Journal of Rheology

RESEARCH ARTICLE | OCTOBER 08 2025

Using neural networks to deduce polymer molecular weight distributions from linear rheology

Robert J. Elliott ; Luisa Cutillo ; Chinmay Das ; Johan Mattsson ; Daniel J. Read

J. Rheol. 69, 955–972 (2025)
https://doi.org/10.1122/8.0001063

Using neural networks to deduce polymer molecular weight distributions from linear rheology

Robert J. Elliott,[1,2,a)] Luisa Cutillo,[1] Chinmay Das,[1] Johan Mattsson,[3] and Daniel J. Read[1,b)]

[1]*School of Mathematics, University of Leeds, Leeds LS2 9JT, United Kingdom*
[2]*DPI, P.O. Box 902, 5600 AX Eindhoven, The Netherlands*
[3]*School of Physics and Astronomy, University of Leeds, Leeds LS2 9JT, United Kingdom*

(Received 17 June 2025; final revision received 15 September 2025; published 8 October 2025)

Abstract

We present a methodology for inferring the molecular weight distribution (MWD) of polydisperse linear polymers from their linear rheology using machine learning techniques. Specifically, we use a state-of-the-art tube model to generate large datasets of artificially produced rheology data. These are used to train neural networks (NNs) to make accurate MWD predictions from frequency-sweep rheology measurements. We target distributions relevant to commercial polymers, so broad polydisperse MWDs are prioritized. To simplify the data format for the NN, we fit Maxwell modes to the rheology with predefined relaxation times and, hence, parameterize the rheology using the mode amplitudes; correspondingly, we propose MWD parameterization using the sum of several log-Gaussian subdistributions with logarithmically spaced mean molecular weights and identical dispersities. We assess the methodology's performance by predicting MWDs using experimental polystyrene rheology data from the literature. Good agreement with gel permeation chromatography data is found where available, and where it is not, the prediction captures known molecular weight statistics (such as weight-average molecular weight and dispersity) even if the precise shape of the MWD is not known. The findings here lay the groundwork for future developments concerning the inversion of this tube model for other polymeric materials. The ability to infer the MWD from rheology would traditionally be prohibited by the mathematical complexity of state-of-the-art tube models, but we bypass this issue with our machine learning methodology. https://doi.org/10.1122/8.0001063

I. INTRODUCTION

All commercial polymeric materials are polydisperse, and their molecular weight distribution (MWD) is a key factor in determining material properties and commercial use. Gel permeation chromatography (GPC) is the standard technique used to measure the MWD, but this has some drawbacks: GPC can be expensive, time-consuming, and occasionally require the use of toxic solvents. Also, very high molecular weight components of a polymer melt are sometimes not detected by GPC despite having profound effects on the material's mechanical properties. These factors limit the use of GPC for high-volume tasks, such as rapid characterization of varied feedstock, e.g., as might be required during mechanical recycling. Therefore, we present a new method for acquiring the MWD from the melt rheology of the polymer, bypassing the need to perform a GPC measurement. Rheology data are rich in molecular information and can be acquired more easily than GPC data in most cases. Specifically, the data we use are the dynamic moduli acquired in a frequency-sweep experiment.

The possibility of inferring molecular weight information from the melt rheology is greatly assisted by the extensive work that has been done to predict rheology from a molecular structure. The tube model of de Gennes [1] and Doi and Edwards [2], leading to the modern forms of the tube model [3], has achieved great success in modeling linear rheology from material parameters and the MWD. This was achieved by simplifying the complex dynamics of polymer melts to a model of the escape of a test polymer chain within a tubelike confinement, representing entanglement interactions with the surrounding chains. This allows modeling of the viscoelastic behavior of polymer melts and extraction of the storage and loss moduli (G' and G''). There are three dominant mechanisms of chain motion that lead to stress relaxation, all resulting from the random thermal motion by which the chains explore the surrounding space. The first mechanism is reptation, a lateral movement of the chain along its path to allow the chain to exit its confinement tube. The time scale of reptation is dependent on the one-dimensional diffusion along the tube and scales with the number of chain segments cubed. This sensitivity underpins a significant component of the link between rheology and the MWD. The second mechanism is contour length fluctuation (CLF) [4,5], where the ends of the chain escape the tube by Brownian motion-induced fluctuation of path length along the tube. The chain coils and uncoils, repeatedly exiting and entering the end of the tube, hence releasing stress from the chain ends, and also reducing the distance the chain needs to diffuse during reptation. The third mechanism is constraint release (CR), where the motion of the chains surrounding a

a)Electronic mail: py19rje@leeds.ac.uk
b)Author to whom correspondence should be addressed; electronic mail: d.j.read@leeds.ac.uk

given test chain releases entanglements, allowing stress relaxation via local rearrangements of the tube [6].

The early tube models neglected some complexity of the interactions between reptation, CLF, and CR. Nevertheless, these models yield accurate results for the linear rheological behavior of monodisperse polymers. Here, CLFs and CR modify the scaling of the reptation time with chain length, typically giving a scaling of $N^{3.4}$ for well-entangled polymers with a detailed, quantitative model being provided by Likhtman and McLeish [7].

Polydisperse melts are more challenging to model since relaxation times between chains can vary significantly, which strongly enhances the effects of CR. Tuminello and others [8–11] developed the "double reptation" model to account for the direct effects of CR on stress relaxation. Double reptation has significant advantages over earlier models for polydisperse polymers and performs well for well-entangled polymers with smooth MWDs. However, it fails to fully capture the entire relaxation behavior as it does not fully take into account the effect that vastly different chain lengths can have on each other. Double reptation assumes that the reptation time of chains of a given length in a polydisperse polymer is unchanged from that in a monodisperse melt of the same chain length. In practice, the fast relaxation of short chains often enables an acceleration in the relaxation of adjacent long and short chains, leading to a propagation of additional relaxation acceleration within the stressed polymer melt. This is seen experimentally where long chains are mixed with those of significantly lower molecular mass; the relaxation time is often decreased in these binary blends [12–15], contradicting double reptation results. Recent studies focusing on relaxation mechanisms in such bidisperse melts [14–19] have culminated in the nested-tube model of Das and Read [3], designed to encode the previous work in a predictive algorithm for the linear rheology of fully polydisperse linear polymers. We make use of this model in the present work.

Previous work has also addressed the problem we target here, predicting the MWD from the rheology, with some studies yielding excellent results. Early studies used the viscosity of the polymer melt to infer the average molecular weight [20]. Later work aimed to use the predictive capabilities of various tube models by reversing the mathematical machinery, for instance, in the case of double reptation inverting the integral over the relaxation functions of the component chain lengths. This has been attempted with double reptation and its subsequent variations [8,21–27]. Although the integral inversion is a strictly ill-posed problem, analytical solutions can be found through regularization techniques [28]. However, as discussed, there are fundamental weaknesses to many of the models that underpin these inferences. For instance, the previously discussed weaknesses of double reptation pose issues for binary MWDs and those with well-separated bimodal shapes. Therefore, despite good agreement with experimental MWD findings in some cases, there is a limit to the generality and flexibility of these methods. Other models, such as the "dual-constraint" model of Pattamaprom and Larson [29], do consider the more complex behavior of CR in the relaxation. The use of this model has enabled good progress in predicting MWDs with large polydispersity [30]. However, this study also has limited generalizability because it requires predefining the shape of the MWD peaks as either log-Gaussian or generalized exponential (GEX), a necessity introduced by the mathematical complexity that comes with increased model accuracy. This is a symptom of the nonfeasibility of the direct inversion of the model and the nonlinearity of the relationship between the rheology and the MWD [28,30].

Here, we propose an alternative inference method allowed both by new developments in rheological modeling and by the recent acceleration of so-called "artificial intelligence" methods. Specifically, we suggest combining the nested-tube model of Das and Read [3] with the use of machine learning for complex pattern recognition and inference. This state-of-the-art tube model takes all three relaxation mechanisms into account as well as the complex interplay between them. The model considers a nested-tube structure with a detailed approach to the dynamic dilution of the tube within which the polymer chain relaxes, and how the tube interacts with the rates of the reptation, CR, and CLF relaxation mechanisms. Testing of this model has yielded excellent agreement with linear rheological data for a wide range of polydisperse melts, and it is designed to be flexible to varied materials and MWDs. It is also usefully encoded in an open-access software tool[1] named LP2R, so we can now rapidly predict G' and G'' accurately for a wide range of polymers of arbitrary polydispersity. Further evidence that the LP2R tool can accurately model the rheology of the polystyrene (PS) samples included here can be found in Fig. S3 of the supplementary material, where the tube-model output matches the experimental rheology very well. The model cannot be mathematically inverted but does allow the *in silico* production of large quantities of accurate rheological data from arbitrarily polydisperse MWDs, allowing the production of large training datasets for machine learning. Hence, we rely on machine learning as a tool to invert the multidimensional nonlinear relationship between MWD and linear rheology.

We present results obtained from an initial test of our methodology to validate the use of neural networks for this task. Results are shown for several PS samples obtained from the literature. PS is selected as an initial target due to the availability of testing data, its rheology being well-characterized, and the large frequency range that can be observed due to the material's compatibility with time-temperature superposition (TTS) [31] in a part of its dynamic regime. However, the only restriction on the use of this method for a much wider set of polymers is the accuracy of the Das and Read [3] tube model. The prediction of rheology for PS has been well tested, and fortunately for polyethylene (PE) and polypropylene (PP), as well as many others, the rheology is also predicted accurately. This enables a simple

[1]Das, C., and D. J. Read, "Linear rheology of linear polydisperse polymers"; https://chinmaydaslds.github.io/LP2R/ (accessed 18 April 2024).

translation of this methodology to these materials, which will be a goal of future work.

In contrast to previous attempts to predict the MWD from rheology, we use neural networks (NNs) to make the prediction. This is partly out of necessity due to the infeasibility of the inversion of the rheological model we use. It is also partly out of an opportunity for flexibility and efficiency. Once a NN model is trained, predictions can be made very quickly, usually in a few seconds or less on a standard computer. This avoids using time to perform a potentially complex mathematical procedure. Also, many pretrained NNs, specialized for specific tasks (e.g., certain molecular weight ranges), can be used in parallel. This offers flexibility and cross-examination methods not available in previous attempts, where the mathematical inversion of the model would usually be deterministic given some particular rheological data. The training of a NN requires large quantities of labeled data to make predictions, which is a limiting factor in their use for physical science applications. This work uses the discussed developments in rheological modeling to bypass this issue by relying entirely on artificial training data. The absence of experimental training data requires trust in the underlying model, but it allows for the use of an arbitrarily large dataset with minimal additional labor.

It is worth noting here that as with many numerical minimization procedures, the result of NN training is not deterministic, and repeated training with identical parameters does not necessarily lead to identical models. Therefore, there is some variation in the accuracy of MWD prediction among the different NNs. The implications of the extent to which this variation appears are discussed further when results are presented.

Convolutional neural networks (CNNs) are NNs that employ the use of one or more convolutional layers and are the type of network used here [32]. These layers extract feature maps from the input data, which are then fed into dense fully connected layers. In these layers, the output from each node in a layer is fed into each node in the subsequent layer, allowing a highly nonlinear mapping to be made. In the past, CNNs have been successfully applied in various tasks, most notably image recognition [33,34] because they offer a more efficient method for extracting the most crucial data in an image, such as spatial information, without the complexity of a huge number of inputs and therefore connections in a fully connected network. Our application is very different from these, but we have found success in implementing two-dimensional convolutional layers in our models, with performance improvements compared to the use of only fully connected dense layers.

The result of this work is a methodology that allows the use of NN's to bypass the difficulties in reversing forward-prediction rheology models and remains robust to challenges faced with experimental data. The details of how datasets were produced, how MWDs and rheology have been parameterized, and how models have been trained are detailed in Sec. II. The experimental data acquired for testing of the trained models are detailed in Sec. III, and the comparison of predictions with these data is shown in Sec. IV. We conclude with a summary and discussion of the possible future developments that this work may allow.

