Harsh hotala 22EEB0A64
Vivek kumar 21BTB0A81

Machine learning-assisted exploration of thermally conductive polymers based on high-throughput molecular dynamics simulations

A B S T R A C T

Finding amorphous polymers with higher thermal conductivity is important, as they are ubiquitous in a wide range of applications where heat transfer is important. With recent progress in material informatics, machine learning approaches have been increasingly adopted for finding or designing materials with desired properties. However, limited effort has been put on finding thermally conductive polymers using machine learning, mainly due to the lack of polymer thermal conductivity databases with reasonable data volume. In this work, we combine high-throughput molecular dynamics (MD) simulations and machine learning to explore polymers with relatively high thermal conductivity (>0.300 W/m-K) – a statistically important threshold as most neat polymers have thermal conductivity lower than this value under normal conditions. We first randomly select 365 polymers from the existing PoLyInfo database and calculate their thermal conductivity using MD simulations. The data are then employed to train a machine learning regression model to quantify the structure-thermal conductivity relation, which is further leveraged to screen polymer candidates in the PoLyInfo database with thermal con- ductivity >0.300 W/m-K. 121 polymers with MD-calculated thermal conductivity above this threshold are eventually identified. Polymers with a wide range of thermal conductivity values are selected for re-calculation under different simulation conditions, and those polymers found with thermal conductivity above 0.300 W/m-K are mostly calculated to maintain values above this threshold despite fluctuation in the exact values. Given the observed uncertainties in the MD-calculated TC, we have also constructed a Bayesian neural network to evaluate the epistemic and aleatoric prediction uncertainties, where a state-of-the-art approximate Bayesian inference algorithm is used for scalable training. The strategy and results from this work may contribute to automating the design of polymers with high thermal conductivity.

Introduction

Bulk amorphous polymers are usually regarded as thermal insulators despite that they are ubiquitous in heat transfer applications, such as thermal interface materials, heat exchangers, and thermal energy storage mediums [1–4]. To improve thermal conductivity (TC) of bulk polymers, efforts have focused on compositing thermally conductive fillers with amorphous polymer matrices [5–7]. Different fillers, such as metals [8,9], ceramics [10,11], and carbon [12,13], have been intensively investigated to enhance the TC of polymer composites. However, it has been pointed out that the polymer matrix TC has a major impact on the composite TC, with higher matrix TC leading to greater enhancement in composite TC [1,4]. As a result, finding amorphous polymers that have intrinsic high TC will be of great importance to their heat transfer applications.

Despite the progress in understanding thermal transport physics in

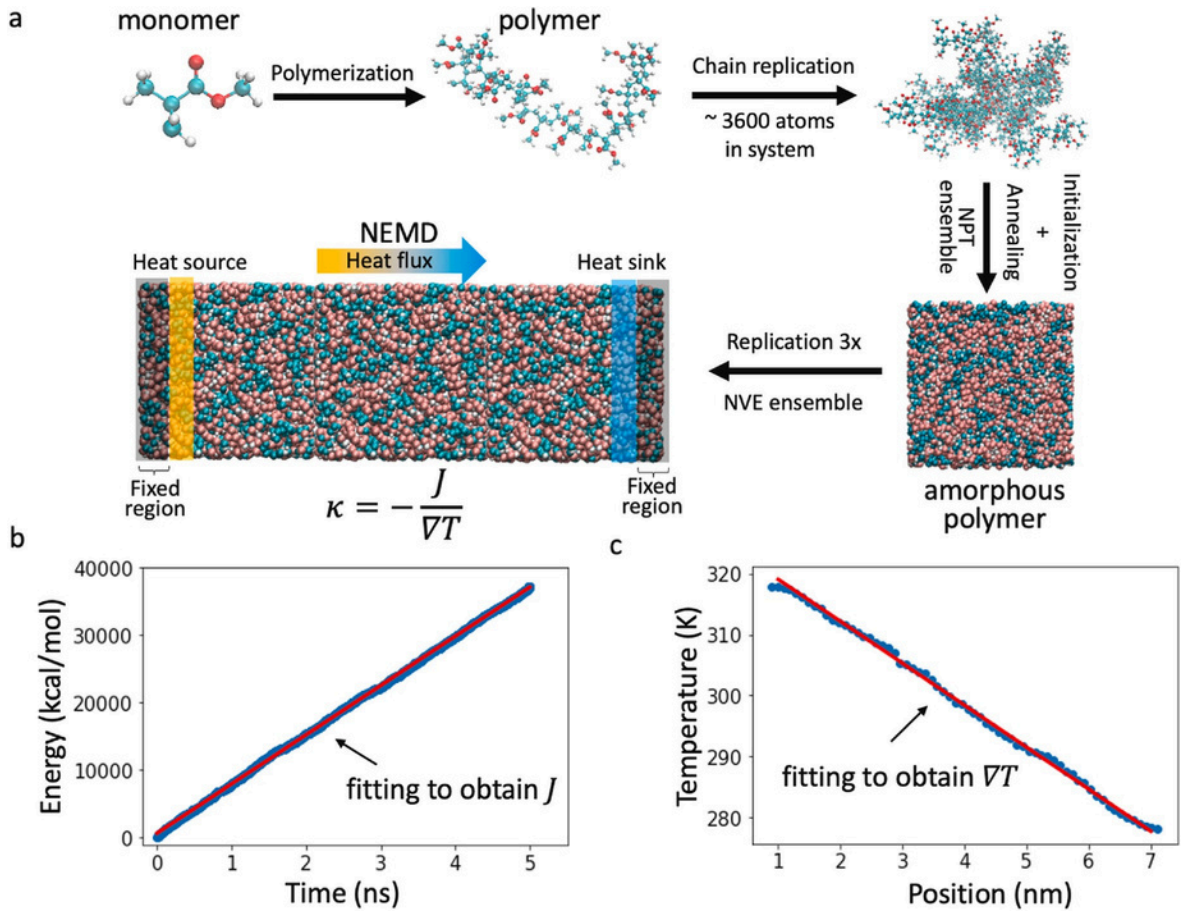

Fig. 1.(a) Diagram of amorphous polymer generation and TC calculation using MD simulations: the amorphous polymer generation includes the polymerization of monomers, chain replication, fully structural relaxation, and TC calculation using NEMD via Fourier's law; (b) an example of energy added to or subtracted from the thermostated regions (i.e., heat source and sink) in the NEMD simulation; (c) an example steady-state temperature profile of the polymer system in the NEMD simulation.

polymers [15–20], relatively limited information is known about which type of amorphous polymer is better in conducting heat, and experimentally screening a large number of polymers is very challenging. The current largest polymer database, PoLyInfo [21], has a very limited number of pure polymers with TC values, and among them, many have noisy data due to different polymer synthesis, processing and TC measurement methods and conditions used in different studies. It would be ideal if a structure-TC relation can be established for quickly screening a large number of polymers to identify promising candidates for further experimental exploration. Recently, data-driven approaches leveraging machine learning have been developed to establish structure-property relations for a variety of materials [22–26,53]. However, such efforts for polymer TC are very limited up to date. Wu et al. leveraged a technique called transfer learning [27] to establish a structure-TC surrogate model for polymers with a predictive accuracy of $R2 = 0.732$ based on only 28 experimental TC data from PoLyInfo [28]. They were able to design a polymer with the TC of 0.410 W/m-K using this surrogate model, which is notably high for amorphous polymers. However, while the transfer learning technique helped to solve the small data issue [28], it is always favorable that big data are available for machine learning studies, which can potentially increase the accuracy of the surrogate models for diverse polymers.

In this work, we leveraged high-throughput molecular dynamics (MD) simulations to generate TC data for quantifying the structure-TC relation in polymers. 365 polymers from the PoLyInfo database were randomly selected for MD simulations, with 324 (89%) of them having MD-labeled TC below 0.300 W/m-K. A random forests (RF) model, which is usually more robust and easier to train compared to neural networks when facing small datasets, was employed to learn the structure-TC relation on the 365 data. The trained RF model was in turn used to screen polymer candidates with TC >0.300 W/m-K in the whole PoLyInfo database. 89 polymer candidates with RF-predicted TC exceeding 0.300 W/m-K were then labeled using MD simulations, with 80 (89.9%) of them having MD-labeled TC above this threshold. While it is found that the MD-calculated TC can vary from one simulation con- dition to another due to the stochastic nature of amorphous polymer morphology, those predicted to have TC >0.300 W/m-K mostly remained to be above this threshold in different simulation conditions. To analyze the epistemic and aleatoric uncertainties of the ML model, we also constructed a Bayesian neural network (BNN) based on the MD data. It was found that the trained BNN can reflect reasonable uncer- tainty estimation without introducing overwhelming training overhead.
Our work may provide useful guidance and an integral component to the experimental exploration of high TC polymers in the future.

2.Methods

The general procedure of the TC data generation using high-throughput MD simulation is depicted in Fig. 1a. It consists of two main steps including amorphous structure optimization and non-equilibrium MD (NEMD) TC calculation as detailed below.

