ESTIMATION OF MOLECULAR WEIGHT DISTRIBUTION OF *CIS* -POLYISOPRENE MELTS FROM DIELECTRIC LOSS SPECTRA USING NEURAL NETWORKS

Martin Kozek†, Denis Miklau, H. Peter Jörgl

† Institute for Machine- and Process-Automation
Vienna University of Technology
fax: +43 1 58801 32899
e-mail: kozek@impa.tuwien.ac.at

Keywords: Neural Networks, Signal and Systems, Process Automation.

Abstract

While the prediction of dielectric loss spectra (DLS) from molecular weight distributions (MWD) is relatively straightforward the inversion is known to be an intrinsically ill-posed problem with high sensitivity to measurement noise. We propose artificial neural networks to solve this problem in two steps: First, the measured DLS is approximated by a special basis function network (BFN), thus reducing the data considerably and inherently smoothing the spectra. Second, a group of simple feedforward networks is employed to estimate the parameters of another BFN. The output of this second BFN is the estimate of the MWD. A simulation demonstrates the performance of the new method.

1 Introduction

The molecular weight distribution (MWD) of polymer melts is the most important parameter to monitor product composition. The direct measurement via size exclusion chromatography is time consuming and expensive and therefore not suitable for on-line assessment. With the availability of dielectric measurements the analysis of dielectric relaxation in such melts moved into focus. The normal-mode relaxation has a distinct dielectric spectrum of a marked low-frequency loss peak with shape and magnitude strongly depending on sample molecular weight. Smaller side-chains contribute to a high-frequency segmental loss peak that is insensitive to sample composition and is not considered for inversion. There exist a host of articles covering dielectric normal-mode relaxation [1, 2, 3, 4] which form a solid basis for the inversion of DLS.

The problem of inverting the DLS is difficult, since the solution may not be unique for a given stochastic measurement error. The strong sensitivity of the inversion can be explained by a crude relation of the MWD to the second derivative of viscosity with respect to frequency [3]. Therefore, extraction of the MWD from noisy data is clearly an ill-posed problem which needs some form of regularization. In the past, methods such as maximum entropy [5], Tikhonov regularization [6], and iterative inversion [7, 3] have been applied to this problem. While Tikhonov regularization is applicable only to systems with linear mixing rule, the iterative algorithm may also be applied to systems with nonlinear mixing.

Regardless of the inversion method, care must be taken to properly extract the DLS from experimental data. This includes correct treatment of irregularly spaced data with respect to the frequency axis ω as well as correct asymptotic behaviour for $\omega \to 0$ and $\omega \to \infty$.

We propose a new method for the inversion by using artificial neural networks (ANN). Basically, a small group of simple feedforward ANNs performs a general (nonlinear) mapping of the DLS on the MWD. The method is structured in two steps: In the first step experimental or simulated data of the DLS are approximated by radial basis functions to provide a smooth representation of the raw data; then a basis function network (BFN) with specifically designed basis functions is trained in order to minimize the number of parameters. The result of this procedure is an inherent regularization [8, 9] and a strong data reduction, since the parameters of the BFN are considered a unique representation of the DLS. In a second step the parameters of another BFN with different basis functions are estimated by ANNs. To keep the ANNs simple and to enable efficient training each of these networks has only a single output corresponding to one parameter of the second BFN. The output of the second BFN represents the MWD.

The remainder of this paper is structured as follows: In section 2 polymer descriptions in formal terms are given. The approximation of the MWD and DLS by basis function networks is given in section 3. A nonlinear mapping between the two approximations is developed in section 4.

A simulation using data with measurement noise from theoretical models is performed in section 5 and results of the overall inversion are presented. A discussion of the advantages and limitations of the new method concludes the paper.

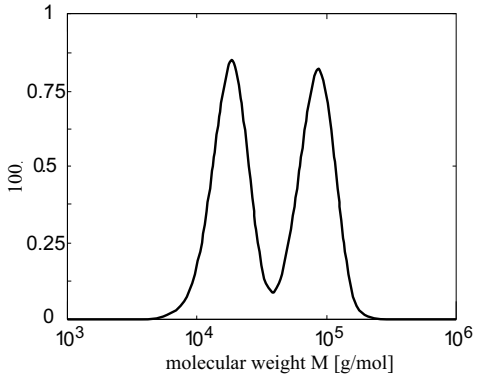

Figure 1: Theoretical molecular weight distribution (MWD) of a mixture of two polymers: $M_1 = 18500\ g/mol$, $M_2 = 86500\ g/mol$, $M_0 = 68\ g/mol$, $a_1 = a_2 = 10$, and $\gamma = 0.18$.

