

Deep Learning for Determination of Properties of Semidilute Polymer Solutions

Ryan Sayko, Michael Jacobs, Marissa Dominijanni, and Andrey V. Dobrynin*

Cite This: <https://doi.org/10.1021/acsapm.3c01282>

Read Online

ACCESS |



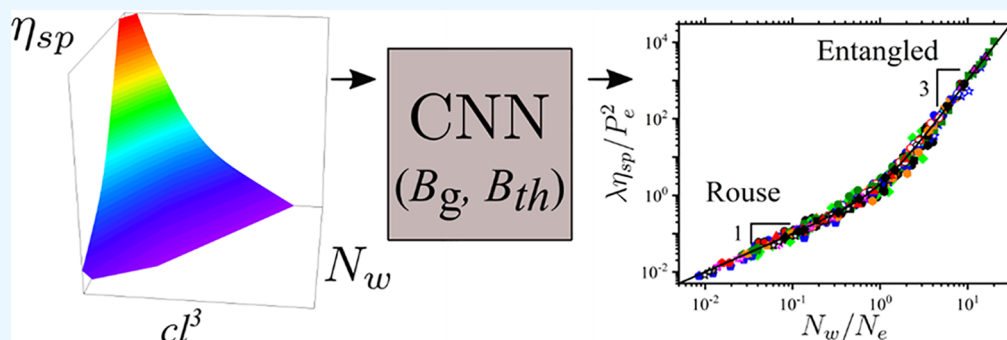
Metrics & More



Article Recommendations



Supporting Information



ABSTRACT: The ability to predict polymer solution viscosity is essential for polymer characterization and processing. Here, we use a synergistic approach combining the scaling theory of polymer solutions and convolutional neural network (CNN) models to obtain system-specific parameters and describe semidilute solution viscosity in the unentangled and entangled solution regimes. The scaling approach relies on the existence of a characteristic microscopic length scale—the solution correlation length (correlation blob size) ξ —which uniquely defines macroscopic solution properties. It is based on a relationship between the solution correlation length $\xi = l g^\nu / B$ and the number of monomers g per correlation volume for polymers with the monomer projection length l . The system-specific set of parameters B_g , B_{th} , and 1 in the corresponding solution regime with the scaling exponent $\nu = 0.588$, 0.5, and 1, respectively. Applying two CNN models, we obtained the sets B_g and B_{th} from the solution specific viscosity, η_{sp} , as a function of concentration, c , and weight-average degree of polymerization, N_w . The CNN was trained on theoretically generated datasets converted to sparse images representing the normalized specific viscosity $\eta_{sp}/N_w (cl^3)^{1/(3\nu-1)}$ in the unentangled Rouse regime. The trained CNN was utilized in automated data analysis of the solution viscosity of polystyrene, poly(ethylene oxide), poly(methyl methacrylate), poly(acrylonitrile-co-itaconic acid), cellulose, sodium hyaluronate, hydroxypropyl methyl cellulose, methyl cellulose, hydroxypropyl cellulose, cellulose tris(phenyl carbamate), xanthan gum, galactomannan, and sodium κ -carrageenan in water, organic solvents, and ionic liquids. This approach produced values of the B -parameters with mean absolute percentage differences of less than 6% from the corresponding values determined by the manual data analysis. The B -parameters are then used to obtain the packing number P_e defining the onset of entanglements in polymer solutions and to describe semidilute solution viscosity as a function of concentration and N_w .

KEYWORDS: AI approach, scaling, solution viscosity, polymers, polymer solutions

INTRODUCTION

Over the past decade, the discovery of new polymeric materials with desired properties has been transformed by advancements in machine learning techniques and algorithms,^{1–4} polymer property databases,^{5–8} and automated polymer synthesis and characterization.^{9–11} This has resulted in a generation of data banks composed of comprehensive molecular parameters, including constituting atoms, chemical bonds, bond angles, interatomic interactions, and electronic structures. However, the application of this approach faces challenges in describing the properties of polymer solutions due to their hierarchical organization, conformational complexity, and inherent polydispersity, where properties do not always have direct

correlations with a chemical structure. In particular, the chemical specificity of solution constituents manifests itself in coarse-grained parameters such as the chain's Kuhn length, excluded volume, and packing density required for formation of entanglements.

Received: June 14, 2023

Accepted: August 8, 2023

The application of artificial intelligence (AI) in establishing structure–property correlations opened a new direction in polymer research.^{12–15} The majority of research in this direction is focused on predicting polymer properties from a monomer chemical structure or their distribution along the polymer backbone.^{16,17} In particular, data-driven approaches have utilized unsupervised learning¹⁸ and reverse engineering analysis methods^{19,20} to expand the analysis of simulation and experimental data.^{21–24} However, leveraging artificial intelligence to accelerate data analysis remains largely unexplored.^{25,26} To address these challenges and the big data requirements needed for neural network training, we developed an approach that utilizes a scaling theory of semidilute polymer solutions^{27–29} in conjunction with a convolutional neural network (CNN).^{30–33} By generating large theoretical datasets using the confirmed scaling relationships, we eliminated the reliance on extensive experimental datasets. The CNN viability was tested on experimental datasets for the concentration dependence of specific viscosity in solutions of polymers with different molecular weights. By determining the sets of polymer/solvent-specific parameters defining chain statistics on different length scales, the onset of entanglements, and crossover concentrations between different solution regimes, the CNN was able to reconstruct a universal viscosity representation across the Rouse and entangled solution regimes. This data-driven approach to the determination of polymer solution properties holds the potential to enable the rapid automated analysis of the solution viscosity and can be used in polymer characterization and processing.

The rest of the article is organized as follows. First, we briefly overview the scaling model of semidilute polymer solutions. After that, we discuss the CNN design and training on the theoretically generated datasets of specific viscosity in semidilute polymer solutions in a broad range of polymer concentrations and molecular weights. A trained CNN is then utilized for automated analysis of 37 polymer solutions^{34–54} to obtain polymer/solvent-specific parameters defining chain statistics on the different length scales and onset of entanglements. The CNN-generated system parameters are compared with those determined from manual analysis of the viscosity data to establish the accuracy of the developed CNN.

SCALING MODEL OF SEMIDILUTE POLYMER SOLUTIONS

We begin the discussion with a brief overview of the scaling theory of semidilute polymer solutions.^{27–29} In the framework of the scaling approach, solution properties are defined by the solution correlation length (blob), ξ , the largest length scale within which there is a hierarchical blob structure reflecting a scale-dependent polymer/solvent affinity (Figure 1a). In a good solvent, the correlation blobs are made of thermal blobs with size D_{th} and each containing g_{th} monomers, within which chain statistics are similar to a random walk on the length scales larger than the Kuhn length b . On the length scales shorter than a chain Kuhn length but larger than the monomer (repeat unit) projection length in an all-trans conformation, l , the root-mean-square end-to-end distance r of the section of a polymer chain scales linearly with the number of monomers g , in it

$$r = l g_r, \quad \text{for } l \leq r \leq b \quad (1)$$

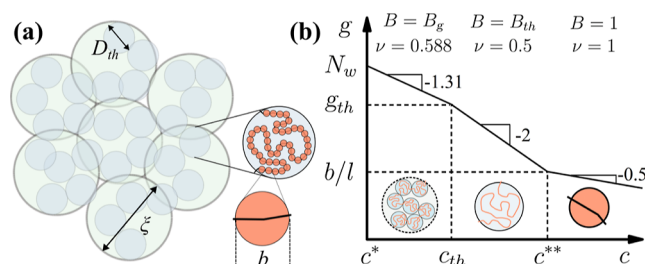


Figure 1. (a) Schematic representation of different length scales of semidilute polymer solutions in a good solvent. ξ is the solution correlation length, D_{th} is the thermal blob size, and b is the Kuhn length. (b) Concentration dependence of the number of repeat units per correlation blob, g , in semidilute polymer solutions $c > c^{**}$ with crossover concentrations c_{th} —thermal blob overlap concentration and c^{**} —crossover concentration to concentrated solution regime. g_{th} is the number of repeat units per thermal blob. Insets show chain conformations at the correlation length scale in different concentration regimes. Logarithmic scales.

with the corresponding scaling exponent $\nu = 1$ and coefficient $B = 1$.