Amorphous polymer structure generation. The structures of 12,777 homopolymers were manually collected from the PoLyInfo database [21], which were stored in their monomer format and

represented using the SMILES [29] (Simplified Molecular Input Line Entry System) strings. We built a Python pipeline based on PYSIMM [30] to automatically generate amorphous polymer structures taking the SMILES strings of monomers as the input. The pipeline first polymerized the monomer into a polymer chain. The number of atoms per chain was controlled to be around 600 for all simulated polymers, and the resultant chain length of all the simulated polymers ranged from ~6 to ~40 nm. The variation in chain length is because some polymers have branches, but some others do not. A previous study has shown that TC does not change significantly after the size of atoms in a chain is greater than 600 [14]. In the meantime, PYSIMM assigned the GAFF2 (General AMBER Force Field 2) forcefield [31] parameters to the polymer, which also generates input scripts for MD simulations using the large-scale atom- ic-molecular massively parallel simulator (LAMMPS) [32]. Polymers with missing GAFF2 forcefield parameters were skipped in the data generation process. We note that skipping some polymers in the data generation process impacts the uniformity of the chemical space sampling (see t-SNE plots in Fig. S1 in the supplementary information, SI). However, this does not influence the conclusion of this study which shows that machine learning models can help increase the possibility of finding high TC polymers. In all simulations, periodic boundary conditions in all spatial directions were used. Each polymer chain was then duplicated by 6 copies and enclosed in a simulation box. The structure was optimized via several steps. In the first few steps, the polymer system was simulated with the electrostatic interactions turned off, and for the Lennard-Jones interactions, a cutoff of 0.300 nm was used. The system was first simulated under the NPT (constant number of atoms, pressure, and temperature) ensemble at 100 K for 2 ps, with a timestep of 0.1 fs applied. After that, the system was heated up from 100 K to 1000 K in 1 ns under the NVT ensemble and then simulated under the NPT (constant number of atoms, pressure, and temperature) ensemble for 50 ps at 0.1 atm and 1000 K. This step was conducted to further relax
the structure and ensure the complete elimination of close contacts. The system was further simulated under NPT at 1000 K for 1 ns, where the pressure was allowed to increase from 0.1 atm to 500 atm, with a timestep of 1 fs and SHAKE constraints [33] applied. The use of SHAKE constraint allows larger timestep of 1 fs to be used in the presence of light hydrogen atoms in the system. In the above steps, turning off electrostatic interactions and employing a series of ad-hoc simulation procedures were implemented to eliminate close contacts between atoms, which was critical to avoiding the system to blow up in the subsequent simulation steps. We call the above-mentioned simulation procedures as the initialization process. We note that one can vary the simulation parameters in the initialization process as long as the resulting polymer structure does not lead to failure in the following annealing simulations.

The obtained polymer system was then annealed with electrostatic interactions turned on, and the PPPM (Particle–Particle–Particle–Mesh)-based Ewald sum method was used to calculate the electrostatic interactions. For the Lennard-Jones interactions, a cutoff of 0.800 nm was used. In the annealing process, the system was first simulated in an NPT ensemble at 1 atm and 1000 K (i.e., the annealing temperature) for 2 ps with a timestep of 0.1 fs? After that, the system was cooled from 1000 K to 300 K with a cooling rate of 140 K/ns in an NPT ensemble at 1 atm (referred to as the annealing process) followed by another NPT run at 300 K and 1 atm for 8 ns (referred to as the relaxation process) to achieve the final amorphous state, with a timestep of 1 fs and SHAKE constraints [33] applied. During annealing, the simulation box size was allowed to relax but constrained to have a cubic shape. These procedures were not designed to mimic experimental annealing but were used as a means to reliably achieve amorphous polymer states. Similar methods have been used extensively for other polymer studies [17,34]. Such a pipeline enables high-throughput amorphous polymer structure generation with little to no human curation. Simulation conditions (e.g., annealing temperature and time) were also varied to examine how they would influence the calculated TC. An example optimization input LAMMPS

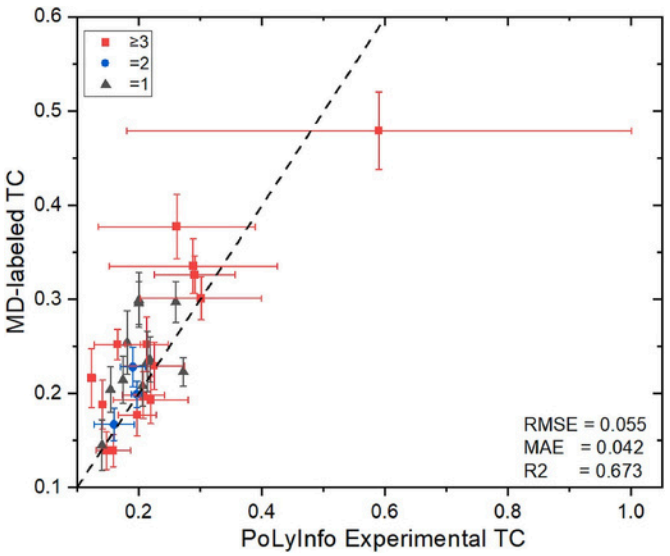

Fig. 2. Pair plot between experimentally measured TC (from PoLyInfo) and MD-labeled TC with error bars. The error bar shows standard deviation of experimental polymer samples (x-direction) or time-averaged TC by MD simulation.

script can be found in the supplemental information (SI).

TC calculation. Each relaxed amorphous polymer structure, which was in a cubic box, was then duplicated in three copies to form a cuboid (see Fig. 1a). Depending on the specific polymers simulated, the dimensions varied a little due to density differences, but they were around 9.900 ×3.300 ×3.300 nm3. The extended system was then used to calculate the TC via NEMD simulations. In the NEMD simulation, the system ran in an NVE (constant number of atoms, volume, and energy) ensemble for 5 ns with a timestep of 0.25 fs? The timestep of 0.25 fs was used to capture the vibration dynamics of the light hydrogen atoms. Larger timesteps would lead to energy drifting when the system includes hydrogen atoms. In these simulations, no SHAKE constraints were used since they would interfere with the natural vibration of atoms, which is key to thermal transport. Heat source (320 K) and sink (280 K) were imposed using Langevin thermostats near the ends of the system (Fig. 1a). The thickness of the source and sink regions were 0.500 nm. A fixed region at each end was implemented as commonly seen in NEMD simulations [55,56]. The fixed regions prevent the drifting of the system as well as heat flow across the periodic boundary. While fixing the atoms at the ends may indirectly influence the dynamics of the nearby atoms within the force interaction range, the Langevin thermostats of heat source and sink regions already perturb the dynamics of these atoms. The strong structural disorder scattering in amorphous polymers should be able to quickly recover the atoms' vibration to their true nature. The recorded heat flux, which was calculated from the heat addition/subtraction from the Langevin heat baths (Fig. 1b), and the temperature profile (Fig. 1c) were averaged over the last 4 ns of the production run, and they were used to calculate the TC via Fourier's law,

$$\kappa = -\frac{J}{\nabla T} \tag{1}$$

where is TC, J is heat flux, ∇T is temperature gradient obtained from the linear fit of the temperature profile. In the production run, the data of the last 4 ns was divided into 8 time periods with each producing a TC value. The standard deviations of these values were calculated (Table S1 in the SI). NEMD can have finite size effect for TC calculations if the system length is shorter than the heat carriers' mean free paths. Recently, Feng et al. [57] found that some polymers with strong electrostatic interactions can have finite size effect on TC in NEMD stimulation. They also found that such an effect is not obvious for polymers

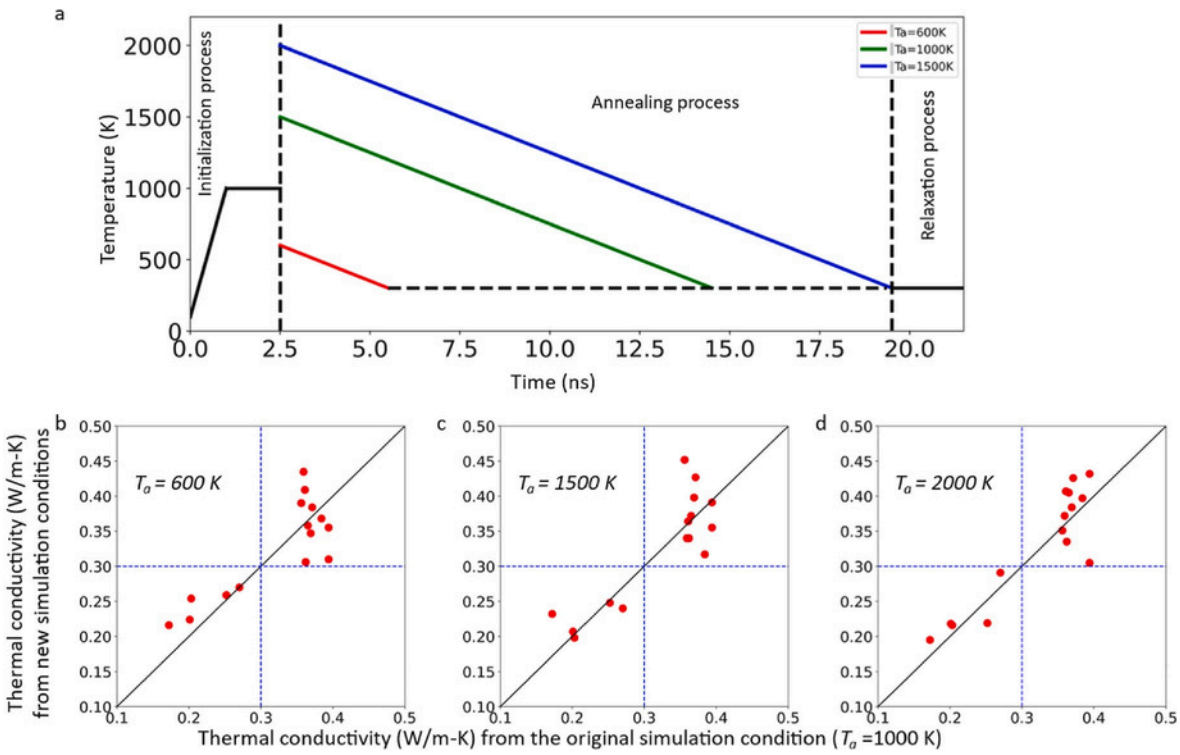

Fig. 3.(a) Illustration of different simulation conditions with different annealing temperatures (*Ta*); (b–d) MD-labeled TC in new testing simulation conditions with different *Ta* (600, 1500 and 2000 K) vs. MD-labeled TC in the original simulation condition (*Ta* =1000 K). The x-axis is for the original simulation condition, and the y-axes are for new testing simulation conditions.

without strong electrostatic interactions. However, other studies have indicated that electrostatic interaction's impact on TC is indirect and short-ranged, and thus size effect should not be significant [17]. If the size effect is significant in our NEMD simulations, we should have under-predicted the TC uniformly compared to benchmark experimental results, which however is not the case as shown in Fig. 2.