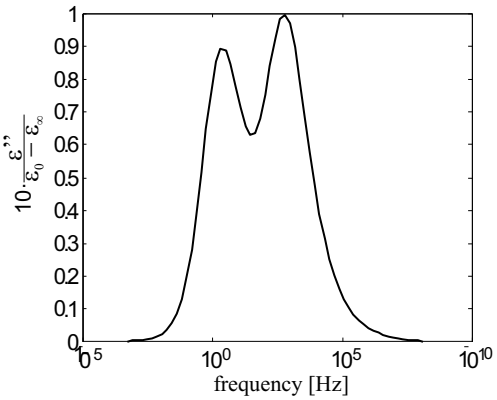

Figure 2: Theoretical dielectric loss spectrum (DLS) of a mixture of two polymers: $M_1 = 18500\ g/mol$, $M_2 = 86500\ g/mol$, $M_0 = 68\ g/mol$, $a_1 = a_2 = 10$, and $\gamma = 0.18$. Only the normal-mode peaks are represented by the model.

2 Background

2.1 Theoretical Molecular Weight Distribution and Linear Mixing

The mathematical model of theoretical MWDs $\Phi_\beta(M)$ of a single polymer is given by [3, 4]

$$\Phi_\beta(M) = \frac{M_0 \cdot a_\beta}{M_{n\beta} \cdot \Gamma(a_\beta + 1)} \cdot a_\beta \cdot \left(\frac{a_\beta \cdot M}{M_{n\beta}}\right)^{a_\beta} e^{-\left(\frac{a_\beta \cdot M}{M_{n\beta}}\right)}, \tag{1}$$

$$\left(\frac{M_{w\beta}}{M_{n\beta}}\right) = \frac{(a_\beta + 1)}{a_\beta}, \tag{2}$$

where M_0, $M_{n\beta}$, $M_{w\beta}$ are the monomer, number-average, and weight-average molecular weights, respectively, and Γ is the gamma function.

The resulting MWD $\Phi(M)$ from the mixture of two polymers is given by a linear mixing rule

$$\Phi(M) = \gamma \cdot \Phi_1(M) + (1 - \gamma) \cdot \Phi_2(M), \tag{3}$$

where γ is the fractional contribution of Φ_1. This model has been experimentally verified for type-A polymers [2]. A possible shape from the mixture of two polymers predicted by this model can be seen in fig. 1.

2.2 Normal-Mode Dielectric Loss Spectrum

Normal-mode dielectric relaxation in type-A polymers with linear mixing can be described by the DLS $\varepsilon''(\omega)$

$$\varepsilon''(\omega) = (\varepsilon_0 - \varepsilon_\infty) \cdot \sum_{j=1}^{L} \Phi_j \cdot \frac{8}{\pi^2} \cdot \sum_{\substack{p=1 \\ p:\ odd}}^{\infty} \frac{\omega \cdot \tau_j}{\left(p^4 + (\omega \cdot \tau_j)^2\right)}, \tag{4}$$

where

$$\tau_j = \tau_0 \cdot \left(\frac{M_j}{M_0}\right)^n \tag{5}$$

is the relaxation time of species with molecular weight M_j, Φ_j is the weight fraction of this species, $(\varepsilon_0 - \varepsilon_\infty)$ is the dielectric strength, n an experimentally determined exponent, τ_0 and M_0 are the reference relaxation time and molecular weight respectively. It should be noted that this definition of the DLS $\varepsilon''(\omega)$ is based on the MWDs Φ_j of the components. The prediction of the DLS from this model corresponding to the polymer mixture from above is depicted in fig. 2.

