ESTIMATING THE PARTICLE SIZE DISTRIBUTION OF A DILUTE LATEX FROM COMBINED ELASTIC AND DYNAMIC LIGHT SCATTERING MEASUREMENTS. A METHOD BASED ON NEURAL NETWORKS

G.S. STEGMAYER†, V.D. GONZALEZ‡, L.M. GUGLIOTTA‡, O.A. CHIOTTI† and J.R. VEGA†,‡

†CIDISI (UTN – Facultad Regional Santa Fe), Santa Fe, Argentina
‡INTEC (UNL – CONICET), Santa Fe, Argentina

Abstract— A method for estimating the particle size distribution (PSD) of a dilute latex from light scattering measurements is presented. The method utilizes a general regression neural network (GRNN), that estimates the PSD from 2 independent sets of measurements carried out at several angles: (i) light intensity measurements, by elastic light scattering (ELS); and (ii) average diameters measurements, by dynamic light scattering. The GRNN was trained with measurements simulated on the basis of typical asymmetric PSDs (unimodal normal-logarithmic distributions of variable mean diameters and variances). First, the ability of the method was tested on the basis of two synthetic examples. Then, the obtained GRNN was used for estimating the PSD of a narrow polystyrene (PS) latex standard of nominal diameter 111 nm. The standard was also characterized by 2 independent techniques: capillary hydrodynamic fractionation, and transmission electron microscopy (TEM). The PSD predicted by the GRNN resulted close to that obtained by TEM. The estimated PSDs were better than those obtained through standard numerical techniques for ‘ill-conditioned’ inverse problems.

Keywords— Elastic Light Scattering – Dynamic Light Scattering – Particle Size Distribution – Neural Network – Inverse Problems

I. INTRODUCTION

The particle size distribution (PSD) of a polymer latex is an important morphological characteristic that determines the processability and end-use properties of the material when used as an adhesive, a coating, an ink, or a paint. Transmission electron microscopy (TEM) is the main reference technique for directly observing and measuring the PSD. However, it is experimentally expensive, time-consuming, and difficult, mainly when analyzing soft latexes and/or broad PSDs. Also, the electron beam can produce a sample damage or a size contraction; and the PSD evaluation may involve counting thousands of particles (Llosent et al., 1996).

Several fractionation techniques (such as capillary hydrodynamic fractionation –CHDF–, field flow fractionation, hydrodynamic chromatography, and disc centrifugation), are also employed for measuring the PSDs. In particular, CHDF separates the particles according to their size, and employs a turbidity detector for determining the number of particles of each eluted fraction. CHDF is an indirect technique, since it requires a particle diameter calibration usually based on the analysis of narrow standards. Even though it is presented as a high resolution technique, it normally produces broad PSDs, as a consequence of the instrumental broadening that mainly occurs in the capillary tube (Silebi and Dos Ramos, 1989).

The so-called ensemble techniques are based on simultaneously measuring all particles in their media without previous fractionation. They include, acoustic-attenuation spectroscopy and focused-beam reflectance measurements, for concentrated systems (e.g., Scheffold et al., 2004; Li and Wilkinson, 2005); and several optical techniques such as turbidimetry, elastic light scattering (ELS), and dynamic light scattering (DLS) (Pecora, 1985; Llosent et al., 1996; Vega et al., 2003a,b). Particles in the sub-micrometer range are frequently measured by ELS and DLS (Bohren and Huffman, 1983; Chu, 1991). The instruments are similar, and basically consist of a monochromatic laser light falling onto a dilute latex sample. A photometer placed at a given detection angle,  \( \theta \), with respect to the incident light, collects the light scattered by the particles over a small solid angle. In ELS, the light intensity, \( I(\theta) \), is measured at each \( \theta \) \( (r = 1, 2, ..., R) \). The ELS measurement model is given by:

\[
I(\theta_r) = \int_0^\infty C_f(\theta_r, D) f(D) dD
\]

where \( f(D) \) is the (unknown) PSD (represented as number of particles vs. diameter \( D \)); and \( C_f(\theta_r, D) \) is the light intensity scattered by a particle of diameter \( D \) at \( \theta_r \), and it is calculated through the Mie scattering theory (Bohren and Huffman, 1983; Glatter et al., 1985).