Within a thermal blob, a section of a chain follows ideal chain statistics described by a scaling exponent $\nu = 0.5$, such that

$$r = (l b g_r)^{0.5} = \frac{l g_r^{0.5}}{B_{th}}, \quad \text{for } b \leq r \leq D_{th} \quad (2)$$

where we introduced

$$B_{th} = (l/b)^{0.5} \quad (3)$$

characterizing the chain flexibility.

On the length scales $D_{th} \leq r \leq \xi$, the effective excluded volume interactions between monomers in the presence of the solvent result in swelling. The statistics of the chain section on these length scales can be described by a self-avoiding walk of thermal blobs with scaling exponent $\nu = 0.588$, resulting in

$$r = D_{th} (g_r / g_{th})^\nu = \frac{l g_r^\nu}{B_g}, \quad \text{for } D_{th} < r \leq \xi \quad (4)$$

The value of the B_g parameter is determined by the chain flexibility and the excluded volume ν ⁵⁵

$$B_g = l(lb)^{3\nu-2} \nu^{1-2\nu} \quad (5)$$

Finally, on the length scale larger than solution correlation length, $\xi < r$, all interactions are screened, and a chain with the weight-average degree of polymerization N_w behaves as an ideal chain of N_w/g correlation blobs (Figure 1a), each containing g repeat units. The size of such a chain is estimated as

$$R = \xi (N_w/g)^{0.5} \quad (6)$$

The concentration dependence of the correlation blob ξ and the number of monomers in it g are derived by imposing the space filling condition on the correlation blobs, $c = g/\xi^3$. After some algebra, we arrive at the following expression for the correlation blob size

$$\xi = l B^{1/(3\nu-1)} (c l^3)^{\nu/(1-3\nu)} \quad (7)$$

and the number of monomers per correlation volume

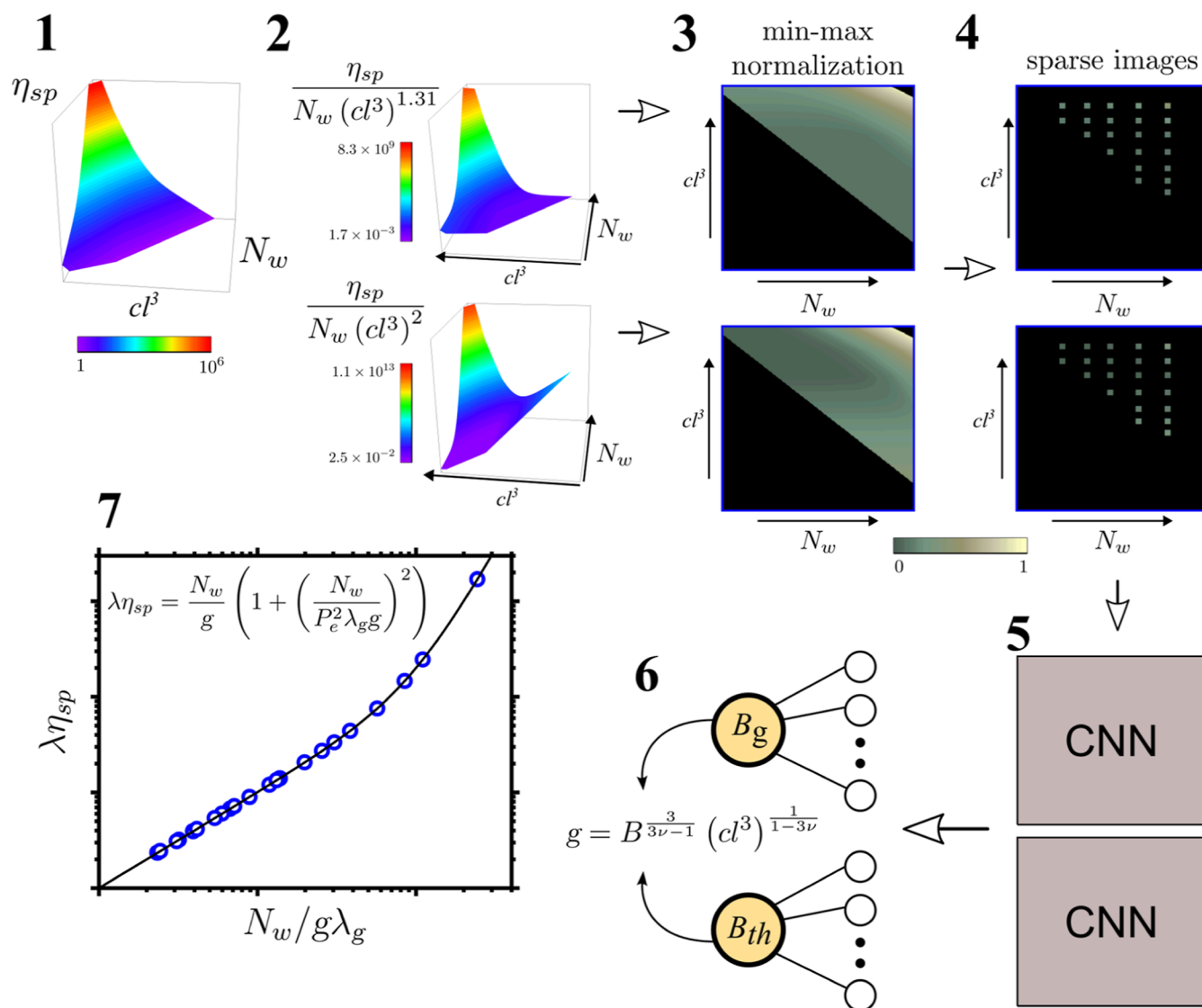


Figure 2. Workflow of obtaining the system-specific parameters B_g , B_{th} , and P_e . Generation of sparse images representing surfaces of normalized specific viscosity for determination of the B_g and B_{th} in the Rouse regime (1–4). Input of sparse images into CNNs to obtain B_g and B_{th} (5–6). Calculation of η_{sp} vs N_w/g for determination of P_e (7), where the blue circles represent values of N_w/g calculated using the obtained B_g and B_{th} parameters and the solid line represents the least-squares fit to the equation in step 7. The scaling parameters λ_g and λ are calculated from the crossover concentrations c_{th} and c^{**} . Multiplication factor $\lambda_g = 1$ for $c \leq c^{**}$ and $\lambda_g = (c/c^{**})^{-1.5}$ for $c^{**} < c$. Multiplication factor $\lambda = \lambda_g^{-1}$ for $c \leq c^{**}$ and $\lambda = \lambda_g^{-1} (c/c^{**})^{-0.5}$ for $c^{**} < c$. See the [Supporting Information](#) for details.

$$g = B^{3/(3\nu-1)} (cl^3)^{1/(1-3\nu)} \quad (8)$$

Figure 1b shows the concentration dependence of the number of monomers per correlation blob in different solution regimes.

