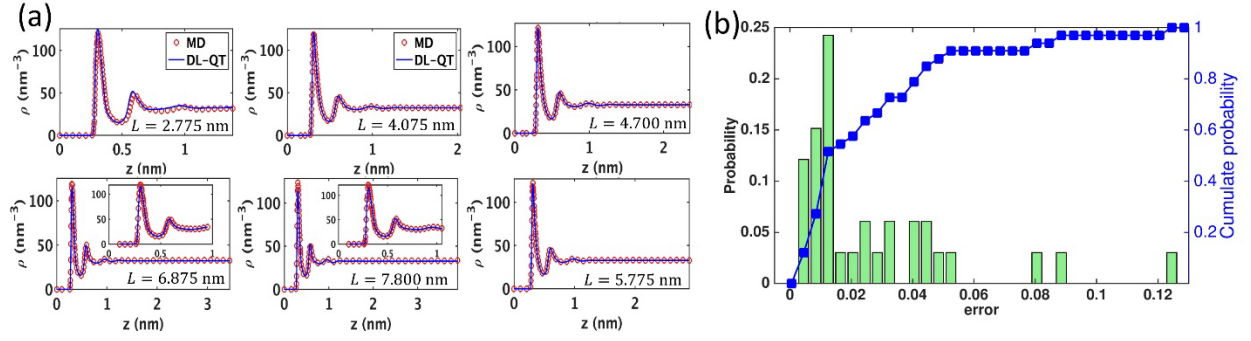


Fig. 8. The predictive performance of the DL-QT model for confined water. (a) Comparison between the solution of the DL-QT model and the ground-truth (MD results) for water confined in a nanochannel of different widths. (b) The histogram of the average relative error between the DL-QT predicted and ground-truth fluid concentration in the near wall region ($z < 2$ nm) over the entire testing dataset.

Table 1. Lennard-Jones parameters for interactions between atoms.

atom	Parameters		
	σ (nm)	ε (Kcal)	Charge (e)
Ar	0.3405	0.2379	0.0
C (graphene)	0.3566	0.0698	0.0
O	0.3166	0.1553	-0.8476
H	0.0	0.0	0.4238