II. METHODOLOGY

The challenge of training a NN to perform the inversion task required for this study is centered around optimal methods of dataset generation and parameterization. We aim to train a model that quantifies the mapping between the melt rheology and MWD, which requires careful consideration of how the data are presented to the algorithm. Here, we begin by discussing the method used to represent the shape of the MWD, which we require to be computationally simple yet flexible. This is followed by a similarly motivated parameterization of the viscoelastic response of the polymer melt using discrete relaxation spectra. The relationship between melt rheology and the MWD is complex and nonlinear, and we do not wish to further this complexity with a poor data structure. This is followed by a discussion of the composition of the training dataset and the specifications of NN training used to produce the results presented here.

A. MWD parameterization

The priority for a MWD parameterization is to avoid needless complexity while maintaining flexibility. As an example, the MWD could be represented by a large number of points forming a curve, as with GPC data, but this would be an inefficient use of computational resources, and the accuracy of each predicted parameter would likely suffer.

The alternative method we have implemented involves reconstructing the MWD as the sum of several log-Gaussian subdistributions, each with the same small dispersity. Each subdistribution is fixed to a specific mean molecular weight, evenly spaced across the logarithmic mass axis. As a result, the only variable parameter for each subdistribution is its relative weight, simplifying the parameterization to a set of variables ϕ_i, representing the volume fraction of each subdistribution. Hence, the MWD is represented as

$$\frac{\mathrm{dW}}{\mathrm{d}\log M} = \sum_{i}^{N_\phi} \phi_i \left[\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log M - \mu_i)^2}{2\sigma^2} \right) \right], \tag{1}$$

where N_ϕ is the number of subdistributions and σ is the standard deviation of the subdistribution (the same value for all i). μ_i is the mean of $\log M$ for the ith subdistribution. The subdistributions are uniformly spaced in logarithmic molecular weight. These two parameters can be calculated from the desired polydispersity index (PDI) and weight-average molecular weight $\overline{M}_{W,i}$ of the subdistribution as

$$\sigma = \sqrt{\log(PDI)}, \tag{2}$$

$$\mu_i = \log\left(\overline{M}_{W,i}\right) - \frac{\sigma^2}{2}. \tag{3}$$

As an initial test, we first check whether this parameterization is able to represent a reasonable range of candidate MWDs. Two examples of GPC MWD data fitted by this method (using a simple least squares fit) are shown in Fig. 1.

As illustrated in this example, if a sufficient number of subdistributions are used to cover the target mass range

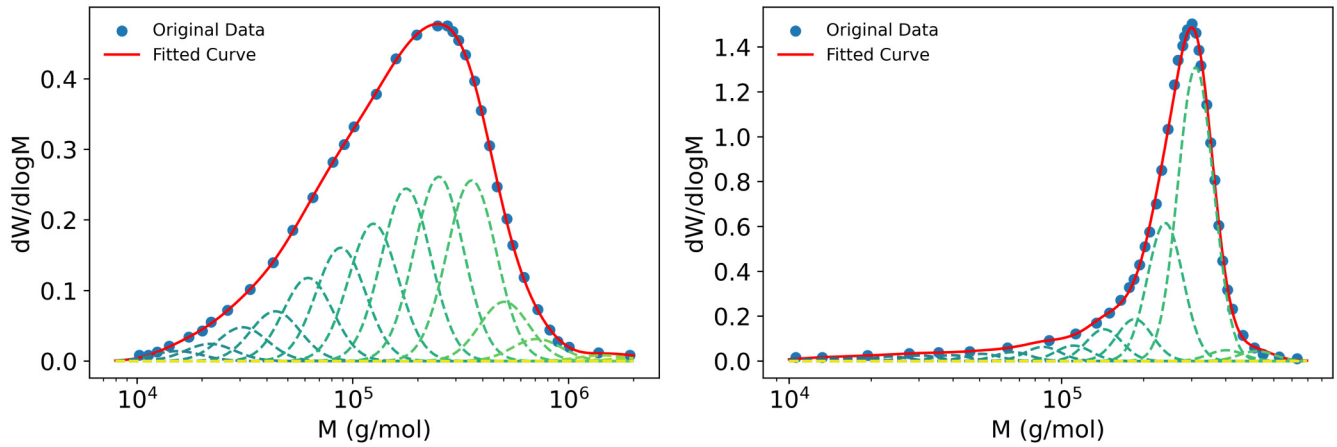

FIG. 1. Examples of fitting a sum of log-Gaussian distributions to GPC MWD data. These example MWDs are nonlog-Gaussian unimodal distributions.

completely, various MWDs can be represented simply by adjusting the ϕ_i values.

It is relevant to discuss the PDI of each subdistribution, as this highlights a known limitation of this method. There is a minimum level of dispersity that can be represented with this system; when a melt is near-monodisperse, or is a blend of a small number of monodisperse components, each peak in the true MWD has a lower PDI than the subdistributions of Eq. (1). In practice, the effective minimum peak PDI we can reliably fit is slightly above the PDI of the subdistributions —1.165 here—because the lack of flexibility in the positioning of individual peaks becomes an issue. To be specific, if the true MWD has a narrow peak sitting in between the values of μ_i for two adjacent values of i, then our method will attempt to represent the peak using a weighted sum of those two subdistributions, which naturally gives a peak with larger dispersity. However, we do not consider this a major drawback of this method as we are targeting industrially relevant MWDs, which are not produced in a manner that results in monodisperse distributions (e.g., through polymerization with well-controlled termination processes), but instead are usually more broad, with a PDI value of at least 2.

Moreover, the dispersity of each subdistribution provides a natural regularization of our method. For broadly polydisperse materials, it is doubtful whether rheology data contain sufficient information to distinguish fine-scale variations in MWD. Evidence for this can be found in Figs. S1 and S3 of the supplementary material. It is shown that reducing the dispersity of the subdistributions so that the overall MWD is more "spiky" has very little effect on the rheology. Therefore, if a greater number of more narrow subdistributions are used, many different combinations of the ϕ_i variables could produce very similar rheology, making their inference from rheology more mathematically challenging. The chosen dispersity of the subdistributions, therefore, provides a natural smoothing of the predicted MWDs, i.e., a regularization.

Through trial-and-error testing of different subdistribution dispersities and molecular weights, a set of 28 log-Gaussian subdistributions was chosen. The decided parameters resulted in $\overline{M}_W$ values uniformly distributed in logarithmic molecular weight from 8.85×10^1 g/mol to 8.85×10^6 g/mol and each with a PDI of 1.165. This set provides the flexibility to fit a wide variety of MWDs, and it covers the range of molecular weights that are most commonly seen in commercial polymers, such as PS, PE, and PP. Once these specific choices are fixed, each MWD is parameterized fully by the set of values of ϕ_i. These are, therefore, the parameters used to represent the MWD to the NN.

It would be possible to alter the parameterization of the subdistributions (e.g., number of distributions, their PDI, and the covered molecular weight range) for different or more specific applications. However, care is required to ensure a sensible relationship between the number of subdistributions, the dispersity of each, and the molecular weight range. The goal is to have a suitable minimum dispersity for the chosen application, while maintaining an overlap between adjacent subdistributions such that their sum can produce a smooth MWD. As discussed, we also anticipate that it would not provide any benefit to increase the number of subdistributions, and reduce their dispersity much beyond the values used here, due to the resolution of the information it is possible to extract from the rheology of broadly polydisperse materials.

B. Relaxation spectrum for input data

The data upon which we wish to make predictions are the linear rheology data of a polymer melt, represented by the storage and loss moduli over a particular frequency range. A key issue foreseen in the practical application of this work is that the rheology data from different experimental setups is unlikely to be in a consistent format (e.g., spacing of data points on the frequency axis may vary between different laboratories). Hence, the formatting of data must be standardized for presentation to the NN. We propose a solution where the rheology is fitted with a discrete relaxation spectrum given by the multimode Maxwell model, using

$$G(t) = \sum_{j=1}^{N_\tau} g_j e^{-t/\tau_j}, \tag{4}$$

$$G'(\omega) = \sum_{j=1}^{N_\tau} \frac{g_j(\omega\tau_j)^2}{\left[1 + (\omega\tau_j)^2\right]}, \tag{5}$$

$$G''(\omega) = \sum_{j=1}^{N_\tau} \frac{g_j(\omega\tau_j)}{\left[1 + (\omega\tau_j)^2\right]}. \tag{6}$$

By using a fixed spectrum of N_τ relaxation times τ_i, i.e., the same values of τ_j for all input data, we can fit the logarithmic form of the data with the magnitude of each mode g_j as variables. In practice, we choose the relaxation times τ_j to be uniformly spaced on a logarithmic axis. We use a least-squares fit, using the logarithm of the mode magnitudes, $\log(g_j)$, as fitting parameters. This avoids prioritizing different regions of the rheology preferentially as the loss and storage moduli cover many orders of magnitude. Then, the rheology can be represented simply in the form of these mode magnitudes ($\log(g_j)$), and the frequency domain is made constant for all data, greatly simplifying the input to the NN model.

In order to capture shifts in the relaxation spectrum of polymer melts as MWD is varied, we have found it advantageous to use a high density of τ_j values, typically with a large number of τ_j per decade. One potential issue when using large densities of Maxwell modes such as this is that adjacent modes can become redundant when the rheology data of interest could be accurately fit with a lower density of modes. This leads to oscillatory behavior where the redundant modes are given very small amplitudes compared to their neighbors. This behavior is illustrated in Fig. S4 of the supplementary material. This results in large differences in the set of variables $\log(g_j)$ from small deviations in the rheology, hindering the NN's ability to easily recognize patterns. To counteract this, we have implemented a smoothing regularization in the cost function for the fitting of the Maxwell modes to penalize solutions with oscillating $\log(g_j)$ values. The cost function includes the discrete approximation of the second derivative of the logarithm of the modal magnitudes, $\log(g_j)$, as

$$\chi_R = \sum_{k=1}^{N_\omega} \left[\left(\log G'_{inp}(\omega_k) - \log G'_{fit}(\omega_k) \right)^2 + \left(\log G''_{inp}(\omega_k) - \log G''_{fit}(\omega_k) \right)^2 \right] + \lambda^2 \sum_{j=2}^{N_\tau - 1} (x_j)^2, \tag{7}$$

$$x_j = \log(g_{j-1}) + \log(g_{j+1}) - 2\log(g_j), \quad j \neq 1, N_\tau, \tag{8}$$

where N_ω is the number of discrete ω_k frequency values for which there are rheology data. λ is a regularization parameter used to prioritize smoothness relative to the most accurate fit. λ is set to be proportional to the square root of the density of rheology data points on the logarithmic frequency space, ρ_ω, to ensure that the priority is maintained for different data sources. This is because as ρ_ω increases, the number of terms in the first sum, and therefore the priority of reducing the least-squares fit of the data, increases proportionally. To counteract this, λ^2 also increases proportionally to ρ_ω to balance the priorities of the two components of the cost function. It is worth noting that there are multiple ways to regularize the Maxwell modes so that the oscillatory behavior is diminished and a solution can be found quickly. However, once a suitable regularization is found, the most important factor is consistency, i.e., to keep using the same method both across the training dataset and when fitting a spectrum to experimental data. The NN must only be made to interpolate within its training parameter space, so always maintaining the same regularization method to fit the Maxwell modes is essential.

For the present study, we select the N_τ relaxation times to be between a maximum of 10^4 s and a minimum of 10^{-6} s with a density of eight modes per decade, evenly distributed on a logarithmic scale. This was selected to account for the maximum frequency range of the data that we will investigate here of approximately $10^{-3} \leq \omega \leq 10^5$ (rad/s).

For this choice of Maxwell mode density, we set the regularization parameter λ to be

$$\lambda = 0.4\sqrt{\rho_\omega}, \tag{9}$$

where the density of rheology data points is

$$\rho_\omega = \frac{N_\omega}{\log_{10}\omega_{\max} - \log_{10}\omega_{\min}}. \tag{10}$$

A proportionality constant of approximately 0.4 was found to yield good results and is used throughout this work.

Once these specific choices are fixed, the rheology data are parameterized fully by the set of values of $\log(g_j)$. These are, therefore, the parameters used to represent the rheology data to the NN.