The chain overlap concentration, c^* , corresponding to a crossover to the semidilute solution regime, is derived by setting the number of monomers per correlation volume (eq 8) to the weight-average chain degree of polymerization, $g = N_w$

$$c^* l^3 = B_g^3 N_w^{1-3\nu} \quad (9)$$

The crossover concentrations c_{th} and c^{**} between different solution regimes (Figure 1b) are determined by comparing the number of monomers per correlation blob g with the number of monomers per thermal blob g_{th} and the number of monomers per chain Kuhn length b/l . In particular, thermal

blobs start to overlap at monomer concentrations $c = c_{th}$ such that $g(c_{th}) = g_{th}$. This results in the following expression

$$c_{th} l^3 = B_{th}^3 (B_{th}/B_g)^{1/(2\nu-1)} \quad (10)$$

At monomer concentration $c = c^{**}$, the number of monomers per correlation blob becomes comparable to the number of monomers per chain Kuhn length $g(c^{**}) = b/l$. Rewriting this condition in terms of B_{th} , we have

$$c^{**} l^3 = B_{th}^4 \quad (11)$$

In the monomer concentration range such that $c > c^{**}$, chain sections within the correlation length are rod-like, which is reflected in the concentration dependence of $g = (cl^3)^{-0.5}$ (Figure 1b).

The scaling approach to the chain dynamics is built on the assumption of screening the hydrodynamic interactions on length scales on the order of the solution correlation length ξ .

This allows the derivation of a crossover expression for the solution specific viscosity η_{sp} of polymer chains with the weight-average degree of polymerization N_w and covering both the Rouse and entangled solution regimes

$$\eta_{sp} = N_w \left(1 + \left(\frac{N_w}{N_e} \right)^2 \right) \begin{cases} g^{-1}, & \text{for } c \leq c^{**} \\ cbl^2, & \text{for } c^{**} < c \end{cases} \quad (12)$$

where N_e is the number of repeat units per entanglement strand. For a polymer solution in a good solvent ($B_g \leq B_{th}^{4-6\nu}$), the concentration dependence of N_e is calculated according to the Kavassalis–Noolandi conjecture^{56–58}

$$N_e = P_e^2 \begin{cases} g, & \text{for } c \leq c^{**} \\ c^{-2}(lb)^{-3}, & \text{for } c^{**} < c \end{cases} \quad (13a)$$

For marginally good solvents with $B_g > B_{th}^{4-6\nu}$, the concentration dependence of N_e is calculated according to the Rubinstein–Colby conjecture^{59–61}

$$N_e = P_e^2 \begin{cases} g(c_{th}b^3)^{2/3}, & c \leq c_{th} \\ B_{th}^2(c l^3)^{-4/3}, & c_{th} < c \leq b^{-3} \\ g, & b^{-3} < c \leq c^{**} \\ B_{th}^{-2}(c^{**}/c)^2, & c^{**} < c \end{cases} \quad (13b)$$

The numerical constant P_e appearing in eqs 13a and 13b represents the packing number, describing the required number of overlapping strands per entanglement.^{27–29}

The scaling approach for neutral polymer solutions has been successfully applied to solutions of poly(ethylene oxide),^{41,50} polystyrene,^{34,51,52} poly(methyl methacrylate),⁴¹ poly(acrylonitrile-*co*-itaconic acid),⁴⁶ sodium hyaluronate,^{35–37} galactomannan,^{47,48,53} xanthan gum,⁴² sodium κ -carrageenan,^{38,54} cellulose,^{43–45,49} and various cellulose derivatives such as hydroxypropyl methyl cellulose,³⁸ methyl cellulose,³⁸ hydroxypropyl cellulose,³⁹ and cellulose tris(phenyl carbamate).⁴⁰ The method takes advantage of the concentration dependence of the solution specific viscosity in the Rouse (unentangled) solution regime ($N_w < N_e$) in extracting the set of B -parameters from plots of normalized specific viscosity, $\eta_{sp}/N_w(c l^3)^{1/(3\nu-1)}$.^{27–29} From this set, one can replot the solution specific viscosity in the universal form by using eqs 12 and 13a and obtain the packing number P_e through a simple least-squares fit.^{27,56–61} However, this manual method of data analysis is tedious (several hours can be spent analyzing experimental data by hand to determine if entanglements are present in a system), complex (understanding the calculation of different concentration regimes is important as this determines how to look at a particular system), and subject to human bias. In the next section, we show how we use the scaling method to procedurally generate input data for a CNN to streamline the automated analysis of solution viscosity and extract the polymer/solvent-specific B_{th} , B_g , and P_e parameters.

■ CNN APPLICATION TO ANALYSIS OF SOLUTION VISCOSITY

Building on the successful implementation of the CNN in analyzing images,^{62–65} we applied the CNN to the analysis of semidilute solution viscosity. Figure 2 outlines the workflow of the CNN implementation in determining the polymer/solvent-

specific parameters $\{B_g, B_{th}, P_e\}$ from the solution viscosity represented as a surface in the 3D space of (cl^3, N_w, η_{sp}) (step 1). The concentration and weight-average degree of polymerization data are logarithmically distributed within a fixed range. For example, row i of the data represents a reduced concentration of $\min(cl^3) e^{i\Delta}$, where Δ is determined by the maximum value and the desired number of rows. Thus, the viscosity as a function of cl^3 and N_w is represented as a matrix of fixed size. The analysis begins with the representation of the concentration dependence of the viscosity surface in the Rouse regime ($\eta_{sp} < 20$)

$$\eta_{sp} = N_w/g = N_w B^{3/(1-3\nu)} (cl^3)^{1/(3\nu-1)} \quad (14)$$

in terms of the ratios $\eta_{sp}/N_w(c l^3)^{1.31}$ (where $\nu = 0.588$ and $B = B_g$ for good solvent) and $\eta_{sp}/N_w(c l^3)^2$ (where $\nu = 0.5$ and $B = B_{th}$ for θ solvent) as functions of cl^3 and N_w (step 2). This is required for the determination of the values of the B -parameters in the corresponding solution regimes (Figure 1b). The obtained ratios and the values of cl^3 and N_w are transformed into the interval $[0, 1]$ through a log-normalization

$$x_{\text{norm}} = \frac{\ln x - \ln(\min x)}{\ln(\max x) - \ln(\min x)} \quad (15)$$

such that the minimum and maximum values of each axis of the 3D surfaces are mapped to 0 and 1, respectively (step 3). At this stage, the data are converted to images that are composed of 224×224 pixels. The location of each pixel is determined by the value of the normalized concentration cl^3 and weight-average degree of polymerization N_w , and each pixel has a value of the log-normalized specific viscosity. All values that lie outside the ranges of (cl^3, N_w, η_{sp}) are set to zero and are shown in black on the images in Figure 2. Each image is transformed into a sparse image (step 4) by randomly selecting between 1 and 13 different columns representing weight-average degrees of polymerization and between 5 and 13 different rows representing concentrations. All values of η_{sp} that are not located in the selected rows and columns are set to zero. This transformation is performed to represent the measured specific viscosity of a polymer solution with one or more molecular weights at discrete concentrations.