3.Results

MD-Calculated TC. We first randomly selected 365 polymers out of the 12,777 in the PoLyInfo database to calculate TC. These include 29 polymers that have experimental TC in the PoLyInfo database for comparison, and our MD-labeled TC agreed generally in trend with the available experimental data in PoLyInfo (Fig. 2). With overall coefficient of determination (*R2*) =0.673, the agreement between experiment and simulation is reasonable. In the PoLyInfo database, there are 85 neat polymers reported with TC measured by experiments. However, careful data cleaning was needed for these experimental data. Since our MD simulation computes the TC at room temperature, we only considered TC measured between 0 and 30 ◦ C – a range in which TC usually does not change significantly [4]. This reduces the number of viable TC data to 53. For these 53 polymers, our MD pipeline could successfully simulate 29 of them, and the rest could not be simulated due to lack of forcefield parameters or failed structural optimization which was likely due to inaccurate forcefield. The MD-calculated TC values for these 29 polymers are compared to the experimental data with error bars for both included (Fig. 2). We note that in Fig. 2, we have labeled TC data according to how many individual data points there are available for the specific polymer in PoLyInfo. We have indeed noticed large variations among individual TC data points for many polymers. If we only consider polymers that have three or more TC data points in PoLyInfo the agreement between our MD-labels and experimental TC increases to *R2* =0.742. There are different reasons for the diversity in reported TC values for a certain polymer in PoLyInfo, but they are generally due to the different sample preparation and/or measurement methods. While the data cleaning process is ad-hoc for different polymer, here, we give examples of the data cleaning process using two polymers: P130085 (poly [(1,1′,3,3′-tetraoxo [5,5′-biisoindoline-2,2′-diyl])-1,4-phenylene]) and P460064 (polyaniline). For P130085, PoLyInfo collected the neat polymer TC values as 0.11, 0.20 and 0.21 W/m-K from Refs. [35,36]. In Ref. [35], Hu et al. mentioned two TC values as 0.115 and 0.21 W/m/K. The sample chosen was Biphenyltetracarboxylic dianhydride-*p*-phenylenediamine (BPDA–PDA) (DuPont PI2611), which is an anisotropic semicrystalline polyimide and has a rigid rod-like chain backbone. 0.115 W/m-K was the TC measured by a 3Ω technique on a 125 Å-thick PI2611 film coated on a 633 μm dielectric silicon wafer, which was a thin film instead of bulk, and the measured TC was for the cross-plane direction. It is known that rigid polyimide polymers, when cast into a thin film, can have highly anisotropic TC after coating with an anisotropy factor of 4–8 [37,38]. Therefore, the in-plane TC of this polymer can be 0.460–0.920 W/m-K. The same goes to the reported value of 0.21 W/m-K in the same paper [35]. In Ref. [36], Hu et al. also measured the BPDA–PDA sample with a density of 1.45 g/cm3 and the specific heat of 1.05 J/g-K, giving a calculated TC of 0.215 W/m-K from the measured thermal diffusivity. However, this value was for the cross-plane TC, and we calculated the in-plane TC based on the reported in-plan thermal diffusivity from the paper to find a value of ∼0.445 W/m-K. Considering these two articles have the same main authors and all of the BPDA–PDA samples are anisotropic but only one in-plane TC was reported, we did not regard the reported data reliable and thus excluded them from the comparison.

For P460064, PoLyInfo collects seven neat polymer TC from 0.16 to 0.35 W/m-K from Refs. [39–44]. Yan et al. [39] measured the thermal diffusivity and specific heat of undoped polyaniline film (insulative) using a laser flash method and a differential scanning calorimeter, respectively. However, they did not show any fabrication details of the polyaniline film. The reported TC of this film was ∼0.16 W/m-K. In Ref. [40], Avlyanov and Mavlyanov showed the TC several polyaniline

Table 1

MD-labeled TC in new testing simulation conditions with different *Ta* (600, 1500 and 2000 K) vs. MD-labeled TC in the original simulation condition (*Ta* = 1000 K). All data units are W/m-K. Data is plotted in Fig. 3.

PID	Name	*Ta* = 1000 K (original)	*Ta* = 600 K	*Ta* = 1500 K	*Ta* = 2000 K	mean	STD
P010002	Poly (prop-1-ene)	0.252	0.259	0.248	0.219	0.245	0.020
P010014	Poly (2-methylprop-1-ene)	0.201	0.224	0.207	0.218	0.213	0.018
P010021	Poly (vinylcyclopropane)	0.172	0.216	0.232	0.195	0.204	0.019
P010026	Poly (allylcyclohexane)	0.203	0.254	0.198	0.216	0.218	0.025
P010042	Poly (dodec-1-ene)	0.362	0.306	0.340	0.335	0.336	0.036
P010044	Poly (octadec-1-ene)	0.365	0.358	0.372	0.405	0.375	0.027
P010077	Poly (cyclodecene)	0.369	0.347	0.398	0.384	0.375	0.036
P010078	Poly (docos-1-ene)	0.371	0.384	0.427	0.426	0.402	0.040
P010080	Poly (hexadec-1-ene)	0.394	0.310	0.355	0.305	0.341	0.030
P010098	Polycyclopentene	0.270	0.270	0.240	0.291	0.268	0.023
P010106	Poly (cyclododecene)	0.394	0.355	0.391	0.432	0.393	0.015
P020056	Poly (4-octadecylstyrene)	0.359	0.435	0.340	0.372	0.364	0.009
P034502	Poly (octadecyl vinyl ether)	0.356	0.390	0.452	0.351	0.387	0.023
P040135	Poly (*N*-docosylacrylamide)	0.384	0.368	0.317	0.397	0.367	0.022
P070101	Poly (pentamethyleneoxydecamethyleneoxydecamethylene oxide)	0.361	0.409	0.364	0.407	0.385	0.018

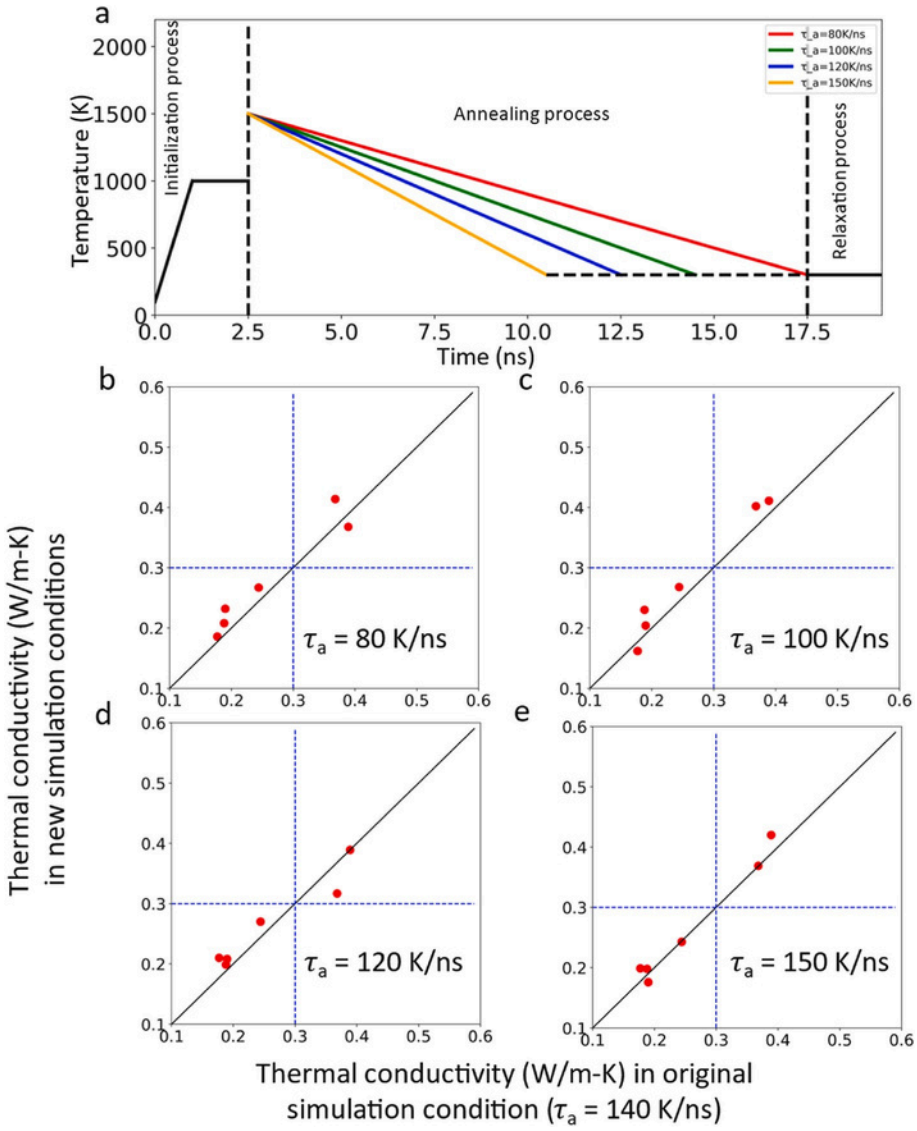

Fig. 4.(a) Illustration of simulation conditions with different annealing rates (τa); (b–e) MD-labeled TC in new simulation conditions vs. MD-labeled TC in the original simulation condition. The TC at the x-axis corresponds to those calculated in the original simulation condition, and the TC at the y-axes corresponds to those calculated in the new simulation conditions.