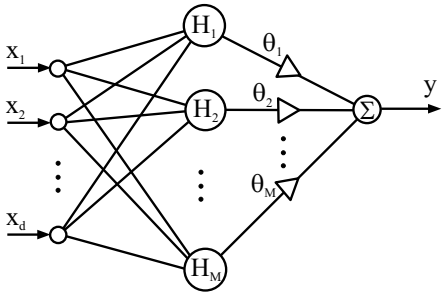

Figure 3: Diagram of a single layer basis function network with a linear combination of nonlinear basis functions H_i.

3 Basis Function Networks (BFN)

3.1 Basic Structure

In order to approximate both the MWD and the DLS by smooth functional representations with a small number of parameters we propose the use of special BFNs.

The output y of a general single layer BFN (fig. 3) may be represented by the continuous mapping f: $\mathbb{R}^d \to \mathbb{R}$

$$y = f(\boldsymbol{\Theta}, x) = \sum_{i=1}^{M} \theta_i H_i(\mathbf{x}), \tag{6}$$

where $\mathbf{x}$ is the vector of input values and $\boldsymbol{\Theta}$ is the parameter vector. It is important to note that the single layer BFN is linear in the parameters θ_i as can be seen from eq.(6). In an approximation problem it is therefore possible to compute an optimal parameter vector $\boldsymbol{\Theta}^*$ using linear standard algorithms.

The main problem of approximation is then posed by the proper choice of (nonlinear) basis functions $H_i(\mathbf{x})$.

3.2 Smooth Approximation

A possible choice of $H_i(\mathbf{x})$ are the so called radial basis functions (RBF)

$$H_i(\mathbf{x}) = \phi\left(\|\mathbf{x} - \mathbf{x}_i\|\right), \tag{7}$$

where ϕ is a continuous function from $\mathbb{R}^+$ to $\mathbb{R}$, $\|\cdot\|$ denotes the Euclidian norm, and the $\mathbf{x}_i$ are the centers of the basis functions. The choice of the function ϕ and the proper location of the centers $\mathbf{x}_i$ are the main design problems. There exist theoretical results which draw a connection between regularization theory and an optimal choice of these parameters [8, 9]. Therefore, we use RBF networks with Gaussian basis functions

$$\phi(\mathbf{x}) = e^{-\left(\frac{\mathbf{x}-\mathbf{x}_i}{a}\right)^2}, \tag{8}$$

to ensure smooth approximations of data corrupted by measurement noise. In eq.(8) $\mathbf{x}_i$ represents the center of the basis function and a is the spread (analogous to the variance of a Gaussian distribution) which was identical in all basis functions. Due to the shape of the basis function the correct asymptotic behaviour of the approximated function is always guaranteed, as opposed to [3] where a polynomial fit required additional measures.

Using the function *newrb* from MATLAB's neural network toolbox we obtained good results typically using $M = 9$ basis functions. The optimal spread a of the basis function was found by minimizing the quadratic validation error.

The measured data are now represented by approximations

$$\begin{aligned} y &= \sum_{i=1}^{M} \theta_i \phi(\mathbf{x} - \mathbf{x}_i) \\ &= \sum_{i=1}^{M} \theta_i e^{-\left(\frac{\mathbf{x}-\mathbf{x}_i}{a}\right)^2}, \end{aligned} \tag{9}$$

with M linear parameters θ_i (weights), M nonlinear parameters $\mathbf{x}_i$ (centers), and the nonlinear parameter a (spread) totalling in $2M + 1$ parameters.

3.3 Molecular Weight Distribution (MWD)

In order to reduce the number of parameters even more significantly we tried to replace the smooth approximation from eq.(9) by a basis function network with only two specially suited basis functions.

The theoretical shape of the MWD of a single polymer is given in eq.(1). It is therefore obvious to choose a basis function $\tilde{\Phi}$ of the same structure for efficient approximation of measured data. A suitable function is

$$\tilde{\Phi}(M, [a, M_n]) = \left(\frac{M}{M_n}\right)^a \cdot e^{a\cdot\left(1-\frac{M}{M_n}\right)}, \tag{10}$$

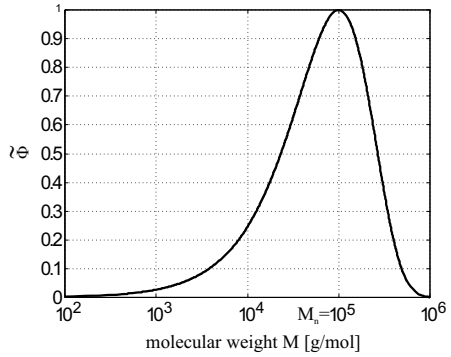

Figure 4: Basis function for approximation of MWD. $a = 1$, $M_n = 10^5\ g/mol$, and $\theta = 1$.