In DLS, a devoted digital correlator measures (at each \( \theta_r \)) the second-order autocorrelation functions of...
the scattered light, \( G^{(2)}_{\theta_i}(\xi) \), for different values of the time delay, \( \xi \). This measurement is related to the (first-order and normalized) autocorrelation function of the electric field, \( g^{(1)}_{\theta_i}(\xi) \), through:

\[
G^{(2)}_{\theta_i}(\xi) = G^{(2)}_{\infty,\theta_i} \left[ 1 + \beta \left( g^{(1)}_{\theta_i}(\xi) \right)^2 \right] 
\]

where \( G^{(2)}_{\infty,\theta_i} \) is the measured baseline; and \( \beta (< 1) \) is an instrumental parameter. Then, the DLS measurement model is given by (see e.g., Vega et al., 2003b):

\[
g^{(1)}_{\theta_i}(\xi) = \int_0^\infty e^{-x/D} C_f(\theta_i, D) f(D) dD 
\]

with

\[
\Gamma_0(\theta_i) = \frac{16\pi}{3} \frac{n_m}{\lambda} \frac{k T}{\eta} \sin^2(\theta_i/2)
\]

where \( \lambda \) (nm) is the in-vacuo wavelength of the incident laser light; \( n_m \) is the refractive index of the non-absorbing medium; \( k (= 0.0138 \, \text{g nm}^2/\text{s}^2 \, \text{K}) \) is the Boltzmann constant; \( T \) (K) is the absolute temperature; and \( \eta \) (g / nm s) is the medium viscosity at the given \( T \).

The two measurement models of Eqs. (1) and (2) are represented by Fredholm equations of the first kind; and the estimation problem consists in finding \( f(D) \) through the inversion of: a) Eq. (1) from the knowledge of \( \Gamma(\theta) \), in the ELS problem; and b) Eq. (2.b) from the knowledge of \( g^{(1)}_{\theta_i}(\tau) \), in the DLS problem. Such inverse problems are typical of most measurement systems where only indirect measurements are available. Inverse problems are normally ‘ill-conditioned’; i.e., small errors in the measurement (for example, small perturbations due to the measurement noise), can originate large changes in the \( f(D) \) estimate. Regularization methods aim at improving the numerical inversion by including adjustable parameters, a priori knowledge of the solution, or some smoothness conditions (Engl et al., 1996; Tikhonov and Arsenin, 1977). While a strong regularization will produce an excessively smoothed and wide PSD, a weak regularization will originate oscillatory PSD estimates. Then, a trade-off solution must be selected.

Unfortunately, all light scattering measurements are recognized to have low information content on the PSD; and consequently only a rather poor resolution is expected. Combination of two or more independent sets of measurements allows increasing the information content, and can contribute to improving the quality of the PSD estimate (Eliçabe and Frontini, 1996; Frontini and Eliçabe, 2000; Vega et al., 2003a).

Several neural network (NN) methods have also been applied for solving inverse problems in light scattering systems. For instance, radial basis function (RBF) NNs proved adequate for simultaneously estimating the average radius and the particle refractive index in a system with homogeneous spheres (Ulanowski et al., 1998). Multilevel NNs with linear activation functions have been used for simultaneously estimating the size and parameters representative of the PSD shape, in binary mixtures of particles (e.g., spheroids and cylinders). However, as far as the authors are aware, no NN-based method has yet been published for estimating the complete PSD.

In this work, a novel method that combines ELS and DLS measurements is proposed to estimate the PSD of a latex. The estimation model is implemented through a NN, that is trained with a large set of typical asymmetric PSDs and their corresponding simulated measurements. The method is then evaluated through both synthetic and experimental examples.