Reference

1. K. P. Lee, T. C. Arnot and D. Mattia, *Journal of Membrane Science* **370** (1-2), 1-22 (2011).
2. M. Suk, A. Raghunathan and N. Aluru, *Applied Physics Letters* **92** (13), 133120 (2008).
3. M. Heiranian, A. B. Farimani and N. R. Aluru, *Nature communications* **6** (1), 1-6 (2015).
4. H. Wu, Y. He and R. Qiao, *Energy & Fuels* **31** (8), 7932-7940 (2017).
5. L. Chen, L. Zhang, Q. Kang, H. S. Viswanathan, J. Yao and W. Tao, *Sci. Rep.* **5** (1), 1-8 (2015).
6. A. Kushima, T. Koido, Y. Fujiwara, N. Kuriyama, N. Kusumi and J. Li, *Nano letters* **15** (12), 8260-8265 (2015).
7. S. Attinger, P. D. Koumoutsakos and S. P. i. Multi-Scale, *Multiscale modelling and simulation*. (Springer, 2004).
8. B. Guillot, *Journal of molecular liquids* **101** (1-3), 219-260 (2002).
9. Y. Gao, M. Li, H. Zhang, Y. Zhang, W. Lu and B. Xu, *Matter* **5** (1), 266-280 (2022).
10. T. Sanghi and N. Aluru, *J. Chem. Phys.* **132** (4), 044703 (2010).
11. J. Wu and Z. Li, *Annu. Rev. Phys. Chem.* **58**, 85-112 (2007).
12. A. Raghunathan, J. Park and N. Aluru, *J. Chem. Phys.* **127** (17), 174701 (2007).
13. S. Izvekov and G. A. Voth, *J. Chem. Phys.* **123** (13), 134105 (2005).
14. I. Vattulainen, M. Karttunen, G. Besold and J. M. Polson, *J. Chem. Phys.* **116** (10), 3967-3979 (2002).
15. Y. Tang, Z. Tong and B. C.-Y. Lu, *Fluid Phase Equilibria* **134** (1-2), 21-42 (1997).
16. M. Motevaselian and N. Aluru, *J. Chem. Phys.* **146** (15), 154102 (2017).
17. T. Sanghi and N. Aluru, *J. Chem. Phys.* **136** (2), 024102 (2012).
18. S. Mashayak and N. Aluru, *J. Chem. Phys.* **137** (21), 214707 (2012).
19. S. Mashayak and N. Aluru, *Journal of chemical theory and computation* **8** (5), 1828-1840 (2012).
20. S. Mashayak and N. Aluru, *J. Chem. Phys.* **148** (21), 214102 (2018).
21. H. Wu and R. Qiao, *Energy and AI* **3**, 100044 (2021).
22. S. M. Moosavi, K. M. Jablonka and B. Smit, *Journal of the American Chemical Society* **142** (48), 20273-20287 (2020).
23. Q. He, D. Barajas-Solano, G. Tartakovsky and A. M. Tartakovsky, *Advances in Water Resources* **141**, 103610 (2020).
24. Y. Zhu, N. Zabarar, P.-S. Koutsourelakis and P. Perdikaris, *J. Comput. Phys.* **394**, 56-81 (2019).
25. A. Moradzadeh and N. R. Aluru, *The journal of physical chemistry letters* **10** (6), 1242-1250 (2019).
26. E. Putin, A. Asadulaev, Q. Vanhaelen, Y. Ivanenkov, A. V. Aladinskaya, A. Aliper and A. Zhavoronkov, *Molecular pharmaceutics* **15** (10), 4386-4397 (2018).
27. J. P. Janet, L. Chan and H. J. Kulik, *The journal of physical chemistry letters* **9** (5), 1064-1071 (2018).
28. K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev and A. Walsh, *Nature* **559** (7715), 547-555 (2018).
29. L. Ward, A. Agrawal, A. Choudhary and C. Wolverton, *npj Computational Materials* **2** (1), 1-7 (2016).
30. J. Kadupitiya, N. Anousheh and V. Jadhao, *arXiv preprint arXiv:2110.14714* (2021).
31. G. T. Craven, N. Lubbers, K. Barros and S. Tretiak, *J. Chem. Phys.* **153** (10), 104502 (2020).
32. J. Kadupitiya, G. C. Fox and V. Jadhao, presented at the International Conference on Computational Science, 2019 (unpublished).
33. M. Raissi, A. Yazdani and G. E. Karniadakis, *Science* **367** (6481), 1026-1030 (2020).
34. P. Perdikaris, M. Raissi, A. Damianou, N. D. Lawrence and G. E. Karniadakis, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **473** (2198), 20160751 (2017).

35. G. C. Peng, M. Alber, A. Buganza Tepole, W. R. Cannon, S. De, S. Dura-Bernal, K. Garikipati, G. Karniadakis, W. W. Lytton and P. Perdikaris, Archives of Computational Methods in Engineering **28** (3), 1017-1037 (2021).
36. M. Alber, A. Buganza Tepole, W. R. Cannon, S. De, S. Dura-Bernal, K. Garikipati, G. Karniadakis, W. W. Lytton, P. Perdikaris and L. Petzold, NPJ digital medicine **2** (1), 1-11 (2019).
37. H. Berendsen, J. Grigera and T. Straatsma, Journal of Physical Chemistry **91** (24), 6269-6271 (1987).
38. T. Darden, D. York and L. Pedersen, J. Chem. Phys. **98** (12), 10089-10092 (1993).
39. S. Plimpton, J. Comput. Phys. **117** (1), 1-19 (1995).
40. I.-C. Yeh and M. L. Berkowitz, J. Chem. Phys. **111** (7), 3155-3162 (1999).
41. S. Nosé, Molecular Physics **57** (1), 187-191 (1986).
42. C. Grigo and P.-S. Koutsourelakis, J. Comput. Phys. **397**, 108842 (2019).
43. H. Wu, W.-Z. Fang, Q. Kang, W.-Q. Tao and R. Qiao, Sci. Rep. **9** (1), 1-12 (2019).
44. V. Dumoulin and F. Visin, arXiv preprint arXiv:1603.07285 (2016).
45. M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving and M. Isard, presented at the 12th Symposium on Operating Systems Design and Implementation (16), 2016 (unpublished).
46. D. P. Kingma and J. Ba, arXiv preprint arXiv:1412.6980 (2014).
47. W. A. Steele, Surface Science **36** (1), 317-352 (1973).