Table 2
MD-labeled TC in new simulation conditions with different τa (80, 100, 120, 150 K/ns) vs. MD-labeled TC from the original simulation condition (τa = 140 K/ns). The last two columns show TC calculated from simulations with different high temperature relaxation simulation times. All data units are W/m-K. Data is plotted in Fig. 4.

PID	Name	τ_a =140 K/ns (original)	τ_a (K/ns) (T_a = 1500 K)				mean	STD	t_h (ns)	
			80	100	120	150			3	30
P010009	Poly(hept-1-ene)	0.244	0.267	0.268	0.270	0.243	0.258	0.012		
P010052	Poly(4-methylhex-1-ene)	0.190	0.232	0.204	0.208	0.176	0.202	0.019	0.233	0.218
P010006	Poly(3-methylbut-1-ene)	0.177	0.186	0.162	0.210	0.199	0.187	0.017		
P020001	Polystyrene	0.188	0.208	0.230	0.199	0.198	0.205	0.014	0.218	0.222
P090099	Poly(hexadecanolactone)	0.389	0.368	0.411	0.389	0.420	0.395	0.018	0.364	0.427
P100233	Poly[(dodecane-1,12-diamine)-alt-(decanedioic acid)]	0.368	0.414	0.402	0.317	0.369	0.374	0.034		
P100235	poly(iminododecane-1,12-diyliminodoccosanedioyl)	0.423	0.442	0.411	0.410	0.379	0.411	0.022	0.359	0.411

derivatives at 300 K to be 0.90, 1.10, 1.00 W/m-K. In Ref. [41], Pilla et al. showed that polyaniline solution in dimethyl sulfoxide (DMSO) after pouring into a Petri dish and dried had TC of 0.23 W/m-K measured by a time-resolved mode-mismatched thermal lens technique. Ref. [42] by Yan et al. found that a polyaniline film had TC of 0.16 W/m/K, which was measured by a combination of a laser flash method and a differential scanning calorimeter. However, there are some other TC values of P460064 in PoLyInfo, but we did not choose them mostly due to the material type or measurement temperature. Ref. [43] by Yan et al. mentioned a TC of 0.25 W/m-K while their study was on polyaniline composites with polyaniline-coated carbon nanotubes. This TC was actually from Ref. [44] by Zeng et al. which only showed TC curves of

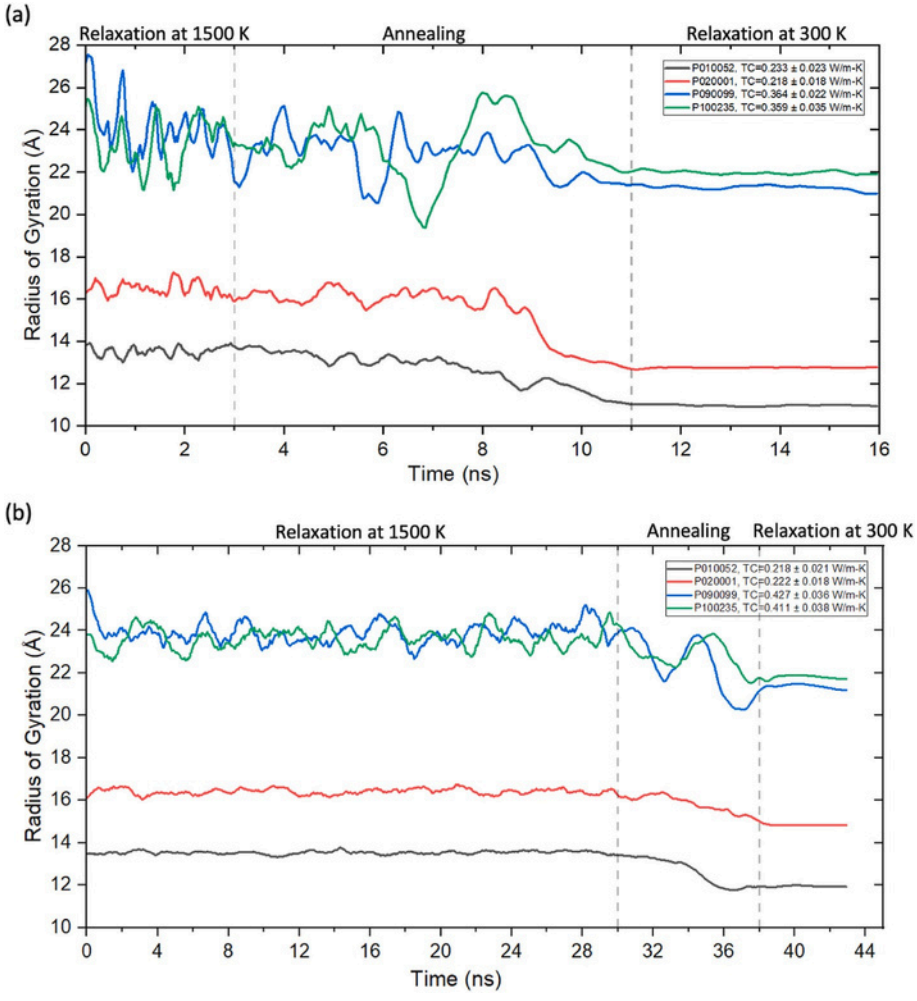

Fig. 5. Radius of gyration (*Rg*) in the relaxation and annealing simulations of four different polymers. (a) th = 3 ns and (b) th = 30 ns, where th is the relaxation simulation time at high temperature. For all simulations, Ta = 1500 K, τa = 150 K/ns and room temperature relaxation time = 5 ns.

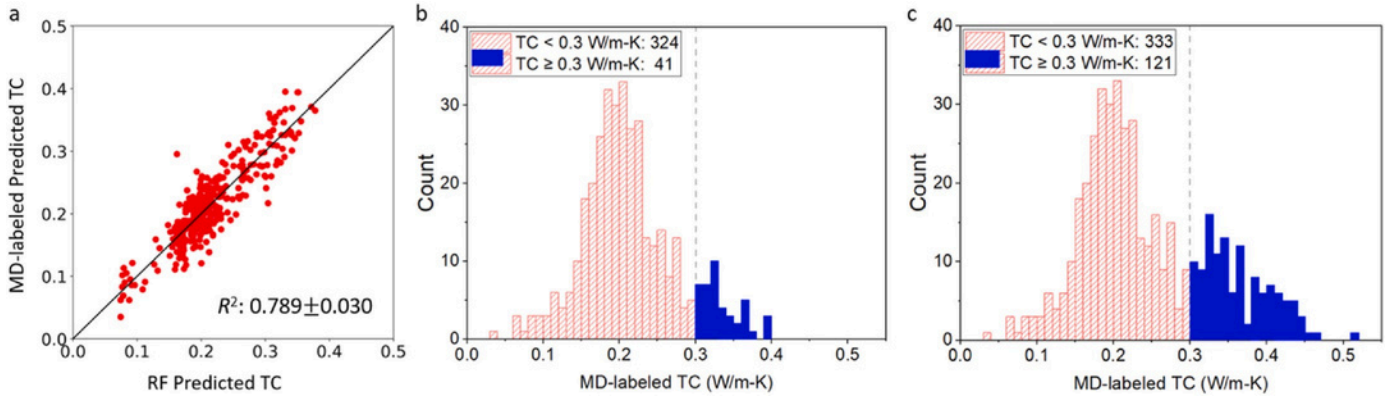

Fig. 6.(a) Pair plot between RF-predicted TC and MD-labeled TC on the validation sets of 365 training data; (b) distribution of the initial 365 MD-labeled TC, with 324 equal to or less than 0.300 W/m-K and 41 above 0.300 W/m-K; (c) distribution of the total 454 MD-labeled TC, with 333 below 0.300 W/m-K and 121 above or equal to 0.300 W/m-K.

polyaniline composite instead of neat polymer. Hence, the TC of P460064 was recorded to be 0.59 ±0.41 W/m-K by averaging the above discussed valid values, which has the largest error bar shown in Fig. 2. Our MD-calculated TC for this polymer is 0.479 ±0.041 W/m-K. Therefore, it is evident that the experimental data from PoLyInfo contains large noise, but the agreement between our MD-calculated TC and these experimental data is reasonable. It is worth noting that all but two polymers collected from PoLyInfo have their mean TC values smaller than 0.3 W/m-K, which is used as a threshold for the following screening using machine learning models. We also note that for some polymers (e.g., polyethylene), there are studies reporting values different from those in PoLyInfo. However, we limited our TC comparison to the PoLyInfo data to maintain consistency throughout this work.