The prepared images are fed into CNNs (steps 5–6). We implement two separate networks with the input of $\eta_{sp}/N_w(c l^3)^{1.31}$ to predict B_g and the input of $\eta_{sp}/N_w(c l^3)^2$ to predict B_{th} . The calculated values of B_g and B_{th} parameters are used for the representation of the solution-specific viscosity in terms of a number of correlation blobs per chain (eqs 8, 11, 12, and 13b) to collapse datasets with different N_w values and to obtain the packing number P_e (step 7).

For training and optimization of the CNN, the data for the concentration dependence of specific viscosity were generated by using the scaling theory for known values of B_g , B_{th} , and P_e in the following parameter space: $3 \times 10^{-5} \leq cl^3 \leq 2 \times 10^{-2}$, $10^2 \leq N_w \leq 10^5$, and $1 \leq \eta_{sp} \leq 10^6$. This range of polymer solution parameters covers experimental datasets.^{27–29} Specifically, we generated three different types of systems: (1) systems that show only a value of B_g , (2) systems that show only a value of B_{th} , and (3) systems that show both values of B_g and B_{th} . The ranges of B_g , B_{th} , and P_e utilized to generate these systems are found in Table S4 of the Supporting Information.

We compare two different architectures of CNN: Visualization Geometry Group (in particular, the VGG13 model)³¹ and Inception (also known as GoogLeNet).³⁰ The VGG

architecture series uses sequential steps to extract information from images, while the Inception architecture utilizes several different elements in parallel and combines them together in an effort to learn more complex patterns. The layers along with the number of parameters learned in each layer of the two architectures are found in the [Supporting Information](#).

Each CNN architecture model used to predict either B_g or B_{th} is optimized and trained three separate times to demonstrate reproducibility. PyTorch⁶⁶ is utilized, and the two models used in this work are adapted from previously developed models from refs 30 and 31. The Adaptive Moment Estimation (Adam) optimization algorithm⁶⁷ with the hyperparameters of learning rate (step size of an optimization through a parameter space), β_1 and β_2 (decay rates of the moving average of the gradient and gradient squared, respectively), and ϵ (parameter for numerical stability in each optimization step) are tuned using Optuna.⁶⁸ This tuning process is described in the [Supporting Information](#). The trial with the hyperparameters showing the lowest average training error for the last epoch is used. The training, validation, and testing of the architectures are run for 300 epochs. For each epoch, we use 51,200 images generated through steps 1–5 shown in [Figure 2](#) to train the models to predict B_g and B_{th} . After the training, 21,952 images, also generated through steps 1–5, are used to validate the models that are trained to predict B_g and B_{th} . After the validation step, we tested the models on 37 images that represent the experimental data we manually analyzed and obtained the B_g and B_{th} parameters.^{27–29} The loss function utilized for training the models is the mean-squared error

$$MSE = \frac{1}{m} \sum_{i=1}^m (h_i - y_i)^2 \quad (16)$$

where m represents the number of datasets, h_i is the predicted value of a parameter, and y_i is the true value of the parameter. To establish how close the predictions of the B -parameters are to our previous calculations of the system-specific B -parameters, we quantify the accuracy of the B -parameter predictions as the mean absolute percentage difference (MAPD)

$$MAPD = \frac{100\%}{m} \sum_{i=1}^m \frac{|h_i - y_i|}{y_i} \quad (17)$$

where m represents the number of datasets, h_i is the predicted value of a parameter, and y_i is the hand-determined value of the parameter.

CNN RESULTS

[Table 1](#) displays the results of three separate trials of optimization and training of both VGG and Inception-based

Table 1. Mean Absolute Percentage Differences of B_g and B_{th} in Validation Step of Theoretically Generated Data

architecture	MAPD, B_g	MAPD, B_{th}
VGG, trial 1	2.29	7.47
VGG, trial 2	2.45	7.13
VGG, trial 3	2.03	4.46
Inception, trial 1	5.01	5.88
Inception, trial 2	5.70	5.29
Inception, trial 3	5.28	5.10

models, all performing with MAPD less than 10%. The VGG model performs well in predicting the B_g -parameter values with what we have found previously with values less than 3% of the MAPD. However, this model does not perform as well at predicting B_{th} parameters from the sparse images. This is not surprising due to the possibility of entanglements covering the thermal blob regime with polymer solutions that have a higher molecular weight.

To visualize how close the VGG and Inception models perform in agreement with our previous analyses of the same systems (which we will term “manual” analysis), we use parity plots to track the predicted values of B -parameters versus the manually obtained B -parameters. The parity models of predicted values of B_g versus observed values of B_g are shown in [Figure 3](#), where points along the dashed line reflect predictions of B -parameters in perfect agreement with manual analysis. The results of the VGG model in [Figure 3a](#) and the Inception model in [Figure 3b](#) highlight two types of systems: (1) systems with only B_g parameters (black circles) and (2) systems with both B_g and B_{th} parameters (red squares). Both models predicted B_g well in agreement for systems in which we only determined a B_g parameter previously, with the VGG model producing a MAPD of 1.61% and the Inception model generating a MAPD of 3.06%.

Regarding systems that have both B_g and B_{th} , the VGG model's accuracy was reflected by a MAPD of 2.74%, while the Inception model's accuracy resulted in a MAPD of 8.25%. The B_g -parameter is extracted by the plateau of $\eta_{sp}/N_w(c)^{1.31}$ versus concentration; this plateau can be found at lower concentrations for all molecular weights in the semidilute Rouse regime. The VGG network is more adept at interpreting this data than the Inception network, as it may be too simple a trend to pick up for such a complex architecture. Overall, both models have higher performance in predicting B_g parameters for systems that are shown to only have a B_g parameter. One example highlighting the behavior of these predictions is for aqueous poly(ethylene oxide). For the system with the manually obtained B_g value of 1.111, the VGG model predicted a B_g of 1.105, and the Inception model predicted a B_g of 1.089. However, if we take a system with more molecular weights such that we highlight a thermal regime (having both B_g and B_{th} values), the models both overestimate B_g (VGG model and Inception model predictions of 1.245 and 1.345, respectively, vs the manually obtained value of 1.119).

The results of the predictions of the B_{th} parameters are shown in [Figure 4a](#) for the VGG model and [Figure 4b](#) for the Inception-based model. For systems with only B_{th} parameters, the VGG model obtained a MAPD of 2.05% compared to the Inception model's MAPD from the hand-determined B_{th} value of 2.75%. Overall, these results show that both models may be used to determine the B -parameters, which is well in agreement with the results that one can obtain manually.