In summary, the data we present to the NN model are of the following form. MWD data are represented by volume fractions ϕ_i of the log-Gaussian subdistributions. Rheology data are represented by a set of Maxwell mode amplitudes $\log(g_j)$. The task of the NN is to learn the nonlinear relationship between the values of ϕ_i and $\log(g_j)$. Once trained, the NN should be able to predict a set of ϕ_i, given the $\log(g_j)$ provided as input, i.e., to predict MWD from rheology.

C. Training dataset

For training the NN, each piece of data represents one material with a given MWD and linear rheology, and so comprises the set of Maxwell mode amplitudes labeled by the MWD volume fractions. NNs are excellent interpolation machines but less reliable in extrapolation beyond the space of the training set. Hence, the challenge in dataset production is to ensure enough variability in "seen" examples for the model to reliably interpolate for "unseen" data. We wish to do this while maintaining flexibility toward, for instance, abnormal MWD shapes.

For a general-purpose training dataset, we wish to include distributions spanning the expected mass range we are targeting and a variety of forms of MWD that can be represented

with the parameter system. To do this, we combine data from MWDs that have been artificially generated in multiple ways.

The first components are unimodal and bimodal log-Gaussian distributions with a range of random mean molecular weights and dispersities. For these distributions, the lower limit for dispersity was selected to be the dispersity of each subdistribution in Eq. (1), which was 1.165, as otherwise the near-monodisperse distributions would not be accurately represented by the parameter system. The upper limit for the dispersity was chosen to be 10.0, to allow most distributions in the dataset to be broad, as seen with commercial polymers, with some being very broad. The molecular weight ranges were set in terms of the number of entanglements Z, as $3 \leq Z \leq 500$, which for the PS parameters we have used is approximately equivalent to $3.9 \times 10^4 \leq \overline{M}_W \leq 6.4 \times 10^6$ (g/mol). For the bimodal distributions, an additional restriction was placed on the relationship between the molecular weight of the short and long peaks, where $\overline{M}_W^{\text{long}} \geq 2\overline{M}_W^{\text{short}}$ to ensure that the two peaks are reasonably separated.

We also wish to include other MWD shapes to train the model on less commonly encountered distributions, to improve the parameter space the NN is trained on. To achieve this, we have used a Monte Carlo (MC) algorithm to randomly produce MWDs. This approach is based on a Metropolis–Hastings algorithm and is covered in detail in Appendix A. Here, we design a pseudoenergy function $E(\{\phi_i\})$ for the MWD, based on the weights ϕ_i in Eq. (1), such that the probability of randomly generating a given MWD is proportional to $e^{-E(\{\phi_i\})}$. Terms in $E(\{\phi_i\})$ ensure smoothness of the MWD, enforce tails decaying toward zero, and favor distributions in which a given number of the ϕ_i are small. By tuning the corresponding parameters, the random generation can be tuned to generate distributions with certain desired characteristics and avoid unrealistic ones. This method allows us to control the general characteristics of the distributions while maintaining the variability that is required to train a capable NN.

The design of the dataset used to train the 28-log-Gaussian system comprises the uni- and bimodal fitted MWDs, along with three datasets generated using the MC algorithm according to different shape characteristics, simply named "medium," "narrow," and "broad"; see Appendix A for more details on these divisions. An illustration of the dataset composition is shown in Fig. 2. The final dataset, therefore, consists of these approximately 800 000 entries of MWD data, each given by the 28 ϕ_i values, accompanied by the relaxation spectra of the related rheology. For each of the generated MWDs, rheology was predicted using the LP2R software tool using parameters appropriate for polystyrene. The parameters used are detailed in Table I.

Another factor that was found to be important in the creation of the training datasets was including a level of noise to improve the robustness toward experimental data. In the early stages of development, NNs were trained using datasets with no such noise added. These models could accurately predict the MWD from tube-model (artificial) rheology, but had extremely poor performance when tested on experimental

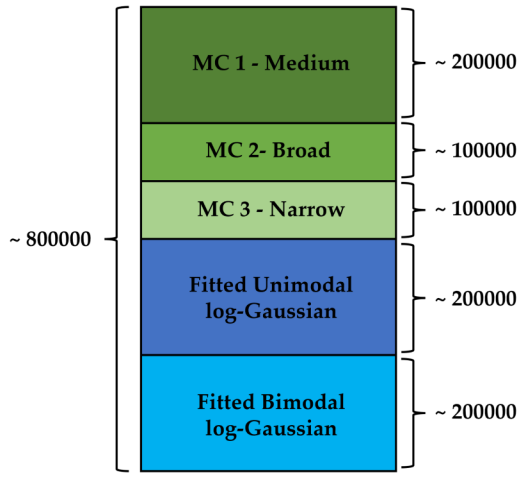

FIG. 2. Illustration showing the general composition of the training datasets used for the training of models on the sum-of-log-Gaussian parameter system.

data. This is illustrative of the general point that it is important to ensure that the trained NN has encountered variability that may be present in real data. NNs typically interpolate well within the parameter space of the training set, but are poor at extrapolation: hence, the presented parameter space should include degrees of freedom that represent the full extent of the expected noise level. Rheology experiments, even when performed expertly, are vulnerable to noise, and here, we consider three types of noise. One is the random fluctuation of each individual data point as can be seen in any experiment. To account for this, we add random values distributed normally with mean zero and some standard deviation to the logarithm of the storage and loss moduli. Standard deviation values between 0.005 and 0.1 were tested, with a value of 0.07 being used in the final training datasets. The second is vertical displacement of the rheology curves on the logarithmic axis due to variation in rheometer calibration and sample loading. To account for this, we add a random number with zero mean and nonzero standard deviation (a value of 0.2 is used here) to the logarithm of the storage and loss moduli, but this time choosing the same random number for all data points.

Finally, we wished to introduce robustness toward variations in the range of frequency measurements and the TTS

TABLE I. Input parameters used for generating the full rheology of PS at 180 °C using LP2R (see footnote 1).

Parameter	Value
Frequency ratio	$\sqrt{2}$
Maximum ω range	$10^{-8} \leq \omega \leq 10^6$ rad/s
Kuhn segment mass, M_K	720.0 g/mol
Entanglement mass, M_e	12 870.0 g/mol
Plateau modulus, G_N^0	2.2×10^5 Pa
Entanglement time, τ_e	2.2×10^{-4} s
Glassy modulus, G_∞	1.2×10^9 Pa
Glassy relaxation time, τ_g	1.3×10^{-9} s
Stretching exponent, β_g	0.390

procedure. The fitting of Maxwell modes does present the rheology in a regular format to the NN, but the fit is sensitive to the range of experimentally measured frequencies. This is to be expected, as when the relaxation time of a certain Maxwell mode falls outside the time scale of the deformation of the material, that mode becomes redundant to the fit. Therefore, for each MWD in the training set, we randomly select the frequency range of the rheology data that the Maxwell modes are fitted to, so the training set contains examples of Maxwell modes for a full spectrum of frequency ranges. To achieve this, we randomly select the minimum and maximum frequency for the rheology data so that $3 \times 10^{-4} \leq \omega_{\min} \leq 7 \times 10^{-3}$ (rad/s) and $4 \times 10^{1} \leq \omega_{\max} \leq 5 \times 10^{5}$ (rad/s). This often leaves many Maxwell modes where the inverse of the relaxation time τ_i falls outside the range of frequencies of the rheology curve. When these modes fall considerably outside the experimental data range, they will not influence the fit and are to some extent not meaningful. However, the smoothness regularization ensures that their behavior is predictable, so the NN can learn to distinguish these modes from the ones representing more significant time scales.

For each MWD in our training set, it would be possible to apply several different realizations of the above three types of "noise" and so generate multiple entries in the training set for each individual MWD. However, when adding each new entry to the training set, if there is a choice between repeating a previously used MWD with a new realization of noise, or instead producing a wholly new MWD with noise, we believe that it is likely better always to use a new MWD because this will maximize the span of the multidimensional parameter space that the NN is trained on. Hence, each entry in our training set is for a unique MWD with a single realization of the noise.

Although the training and application of the resulting NNs are quick, training dataset production is more demanding. To produce the 800 000 entries of raw rheology data (before the addition of noise or Maxwell mode fitting), 160 parallel jobs required approximately 24 h each. Noise application was completed by a single job taking approximately 30 min. Maxwell mode fitting then required a wall-clock time of 4 h when run across 160 processors. It is likely that the size of the dataset used here is larger than that needed to achieve comparable results, but our philosophy for this initial test was to ensure that the dataset size was not a restriction to performance. Although this process is resource intensive, the bulk of the computational cost is in producing the raw rheology data. This means that once this dataset was complete, it can be used to optimize noise levels, Maxwell mode parameters, and NN specifications, which are comparatively cheaper.

D. Neural network specifications

Results shown here have been produced with nine CNNs, each trained with the same architecture, parameters, and dataset. We have used models with two convolutional layers, each with 32 kernels. A batch normalization is applied to the output of each convolutional layer to stabilize training and accelerate convergence. The convolutional layers are followed by several dense layers (ten are used in the models shown here), with 256 or 512 neurons per layer. We use the ReLU (Rectified Linear Unit) activation function for each layer and the Adam optimization function. The other network parameters [35,36] found to give the best results were: an initial learning rate of 10^{-3}, a learning rate reduction factor of 0.75, a learning rate reduction patience of 10 epochs, a minimum learning rate of 10^{-8}, an early stop delta (minimum required change in loss required such that early stop is not triggered) of 10^{-7}, an early stop patience of 50 epochs, and a batch size of 256. We have used the Tensorflow package [45] in Python to create and train these models. Training was undertaken on ARC4, part of the High Performance Computing facilities at the University of Leeds, UK. Each NN model took approximately 20 min to train using maximum requested compute resources of a single NVIDIA V100 32Gbytes card, 10 CPU cores, and 48 GB system memory. Once trained, NNs can be deployed on non-specialized computers in a matter of seconds.

III. EXPERIMENTAL DATA

Data have been acquired from freely available sources, such as the Reptate [37] data files, and by extracting data from figures in the literature. The temperatures for which the rheology data are obtained vary between samples, but the NN models have been trained exclusively on PS data at 180 °C. Therefore, we shift all the experimental data to the same temperature; we do this using the Williams–Landel–Ferry (WLF) time-temperature shifting (see Appendix B) [38]. This simple shift would also be required to make a NN prediction for any new rheology data with any measurement temperature other than 180 °C. The measurement temperatures and frequency ranges observed for each of the polydisperse PS samples are shown in Table II.

A. Initial comparison of rheology data

When we shift the rheology curves from the measurement temperatures associated with each rheology dataset to the reference temperature of 180 °C, we can compare the curves for the storage and loss modulus, as seen in Fig. 3. We see a range of viscoelastic responses displaying the variability in the samples. Also evident is that the frequency ranges of the rheology for the various samples are not consistent, which is hardly surprising for data acquired from multiple laboratories. As noted above, we can account for this variability in the data by introducing the same variability in the NN training data. However, we also note that this variability may also influence the quality of the MWD predictions, as chains of different lengths relax on different time scales and, therefore, have a greater influence at some frequencies than others. Hence, the measured frequency range dictates, to some extent, the range of molecular weights for which the rheology data carry information so that inference of MWD can be made, as explored by Wasserman [22].

Figure 3 also shows the high-frequency response of the samples (for those with sufficient frequency ranges). This region shows the response of the material related to the

TABLE II. List of PS samples used for model validation and source where data were acquired. Also, the recorded rheology measurement temperatures are stated, along with the frequency ranges when the data were shifted to the reference temperature of 180 °C.