To further assess the robustness of our MD simulations, we chose a number of polymers with different MD-labeled TC values and used several different simulation conditions to re-calculate their TC (Fig. 3b–d). In one set of tests, we used different annealing temperatures, *Ta*, while the simulation settings in other processes (initialization, relaxation, and TC calculation processes) and cooling rate (100 K/ns) were fixed (Fig. 3a). Three new annealing temperatures, 600, 1500 and 2000 K were used, and 15 polymers were tested. Polymers selected for these tests had a wide range of MD-labeled TC (0.172–0.394 W/m-K). The results are shown Fig. 3b–d, which show the parity plots between the TC from the simulations with the additional annealing temperatures and that from the original calculations with Ta =1000 K. The variation in TC, characterized by the standard deviation of data shown in Table 1, can be up to 0.040 W/m-K, and such variation is generally larger for those polymers with higher TC, as also can be seen from Fig. 3b–d. However, there does not seem to be a systematic change of TC with respect to the *Ta*, i.e., the variation is random. Thus, it is not conclusive that one *Ta* can yield more accurate results than another. However, there were overall positive correlation between TC data from simulations with different *Ta* values. The numerical data are shown in Table 1.

One potential influencer of the TC variation is density [4]. The densities of those optimized polymer structures were recorded in their TC calculation processes, for both the original simulation condition and the three new simulation conditions. As shown in Table S2 of SI, the variation in density was small in different simulation conditions and did not correspond to the variation in the TC data, suggesting that the variation in TC is likely from the randomness of morphology from different simulations. In addition, we have tested a number of other polymers with diverse TC and compared their standard deviation within one simulation and that from five ensembles (Table S1 and Fig. S1 in SI). It was found that within each NEMD simulation, the TC uncertainty is more consistent for all the tested polymers (7–10%), but the ensembles have more spread percentile uncertainty (7–20%). The disparity between these two uncertainties increased as TC increased. This finding also suggests that it is the randomness in morphology that leads to TC variation, as rigid polymers usually have larger TC [4] and are usually more difficult to reach the global minimal structure and thus more structural diversity in MD simulations. As shown later, high TC polymers indeed turn out to be more rigid polymers.

Besides the annealing temperature, the cooling rate in the annealing process was also tested to examine their impact on the MD-calculated TC. Thus, we also re-calculated the TC for 6 different polymers with different annealing rates (τa) but with the same Ta =1500 K. These polymers were selected based on their diversity in the MD-labeled TC from the original calculations. These additionally tested τa, besides the original 140 K/ns, were 80, 100, 120 and 150 K/ns (Fig. 4a), respectively. The simulation settings in other processes (initialization, relaxation, and TC calculation processes) and annealing temperature (1500 K) were fixed. From Fig. 4b–e, the different annealing rates led to little variations in the predicted TC, and there was again no systematic trend of TC dependency on the annealing rate observed. Still, those polymers with TC >0.300 W/m-K from the original calculation are all above this threshold in the new test simulation conditions. The data used for Fig. 4 are shown in Table 2.

We have also tested how the structural relaxation at high temperature before annealing would influence TC. We simulated four polymers from those shown in Table 2 by varying the simulation time (*th*) where the system was maintained at Ta =1500 K. Two of the selected polymers have MD-labeled TC close to 0.20 W/m-K, and the other two have TC greater than 0.35 W/m-K. Simulations with th =3 ns and 30 ns were performed and compared to the original simulation process where th = 0. Other simulation conditions, the cooling rate and relaxation time at room temperature, were not changed. Fig. 5 shows the radius of gyration (*Rg*) of these simulations during the MD simulation before NEMD TC calculation. *Rg* is a measure of the spatial extension of the polymer chains and can shed light on the structural relaxation. The trends and the end values of *Rg* are very similar for the th =3 ns and 30 ns simulations. Compared with the previous calculations in Table 2, the TC values from these two sets of calculations (last two columns in the table) are mostly within the error bar of previous TC values. There is also no systematic trend of TC as a function of *th*. These results indicate that no additional relaxation at high temperature is needed to achieve the relaxed structures since the annealing process already allowed such structural relaxation. We also note that any of the relaxation strategies tested did not result in obvious crystallization of the polymers, which happens in much longer time scales than the MD-simulated annealing time if a polymer can indeed crystalize in reality [45].

Regression models. According to these tests, it is understood that the TC data generated from MD simulations have uncertainties, but such uncertainties do no prevent one from using the simulations to predict a trend in TC or tell if a polymer is likely to have relatively high TC

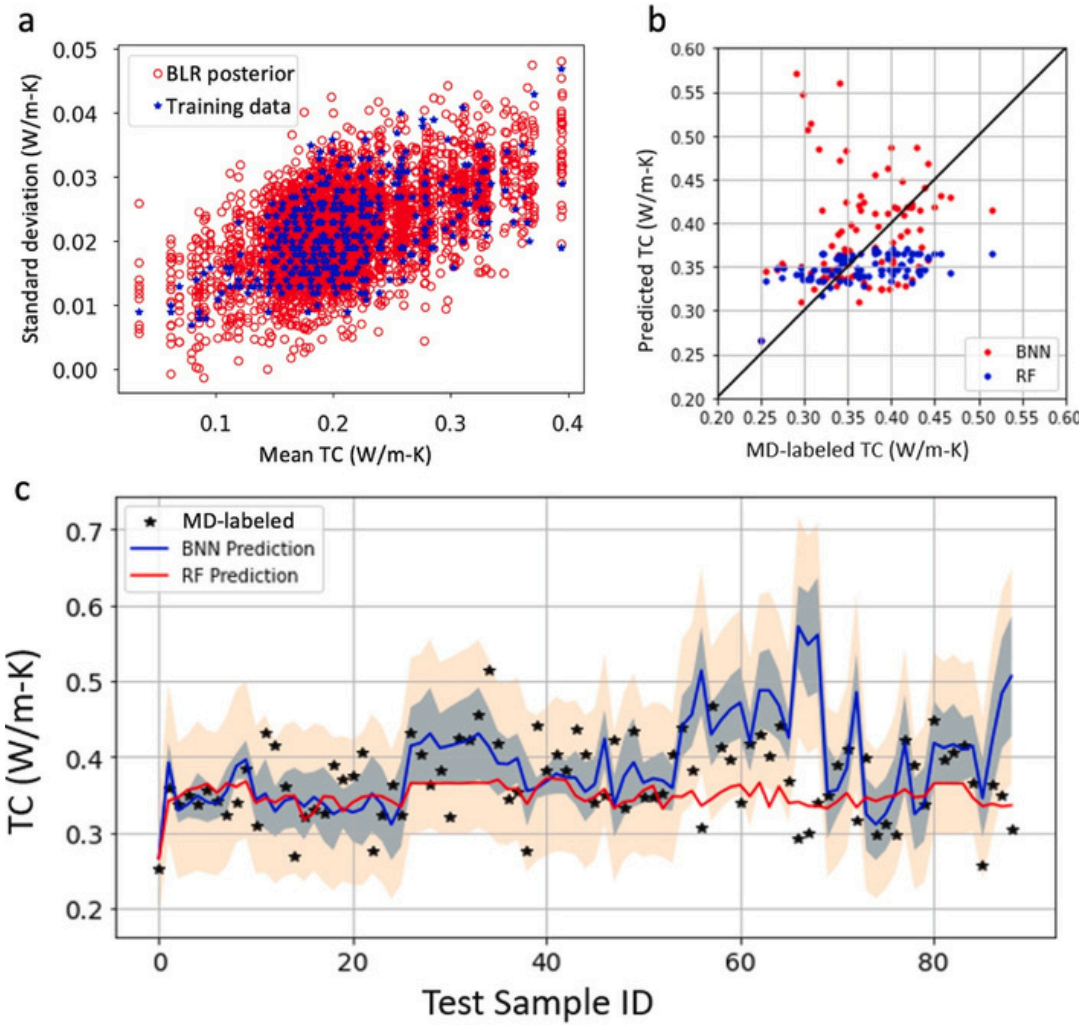

Fig. 7.(a) Mean and standard deviation of TC from training data set (blue star) and the standard devia- tion sampled from the posterior of the trained BLR model; (b) Parity plot of the predicted TC from BNN and RF against the MD-labeled TC for the test dataset; (c) RF and BNN models predicted TC values for the 89 test data, and uncertainty range from BNN. The shaded blue is the epistemic uncertainty and the shaded yellow is the aleatoric uncertainty of BNN. Both shaded areas correspond to 99.7% confidence (i. e., three times the standard deviation).

(>0.300 W/m-K) or low TC. As a result, we used the generated 365 TC data from the original simulation conditions to perform two machine learning regression tasks.

In one of the regression tasks, we established a random forests (RF) surrogate model for the polymer chemistry-TC relation. Each polymer was first represented using the polymer embedding (a continuous-valued vector in the length of 300), which is a machine-learned polymer representation trained using a combination of PoLyInfo and PI1M databases [26], totaling about 1,008,576 polymer structures. PI1M is a virtual library of polymer SMILES sampled from a recurrent neural network (RNN) model trained based on the PoLyInfo polymer structures. The polymer embedding representation was obtained by training the mol2vec [46] model on the monomer structures. In mol2vec, monomers were first decomposed into sequences of substructures, where the center substructure was used as input of a single-hidden-layer neural network to predict its surrounding substructures. Each substructure in the monomer was used as a center substructure for once, and the polymer embedding was derived from the weights of the hidden layer after training was done. More details of polymer embedding can be found in Ref. [26]. We then trained a RF model [47] to quantify the relationship between the polymer embedding and the MD-calculated TC. The number of trees in the RF was set to 1,500 based on grid search in hyperparameter optimization. The RF was trained in the manner of 10-fold cross-validation. The prediction accuracy, estimated by *R2*, is 0.789 ± 0.030 on the validation sets, and the pair plot between RF-predicted TC and MD-calculated TC is shown in Fig. 6a. The distribution of MD-labeled TC is shown in Fig. 6b, the range of which is between 0.035 and 0.395 W/m-K, with the majority below 0.300 W/m-K (324 out of 365, or 89%). We believe that there should be more polymers with TC greater than 0.300 W/m-K in the PoLyInfo database.