**II. THE PROPOSED METHOD**

In principle, Eqs. (1, 2) can be combined in a single inversion problem for estimating the PSD from both sets of measurements. To this effect, a scalar value \( \{ I(\theta_i) \} \), at each \( \theta_i \), since each individual inversion problem (ELS and DLS) is ‘ill-conditioned’, their combination in a single inversion problem will be ill-conditioned too. As an alternative, we suggest to replace Eqs. (2.b,c) by the expression corresponding to the average diameter typically calculated in a DLS measurement at each \( \theta_i \), \( D_{\text{DLS}}(\theta_i) \). These diameters can be calculated on the basis of \( g^{(1)}_{\theta_i}(\tau) \) computed from Eq. (2.a), through the cumulants method by Koppel (1972). At each \( \theta_i \), \( D_{\text{DLS}}(\theta_i) \) is usually reported by the standard software included in the commercial equipments.

For a given PSD, \( D_{\text{DLS}}(\theta_i) \) can be theoretically calculated through the following expression:

\[
D_{\text{DLS}}(\theta_i) = \frac{\int_0^\infty C_f(\theta_i, D) f(D) dD}{\int_0^\infty f(D) dD} 
\]

For a monodisperse PSD (i.e., all particles of the same diameter, \( D_0 \)), then \( D_{\text{DLS}}(\theta_i) = D_0 \), for all \( \theta_i \). In contrast, for a wide PSD with distributed diameters, \( D_{\text{DLS}}(\theta_i) \) varies with \( \theta_i \) according to Eq. (3).

Assume now that the \( D \)-axis is discretized in \( N \) points \( \{ D_1, ..., D_N \} \) at regular intervals \( \Delta D \), in a given diameter range \( [D_{\text{min}}, D_{\text{max}}] \). Similarly, the detection angle \( \theta \) is discretized in \( R \) points \( \{ \theta_1, ..., \theta_R \} \) at regular intervals \( \Delta \theta \), in the diameter range \( [\theta_{\text{min}}, \theta_{\text{max}}] \). In the defined grids for \( \theta \) and \( D \), Eqs. (1) and (3) can be rewritten as:

\[
I(\theta_i) = \sum_{i=1}^N C_f(\theta_i, D_i) f(D_i) \quad (4.a)
\]

\[
D_{\text{DLS}}(\theta_i) = \frac{\sum_{i=1}^N C_f(\theta_i, D_i) f(D_i)}{\sum_{i=1}^N f(D_i)} \quad (4.b)
\]

Then, the estimation problem consists in finding the PSD ordinates \( f(D) \) [in Eqs. (4)], from the knowledge of
both indirect measurements: \( I(\hat{\theta}) \) and \( D_{\text{DLS}}(\hat{\theta}) \). In this work, such inverse problem will be solved through a NN-based model.

A. The NN Inverse Model

Consider an arbitrary system with ‘n’ input variables and ‘m’ output variables; and call 
\[ x = [x_1, \ldots, x_n]^T \]
and 
\[ y = [y_1, \ldots, y_m]^T \]
the input and output vectors, respectively. A NN model is useful for describing the input/output performance of such a system, from the knowledge of several \{x, y\} pairs.

A RBF NN has a simple architecture with only one hidden layer (Cybenko, 1996). The activation function of the j-th hidden unit depends on the Euclidean distance \( d \) between the input vector \( x \) and a reference vector \( c_j \), defined as the center of the neuron. The most common shape of such activation functions is a Gaussian that exhibits a radial symmetry around \( c_j \), and often a single width is used. When the units RBF (also called "smoothing parameters" or "spreads"), and often a single width is used. When the target values of training cases close to the given input vectors are closest to the new input vector. As spread becomes larger, more neurons will contribute to the average, and therefore the network function becomes smoother (Wasserman, 1993).

Before training the network, the centers of the hidden neurons must be chosen. To maximize accuracy, each training input is made to be a center, and therefore the number of hidden nodes in the GRNN model coincides with the number of training vectors.