Label	Source	Temp (°C)	$\omega_{\min}$ (rad/s)	$\omega_{\max}$ (rad/s)	$\log_{10}\omega_{\max} - \log_{10}\omega_{\min}$
PS1	BASF Laboratory via Reptate files [37]	170	3.72×10^{-3}	1.85×10^{5}	7.70
PS2	BASF Laboratory via Reptate files [37]	170	1.18×10^{-3}	1.85×10^{5}	8.20
PS3	BASF Laboratory via Reptate files [37]	170	5.46×10^{-4}	1.85×10^{5}	8.53
M1	Wasserman and Graessley [39]	150	8.61×10^{-4}	1.16×10^{5}	8.13
M2	Wasserman and Graessley [39]	150	9.60×10^{-4}	1.82×10^{5}	8.28
PSA	Sugimoto *et al.* [40]	160	5.70×10^{-3}	1.65×10^{1}	3.46
A1PS	Ferri and Lomellini [41]	200	1.20×10^{-2}	7.86×10^{2}	4.82
PScom	Wasserman and Graessley [39]	150	2.24×10^{-3}	1.06×10^{5}	7.68
PS8	Montfort *et al.* [42]	160	1.72×10^{-3}	9.02×10^{2}	5.72

transition to subtube-diameter Rouse relaxation and the tail of the glassy response. This should be MWD-independent [28] for molecular weights sufficiently larger than the entanglement molecular weight M_e (here taken to be $M_e = 12\,870$ g/mol). This criterion is satisfied by all the samples.

However, we see that there is variability of rheology data in this frequency range between the samples, especially for the samples PS2 and PScom in the G'' curves. The cause of this lack of agreement is not known to us, but it could be due to polymer degradation; microstructure differences, such as tacticity; contaminants, such as residual solvents or additives (which might particularly be present in industrial grade samples); or possibly a small offset between the actual and recorded measurement temperature. Although the differences observed in this high-frequency region are in themselves small and on their own will not have a large direct influence on MWD prediction, the variability could be a symptom of deeper issues with the rheology data. It is plausible that the rest of the rheology curve, where the moduli have a greater influence on the inferred MWD, could be influenced by the same factors that give rise to the observed discrepancies at high frequency. We do not know the cause of the discrepancy, and it seems reasonable to assume that such variations will realistically be present in practical data. Hence, we

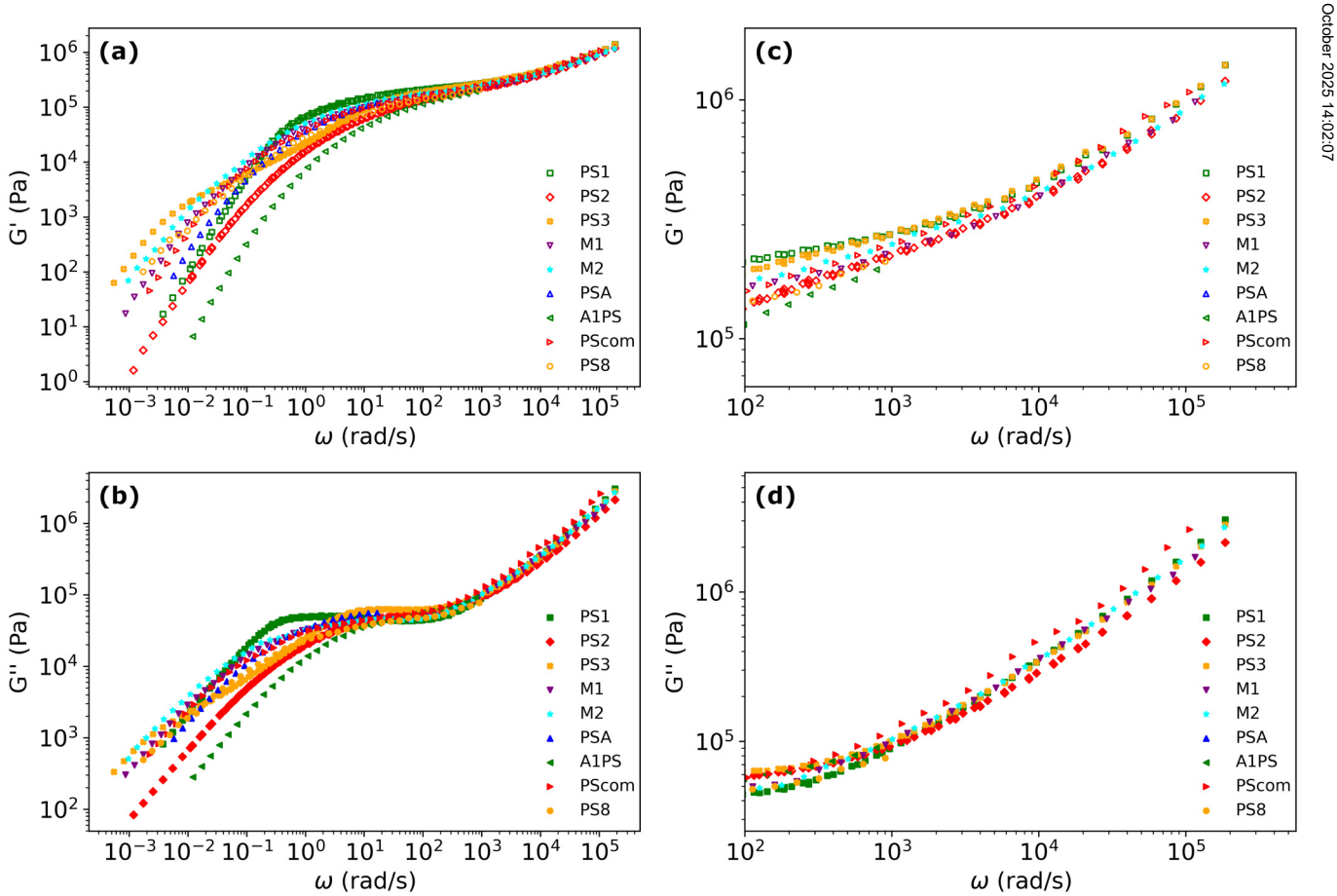

FIG. 3. (a) Storage and (b) loss modulus data, respectively, for PS samples shifted to 180 °C, with (c) and (d) high-frequency glassy response, which should be independent of MWD.

believe that the data are of high enough quality to provide insights into the performance of the models we have trained and the effects that such differences may have on their predictions.

B. Molecular weight distributions

We wish to predict the MWD of the PS samples and use GPC measurements to evaluate these predictions. GPC data are available for all but three of the samples (M1, M2, and PScom). The summary statistics for each of the MWDs are shown in Table III. For M1 and M2, GPC data are not available because these two samples were created as mixtures of near-monodisperse PS standards [39]. The mean molecular weights and the relative weights of each component in the mixtures are known so we can infer the overall MWD from these components. We do this by assuming that the total MWD is the sum of narrow log-Gaussian distributions with $\overline{M}_W$ values and relative weights as given for each of the PS standards in the mixture by Wasserman and Graessley [39].

There is a choice to be made as to what value of PDI to assume for each of the component PS standards. In reality, it is likely that each component will have a narrow dispersity, and that the true MWD is "spiky" with multiple peaks. For example, if we assume a PDI of 1.05 for each component, this gives rise to the gray-shaded MWDs with multiple peaks shown in Fig. 4. However, we recognize that, for many-component mixtures, the rheology is wholly insensitive to the fine-grained details of the MWD: it is impossible to distinguish the rheology of such a "spiky" MWD from that of a smoothed-out distribution for which we assume the PDI of each of the component PS standards to be broader. Evidence for this can be found in Fig. S3 of the supplementary material. Certainly, our methodology, in which the MWD is assumed to be the sum of log-Gaussian subdistributions with PDI = 1.165 for each, is not capable of resolving such narrow peaks. We do not believe that it is possible to reliably infer such peaks from rheology for multicomponent mixtures where the components are closely spaced. Hence, for a fair comparison of the output of our method with the "true" MWDs of M1 and M2, we consider a "smoothed" effective MWD for each sample in which the PDI of each of the component PS standards is chosen so as to eliminate oscillations in the MWD curve. We find that choosing a PDI of 1.165 for each of the PS standards [the same as we use in Eq. (1)] gives rise to a suitably smoothed curve. This smoothed curve is also shown as the solid red line in Fig. 4. The calculated PDI of each of the two distributions is also dependent on the assumed dispersity of the components. By assuming a component dispersity of 1.165, M1 is calculated to have an overall PDI = 2.70, and for M2, the new value is 2.99. The weight-average molecular weights of the distributions are unchanged from the values presented in Table III to the given level of precision. We see that the MWDs of M1 and M2 are almost identical, except for a slightly larger high molecular weight tail for M2.

No GPC data are available for PScom, making direct verification of the predictions difficult. However, the sample is known to have $\overline{M}_W = 321$ kg/mol and PDI = 1.87 and is assumed to be approximately characterized by a log-Gaussian MWD. Das and Read showed that this assumption can produce good agreement with the experimental rheology using their tube model [3].

The MWD for each of the samples can be seen in Fig. 5. Most of the samples have quite similar broad unimodal MWDs. The exceptions are PS1 and PS3, with a narrow unimodal shape for PS1 and a trimodal shape for PS3. These two samples do not represent the type of broad MWD that is the focus of this work, but they still provide useful information about the limitations of the NN system developed. Our methodology is designed to be flexible to much more variable and complex MWD shapes than are shown here, but the availability of testing data is a limiting factor. Due to the lack of good testing data for these distributions, we cannot test the full capabilities of our methodology, but the NN models are trained to be suitable for, e.g., bimodal polydisperse MWDs and those with lower $\overline{M}_W$ values. Similar results should be possible for these distributions as for unimodal polydisperse MWDs, provided the frequency range of the rheology is suitable.

TABLE III. Weight-average molecular weight ($\overline{M}_W$) and the PDI for each of the PS samples, and whether GPC data exist for the sample in question. Where GPC data do not exist, the method of MWD comparison is discussed in the main body.

Label	$\overline{M}_W$ (g/mol)	PDI	GPC exists
PS1	3.20×10^5	1.18	Y
PS2	2.74×10^5	2.72	Y
PS3	4.07×10^5	2.82	Y
M1[a]	3.57×10^5	2.32	N
M2[a]	3.99×10^5	2.57	N
PSA	2.56×10^5	2.16	Y
A1PS	1.64×10^5	1.59	Y
PScom	3.21×10^5	1.87	N
PS8	4.03×10^5	2.70	Y

[a]For M1 and M2, the summary statistics are here calculated using only the weight-average molecular weights and the relative weights of the discrete PS standards' of each, without assuming a dispersity for each component.

IV. RESULTS AND DISCUSSION

To evaluate the performance of the NN models in predicting the MWD of the PS samples, a selection of NN models trained with identical parameters and training datasets was used. Because of the random nature of NN training, each model gives different results. For each sample, the best and worst predictions (as measured by the root mean square error, RMSE, detailed below) will be shown from the nine total models trained on the same dataset.

The results for PS2, PSA, A1PS, and PS8 are shown in Fig. 6. It is worth noting at this point that the rheology data for the PSA, A1PS, and PS8 samples include considerably smaller frequency ranges than the other samples, so a lower performance and a greater variability in the output prediction could be expected. However, despite this, all three of these samples were well predicted by the NNs and did not show significant differences between the nine models. Errors in the

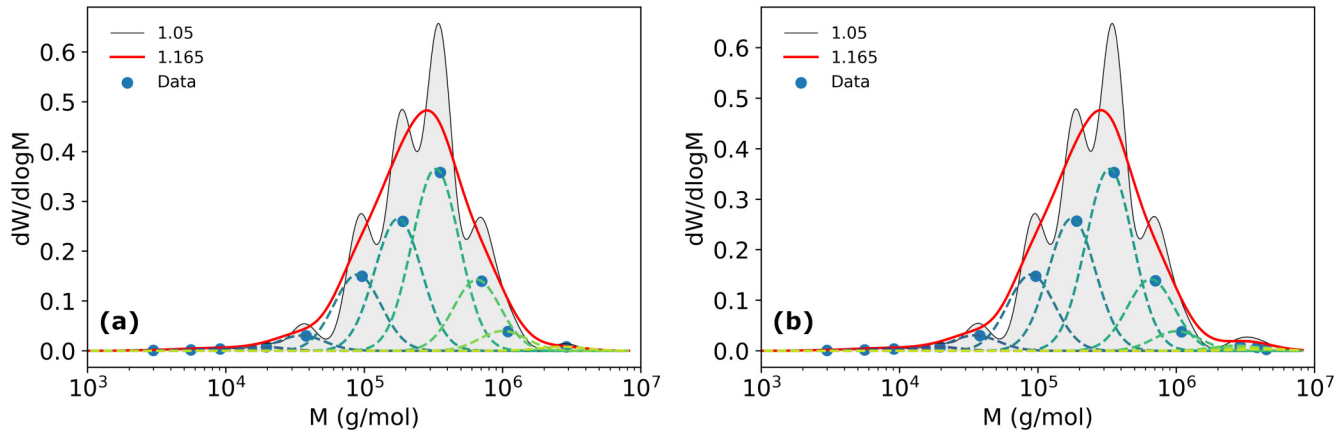

FIG. 4. Composite approximated MWDs for the two PS mixtures (a) M1 and (b) M2, with the MWD that was used shown as a solid red line and the component log-Gaussian curves shown as dashed lines, each with PDI = 1.165. Dots represent the M_W values and relative weights of the PS standards in the mixture. The gray shaded region shows MWD if each PS standard is assumed to have a lower PDI = 1.05.