which follows directly from eq.(1) by normalizing the maximum to 1. The basis function $\tilde{\Phi}(M, [a, M_n])$ depends on 2 parameters: a, which defines the spread of the function ($\frac{1}{a}$ is analogous to the variance of a Gaussian distribution function) and M_n, which is the center of the basis function (weight fraction where maximum occurs). The weighting factors θ_i of the feed-forward network (fig. 3) equal the maximum values of the basis functions given in fig.4. While the network is linear in the parameters θ, the quantities a and M_n represent nonlinear parameters of the basis function.

The use of a BFN with a special basis function $\tilde{\Phi}$ can be viewed as a grey model. Although the parameter *values* are not known in advance the knowledge of the special functional *structure* can be incorporated in the model.

Since the basis function $\tilde{\Phi}$ is of the same shape as the theoretical MWD (1) it should be possible to approximate the MWD of a mixture of two polymers with only two basis functions $\tilde{\Phi}_1$ and $\tilde{\Phi}_2$ thus reducing the total number of parameters necessary for describing the MWD to 6: $a_1, a_2, M_{n1}, M_{n2}, \theta_1, \theta_2$. Among these parameters θ_1 and θ_2 are easily determined from a least squares optimization while both spread and centers of the two basis functions are determined by an iterative algorithm.

In order to obtain suitable starting values for weight and centers characteristic points in the smooth function approximation such as extrema of the function and its derivatives up to third order were calculated. The correlations between these points and the true parameter values were evaluated for varying fractional contributions γ. Overall, 6 characteristic shapes of the MWD could be found (with increasing γ): Single maximum left, single maximum plus shoulder left, local maximum left, local maximum right, single maximum plus shoulder right, and single maximum right. By building a table from these cases and the characteristic points simple decisions based on the correlations were used to determine the starting values for weights and centers.

Then, the best spread for these parameter values was determined by a simple gradient search. The performance criterion for this optimization is the quadratic error only in the data left

of the maximum for the left basis function and right of the maximum for the right basis function. This choice ensures a minimum of interference between the two basis functions and a correct shape of the distribution at the upper and lower ends.

Next, optimal centers and weights of $\tilde{\Phi}_1$ and $\tilde{\Phi}_2$ were found iteratively by alternatingly optimizing the $M_n i$ and θ_i of both basis functions. In this step the complete data are used to form the quadratic error. The iteration is terminated when the decrease in the error criterion is smaller than a given constant.

3.4 Dielectric Loss Spectrum (DLS)

The functional representation of the DLS (4) is complex and inherently requires the knowledge of the corresponding MWD. Therefore, it is not a suitable choice for a basis function. Instead, the comparatively simple Gaussian distribution function works very well in approximating a DLS. The basis function is given by

$$\tilde{\Psi}(\omega, [a, \omega_n]) = e^{-\left(\frac{\omega-\omega_n}{a}\right)^2}, \tag{11}$$

where a defines the spread of the function ($\frac{1}{a}$ is the variance of the Gaussian), ω_n is the center of the basis function (corresponding to the mean of the Gaussian). Since the function value of $\tilde{\Psi}$ depends only on $|\omega - \omega_n|$ the resulting network is a radial basis function network. The favorable properties mentioned in the previous section also exist in this case.

The iterative algorithm outlined in the previous section is also applied on the smooth approximation of the DLS, only the first part of finding characteristic points is skipped. Since the basis functions for the DLS are symmetric Gaussian distributions the choice of the initial values is less critical and the global maximum of the DLS is directly used to obtain the starting values for the first basis function.