Rather than continuing to generate data for randomly selected polymers from PoLyInfo, we used RF to exploit polymers with relatively high TC. More specifically, we predicted the TC (i.e., pseudo label) for all 12,777 polymers in the PoLyInfo database using the RF model to identify polymer candidates that have RF-predicted TC greater than 0.300 W/m- K but have not been MD-labeled. We choose 0.300 W/m-K as the threshold because the tests in the previous sections indicate that the MD simulations tend to correctly predict polymers with TC above or below this threshold despite the calculation uncertainty. We found 880 poly- mers with RF-predicted TC greater than this threshold, and we randomly selected 89 of them to label using MD simulations. Among these 89 polymers, 80 of them (or 89.9%) had TC greater than 0.300 W/m-K in contrast to 41/365 (or 11%) in the original dataset. The TC distribution of the total 454 MD-labeled polymers is shown in Fig. 6c, which shows a much higher population above 0.300 W/m-K.

Since there is inherent randomness in the training dataset, it is natural to evaluate the prediction uncertainty via BNN. By using the BNN, the predicted mean, as well as the prediction uncertainty, can be estimated. In general, the predicted uncertainty is composed of two parts, epistemic uncertainty and aleatoric uncertainty. The epistemic uncertainty arises from incomplete information, e.g., the imperfectness of the model or the sparsity of the data, and it is reducible given more information. On the contrary, the aleatoric uncertainty is due to the stochasticity of the physical process by nature, which is irreducible. In our case, it reflects the inherent uncertainty of the MD results for a given polymer.

A BNN model is constructed to predict the TC of new polymers and, at the same time, quantify the prediction uncertainty of MD-calculated TCs. Similar to the RF model, the BNN model is trained with the initial 365 MD-labeled TC and tested on the 89 new polymers. In terms of the BNN architecture, we use a multilayer perceptron (MLP) with three layers and 100 hidden neurons per layer. To efficiently sample from the intractable posterior, researchers have developed various variational inference (VI) algorithms [48–50]. However, the training

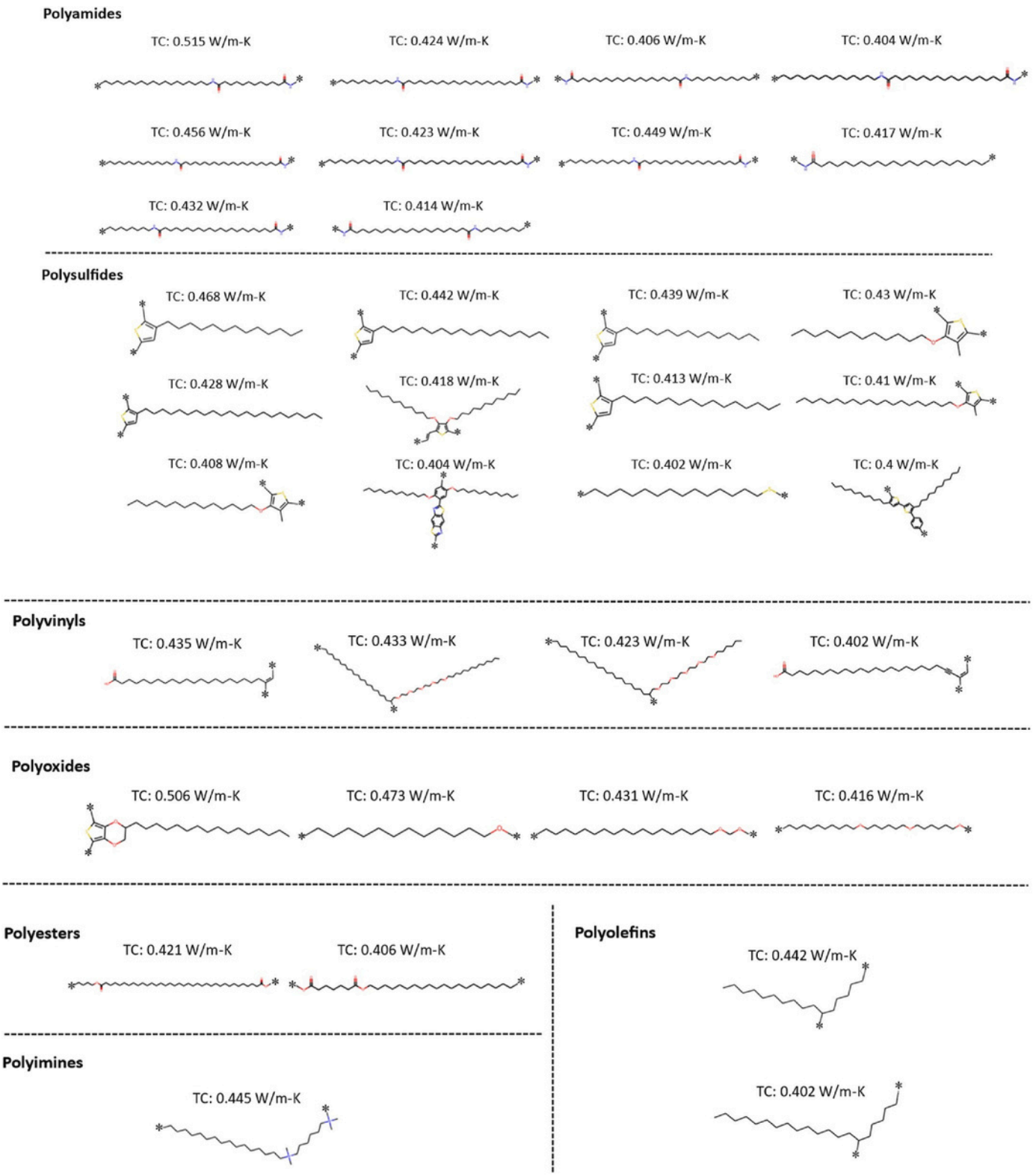

Fig. 8. The polymer structures with MD-labeled TC above 0.400 W/m-K, with their corresponding polymer classes indicated. Red, blue, and yellow colors indicate oxygen, nitrogen, and sulfur atoms respectively.

can still be difficult to converge when the underlying model becomes complex and can introduce significant training overhead. Alternatively, we chose to leverage a Stochastic Gradient Descent (SGD) trajectory-based algorithm called Stochastic Weight Averaging-Gaussian (SWAG) [51]. It is easily scalable to high-dimensional problems and can represent the loss landscape very well. We used the SWAG-based subspace inference algorithm [52] to learn the posterior distributions of the networks. The details of the Bayesian learning algorithm are provided in the SI. In the pre-training stage, we start with a large learning rate of $1 \times 10-2$ that gradually decays to $1 \times 10-3$, where we begin to collect the

Table 3
MD-labeled TC for the 34 high-TC polymers with different initial structures. All data units are W/m-K.

PID	Name	Original	Different initial structure 1	Different initial structure 2	Mean	STD
P070069	Poly (dodecamethylene oxide)	0.473	0.419	0.400	0.431	0.031
P070096	Poly (methyleneoxyoctadecamethylene oxide)	0.431	0.372	0.415	0.406	0.025
P070099	Poly (decamethyleneoxyhexamethyleneoxyhexamethylene oxide)	0.416	0.325	0.344	0.362	0.039
P100018	Poly [(octane-1,8-diamine)-*alt*-(docosanedioic acid)]	0.432	0.448	0.432	0.437	0.008
P100029	Poly [(dodecane-1,12-diamine)-*alt*-(octadecanedioic acid)]	0.404	0.355	0.483	0.414	0.053
P100229	Poly [(decane-1,10-diamine)-*alt*-(docosanedioic acid)]	0.424	0.333	0.408	0.388	0.040
P100235	Poly [(dodecane-1,12-diamine)-*alt*-(dodecanedioic acid)]	0.423	0.477	0.376	0.425	0.041
P100238	Poly [(tetradecane-1,14-diamine)-*alt*-(docosanedioic acid)]	0.456	0.389	0.435	0.427	0.028
P100240	Poly [(octadecane-1,18-diamine)-*alt*-(dodecanedioic acid)]	0.515	0.415	0.371	0.434	0.060
P100696	Poly (22-aminodocosanoic acid)	0.417	0.531	0.383	0.444	0.063
P312010	Poly [(buta-1,3-diene)-*alt*-(1-dodecene)]	0.442	0.375	0.401	0.406	0.028
P312012	Poly [(buta-1,3-diene)-*alt*-(hexadec-1-ene)]	0.402	0.453	0.354	0.403	0.040
P332076	Poly (22-tricosynoic acid)	0.435	0.357	0.417	0.403	0.033
P332081	Poly (22,24-pentacosadiynoic acid)	0.402	0.354	0.447	0.401	0.038
P332304	Poly [1-(2,5,8,11,14-pentaoxaicosan-1-yl)henicosane-1,21-diyl]	0.423	0.349	0.426	0.399	0.036
P332306	Poly [1-(2,5,8,11,14-pentaoxaoctacosan-1-yl)henicosane-1,21-diyl]	0.433	0.375	0.345	0.384	0.037
P373742	Poly(2-hexadecylthieno[3,4-*b*][1,4]dioxane-5,7-diyl)	0.506	0.391	0.437	0.445	0.047
P380064	Poly (tetradecamethylene sulfide)	0.402	0.309	0.313	0.341	0.043
P380068	Poly (3-tetradecylthiophene)	0.439	0.468	0.348	0.418	0.051
P380123	Poly{(2,5-disulfanyl-*p*-phenylenediamine)-alt-[2,5-bis(dodecyloxy)terephthaloyl dichloride]}	0.404	0.336	0.354	0.365	0.029
P382002	Poly (3-tridecylthiophene)	0.468	0.415	0.331	0.405	0.056
P382003	Poly (3-pentadecylthiophene)	0.413	0.406	0.304	0.374	0.050
P382013	Poly{[2,5-dibromo-3,4-bis-(dodecyloxy)thiophene]-alt-[1,2-bis(tributylstannio) ethene]}	0.418	0.328	0.349	0.365	0.038
P382020	Poly (3-docosylthiophene)	0.428	0.327	0.437	0.397	0.050
P382079	Poly [1,4-bis(3-dodecyl-2-thienyl)bezene]	0.400	0.418	0.421	0.413	0.009
P382158	Poly (3-octadecylthiophene)	0.442	0.340	0.314	0.365	0.055
P382279	Poly (3-(dodecyloxy-4-methylthiophene))	0.430	0.377	0.329	0.379	0.041
P382280	Poly [3-methyl-4-(tetradecyloxy)thiophene]	0.408	0.456	0.446	0.437	0.021
P382282	Poly [3-(icosyloxy)-4-methylthiophene]	0.410	0.399	0.315	0.375	0.042
P392423	Poly [(butane-1,4-diol)-*alt*-(hexatriacontanedioic acid)]	0.421	0.513	0.550	0.495	0.054
P402128	Poly (iminododecane-1,12-diyliminoicosane-1,20-dioyl)	0.449	0.443	0.428	0.440	0.009
P402250	Poly [(undecane-1,11-diamine)-*alt*-(icosanedioic acid)]	0.406	0.429	0.447	0.427	0.017
P402252	Poly [(heptane-1,7-diamine)-*alt*-(icosanedioic acid)]	0.414	0.351	0.476	0.414	0.051
P462521	Poly [(dimethylazaniumdiyl)hexane-1,6-diyl (dimethylazaniumdiyl)hexadecane-1,16-diyl]	0.445	0.367	0.412	0.408	0.032