B. The GRNN Training Stage

The GRNN was validated through simulated or "synthetic" examples, since in these cases the solutions are "a priori" known, and therefore the NN performance can be clearly evaluated. Then, the model was tested through an experimental example that involved a polystyrene (PS) latex of narrow PSD and known nominal diameter. In this case, the "true" PSD is unknown; but the best

\[ \text{RBF is wider and several neurons will respond to an input vector. In summary, the network will produce a weighted average between target vectors whose design input vectors are closest to the new input vector. As spread becomes larger, more neurons will contribute to the average, and therefore the network function becomes smoother (Wasserman, 1993).} \]

\[ \text{The } \theta \text{-range } [20^\circ-140^\circ] \text{ at regular intervals of } \Delta \theta = 10^\circ. \text{ Thus, the input vector is obtained by stacking } I(\hat{\theta}) \text{ and } D_{\text{DLS}}(\hat{\theta}), \text{ each represented by } R = 13 \text{ discrete points, and the total number of input points to the model is } 2 \times 13 = 26. \text{ The output variables, } f(D_i), \text{ resulted of } N = 211 \text{ points.} \]

Ideally, training data must be generated with the criterion of covering all possible PSDs. In our case, we have restricted the PSDs to only follow a log-normal distribution, given by:

\[ f(D_i) = \frac{1}{D_i \sigma \sqrt{2\pi}} \exp \left[ -\frac{\ln(D_i/D)}{2\sigma^2} \right] \]

where \( D_i \) represents each discrete diameter; \( \bar{D} \) is the geometric mean diameter; and \( \sigma \) is the standard deviation.

For generating the training set, \( \bar{D} \) was varied in the range [80-1000] nm, each 5 nm. At each \( \bar{D} \), 20 distributions were generated, with standard deviations in the range [0.01-0.20] nm, each 0.01 nm. Hence, 185 different \( \bar{D} \) values were considered, with 20 PSDs at each point, yielding a total of 3700 training patterns. All patterns were normalized to fall in the range [0,1]. The network perfectly learned the training data, with an approximate root mean square error (RMSE) of 10^{-7}.

III. NN-MODEL VALIDATION

Two kind of validations were implemented. First, the GRNN was validated through simulated or "synthetic" examples, since in these cases the solutions are "a priori" known, and therefore the NN performance can be clearly evaluated. Then, the model was tested through an experimental example that involved a polystyrene (PS) latex of narrow PSD and known nominal diameter. In this case, the "true" PSD is unknown; but the best
approximation is given by an independent PSD measurement as obtained from TEM.

A. Validation with simulated examples

Two asymmetric and unimodal PSDs of a PS latex were simulated. The first PSD, \( f_1(D) \), follows a log-normal distribution as given by Eq. (7), with \( D_1 = 205 \) nm, and \( \sigma_1 = 0.115 \) nm. The second PSD was assumed to follow an exponentially-modified Gaussian (EMG) distribution, that is obtained by convoluting a (symmetric) Gaussian distribution (of mean diameter \( D_2 \) and standard deviation \( \sigma_2 \)), with a decreasing exponential function (of decay constant, \( \tau \)), i.e.: 

\[
f_2(D) = \frac{1}{\sqrt{2\pi \sigma_2}} \exp\left[-\frac{(D_2 - D)^2}{2\sigma_2^2}\right] \exp\left(-\frac{D_2}{\tau}\right)
\]  

where the symbol ‘*’ represents the “convolution product”. In this example, the following parameters were adopted: \( D_2 = 340 \) nm; \( \sigma_2 = 20 \) nm; and \( \tau = 10 \) nm.

Both selected PSDs are represented in Fig 1.a). The log-normal PSD belongs to the class of distributions used for training the NN, and therefore it will be useful for evaluating the model interpolation ability. In contrast, the EMG PSD can exhibit a high asymmetry, and was selected to evaluate the ability of the network for estimating a PSD that differs in shape from those used during the training stage.