PS2 prediction are consistent in each of the nine models tested, with a low variability in the prediction relative to some of the other samples. This may be related in some way to the described inconsistency in the high-frequency rheology, which was especially prevalent in PS2. If some systematic error is present in the rheology, then this consistent error in the prediction relative to the GPC would be observed. Pattamaprom *et al.* [30] also made predictions on the rheology of PS2, and a very similar shape of MWD was produced as shown in Fig. S7 of the supplementary material. It is worth noting here that when a different measurement temperature of 180 °C is assumed for PS2, the inferred MWD matches the GPC data much more closely. This does not necessarily indicate that the experimental temperature was actually 10 °C higher, but instead only that the rheology is closer to the tube-model prediction when different material parameters are used, reiterating the possibility of sample contamination, degradation, or some other discrepancy.

The predictions for the M1 and M2 samples are shown in Fig. 7. Among the nine models, there was little variation in performance. The best predictions for both are excellent, although it is not surprising how similar the two predictions are given the similarity of the samples. The main difference between the two samples is at the highest molecular weights: two components with M_W values of approximately 3.8×10^6 g/mol and 4.5×10^6 g/mol are introduced into the M2 sample. These changes give a small shoulder in the true MWD not present in the M1 sample. This feature is partially reflected in the MWD predictions from our methodology, which do not predict a distinct shoulder but do elongate the high M_W tail relative to M1 predictions. Here, we are at the limit of the detail that can be resolved by our method: the shoulder itself is, in fact, a distinct, near-monodisperse component, and as we have noted, our method is not designed to resolve these. This is again comparable to the prediction by Pattamaprom *et al.* [30] for this sample, as shown in Fig. S7 of the supplementary material.

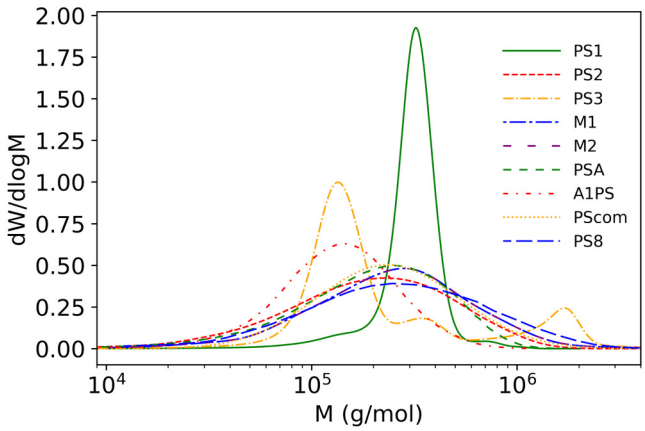

FIG. 5. Overlaid MWD data for all PS samples. The data are acquired from GPC measurements where available and reconstructed from other information for M1, M2, and PScom. A discussion of these three samples is in the main text.

We also recall Fig. 4, which illustrates that the "true" MWDs for these two samples may, in fact, contain multiple peaks, depending on the dispersity assumed for the PS standards from which the samples are constructed. Our prediction results provide insight into the limits of the resolution of information within the rheology. Although our MWD parameterization system could, in principle, produce an MWD with more defined narrow peaks, it does not. This is likely because the rheology of a MWD with multiple narrow, closely spaced, individual peaks is indiscernible from that of the smoother MWD we have used. Setting aside this ambiguity in the true MWD, and assuming the distribution is, in fact, smooth, the predictions for the M1 and M2 samples are the most accurate from among the samples presented in this work. This is notable and encouraging because these samples are the two for which the MW composition is most accurately known since they are constructed from PS standards rather than having MWD measured by GPC.

As shown previously, the PS1 and PS3 samples have uni- and trimodal MWDs, respectively, with near-monodisperse peaks. These narrow distributions are rarely used in industrial polymers and have not been targeted by the work in this study. The MWD parameterization system used here has a minimum possible PDI given by the individual

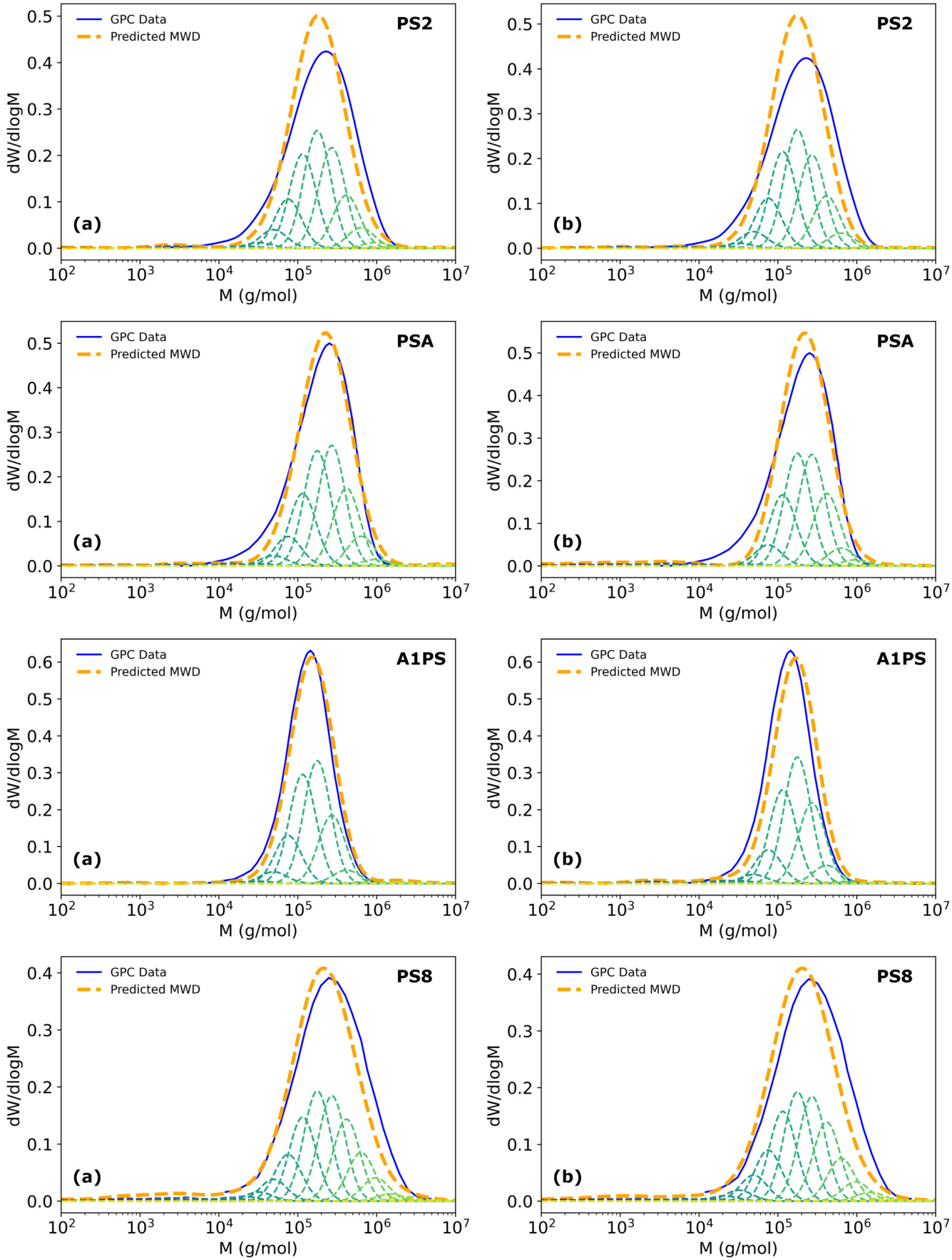

FIG. 6. (a) Lowest and (b) highest RMSE predictions for PS2, PSA, A1PS, and PS8.

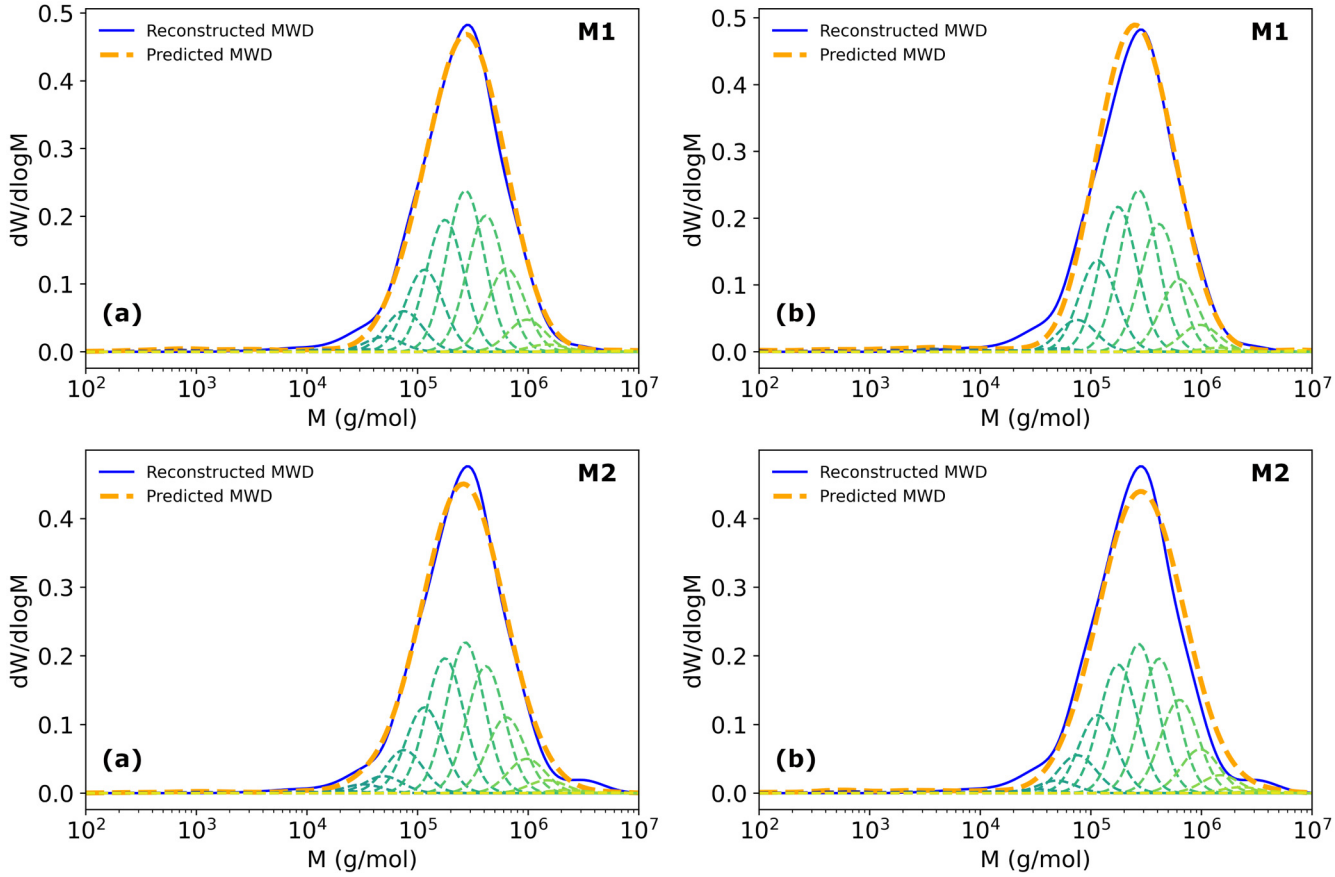

FIG. 7. (a) Lowest and (b) highest RMSE predictions for M1 and M2.

subdistributions. However, it is still prudent to test the ability of the NNs to recognize these MWDs. The results for these samples are shown in Fig. 8. Despite the impossibility of accurately matching the exact MWDs for these samples, the NN produces impressive results. For both samples, the predicted MWDs are in the correct location on the molecular weight axis and approximately follow the best shape feasible for each. Furthermore, the best and worst models predict similar results, showing good consistency across models, implying a true recognition of mean molecular weight and not random chance.