4 Nonlinear Mapping Between Functional Representations

In the previous sections the methods and algorithms for smoothing and reducing the data of measured MWDs and DLSs have been presented. Each function is now uniquely defined by 6 parameters and the inversion problem has been transformed into a continuous mapping G: $\mathbb{R}^6 \rightarrow \mathbb{R}^6$. Instead of trying to identify this MIMO function directly the problem is split up into identification of 6 independent MISO mappings g_i: $\mathbb{R}^6 \rightarrow \mathbb{R}$. Each of these mappings g_i can be represented by a simple single hidden-layer feed-forward network (see fig.5) with L sigmoid activation functions σ where the mathematical description is given by

$$y = g_i(\Theta, \Lambda, x) = \sum_{i=1}^{L} \theta_i \sigma \left(\sum_{j=1}^{6} \lambda_{ij} x_j \right). \tag{12}$$

The justification for the replacement of a MIMO function by 6 independent MISO functions can be given as follows: Each MISO function may be constructed from the original MIMO mapping where only one output is being considered. In the case

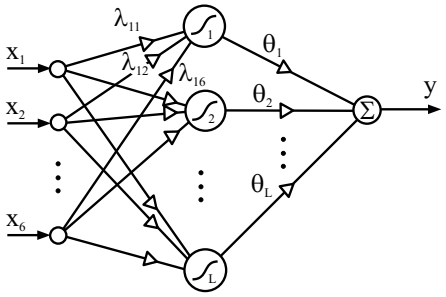

Figure 5: Diagram of a single layer feed-forward network with sigmoid activation functions σ and a linear output neuron.

of feed-forward networks, this approach allows a reduction of parameters since all connections to the discarded outputs and the associated weighting parameters may be omitted.

Again, we are using a linear output neuron which causes the parameters θ_i to be linear while the λ_is are nonlinear parameters.

The hidden layer consisted of only 6 sigmoidal neurons with an activation function

$$y(x) = \frac{2}{1 + e^{-2x}} - 1. \tag{13}$$

Training of the network (12) was done using the Levenberg-Marquard method. We used a batch algorithm where the all training data are presented to the network before the parameters are adapted. Success of the training was evaluated using validation data and the training was terminated when the validation error remained within a small tolerance band or started to rise again. In order to ensure a minimal dependency upon initial conditions several repeated training runs with varying initial conditions were performed and the best parameter set was implemented.

5 Simulation

In a simulation study we tested the ability of the proposed method to correctly predict the unknown MWD for a given DLS from a mixture of two polymers. In a chemical process plant the exact specification (MWD) of the individual components is usually given by size exclusion chromatography. Therefore, the centers M_{ni} and the spread parameters σ_i are well known. The remaining problem is the estimation of the weighting parameters θ_i. We further assume that the mixing parameter γ is varied between 0.2 to 0.3. With our choice of polymers this range covers all generic shapes of the MWD: single peak, left and right shoulders, double peak, and plateau. 21 discrete values of γ were chosen to generate complete data sets (both MWD and DLS) from equations (1), (3) and (4). The parameters were chosen as

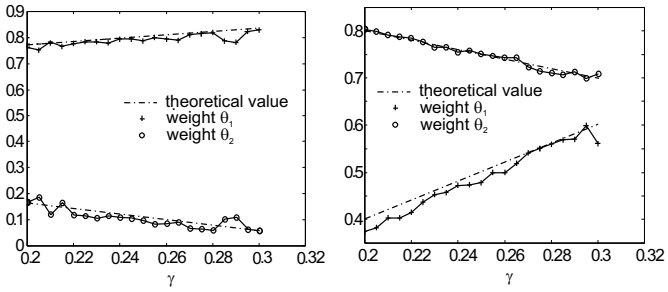

Figure 6: Left: Weighting parameters θ_1 and θ_2 of the approximated DLS for different values of γ. Right: Weighting parameters θ_1 and θ_2 of the approximated MWD for different values of γ.

$$
\begin{aligned}
M_0 &= 68\ g/mol \\
M_{n1} &= 18500\ g/mol \\
M_{n2} &= 37100\ g/mol \\
a_1 &= a_2 = 10 \\
\gamma &= 0.2 \div 0.3
\end{aligned}
$$

Each distribution and spectrum consisted of 61 discrete data points. Normally distributed noise with zero mean and a standard deviation of $s = 0.017$ was added to theoretical data.