ensembles along the SGD trajectories for estimating the statistics of the approximate Gaussian posterior. During the inference phase, we construct a low dimensional subspace based on the PCA decomposition of the SGD trajectory. A ten-dimensional subspace is found to be sufficient to capture the overall feature of the posterior landscape. Moreover, to account for the scenarios where the standard deviation information of testing data is unknown, we learned the relationship between the TC mean and standard deviation based on the training dataset. Due to observed positive relationship between mean TC and standard deviation in the training dataset, we adopted a Bayesian linear regression (BLR) model using PyMC3 [54], a package for Bayesian linear regression, to capture their relationship, which is assumed to hold for the test dataset. As such, we are able to provide a reasonable estimate for the missing standard deviation information, more specifically, the aleatoric uncertainty, of the test data.

Fig. 7a shows the relationship between the mean and standard deviation of TC of the training dataset (blue stars). The red circles are the standard deviation sampled from the posterior of the trained BLR model. As can be seen, the sampled standard deviation covers the range of the standard deviation (blue starts) of the training data well. Fig. 7b shows the parity plot of the BNN-predicted mean TC, the RF-predicted TC and the MD TC. The RF predictions are concentrated around a narrow range between 0.32 and 0.37 W/m-K, while the BNN prediction covers a range (0.31–0.57 W/m-K) closer to the MD-labeled data (0.25–0.53 W/m-K). Fig. 7c shows the uncertainty information. The BNN predicted TC means are marked by the solid blue line, RF predictions are indicated by red lines, and the ground truth values are marked as black stars. It shows that the RF prediction mostly captures the average TC level but fails to differentiate the variations among different test points, generating an approximately horizontal line. In comparison, the BNN prediction shows a much better overall performance, although bias can be observed at certain testing points. However, the predicted uncertainty range can cover true values very well. Among the 89 test data, 94% of the MD labels are within the BNN-predicted uncertainty range, except for five outliers (Fig. 7c). To sum up, the BNN provides a better mean prediction, and meanwhile, the prediction uncertainty range can cover the ground truth of most testing data.

In Fig. 8, we visualize the top 34 polymers with the highest MD-labeled TC and list their corresponding polymer classes. These polymers are MD-labeled to have TC >0.400 W/m-K. 29 out of 34 were among the 89 MD-labeled new polymers, more polymers are validated by our pipeline again to increase the size of our TC dataset and six more polymers are labeled with TC >0.400 W/m-K. The majority of them belong to polyamides (10/34) and polysulfides (12/34), and the rest are
polyvinyls, polyoxides, polyimines, polyolefins, and polyesters. The fact that polyamides are predicted to have high TC is consistent with a prior study [28], where three thermally conductive polymers designed all belong to polyamides. Given the uncertainty in MD calculations as discussed above, to further confirm the high-TC nature of those 34 polymers, we re-simulated them with different initial structures, as we have concluded that morphology randomness is a major cause of the variation in MD-labeled TC. The MD-labeled TC for those two different initial structures is shown in Table 3. When taking the average of MD-labeled TC values, as can be seen from Table 3, 22 out of 34 still had MD-labeled TC above 0.400 W/m-K, and all, except one, had TC above 0.360 W/m-K. However, more in-depth studies in the future on the chemistry-structure-TC relationship are warranted.

4.Conclusion

In summary, we first used high-throughput MD simulations to calculate TC for a large number of polymers in the PoLyInfo database. Tests on annealing temperature, annealing rate and high temperature relaxation time in the simulation procedure were performed to examine their impact on the calculated TC. The test results showed that the MD-calculated TC contains noise, but could still provide meaningful trends. For selected polymers, we also performed ensemble averaging studies. We found that within each NEMD simulation, the TC uncertainty was more consistent (7–10%), but the ensembles showed more spread percentile uncertainty (7–20%). The difference between these two uncertainties increased as TC increased. These findings suggested that it was the randomness in morphology that led to TC variation, as rigid polymers usually had larger TC and are usually more difficult to relax to the ground state and thus more structural diversity in MD simulations. The MD-generated TC data was then used to train a RF regression model, which was used to screen the polymers in the PoLyInfo database to identify candidates with TC >0.300 W/m-K. Selected polymer candidates proposed by the regression model were further MD-labeled, and 121 polymers with TC >0.300 W/m-K were eventually identified. To characterize prediction uncertainty, a BNN was trained by the initial training dataset which covered only a few polymers with TC >0.300 W/m-K, but it established reasonable uncertainty estimation while predicting on the test set containing many polymers with TC >0.400 W/m-K. Also, 34 were found to have TC >0.400 W/m-K, with majority of them belonging to polyamide and polysulfides classes, and the rest are polyvinyls, polyoxides, polyimines, polyolefins, and polyesters. It is worth mentioning that since the initial training data are mostly below 0.3 W/m-K, it is unlikely that the ML model can accurately predict TC values much beyond this limit. However, the iteration between machine learning model training/prediction and MD TC calculation may provide a strategy to gradually extend the upper boundary of polymer TC. Some studies have reported TC values greater than 0.9 W/m-K in recent years [37,58,59], but most of them involve certain level of chain orientation although the overall structures were still amorphous. In this sense, to explore high TC polymers, it would be beneficial if there is a way to include orientation factors in the MD data generation process. Nevertheless, the strategy and results from this work may provide useful guidance and an integral component to the experimental exploration of high TC polymers.

Author statement

Harsh and Vivek We both successfully increased our discovery "hit rate" from 11% (random search) to 89.9% (ML-guided search). We created a new, valuable database of 121 high-TC polymers for future experimental research. We identified Polyamides and Polysulfides as the most promising chemical. structures to investigate next

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

The computation is supported in part by the University of Notre Dame, Center for Research Computing, and NSF through XSEDE resources provided by TACC Stampede II under a grant number TG-CTS100078. TL would like to thank the support by National Science Foundation (2102592).

Appendix A.Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mtphys.2022.100850.
The raw/processed data required to reproduce these findings cannot be shared at this time due to technical or time limitations. However, they will be incorporated into the PI1M database (https://github.com/RUIMINMA1996/PI1M) upon or before the publication of this manuscript.