The sample for which we possess the least information is PScom. The recorded mean molecular weight $\overline{M}_W = 321$ kg/mol and PDI = 1.87 are known. We assume a log-Gaussian distribution when comparing the prediction given by the NN models with the "true" MWD. The results are shown in Fig. 9, with a moderate agreement when the RMSE is the lowest. However, the result with the highest RMSE was noticeably different from the assumed log-Gaussian shape, and as shown in Fig. 10, there is a definite trend in the predictions of all nine models. Although there is some variation, which is expected, the trend is that the models do not predict a simple log-Gaussian MWD shape. It is worth noting that the high-frequency rheology of PScom was, along with PS2, not in agreement with the other samples, so this result may not be sufficient to credibly diagnose an alternative shape for the MWD. It does, however, highlight the power of this methodology in characterization of the MWD from rheology, where the question is now raised of whether the assumption of a log-Gaussian MWD is valid. As we do not have GPC data to compare the prediction to directly, it is worth comparing the statistics of the predicted MWD with those provided. The best prediction for PScom gave $\overline{M}_W = 327$ kg/mol, which matches the true value of 321 kg/mol well within the typical precision of GPC measurements. The PDI for this prediction was 10.46. This is very different to the true value because there is a long nonzero low molecular weight tail present in many of the predictions, which, although it has little effect on $\overline{M}_W$ or the shape of the main peak, significantly affects $\overline{M}_n$. However, when the predictions of any molecular weight below 5 kg/mol are ignored, the predicted PDI becomes 2.35, which is much closer to the given value of 1.87. The predicted $\overline{M}_W$ is unchanged to the level of precision given. This issue with the low molecular weight tail is something we wish to address with future work, as it can give misleading summary statistics that do not represent the true shape of the MWD.

We quantify the quality of our predictions through two measures, the results of which are presented in Table IV. The RMSE is a measure of the deviation of our predictions from the experimentally measured MWD. The RMSE was calculated by first interpolating the distributions onto a common molecular weight axis, identical for each sample. It is then calculated as

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{x}^{N}\left(w_x^{\mathrm{data}} - w_x\right)^2}, \tag{11}$$

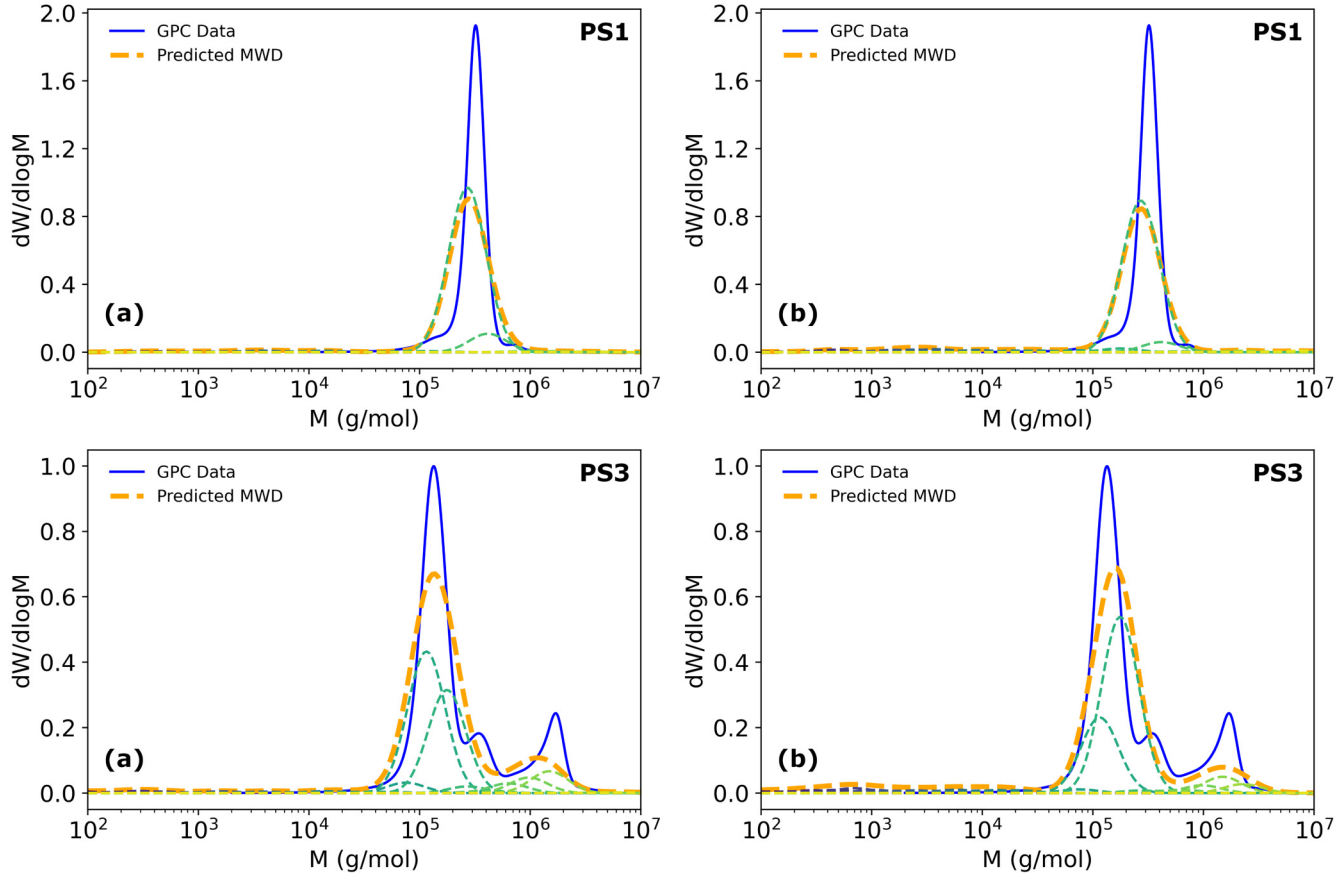

FIG. 8. (a) Lowest and (b) highest RMSE predictions for PS1 and PS3.

where N is the number of data points on the common mass axis, w_x^{data} the experimental value of $\mathrm{dW/d\,log}\,M$, and w_x the value predicted by the NN. Table IV shows the maximum and minimum RMSE across the nine NN models, together with the mean value across all models. The Standard Deviation (Std Dev) is instead a measure of the consistency of prediction between the NN models and is evaluated as

$$\text{Std Dev} = \frac{1}{N}\sum_{x}^{N}\sqrt{\frac{1}{N_{NN}}\sum_{m}^{N_{NN}}\left(w_x^m - \overline{w_x}\right)^2}. \qquad (12)$$

N_{NN} is the number of NN models, w_x^m is a point on the MWD curve for the mth NN model, and $\overline{w_x}$ is the mean value at a point x on the mass axis of the predictions across the models used.

As expected from their sharply peaked MWDs, PS1 and PS3 produced the largest errors as measured by RMSE. It is notable also that these two samples give larger values of Std Dev. This is most likely because MWDs with narrow peaks are not well-represented by the parameter system of log-Gaussian subdistributions. The NN recognizes the approximate location of the peaks on the molecular weight

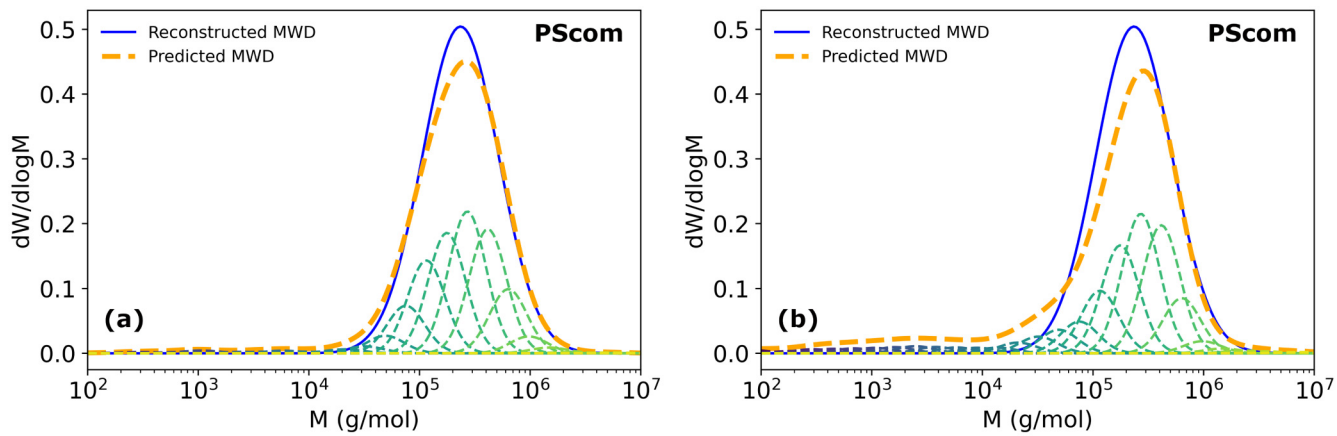

FIG. 9. (a) Lowest and (b) highest RMSE predictions for PScom, measured against the MWD reconstructed from known statistics and an assumed log-Gaussian shape.

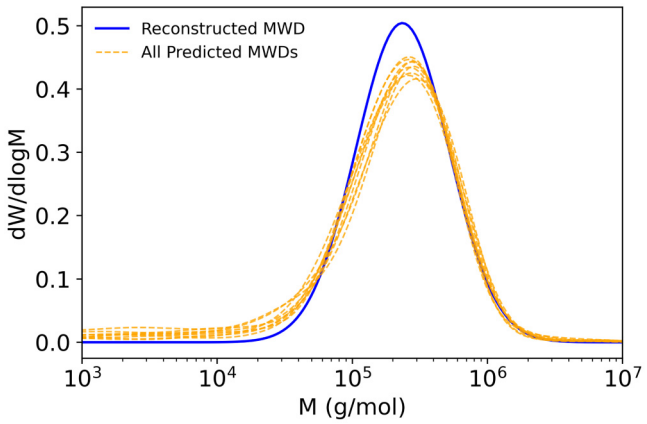

FIG. 10. All nine predictions of the reconstructed MWD for the sample PScom with the assumed log-Gaussian shape. Similar plots for all other samples are shown in Fig. S2 of the supplementary material.

axis, but there are no configurations of the subdistributions that accurately give the MWD shapes for these samples. Therefore, we suppose that this high variability is a result of the exaggeration of small differences in the NN prediction by the discretization of the mean molecular weights of the subdistributions. As a result, there is less consistency in the NN predictions.

PS2 had the largest RMSE of the remaining samples, which, as discussed, is likely due to a systematic error in the GPC or rheology data. A1PS, PS8, PScom, and PSA closely followed with good predictive performance, especially for the lowest RMSE predictions. M1 and M2 gave the lowest errors and some of the lowest standard deviations in their results, indicating an excellent characterization of the sample's MWD from the rheology.

It is a common theme among the predictions we present that the general shape and the mean molecular weight are predicted accurately. However, where the predictions struggle is with the shape and extent of low molecular weight, low-volume-fraction tails of the distribution, as seen with samples PS2, PSA, M1, and M2. These tails (comprising polymer chains only just long enough to be entangled) have a relatively small impact on the rheology and are, therefore, inherently more difficult to detect or quantify. This is noted as a limitation of this methodology.