Smoothing and data reduction is carried out efficiently with about 9 radial basis functions for the smooth approximation and 11 iterations to reduce the data to two special basis functions. The estimated parameters over the whole range of γ are close to the actual values as can be seen in figure 6. A small bias is present when noisy data are approximated, however, the final prediction is not adversely affected. The reason for the larger bias in the parameter θ_1 (see figure 6, right, and table 1) is the weakly pronounced shoulder on the left side of the MWD (see figure 8) which makes the estimation of the corresponding weighting parameter θ_1 more problematic compared to the estimation of θ_2 which is associated with a distinct peak. The situation obviously changes with different shapes of the MWD; in general, the less pronounced the effect of a single basis function is the larger the relative estimation error will be. Nevertheless, the absolute error will be still acceptable since the contribution of the uncertain parameter to the overall prediction is small in this case.

The feed-forward neural network (12) was trained using 16 out of the 21 data sets with normalized data. The network was initialized with random parameters (method of Nguyen-Widrow for hidden layer, uniformly random for output layer). A maximum of 500 epochs of batch training was carried out and the performance of the resulting network was evaluated using the 5 validation data sets.

The resulting prediction of the MWD for $\gamma = 0.255$ from the network, the theoretical distribution, and the prediction error is given in figure 8. The predicted and the actual weight parameters and the respective errors are listed in table 1. It should be noted that the DLS for $\gamma = 0.255$ does not exhibit a double peak but shows a rather distinct single maximum (see fig.7).

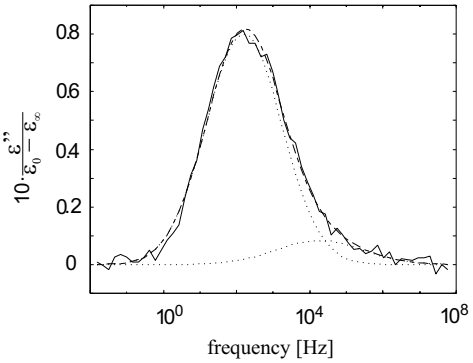

Figure 7: Theoretical DLS with measurement noise for $\gamma = 0.255$. Noisy DLS - solid, theoretical DLS - dash-dotted, basis functions - dotted.

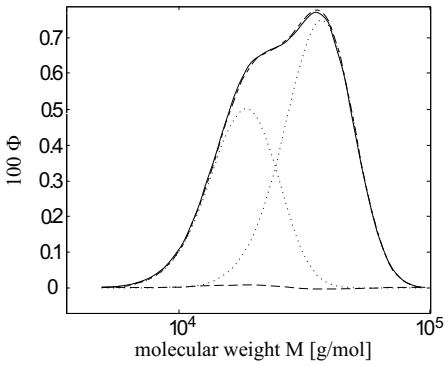

Figure 8: Predicted MWD from inversion for $\gamma = 0.255$. Predicted MWD - solid, theoretical MWD - dash-dotted, basis functions - dotted, prediction error - dashed.

6 Discussion

The simulation studies confirm that the new method works well for smoothing, data reduction, and the inversion problem. In an industrial process the measured MWD and DLS for several mixing parameters γ have to be measured in order to provide the necessary data for training the feed-forward network. This training may be done off-line and the resulting network can be applied on-line to estimate the actual composition of the polymer mixture. Computation of the predicted MWD is less time consuming than the measurement of the DLS.

There are some shortcomings to the proposed algorithm: The data reduction algorithm incorporates some heuristic elements

γ	1st weight θ_1	2nd weight θ_2
theoretical value	0.511	0.745
predicted value	0.495	0.751
abs. prediction error	-0.0163	0.00563
rel. prediction error	3.19 %	0.76 %

Table 1: Theoretical and predicted values for the weights θ_1 and θ_2 of the MWD for $\gamma = 0.255$.

(choice of initial values for iteration) although they are programmed to be executed automatically. It is also not yet optimized and a more sophisticated scheme with tuned parameters is expected to yield even better results. Moreover, we have no analytical proof of convergence and stability of our method although simulation results indicate a robust behaviour with respect to noise and no unstable runs have been observed. Training results of the feed-forward network show a well known dependency upon initial conditions. This requires repeated runs with varying starting parameters. However, the impact on overall performance is negligible since extended experiments have shown that within 4 to 5 training runs a set of parameters very close to the optimal set can be found. Moreover, the most critical part for the accuracy of the overall algorithm is clearly the smooth approximation of the noisy DLS. After a correct approximation of the DLS the training of the feed-forward net poses no problem. This is true even in on-line application since the training is very fast for this small network and 4 to 5 repeated training runs can be done in a few seconds on an average computer.