With systems containing both B_g and B_{th} parameters, the VGG model showed a MAPD of 5.46%, while the Inception model had a MAPD of 6.08%. Determining whether a system has a B_{th} parameter has always been a tedious task. Here, we know that a system would probably have only a B_{th} -parameter if the specific viscosity scales with the square of the concentration in the Rouse regime at lower concentrations. For systems that have both a B_g and B_{th} parameter, the specific viscosity should scale with a concentration power law of $1.31 = 1/(3\nu - 1)$ with $\nu = 0.588$ in the Rouse regime. However, one must determine if the minimum of $\eta_{sp}/N_w(c)^2$ is the result of

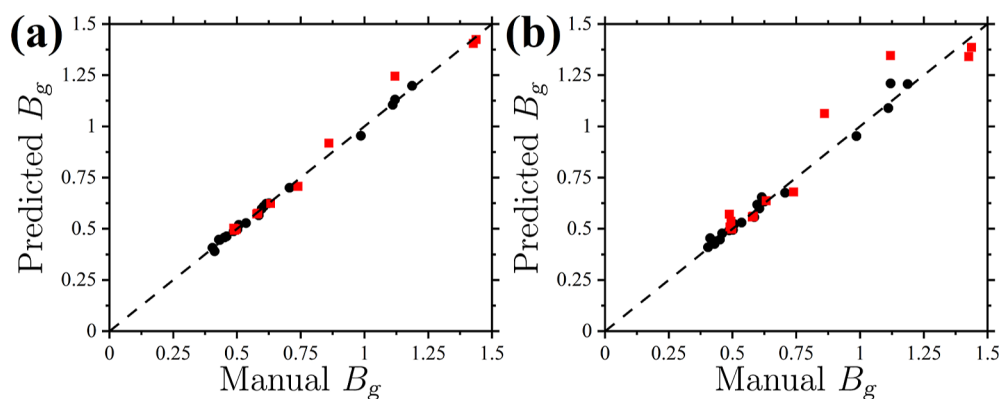


Figure 3. Parity plots for VGG (a) and Inception (b) model training best runs for the B_g -parameter prediction. Symbol notations: systems with only manually obtained B_g parameters (black circles) and systems with both manually obtained B_g and B_{th} parameters (red squares).

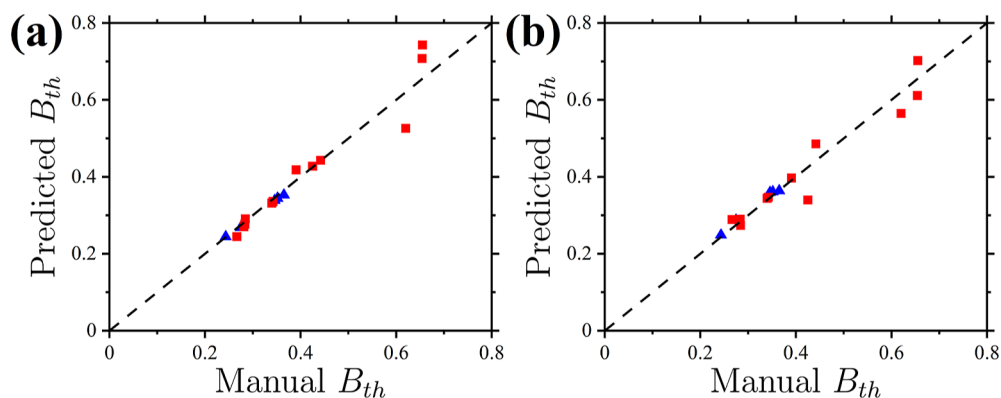


Figure 4. Parity plots for VGG (a) and Inception (b) model training best runs for B_{th} -parameter prediction. Symbol notations: systems with only manually obtained B_{th} parameters (blue triangles) and systems with both manually obtained B_g and B_{th} parameters (red squares).

a thermal plateau or entanglements from the good solvent Rouse regime. For lower molecular weights, it can be seen that the plateau is indeed the real thermal plateau (see the $\eta_{sp}/N_w(cI^3)^2$ surface example in Figure 2), while for higher molecular weights, the minimum is disguised by entanglements before it reaches the absolute minimum of the surface.

Table 2 shows the predicted values of the B -parameters and P_e together with corresponding values obtained by manual analysis^{27–29} of the specific viscosity data.^{34–54} In particular, we present the B_g and B_{th} parameters for the VGG model of trial 3, which had the lowest MAPD. For each predicted value of B_g and B_{th} , we recalculate the number of blobs per chain, N_w/g , and use this information for the representation of the solution viscosity and determination of the packing number, P_e , from curve fitting

$$\lambda\eta_{sp} = \frac{N_w}{g\lambda_g} \left(1 + \left(\frac{N_w}{P_e^2 g \lambda_g} \right)^2 \right) \quad (18)$$

where both λ_g and λ are multiplication factors that are given in the Supporting Information. The deviations of the predicted packing numbers from the manually obtained packing numbers are due to differences in the B_g and B_{th} parameters. For most systems, we observe that systems with a higher B -parameter predicted by the CNN than the manually determined values result in higher values of g (eq 8) and, in turn, lower N_w/g values and smaller values of P_e . The main reason behind this is the sensitivity of packing number P_e to the value of the crossover concentrations determined by B_{th} . Using the

predicted values of packing numbers, in Figure 5, we reconstruct a universal representation of crossover between the Rouse and entangled regimes described by

$$\frac{\lambda\eta_{sp}}{P_e^2} = \frac{N_w}{N_e} \left(1 + \left(\frac{N_w}{N_e} \right)^2 \right) \quad (19)$$

where the slope of 1 represents the Rouse regime transforming to a slope of 3 in the entangled regime. The good collapse of the data provides further justification for the use of the CNN-based approach in the analysis of the solution viscosity.

CONCLUSIONS

We demonstrate the successful application of the CNN, trained and optimized on the datasets generated by the scaling theory of semidilute polymer solutions, to the automated analysis of experimental viscosity data. The developed framework provides a polymer/solvent-specific set of B -parameters (Figures 3 and 4) together with crossover concentrations into different solution regimes (Figure 1b). This information is used to calculate the packing parameter P_e . The obtained set of system-specific parameters $\{B_{th}, B_g, P_e\}$ is validated by comparing it with the manually determined set and by confirming a universal representation of the solution viscosity in the entire concentration range, covering both the Rouse and entangled semidilute solution regimes in terms of the reduced variables (Figure 5).

The developed approach could be extended to analyze experimental data on osmotic pressure^{69,70} and diffusion^{71–73}