TABLE IV. RMSEs and the mean standard deviation of RMSEs for nine models used for each PS sample.

Label	$RMSE_{Mean}$	$RMSE_{Min}$	$RMSE_{Max}$	Std Dev
PS1	0.1777	0.1736	0.1818	0.006 65
PS2	0.0367	0.0330	0.0404	0.004 17
PS3	0.0772	0.0703	0.0830	0.012 15
M1	0.0100	0.0082	0.0134	0.004 07
M2	0.0132	0.0085	0.0170	0.005 15
PSA	0.0222	0.0181	0.0269	0.003 60
A1PS	0.0311	0.0188	0.0378	0.005 77
PScom	0.0262	0.0169	0.0355	0.006 89
PS8	0.0289	0.0256	0.0311	0.004 80

As an additional test of the ability of the NNs to invert a state-of-the-art rheological model, we have checked the reversibility of our predictions. We used the Das and Read tube model to predict the rheology for the highest and lowest RMSE MWDs predicted by the NNs for each of the PS samples. The comparison of these predictions with the experimental rheology data can be found in Figs. S5 and S6 of the supplementary material. In most cases, the tube-model output follows the experimental rheology very closely, with slight differences between the rheology for the two MWDs used for each sample. This is good evidence that the NN predictions are consistent with the tube model, and each NN has "learned" a feasible relationship between the MWD and rheology. The two cases where the model-generated and experimental rheology do not match to the same level are for PS1 and PS3. This is likely due to the discussed incompatibility of these MWDs with the parameter system we have used, where the dispersity of each subdistribution is too large to accurately describe the more narrow peaks of these two MWDs.

V. SUMMARY AND OUTLOOK

We have presented a method for inferring the MWD of polydisperse polymer melts using neural networks trained on large, artificially generated linear rheology data. We use the nested-tube structure model of Das and Read [3] to produce the training data. The predictions match the known MWDs of the nine test PS samples well.

Key developments in this work are (i) the introduction of a new method of MWD prediction, using NNs as a means to "invert" the forward prediction of rheology from MWD using a well tested model; (ii) the treatment of the data input to the NNs, specifically in representing MWDs as a sum over log-Gaussian subdistributions and rheology as a sum over regularized Maxwell modes; and (iii) the creation of suitable datasets, accounting for reasonable variation in MWD, together with expected experimental noise and variability in the frequency range of data. Together, these developments simplify the task demanded of the model and make it robust to different challenges encountered with experimental data.

Our NN method differs from previous approaches to this problem in the ability to be flexible, which is enabled by the above parameterization of the MWD we have developed. Previous approaches often use rigid constraints, such as assumed MWD shapes (e.g., log-Gaussian or GEX), to simplify the inversion. These constraints are often not representative of the true MWD and can cause nonoptimal inference outcomes. Also, although the training of NNs is often resource intensive, this cost is only incurred once during model development. Once trained, inference requires only a single pass of data through the NN, which is far faster than the repetitive solving of a model during a more traditional least-squares optimization process.

The need for the use of machine learning techniques was partly due to the infeasibility of reversing the tube model of Das and Read using analytical mathematical methods. We have shown here that using the inference capabilities of modern NNs, we can forego this challenge and reverse

the model to make accurate MWD inferences. Hence, the methodology detailed here is limited primarily by the forward-predictive capabilities of current models. This is an improvement on previous attempts at predicting the MWD from rheology, where an additional restriction is the ability to reverse the model using more traditional methods.

A further limitation on any MWD inference method is the amount of MWD information that is uniquely inferrable from the rheology data. Some consideration of the limits of molecular weight inference given a particular rheology frequency range has been presented in the past [22]. This involved using the expected relaxation time scales of various molecular weights to estimate the frequency where the signature of certain chain lengths can be detected in the rheology data. A possible direction for future research could be to incorporate these ideas into future NN systems. This would manifest practically as, along with the MWD prediction, providing limits on the "known" molecular weight range with relaxation time scales that are firmly within the frequency range of the input rheology. It may nevertheless still be possible to approximate the distribution of chain lengths outside this range as the effect these chains have on the rheology is not confined to only one frequency.

In this work, we have focused primarily on relatively broad MWDs, and we have noted that for such distributions, there is a limit on the resolution of fine-grained detail of the MWD: we are not able to resolve closely spaced sharp peaks, and we suspect that there is not sufficient information available in the rheology to distinguish such a distribution from a smoothed-out equivalent. Nevertheless, if the MWD comprises a small number of well-separated narrow peaks, there may still be sufficient information present in the rheology to resolve both their position and width, by constructing a different set of NN models specifically designed for this purpose. This (albeit speculative) consideration highlights that the resolution limit of MWD inference from rheology may depend on the nature of the MWD itself.

The advantages of the tube model used here lie primarily in the more precise treatment of melts containing vastly different chain lengths at the extremes of polydispersity. Unfortunately, we have not as yet been able to test our methodology on the type of polymer melt for which this advantage is most clear, for example, bimodal polydisperse MWDs. This is due to the lack of availability of this type of experimental data for testing purposes. Nevertheless, trials on synthetic test data indicate that it should be possible to infer the MWDs for this type of distribution, provided that the relaxation time scales of these chain lengths fall within the frequency domain of the rheology data.

There are also no inherent challenges with translating this technique to other polymers, such as PE or PP, which are the most abundantly used in modern society. A forward prediction of the linear rheology for these materials is well tested, with only small alterations to material parameters [3]. The main complication involved with these materials is the incompatibility with TTS due to nonproportional scaling of different relaxation time scales with temperature. The result is that we observe a more limited frequency range, which will have some adverse effects on prediction accuracy at the extremes of molecular weight. However, we have seen, for example, with PSA, A1PS, and PS8, that high-quality predictions are still made when the frequency range is limited, although the molecular weight range of accurate predictions may be limited.

This work has fulfilled the purpose of a proof-of-concept for the use of NNs to reverse the predictions of advanced rheological models and to extract MWD information. This development allows the inference of molecular weight from rheology using state-of-the-art tube models, which would not be possible using traditional mathematical methods. The noted key developments serve as valuable insights for future work, with the goal of expanding the capabilities of the NN system.

The main priority of future work should be to extend compatibility to other polymeric materials, such as PE and PP, which together represent a large fraction of the global polymer industry. For such polymers, it seems plausible to extract the MWD when the architecture is linear. There are two main limitations in this regard. First, it is typical that for such polymers, the linear rheology is measured over only a few decades since TTS is limited by crystallization. This may reduce the information that can be inferred from the measurements and so reduce the quality of predictions, especially for large PDI. Second, such polymers also often contain comonomers, for example, to introduce short chain branching, which improves performance in final application. The rheology parameters of entanglement spacing, entanglement time, and modulus typically vary with comonomer content; see, e.g., [43]. Thus, the method needs to be flexible with respect to changing such parameters. We defer such considerations to future work.

A further issue is that polymers, such as PE, and to some extent PP, also often contain long chain branching. Here, the forward problem of predicting rheology from molecular architectures is well discussed in the literature; see, e.g., [28] for a summary. Hence, synthetic data for training NNs could readily be generated. The problem is whether there is sufficient information present in linear rheology alone to tease apart the parametric complexity of branched polymers. In our view, it is impossible to uniquely determine branched architectures from linear rheology alone since the same rheology can be produced from multiple combinations of different structures. Extra information must be brought to bear on the problem, for example, by restricting the set of architectures considered within a certain class, or within the distribution implied by a given reaction chemistry such that there are only a small number of unknown parameters to determine. This would, for example, be the case for idealized single site catalysis [28]. This would be an interesting topic for further study.

SUPPLEMENTARY MATERIAL

See the supplementary material that includes figures, which may be of value to the interested reader. This includes plots in the style of Fig. 10, with all predictions shown, for the other PS samples. Also included are supporting data for arguments made here that the rheology of an MWD with closely spaced narrow peaks is extremely similar to that from a "smoothed-out" equivalent.

ACKNOWLEDGMENTS

The work of Robert J. Elliott, Luisa Cutillo, Johan Mattsson, and Daniel J. Read forms part of the research program of DPI (Project No. 861). This work was undertaken on ARC4, part of the High Performance Computing facilities at the University of Leeds, UK.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

DATA AVAILABILITY

The data that support the findings of this study are openly available at https://doi.org/10.5518/1689, [44].

APPENDIX A: MONTE CARLO DATASET GENERATION

Here, we describe the method used to randomly generate MWDs of a variable shape for use in the training dataset. A completely random selection of ϕ_i in Eq. (1) would produce unrealistic and nonrelevant distributions. Instead, we have implemented a MC system based on a Metropolis–Hastings algorithm designed to guide the choice of ϕ_i toward smoother, more realistic distributions. The method is based on a pseudoenergy function $E(\{\phi_i\})$ for the MWD, using the weights ϕ_i in Eq. (1). Here, we have chosen to use a function of the form

$$E(\{\phi_i\}) = A\sum_{i=2}^{N_\phi}\sqrt{(\phi_i-\phi_{i-1})^2} + BN_\phi\left(\phi_1^2+\phi_2^2+\phi_{N_\phi-1}^2+\phi_{N_\phi}^2\right) + CN_\phi\sum_{i=1}^{N_\phi-12}\phi_i^s. \tag{A1}$$

A, B, and C are input parameters. The first term sums over the difference between adjacent ϕ values and, therefore, penalizes "spiky" distributions and prioritizes smoothness in accepted distributions. Note that although A, B, and C are constants, the prefactor on each term is the relevant constant multiplied by N_ϕ; hence, there is no factor of $1/N_\phi$ to normalize the first sum. The second term is a sum of the ϕ values representing the subdistributions with the two smallest, and the two largest, mean molecular weights. This penalizes distributions where there are many polymers at the extremes of the considered molecular weight range. The third term sums over the first $N_\phi - 12$ values of ϕ_i^s, which are the ϕ_i subdistribution volume fractions sorted in order of ascending magnitude. Hence, the third term in $E(\{\phi_i\})$ sums over the ϕ values of smaller magnitude, excluding the largest 12. This term is designed so that the ϕ values quickly move away from their initial condition of identical magnitude, and MWDs are not favored if many of the subdistributions are large and, therefore, similar in magnitude.

Each discrete step of the algorithm works by proposing a change to the MWD, and this change is accepted or rejected with a probability P of the form

$$P = e^{-\left[E(\{\phi_i\}_{\text{proposed}})-E(\{\phi_i\}_{\text{current}})\right]}. \tag{A2}$$

At every subsequent step, two random indices x and y are selected, and a random number δ is generated uniformly between the limits 0 and some maximum step size Δ defined as

$$\Delta = \frac{\text{SF}}{N_\phi}, \tag{A3}$$

where SF is a step fraction, which is set as an input parameter to the algorithm. The random step δ is added to the first selected ϕ value ϕ_x and subtracted from ϕ_y; this preserves normalization. Another input parameter is a "zero move probability" ZMP. When randomly generating δ, there is a probability ZMP that δ will be set as $|\phi_y|$. Thus, ϕ_y will be set to zero, and ϕ_x will be set to $|\phi_x| + |\phi_y|$. This was implemented in the algorithm to make the process more dynamic and prevent stagnation near local minima in the energy function and allow ϕ values to exist at magnitudes of 0, which rarely happens due to the random nature of the proposed change δ. When the proposed state is decided, $E(\{\phi_i\}_{\text{proposed}})$ is calculated and compared to the current energy. If any ϕ_i value is negative, the energy is set to be infinite, and therefore, the state is never accepted.

There is a "burn-in" phase to allow the algorithm to escape its initial condition before MWDs are saved. Following the burn-in period, every tenth MWD is saved. There is also a system for preventing the algorithm from getting stuck. The program running the algorithm caches the last 100 accepted states $\{\phi_i\}$. If proposed moves are not accepted for 50 successive steps, the current state is reverted back to the oldest state in the cache.