Inversion by Tikhonov regularization works only in the case of linear mixing (3) because of the assumption of a linear relation between DLS and MWD. However, the more complex theory of nonlinear mixing [3] definitely calls for a nonlinear mapping as proposed in section 4. Linear mixing may be considered a special case of the general nonlinear framework. Since the equations for nonlinear mixing contain only mild nonlinearities the smooth approximation described in section 3.2 is expected to give acceptable results. The nonlinear mapping between DLS and MWD (section 4) should be working irrespective of the method used for approximation because of the general nonlinear nature of the feed-forward network. We are currently investigating the performance of our method when applied to polymer mixtures with nonlinear mixing rules.

The extension of the proposed method to multi-component mixtures is straightforward: The part of smoothing the measured DLS with radial basis functions can remain unchanged. Data reduction has to be performed with one special basis function for each component instead of fitting only two basis functions. This will certainly call for an adaptation of the heuristic rules for the starting values and additionally, the iterative algorithm to estimate the parameters θ_i will converge more slowly. In order to investigate the feasibility of this approach further simulation studies are required.

As demonstrated by simulations the method is fairly insensitive to measurement noise and inherently provides for a regularization of the inversion problem. The use of specialized basis functions (1) and (4) automatically guarantees correct asymptotic behaviour and efficient data reduction.

In chemical engineering the method outlined here provides a useful tool to implement an automated on-line measurement of the composition of a mixture of two polymers.

7 Conclusions

A new method for inversion of the DLS of a mixture of two polymers from the measured MWD has been developed and the performance has been demonstrated by a simulation. Artificial neural networks are employed to solve the problem in two steps: First, the measured data DLS and MWD are approximated by radial basis functions to ensure correct asymptotic behaviour and a smooth function. To achieve further data reduction these smoothed functions are approximated by 2 specialized basis functions using an iterative algorithm. In the second step simple feed-forward networks with sigmoidal activation functions are used to represent the nonlinear mapping between DLS and MWD parameters. A simulation study demonstrates the performance of the method for prediction of the MWD from DLS data corrupted with measurement noise.

References

[1] Y. Imanishi, K. Adachi, and T. Kotaka, "Dielectric relaxation spectra for the bulk and concentrated solutions of *cis*-polyisoprene," *Journal of Chemical Physics*, vol. 89, no. 12, pp. 7593–7598, 1988.

[2] J. S. Fodor and D. A. Hill, "Predicted effects of polydispersity on the dielectric normal-mode relaxation of entangled *cis*-polyisoprene melts," *Macromolecules*, no. 26, pp. 5379–5388, 1993.

[3] J. S. Fodor and D. A. Hill, "Determination of molecular weight distribution of entangled *cis*-polyisoprene melts by inversion of normal-mode dielectric loss spectra," *Journal of Physical Chemistry*, vol. 98, no. 31, pp. 7674–7684, 1994.

[4] J. S. Fodor, J. R. Huljak, and D. A. Hill, "Dielectric and viscoelastic normal-mode relaxation in entangled, polydisperse *cis*-polyisoprene melts," *Journal of Chemical Physics*, vol. 103, no. 13, pp. 5725–5734, 1995.

[5] S. F. Gull and J. Skilling, "Maximum entropy method in image processing," *IEE Proceedings F*, vol. 131, no. 6, pp. 646–650, 1984.

[6] C. Elster, J. Honerkamp, and J.Weese, "Using regularization methods for the determination of relaxation and retardation spectra of polymeric liquids," *Rheologica Acta*, vol. 31, no. 2, pp. 161–174, 1992.

[7] Y. Imanishi, K. Adachi, and T. Kotaka, "Dielectric relaxation spectra for the bulk and concentrated solutions of *cis*-polyisoprene," *Journal of Chemical Physics*, vol. 89, no. 12, pp. 7585–7592, 1988.

[8] T. Poggio and F. Girosi, "A theory of networks for approximation and learning," *MIT AI Memo*, no. 1140, 1989.

[9] T. Poggio and F. Girosi, "Networks for approximation and learning," *IEEE Proceedings*, vol. 78, pp. 1481–1497, 1990.