References

[1]A. Henry, Thermal transport in polymers, Annual Review of Heat Transfer 17 (2014).
[2]N. Mehra, et al., Thermal transport in polymeric materials and across composite interfaces, Appl. Mater. Today 12 (2018) 92–130.
[3]X. Xu, J. Chen, J. Zhou, B. Li, Thermal conductivity of polymers and their nanocomposites, Adv. Mater. 30 (17) (2018).
[4]X. Wei, Z. Wang, Z. Tian, T. Luo, Thermal transport in polymers: a review, J. Heat Tran. 143 (7) (2021).
[5]Z. Han, Fina Alberto, Thermal conductivity of carbon nanotubes and their polymer nanocomposites: a review, Prog. Polym. Sci. 36 (7) (2011) 914–944.
[6]O. Breuer, Uttandaraman Sundararaj, Big returns from small fibers: a review of polymer/carbon nanotube composites, Polym. Compos. 25 (6) (2004) 630–645.
[7]Y.P. Mamunya, et al., Electrical and thermal conductivity of polymers filled with metal powders, Eur. Polym. J. 38 (9) (2002) 1887–1897.
[8]A. Boudenne, et al., Electrical and thermal behavior of polypropylene filled with copper particles, Compos. Appl. Sci. Manuf. 36 (11) (2005) 1545–1554.
[9]A.S. Luyt, J.A. Molefi, H. Krump, Thermal, mechanical and electrical properties of copper powder filled low-density and linear low-density polyethylene composites, Polym. Degrad. Stabil. 91 (7) (2006) 1629–1636.
[10]S.H. Lee, Yong Choi, Electro-physical properties of composites with nano-sized oxides, J. Nanosci. Nanotechnol. 13 (11) (2013) 7610–7614.
[11]Z. Ge, Feng Ye, Yulong Ding, Composite materials for thermal energy storage: enhancing performance through microstructures, ChemSusChem 7 (5) (2014) 1318–1325.
[12]H.S. Kim, et al., Volume control of expanded graphite based on inductively coupled plasma and enhanced thermal conductivity of epoxy composite by formation of the filler network, Carbon 119 (2017) 40–46.
[13]R.R. Nair, et al., Fine structure constant defines visual transparency of graphene, Science 320 (5881) (2008), 1308-1308.
[14]Xingfei Wei, Tengfei Luo, Chain length effect on thermal transport in amorphous polymers and a structure–thermal conductivity relation, Phys. Chem. Chem. Phys. 21.28 (2019) 15523–15530.
[15]P.W. Anderson, B.I. Halperin, C.M. Varma, Anomalous low-temperature thermal properties of glasses and spin glasses, Phil. Mag. 25 (1) (1972) 1–9.
[16]W.A. Phillips, Tunneling states in amorphous solids, J. Low Temp. Phys. 7 (3–4) (1972) 351–360.
[17]Xingfei Wei, Tengfei Luo, Role of ionization in thermal transport of solid polyelectrolytes, J. Phys. Chem. C 123.20 (2019) 12659–12665.
[18]S. Shenogin, A. Bodapati, P. Keblinski, A.J. McGaughey, Predicting the thermal conductivity of inorganic and polymeric glasses: the role of anharmonicity, J. Appl. Phys. 105 (3) (2009): 034906.
[19]T. Zhang, T. Luo, Role of chain morphology and stiffness in thermal conductivity of amorphous polymers, J. Phys. Chem. B 120 (4) (2016) 803–812.
[20]Q. Xi, J. Zhong, J. He, X. Xu, T. Nakayama, Y. Wang, J. Liu, J. Zhou, B. Li, A ubiquitous thermal conductivity formula for liquids, polymer glass, and amorphous solids, Chin. Phys. Lett. 37 (10) (2020).
[21]S. Otsuka, et al., *PoLyInfo: Polymer Database for Polymeric Materials Design.* Emerging Intelligent Data and Web Technologies (EIDWT), 2011 International Conference on, IEEE, 2011.
[22]D.J. Audus, Juan J. de Pablo, Polymer Informatics: Opportunities and Challenges, 2017, pp. 1078–1082.
[23]S. Wu, H. Yamada, Y. Hayashi, M. Zamengo, R. Yoshida, Potentials and Challenges of Polymer Informatics: Exploiting Machine Learning for Polymer Design, arXiv preprint arXiv vol. 2010, 2020, 7683.
[24]Chen, L., Pilania, G., Batra, R., Huan, T.D., Kim, C., Kuenneth, C. and Ramprasad, R., Polymer Informatics: Current Status and Critical Next Steps. arXiv preprint arXiv, 2020. 2011(00508).
[25]R. Ma, et al., Evaluating polymer representations via quantifying structure-property relationships, J. Chem. Inf. Model. 59 (7) (2019) 3110–3119.

l.
[26]R. Ma, T. Luo, PI1M: a benchmark database for polymer informatics, J. Chem. Inf. Model. 60 (10) (2020) 4684–4690.
[27]H. Yamada, et al., Predicting materials properties with little data using shotgun transfer learning, ACS Cent. Sci. 5 (10) (2019) 1717–1730.
[28]S. Wu, et al., Machine-learning-assisted discovery of polymers with high thermal conductivity using a molecular design algorithm, npj Computational Materials 5 (1) (2019) 1–11.
[29]D. Weininger, SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules, J. Chem. Inf. Comput. Sci. 28 (1) (1988) 31–36.
[30]M.E.a.C. Fortunato, pysimm: A python package for simulation of molecular systems, C.M. Software 6 (2017) 7–12.
[31]D. Vassetti, Marco Pagliai, Piero Procacci, Assessment of GAFF2 and OPLS-AA general force fields in combination with the water models TIP3P, SPCE, and OPC3 for the solvation free energy of druglike organic molecules, J. Chem. Theor. Comput. 15 (3) (2019) 1983–1995.
[32]S. Plimpton, Fast parallel algorithms for short-range molecular dynamics, J. Comput. Phys. 117 (1) (1995) 1–19.
[33]J.P. Ryckaert, G. Ciccotti, H.J. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, J. Comput. Phys. 23 (3) (1977) 327–341.
[34]X. Wei, R. Ma, T. Luo, Thermal conductivity of polyelectrolytes with different counterions, J. Phys. Chem. C 124 (8) (2020) 4483–4488.
[35]Chuan Hu, Michael Kiene, Paul S. Ho, Thermal conductivity and interfacial thermal resistance of polymeric low k films, Appl. Phys. Lett. 79.25 (2001) 4121–4123.
[36]Chuan Hu, et al., Study of the thermal properties of polymeric dielectric materials by photothermal technique, MRS Online Proc. Libr. 511 (1998).
[37]Xingfei Wei, et al., Thermal conductivity of pentiptycene-based poly (o-hydroxyimide) copolymers: a study via integrated experiments and simulations, ACS Applied Polymer Materials 3.6 (2021) 2979–2987.
[38]Katsuo Kurabayashi, et al., Measurement of the thermal conductivity anisotropy in polyimide films, J. Microelectromech. Syst. 8.2 (1999) 180–191.
[39]Hu Yan, Naonori Ohno, Naoki Toshima, Low thermal conductivities of undoped and various protonic acid-doped polyaniline films, Chem. Lett. 29.4 (2000) 392–393.
[40]J.K. Avlyanov, A. Mavlyanov, Low temperature transitions in Polyanilines, in: Electronic Properties of Polymers, Springer, Berlin, Heidelberg, 1992, pp. 268–270.
[41]Viviane Pilla, et al., Thermal lensing in poly (vinyl alcohol)/polyaniline blends, J. Polym. Sci. B Polym. Phys. 40.17 (2002) 1949–1956.
[42]Hu Yan, Norina Sada, Naoki Toshima, Thermal transporting properties of electrically conductive polyaniline films as organic thermoelectric materials, J. Therm. Anal. Calorim. 69.3 (2002) 881–887.
[43]Haiyan Yan, Kaichang Kou, Enhanced thermoelectric properties in polyaniline composites with polyaniline-coated carbon nanotubes, J. Mater. Sci. 49.3 (2014) 1222–1228.
[44]J.L. Zeng, et al., Thermal conductivity enhancement of MWNTs on the PANI/tetradecanol form-stable PCM, J. Therm. Anal. Calorim. 91.2 (2008) 443–446.
[45]S. Chew, J.R. Griffiths, Z.H. Stachurski, The crystallization kinetics of polyethylene under isothermal and non-isothermal conditions, Polymer 30.5 (1989) 874–881.
[46]S. Jaeger, S. Fulle, S. Turk, Mol2vec: unsupervised machine learning approach with chemical intuition, J. Chem. Inf. Model. 58 (1) (2018) 27–35.
[47]L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32.
[48]Charles Blundell, et al., Weight uncertainty in neural network, in: International Conference on Machine Learning, PMLR, 2015.
[49]Qiang Liu, Dilin Wang, Stein variational gradient descent: a general purpose bayesian inference algorithm, Adv. Neural Inf. Process. Syst. 29 (2016).
[50]Luning Sun, Jian-Xun Wang, Physics-constrained bayesian neural network for fluid flow reconstruction with sparse and noisy data, Theoretical and Applied Mechanics Letters 10.3 (2020) 161–169.
[51]Wesley J. Maddox, et al., A simple baseline for bayesian uncertainty in deep learning, Adv. Neural Inf. Process. Syst. 32 (2019).
[52]Pavel Izmailov, et al., Subspace inference for Bayesian deep learning, in: Uncertainty in Artificial Intelligence, PMLR, 2020.
[53]Ruimin Ma, Hanfeng Zhang, Tengfei Luo, Exploring high thermal conductivity amorphous polymers using reinforcement learning, ACS Appl. Mater. Interfaces 14.13 (2022) 15587–15598.
[54]John Salvatier, Thomas V. Wiecki, Christopher Fonnesbeck, Probabilistic programming in Python using PyMC3, PeerJ Computer Science 2 (2016) e55.
[55]Hua Bao, et al., A review of simulation methods in micro/nanoscale heat conduction, ES Energy & Environment 1.39 (2018) 16–55.
[56]Zhen Li, et al., Influence of thermostatting on nonequilibrium molecular dynamics simulations of heat conduction in solids, J. Chem. Phys. 151.23 (2019): 234105.
[57]Tianli Feng, et al., Size effects in the thermal conductivity of amorphous polymers, Physical Review Applied 14.4 (2020): 044023.
[58]Virendra Singh, et al., High thermal conductivity of chain-oriented amorphous polythiophene, Nat. Nanotechnol. 9.5 (2014) 384–390.
[59]Yanfei Xu, et al., Molecular engineered conjugated polymer with high thermal conductivity, Sci. Adv. 4.3 (2018) eaar3031.