Table 2. Summary of B -Parameter and P_e Predictions from the VGG Model

Ref	Polymer	Solvent	Symbol	$B_{g,manual}$	$B_{g,pred}$	$B_{th,manual}$	$B_{th,pred}$	$P_{e,manual}$	$P_{e,fit}$
34	Polystyrene	Tetrahydrofuran	●	0.615	0.622			6.3	5.9
35	Sodium hyaluronate	Water	★	0.605	0.609			3.7	3.1
36, 37	Sodium hyaluronate	Water, PBS	●	0.597	0.597			3.2	2.9
				0.985	0.954				
				0.535	0.527				
				0.626	0.624				
38	Hydroxypropyl methyl cellulose	Water	■	0.429	0.446			3.0	2.6
38	Methyl cellulose	Water	■	0.433	0.445			3.4	3.1
39	Hydroxypropyl cellulose	Water	▲	0.502	0.497			3.4	3.6
39	Hydroxypropyl cellulose	Water	▲	0.487	0.488			2.9	2.9
39	Hydroxypropyl cellulose	Water	▲	0.507	0.519			2.7	2.5
40	Cellulose tris(phenyl carbamate)	Tetrahydrofuran	○	0.460	0.463			3.4	3.4
41	Polyethylene oxide	[BMIm][TFSI]	○	1.187	1.197			3.5	3.6
42	Xanthan gum	Water	▼	1.120	1.130			2.6	3.0
41	Polyethylene oxide	Water	○	0.413	0.390			3.9	3.9
43	Cellulose	[AMIm][Cl]	■	1.111	1.105			3.0	2.6
44	Cellulose	[EMIm][P(OCH ₃) ₂ (H)O ₂]	▼	0.404	0.407			3.6	3.0
45	Cellulose	[BMIm][Cl]	■	0.450	0.457			3.6	3.4
46	Poly(acrylonitrile-co-itaconic-acid)	[BMIm][Cl]	◆	0.585	0.566			3.7	3.6
47	Galactomannan - guar gum	Water	◆	0.706	0.700	0.275	0.268	6.5	8.0
48	Galactomannan - <i>Cassia nodosa</i>	Water	◆			0.243	0.244	7.2	6.8
49	Cellulose	[BMIm][Ac]	★			0.346	0.339	8.3	11.9
						0.352	0.344		
						0.365	0.353		
50	Polyethylene oxide	Water	●	1.119	1.245	0.620	0.526	7.2	9.2
51, 52	Polystyrene	Toluene	●	0.739	0.707	0.425	0.427	4.4	4.6
53	Galactomannan - <i>M. flocculosa</i>	Water	●	0.487	0.503	0.284	0.291	3.3	3.0
53	Galactomannan - <i>M. bimicronata</i>	Water	●	0.494	0.499	0.266	0.244	7.2	11.4
53	Galactomannan - <i>M. aspericarpa</i>	Water	●	0.490	0.493	0.281	0.270	5.8	6.6
53	Galactomannan - <i>M. taimensis</i>	Water	●	0.497	0.497	0.283	0.277	7.6	8.4
38, 54	Sodium kappa-carrageenan	Water	▲	0.632	0.623	0.391	0.418	10.8	7.2
41	Polyethylene oxide	[BMIm][PF ₆]	●	1.426	1.404	0.655	0.743	11.5	9.3
				1.437	1.424	0.654	0.707		
49	Cellulose	[BMIm][Ac]	★	0.578	0.572	0.340	0.332	8.2	8.8
				0.582	0.575	0.342	0.336		
41	Polymethyl methacrylate	[BMIm][TFSI]	▲	0.860	0.918	0.442	0.443	9.9	9.9

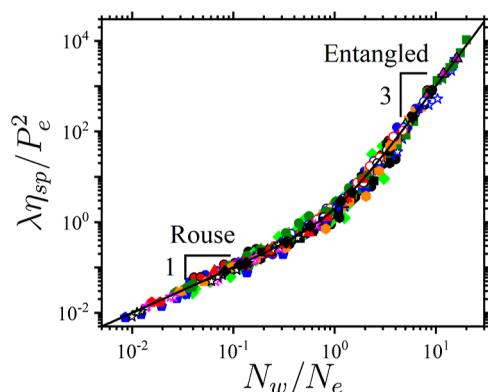


Figure 5. Dependence of normalized specific viscosity $\lambda\eta_{sp}/P_e^2$ vs number of entanglements per chain N_w/N_e for all systems using the VGG model. Symbol notations are listed in Table 2.

in semidilute solutions. The determination of the $\{B_{th}, B_g, P_e\}$ set from other experimental techniques should provide further justification for the approach validity as well as present a complete toolset for solution characterization.

Future research should also focus on connecting the set of B -parameters with polymer/solvent molecular identifiers such as BigSMILES³ or Morgan fingerprinting.⁷⁴ By generating a comprehensive library combining specific viscosity, diffusion, and osmotic pressure data for various polymer/solvent systems, one will be able to use AI to elucidate hidden

correlations between a specific chemical structure of polymers and solvents with their coarse-grained properties exemplified by B sets.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsapm.3c01282>.

CNN architecture parameters, calculation of λ and λ_g , model training, and Inception model results (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Andrey V. Dobrynin – Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina 27599-3290, United States; orcid.org/0000-0002-6484-7409; Email: avd@email.unc.edu

Authors

Ryan Sayko – Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina 27599-3290, United States; orcid.org/0000-0002-5986-4829

Michael Jacobs – Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina 27599-3290, United States; Present Address: Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak

Ridge, Tennessee 37831; orcid.org/0000-0002-7255-3451

Marissa Dominijanni – Department of Computer Science and Engineering, University at Buffalo, Buffalo, New York 14260-1660, United States

Complete contact information is available at:
<https://pubs.acs.org/10.1021/acsapm.3c01282>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under grant DMR 2049518.