Figure 2 shows the composition of the final dataset and that there were three different characteristics of MWD that were included. Table V shows the parameters used to produce these characteristics, the major difference between them being the magnitude of A (smoothness). Figure 11 shows some representative MWDs generated by the algorithm with each of the characteristics included in the final dataset. These distributions give an intuition for the range of the MWD shape and the molecular weight range that the NN will be trained on.

TABLE V. Parameter values used to generate the MC components of the 800 000-entry training dataset.

Dataset component	A	B	C	SF	Burn-in	ZMP
MC1 medium	15.0	20.0	1.0	0.5	1000	0.25
MC2 broad	60.0	10.0	1.0	0.5	1000	0.25
MC3 narrow	5.0	20.0	1.0	0.5	1000	0.25

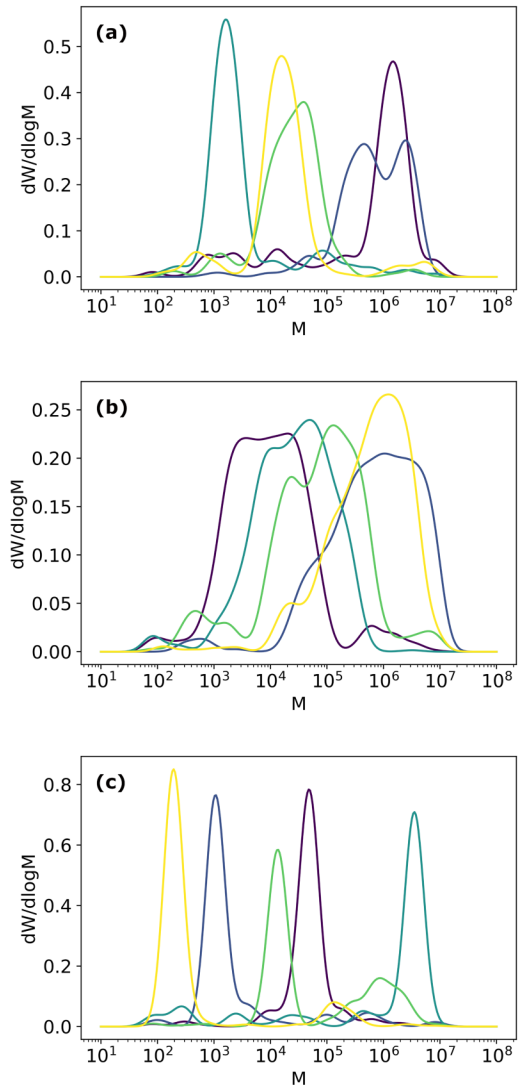

FIG. 11. Representative example distributions for the MC-generated dataset with parameters for (a) medium, (b) broad, and (c) narrow MWDs.

APPENDIX B: WLF SHIFT

To shift PS rheology from a given measurement temperature to the NN operating temperature of 180 °C, we use the WLF equation using the form presented in the Reptate software [37] with parameters suitable for most PS samples. The two shift parameters are calculated as

$$\log_{10}\alpha_T = \frac{-B_1(T - T_r)}{(B_2 + T_r)(B_2 + T)}, \tag{B1}$$

$$b_T = \frac{(1 + \alpha T)(T_r + 273.15)}{(1 + \alpha T_r)(T + 273.15)}, \tag{B2}$$

where $B_1 = 651.9$, $B_2 = -52.24$, $\log_{10}\alpha = -3.161$ from Boudara *et al.* [37], T is the measurement temperature in degrees Celsius, and T_r is the reference temperature, in this case 180 °C. To shift rheology data from T to T_r, we divide the G' and G'' data by b_T and multiply the frequency coordinates by α_T to perform the required vertical and horizontal shift, respectively.

REFERENCES

[1] de Gennes, P. G., "Reptation of a polymer chain in the presence of fixed obstacles," J. Chem. Phys. **55**, 572–579 (1971).
[2] Doi, M., and S. F. Edwards, "Dynamics of concentrated polymer systems. Part 1.—Brownian motion in the equilibrium state," J. Chem. Soc. Faraday Trans. 2 **74**, 1789–1801 (1978).
[3] Das, C., and D. J. Read, "A tube model for predicting the stress and dielectric relaxations of polydisperse linear polymers," J. Rheol. **67**, 693–721 (2023).
[4] Milner, S. T., and T. C. B. McLeish, "Reptation and contour-length fluctuations in melts of linear polymers," Phys. Rev. Lett. **81**, 725–728 (1998).
[5] Doi, M., "Explanation for the 3.4-power law for viscosity of polymeric liquids on the basis of the tube model," J. Polym. Sci. Polym. Phys. Ed. **21**, 667–684 (1983).
[6] De Gennes, P. G., "Dynamics of entangled polymer solutions. I. The Rouse model," Macromolecules **9**, 587–593 (1976).
[7] Likhtman, A. E., and T. C. B. McLeish, "Quantitative theory for linear dynamics of linear entangled polymers," Macromolecules **35**, 6332–6343 (2002).
[8] Tuminello, W. H., "Molecular-weight and molecular-weight distribution from dynamic measurements of polymer melts," Polym. Eng. Sci. **26**, 1339–1347 (1986).
[9] Tsenoglou, C., "Molecular weight polydispersity effects on the viscoelasticity of entangled linear polymers," Macromolecules **24**, 1762–1767 (1991).
[10] des Cloizeaux, J., "Double reptation *vs.* simple reptation in polymer melts," Europhys. Lett. **5**, 437 (1988).
[11] Frischknecht, A. L., and S. T. Milner, "Linear rheology of binary melts from a phenomenological tube model of entangled polymers," J. Rheol. **46**, 671–684 (2002).
[12] Watanabe, H., and T. Kotaka, "Viscoelastic properties and relaxation mechanisms of binary blends of narrow molecular-weight distribution polystyrenes," Macromolecules **17**, 2316–2325 (1984).
[13] Wang, S., S.-Q. Wang, A. Halasa, and W.-L. Hsu, "Relaxation dynamics in mixtures of long and short chains: Tube dilation and impeded curvilinear diffusion," Macromolecules **36**, 5355–5371 (2003).
[14] Read, D. J., K. Jagannathan, S. K. Sukumaran, and D. Auhl, "A full-chain constitutive model for bidisperse blends of linear polymers," J. Rheol. **56**, 823–873 (2012).
[15] Read, D. J., M. E. Shivokhin, and A. E. Likhtman, "Contour length fluctuations and constraint release in entangled polymers: Slip-spring simulations and their implications for binary blend rheology," J. Rheol. **62**, 1017–1036 (2018).
[16] Viovy, J. L., M. Rubinstein, and R. H. Colby, "Constraint release in polymer melts: Tube reorganization versus tube dilation," Macromolecules **24**, 3587–3596 (1991).
[17] Park, S. J., and R. G. Larson, "Tube dilation and reptation in binary blends of monodisperse linear polymers," Macromolecules **37**, 597–604 (2004).
[18] Park, S. J., and R. G. Larson, "Long-chain dynamics in binary blends of monodisperse linear polymers," J. Rheol. **50**, 21–39 (2006).
[19] van Ruymbeke, E., V. Shchetnikava, Y. Matsumiya, and H. Watanabe, "Dynamic dilution effect in binary blends of linear polymers with well-separated molecular weights," Macromolecules **47**, 7653–7665 (2014).
[20] Wu, S., "Polymer molecular-weight distribution from dynamic melt viscoelasticity," Polym. Eng. Sci. **25**, 122–128 (1985).
[21] Mead, D. W., "Determination of molecular-weight distributions of linear flexible polymers from linear viscoelastic material functions," J. Rheol. **38**, 1797–1827 (1994).

[22] Wasserman, S. H., "Calculating the molecular-weight distribution from linear viscoelastic response of polymer melts," J. Rheol. **39**, 601–625 (1995).

[23] Carrot, C., and J. Guillet, "From dynamic moduli to molecular weight distribution: A study of various polydisperse linear polymers," J. Rheol. **41**, 1203–1220 (1997).

[24] Anderssen, R. S., D. W. Mead, and J. J. Driscoll, "On the recovery of molecular weight functionals from the double reptation model," J. Non-Newtonian Fluid Mech. **68**, 291–301 (1997).

[25] Thimm, W., C. Friedrich, M. Marth, and J. Honerkamp, "On the Rouse spectrum and the determination of the molecular weight distribution from rheological data," J. Rheol. **44**, 429–438 (2000).

[26] van Ruymbeke, E., R. Keunings, and C. Bailly, "Determination of the molecular weight distribution of entangled linear polymers from linear viscoelasticity data," J. Non-Newtonian Fluid Mech. **105**, 153–175 (2002).

[27] Léonardi, F., A. Allal, and G. Marin, "Molecular weight distribution from viscoelastic data: The importance of tube renewal and Rouse modes," J. Rheol. **46**, 209–224 (2002).

[28] Dealy, J. M., D. J. Read, and R. G. Larson, *Structure and Rheology of Molten Polymers*, 2nd ed. (Hanser, Munich, Germany, 2018).

[29] Pattamaprom, C., and R. G. Larson, "Predicting the linear viscoelastic properties of monodisperse and polydisperse polystyrenes and polyethylenes," Rheol. Acta **40**, 516–532 (2001).

[30] Pattamaprom, C., R. G. Larson, and A. Sirivat, "Determining polymer molecular weight distributions from rheological properties using the dual-constraint model," Rheol. Acta **47**, 689–700 (2008).

[31] Ferry, J. D., *Viscoelastic Properties of Polymers*, 3rd ed. (Wiley, New York, 1980).

[32] LeCun, Y., Y. Bengio, and G. Hinton, "Deep learning," Nature **521**, 436–444 (2015).

[33] Krizhevsky, A., I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Commun. ACM **60**, 84–90 (2017).

[34] Chen, L.-C., G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," IEEE Trans. Pattern Anal. Mach. Intell. **40**, 834–848 (2018).

[35] Goodfellow, I., Y. Bengio, and A. Courville, *Deep Learning* (MIT, Cambridge, MA, 2016).

[36] Smith, L. N., "A disciplined approach to neural network hyper-parameters: Part 1 – learning rate, batch size, momentum, and weight decay," arXiv:1803.09820 (2018).

[37] Boudara, V. A. H., D. J. Read, and J. Ramírez, "Reptate rheology software: Toolkit for the analysis of theories and experiments," J. Rheol. **64**, 709–722 (2020).

[38] Williams, M. L., R. F. Landel, and J. D. Ferry, "The temperature dependence of relaxation mechanisms in amorphous polymers and other glass-forming liquids," J. Am. Chem. Soc. **77**, 3701–3707 (1955).

[39] Wasserman, S. H., and W. W. Graessley, "Effects of polydispersity on linear viscoelasticity in entangled polymer melts," J. Rheol. **36**, 543–572 (1992).

[40] Sugimoto, M., T. Koizumi, T. Taniguchi, K. Koyama, K. Saito, D. Nonokawa, and T. Morita, "Melt rheology of hyperbranched-polystyrene synthesized with multisite macromonomer," J. Polym. Sci. Part B: Polym. Phys. **47**, 2226–2237 (2009).

[41] Ferri, D., and P. Lomellini, "Melt rheology of randomly branched polystyrenes," J. Rheol. **43**, 1355–1372 (1999).

[42] Montfort, J. P., G. Marin, J. Arman, and Ph. Monge, "Viscoelastic properties of high molecular weight polymers in the molten state: II. Influence of the molecular weight distribution on linear viscoelastic properties," Rheol. Acta **18**, 623–628 (1979).

[43] Chen, X., F. J. Stadler, H. Münstedt, and R. G. Larson, "Method for obtaining tube model parameters for commercial ethene/α-olefin copolymers," J. Rheol. **54**, 393–406 (2010).

[44] Elliott, R. J., L. Cutillo, C. Das, J. Mattsson, and D. J. Read, "Supporting data for 'Using neural networks to deduce polymer molecular weight distributions from linear rheology,'" (2025). University of Leeds, Dataset.

[45] Abadi, M., P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D.G. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: A system for large-scale machine learning,' In *Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation* (USENIX Association, Savannah, GA, 2016), pp. 265–283.