REFERENCES

- (1) Pennington, J.; Socher, R.; Manning, C. D. *GloVe: Global Vectors for Word Representation*; EMNLP, 2014, pp 1532–1543.
- (2) Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient Estimation of Word Representations in Vector Space, 2013. [arXiv:1301.3781](https://arxiv.org/abs/1301.3781).
- (3) Lin, T.-S.; Coley, C. W.; Mochigase, H.; Beech, H. K.; Wang, W.; Wang, Z.; Woods, E.; Craig, S. L.; Johnson, J. A.; Kalow, J. A.; Jensen, K. F.; Olsen, B. D. BigSMILES: A Structurally-Based Line Notation for Describing Macromolecules. *ACS Cent. Sci.* **2019**, *5* (9), 1523–1531.
- (4) Ma, R.; Liu, Z.; Zhang, Q.; Liu, Z.; Luo, T. Evaluating Polymer Representations via Quantifying Structure-Property Relationships. *J. Chem. Inf. Model.* **2019**, *59* (7), 3110–3119.
- (5) Kim, C.; Chandrasekaran, A.; Huan, T. D.; Das, D.; Ramprasad, R. Polymer Genome: A Data-Powered Polymer Informatics Platform for Property Predictions. *J. Phys. Chem. C* **2018**, *122* (31), 17575–17585.
- (6) Otsuka, S.; Kuwajima, I.; Hosoya, J.; Xu, Y.; Yamazaki, M. In *PoLyInfo: Polymer Database for Polymeric Materials Design*; 2011 International Conference on Emerging Intelligent Data and Web Technologies, 7–9 Sept. 2011, 2011; pp 22–29.
- (7) Lafferty, M. CHEMnetBASE. *Issues Sci. Technol. Librariansh.* **2007**, *50*.
- (8) Polymer Property Predictor and Database. <https://pppdb.uchicago.edu/>.
- (9) Gormley, A. J.; Webb, M. A. Machine learning in combinatorial polymer chemistry. *Nat. Rev. Mater.* **2021**, *6* (8), 642–644.
- (10) Reis, M.; Gusev, F.; Taylor, N. G.; Chung, S. H.; Verber, M. D.; Lee, Y. Z.; Isayev, O.; Leibfarth, F. A. Machine-Learning-Guided Discovery of 19F MRI Agents Enabled by Automated Copolymer Synthesis. *J. Am. Chem. Soc.* **2021**, *143* (42), 17677–17689.
- (11) Lee, J.; Mulay, P.; Tamasi, M. J.; Yeow, J.; Stevens, M. M.; Gormley, A. J. A fully automated platform for photoinitiated RAFT polymerization. *Digit. Discov.* **2023**, *2* (1), 219–233.
- (12) Audus, D. J.; de Pablo, J. J. Polymer Informatics: Opportunities and Challenges. *ACS Macro Lett.* **2017**, *6* (10), 1078–1082.
- (13) Le, T.; Epa, V. C.; Burden, F. R.; Winkler, D. A. Quantitative Structure-Property Relationship Modeling of Diverse Materials Properties. *Chem. Rev.* **2012**, *112* (5), 2889–2919.
- (14) Yan, C.; Li, G. The Rise of Machine Learning in Polymer Discovery. *Adv. Intell. Syst.* **2023**, *5* (4), 2200243.
- (15) Seddon, D.; Müller, E. A.; Cabral, J. T. Machine learning hybrid approach for the prediction of surface tension profiles of hydrocarbon surfactants in aqueous solution. *J. Colloid Interface Sci.* **2022**, *625*, 328–339.
- (16) Patel, R. A.; Borca, C. H.; Webb, M. A. Featurization strategies for polymer sequence or composition design by machine learning. *Mol. Syst. Des. Eng.* **2022**, *7* (6), 661–676.
- (17) Webb, M. A.; Jackson, N. E.; Gil, P. S.; de Pablo, J. J. Targeted sequence design within the coarse-grained polymer genome. *Sci. Adv.* **2020**, *6* (43), No. eabc6216.
- (18) Banerjee, A.; Hsu, H.-P.; Kremer, K.; Kukharenko, O. Data-Driven Identification and Analysis of the Glass Transition in Polymer Melts. *ACS Macro Lett.* **2023**, *12* (6), 679–684.
- (19) Heil, C. M.; Patil, A.; Dhinojwala, A.; Jayaraman, A. Computational Reverse-Engineering Analysis for Scattering Experiments (CREASE) with Machine Learning Enhancement to Determine Structure of Nanoparticle Mixtures and Solutions. *ACS Cent. Sci.* **2022**, *8* (7), 996–1007.
- (20) Beltran-Villegas, D. J.; Wessels, M. G.; Lee, J. Y.; Song, Y.; Wooley, K. L.; Pochan, D. J.; Jayaraman, A. Computational Reverse-Engineering Analysis for Scattering Experiments on Amphiphilic Block Polymer Solutions. *J. Am. Chem. Soc.* **2019**, *141* (37), 14916–14930.
- (21) Archibald, R. K.; Doucet, M.; Johnston, T.; Young, S. R.; Yang, E.; Heller, W. T. Classifying and analyzing small-angle scattering data using weighted k nearest neighbors machine learning techniques. *J. Appl. Crystallogr.* **2020**, *53* (2), 326–334.
- (22) Yager, K. G.; Majewski, P. W.; Noack, M. M.; Fukuto, M. Autonomous x-ray scattering. *Nanotechnology* **2023**, *34* (32), 322001.
- (23) Stein, H. S.; Gregoire, J. M. Progress and prospects for accelerating materials science with automated and autonomous workflows. *Chem. Sci.* **2019**, *10* (42), 9640–9649.
- (24) Dong, H.; Butler, K. T.; Matras, D.; Price, S. W. T.; Odarchenko, Y.; Khatry, R.; Thompson, A.; Middelkoop, V.; Jacques, S. D. M.; Beale, A. M.; Vamvakeros, A. A deep convolutional neural network for real-time full profile analysis of big powder diffraction data. *Npj Comput. Mater.* **2021**, *7* (1), 74.
- (25) Chuang, K. V.; Gunsalus, L. M.; Keiser, M. J. Learning Molecular Representations for Medicinal Chemistry. *J. Med. Chem.* **2020**, *63* (16), 8705–8722.
- (26) Martin, T. B.; Audus, D. J. Emerging Trends in Machine Learning: A Polymer Perspective. *ACS Polym. Au* **2023**, *3* (3), 239–258.
- (27) Dobrynin, A. V.; Jacobs, M.; Sayko, R. Scaling of Polymer Solutions as a Quantitative Tool. *Macromolecules* **2021**, *54* (5), 2288–2295.
- (28) Sayko, R.; Jacobs, M.; Dobrynin, A. V. Quantifying Properties of Polysaccharide Solutions. *ACS Polym. Au* **2021**, *1* (3), 196–205.
- (29) Sayko, R.; Jacobs, M.; Dobrynin, A. V. Universality in Solution Properties of Polymers in Ionic Liquids. *ACS Appl. Polym. Mater.* **2022**, *4* (3), 1966–1973.
- (30) Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the Inception Architecture for Computer Vision, 2015. [arXiv:1512.00567](https://arxiv.org/abs/1512.00567). <https://arxiv.org/abs/1512.00567>.
- (31) Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition, 2014. [arXiv:1409.1556](https://arxiv.org/abs/1409.1556). <https://arxiv.org/abs/1409.1556>.
- (32) LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521* (7553), 436–444.
- (33) LeCun, Y.; Boser, B.; Denker, J. S.; Henderson, D.; Howard, R. E.; Hubbard, W.; Jackel, L. D. Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Comput.* **1989**, *1* (4), 541–551.
- (34) Jamieson, A. M.; Telford, D. Newtonian viscosity of semidilute solutions of polystyrene in tetrahydrofuran. *Macromolecules* **1982**, *15* (5), 1329–1332.
- (35) Fouissac, E.; Milas, M.; Rinaudo, M. Shear-rate, concentration, molecular weight, and temperature viscosity dependences of hyaluronate, a wormlike polyelectrolyte. *Macromolecules* **1993**, *26* (25), 6945–6951.
- (36) Morris, E. R.; Rees, D. A.; Welsh, E. J. Conformation and dynamic interactions in hyaluronate solutions. *J. Mol. Biol.* **1980**, *138* (2), 383–400.
- (37) Krause, W. E.; Bellomo, E. G.; Colby, R. H. Rheology of Sodium Hyaluronate under Physiological Conditions. *Biomacromolecules* **2001**, *2* (1), 65–69.
- (38) Potier, M.; Tea, L.; Benyahia, L.; Nicolai, T.; Renou, F. Viscosity of Aqueous Polysaccharide Solutions and Selected

Homogeneous Binary Mixtures. *Macromolecules* **2020**, *53* (23), 10514–10525.

(39) Lopez, C. G.; Voleske, L.; Richtering, W. Scaling laws of entangled polysaccharides. *Carbohydr. Polym.* **2020**, *234*, 115886.

(40) Sato, T.; Hamada, M.; Teramoto, A. Solution Viscosity of a Moderately Stiff Polymer: Cellulose Tris(phenyl carbamate). *Macromolecules* **2003**, *36* (18), 6840–6843.

(41) Liu, F.; Lv, Y.; Liu, J.; Yan, Z.-C.; Zhang, B.; Zhang, J.; He, J.; Liu, C.-Y. Crystallization and Rheology of Poly(ethylene oxide) in Imidazolium Ionic Liquids. *Macromolecules* **2016**, *49* (16), 6106–6115.

(42) Milas, M.; Rinaudo, M.; Tinland, B. The viscosity dependence on concentration, molecular weight and shear rate of xanthan solutions. *Polym. Bull.* **1985**, *14* (2), 157–164.

(43) Kuang, Q.-L.; Zhao, J.-C.; Niu, Y.-H.; Zhang, J.; Wang, Z.-G. Celluloses in an Ionic Liquid: the Rheological Properties of the Solutions Spanning the Dilute and Semidilute Regimes. *J. Phys. Chem. B* **2008**, *112* (33), 10234–10240.

(44) Chen, X.; Liang, S.; Wang, S.-W.; Colby, R. H. Linear viscoelastic response and steady shear viscosity of native cellulose in 1-ethyl-3-methylimidazolium methylphosphonate. *J. Rheol.* **2018**, *62* (1), 81–87.

(45) Chen, X.; Zhang, Y.; Wang, H.; Wang, S.-W.; Liang, S.; Colby, R. H. Solution rheology of cellulose in 1-butyl-3-methyl imidazolium chloride. *J. Rheol.* **2011**, *55* (3), 485–494.

(46) Zhu, X.; Chen, X.; Saba, H.; Zhang, Y.; Wang, H. Linear viscoelasticity of poly(acrylonitrile-co-itaconic acid)/1-butyl-3-methylimidazolium chloride extended from dilute to concentrated solutions. *Eur. Polym. J.* **2012**, *48* (3), 597–603.

(47) Robinson, G.; Ross-Murphy, S. B.; Morris, E. R. Viscosity-molecular weight relationships, intrinsic chain flexibility, and dynamic solution properties of guar galactomannan. *Carbohydr. Res.* **1982**, *107* (1), 17–32.

(48) Kapoor, V. P.; Milas, M.; Taravel, F. R.; Rinaudo, M. Rheological properties of seed galactomannan from *Cassia nodosa* buch.-hem. *Carbohydr. Polym.* **1994**, *25* (2), 79–84.

(49) Lefroy, K. S.; Murray, B. S.; Ries, M. E. Rheological and NMR Studies of Cellulose Dissolution in the Ionic Liquid BmimAc. *J. Phys. Chem. B* **2021**, *125* (29), 8205–8218.

(50) Ebagninin, K. W.; Benchabane, A.; Bekkour, K. Rheological characterization of poly(ethylene oxide) solutions of different molecular weights. *J. Colloid Interface Sci.* **2009**, *336* (1), 360–367.

(51) Kulicke, W. M.; Kniewske, R. The shear viscosity dependence on concentration, molecular weight, and shear rate of polystyrene solutions. *Rheol. Acta* **1984**, *23* (1), 75–83.

(52) Zakin, J. L.; Wu, R.; Luh, H.; Mayhan, K. G. Generalized correlations for molecular weight and concentration dependence of zero-shear viscosity of high polymer solutions. *J. Polym. Sci. B. Polym. Phys.* **1976**, *14* (2), 299–308.

(53) Ganter, J. L. M. S.; Reicher, F. Water-soluble galactomannans from seeds of Mimosaceae spp. *Bioresour. Technol.* **1999**, *68* (1), 55–62.

(54) Croguennoc, P.; Meunier, V.; Durand, D.; Nicolai, T. Characterization of Semidilute κ -Carrageenan Solutions. *Macromolecules* **2000**, *33* (20), 7471–7474.

(55) Rubinstein, M.; Colby, R. H. *Polymer Physics*; Oxford University Press: Oxford, 2007.

(56) Kavassalis, T. A.; Noolandi, J. New View of Entanglements in Dense Polymer Systems. *Phys. Rev. Lett.* **1987**, *59* (23), 2674–2677.

(57) Kavassalis, T. A.; Noolandi, J. A new theory of entanglements and dynamics in dense polymer systems. *Macromolecules* **1988**, *21* (9), 2869–2879.

(58) Kavassalis, T. A.; Noolandi, J. Entanglement scaling in polymer melts and solutions. *Macromolecules* **1989**, *22* (6), 2709–2720.

(59) Colby, R. H.; Rubinstein, M. Two-parameter scaling for polymers in Θ solvents. *Macromolecules* **1990**, *23* (10), 2753–2757.

(60) Heo, Y.; Larson, R. G. Universal Scaling of Linear and Nonlinear Rheological Properties of Semidilute and Concentrated Polymer Solutions. *Macromolecules* **2008**, *41* (22), 8903–8915.

(61) Milner, S. T. Predicting the Tube Diameter in Melts and Solutions. *Macromolecules* **2005**, *38* (11), 4929–4939.

(62) Tao, L.; Varshney, V.; Li, Y. Benchmarking Machine Learning Models for Polymer Informatics: An Example of Glass Transition Temperature. *J. Chem. Inf. Model.* **2021**, *61* (11), 5395–5413.

(63) Tao, L.; Chen, G.; Li, Y. Machine learning discovery of high-temperature polymers. *Patterns* **2021**, *2* (4), 100225.

(64) Yan, X.; Zhang, J.; Russo, D. P.; Zhu, H.; Yan, B. Prediction of Nano-Bio Interactions through Convolutional Neural Network Analysis of Nanostructure Images. *ACS Sustain. Chem. Eng.* **2020**, *8* (51), 19096–19104.

(65) Xu, L.; Xu, L.; Chen, Y.; Zhang, Y.; Yang, J. Accurate Classification of Algae Using Deep Convolutional Neural Network with a Small Database. *ACS EST Water* **2022**, *2*, 1921–1928.

(66) Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; Desmaison, A.; Köpf, A.; Yang, E.; DeVito, Z.; Raison, M.; Tejani, A.; Chilamkurthy, S.; Steiner, B.; Fang, L.; Bai, J.; Chintala, S. PyTorch: An Imperative Style, High-Performance Deep Learning Library, 2019. arXiv:1912.01703. <https://arxiv.org/abs/1912.01703>.

(67) Kingma, D. P.; Ba, J. Adam: A Method for Stochastic Optimization, 2014. arXiv:1412.6980. <https://arxiv.org/abs/1412.6980>.

(68) Akiba, T.; Sano, S.; Yanase, T.; Ohta, T.; Koyama, M. Optuna: A Next-generation Hyperparameter Optimization Framework, 2019. arXiv:1907.10902. <https://arxiv.org/abs/1907.10902>.

(69) Cohen, J. A.; Podgornik, R.; Hansen, P. L.; Parsegian, V. A. A Phenomenological One-Parameter Equation of State for Osmotic Pressures of PEG and Other Neutral Flexible Polymers in Good Solvents. *J. Phys. Chem. B* **2009**, *113* (12), 3709–3714.

(70) Noda, I.; Kato, N.; Kitano, T.; Nagasawa, M. Thermodynamic properties of moderately concentrated solutions of linear polymers. *Macromolecules* **1981**, *14* (3), 668–676.

(71) Von Meerwall, E. D.; Amis, E. J.; Ferry, J. D. Self-diffusion in solutions of polystyrene in tetrahydrofuran: comparison of concentration dependences of the diffusion coefficients of polymer, solvent, and a ternary probe component. *Macromolecules* **1985**, *18* (2), 260–266.

(72) Zettl, U.; Hoffmann, S. T.; Koberling, F.; Krausch, G.; Enderlein, J.; Harnau, L.; Ballauff, M. Self-Diffusion and Cooperative Diffusion in Semidilute Polymer Solutions As Measured by Fluorescence Correlation Spectroscopy. *Macromolecules* **2009**, *42* (24), 9537–9547.

(73) Oostwal, M.; Odijk, T. Novel dynamic scaling hypothesis for semidilute and concentrated solutions of polymers and polyelectrolytes. *Macromolecules* **1993**, *26* (24), 6489–6497.

(74) Rogers, D.; Hahn, M. Extended-Connectivity Fingerprints. *J. Chem. Inf. Model.* **2010**, *50* (5), 742–754.