To simulate the ELS and DLS measurements, we assumed that the light scattering photometer is fitted with a vertically-polarized He–Ne laser. At the laser wavelength (632.8 nm), the particle refractive index is \( n_p = 1.5729 \), and the medium (pure water) refractive index is \( n_m = 1.3319 \). From these parameters, the \( C(I,\theta,D) \) coefficients can be evaluated through the Mie theory (Bohren and Huffman, 1983). The noise-free measurements were simulated through Eqs. (4). Then, random Gaussian noises were added to each measurement. The noise standard deviations were selected as those typically found in measurements taken from a commercial equipment, i.e.: i) 1.0% of the signal maximum, for \( I(I) \); and ii) 0.32% of the signal maximum, for \( D_{DLS}(I) \).

In Fig. 1.a), the PSDs \( f_1(D) \) and \( f_2(D) \), are represented by a continuous thick line. In Figs. 1.b,c), the noise-free measurements (in continuous line), and the noisy measurements (in symbols), are shown. From the noisy measurements, the trained GRNN predicted the PSDs, \( \hat{f}_1(D) \) and \( \hat{f}_2(D) \), represented (in thin trace) in Fig. 1.a). As expected, the prediction of \( f_1(D) \) resulted much better than the prediction of \( f_2(D) \).

B. Experimental validation

A commercial latex standard of PS and nominal diameter 111 nm was measured through the following techniques: 1) DLS; 2) ELS; 3) CHDF; and 4) TEM. The PSDs obtained from TEM and CHDF were previously reported by Gugliotta et al. (2000); and are reproduced in Fig. 2a).

### Table 1. DLS and ELS measurements.

<table>
<thead>
<tr>
<th>( \theta ) (°)</th>
<th>( D_{DLS}(I) ) (nm)</th>
<th>( I(I) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>127.1</td>
<td>1.0000</td>
</tr>
<tr>
<td>30</td>
<td>114.6</td>
<td>0.6498</td>
</tr>
<tr>
<td>40</td>
<td>116.3</td>
<td>0.4923</td>
</tr>
<tr>
<td>50</td>
<td>114.8</td>
<td>0.3957</td>
</tr>
<tr>
<td>60</td>
<td>113.2</td>
<td>0.3331</td>
</tr>
<tr>
<td>70</td>
<td>114.1</td>
<td>0.2941</td>
</tr>
<tr>
<td>80</td>
<td>114.6</td>
<td>0.2658</td>
</tr>
<tr>
<td>90</td>
<td>110.5</td>
<td>0.2438</td>
</tr>
<tr>
<td>100</td>
<td>111.5</td>
<td>0.2362</td>
</tr>
<tr>
<td>110</td>
<td>110.3</td>
<td>0.2396</td>
</tr>
<tr>
<td>120</td>
<td>109.7</td>
<td>0.2478</td>
</tr>
<tr>
<td>130</td>
<td>111.4</td>
<td>0.2771</td>
</tr>
<tr>
<td>140</td>
<td>110.7</td>
<td>0.3265</td>
</tr>
</tbody>
</table>
Alternatively, the PSD can be directly estimated from the ELS measurements, \( i(\theta) \), by numerically inverting Eq. (4.a). To this effect, Eq. (4.a) was first converted into the following vector expression:

\[
\mathbf{i} = \mathbf{C}_1 \mathbf{f}
\]

(9)

where \( \mathbf{i} = [i(\theta_1), \ldots, i(\theta_k)]^T \) is a \((R \times 1)\) vector containing the ordinates of \( i(\theta) \) at the \( R \) selected discrete angles; \( \mathbf{f} = [f(D_1), \ldots, f(D_N)]^T \) is a \((N \times 1)\) vector containing the ordinates of \( f(D) \) at the \( N \) discrete diameters; and \( \mathbf{C}_1 \) is a \((R \times N)\) matrix containing the corresponding Mie coefficients.

To solve Eq. (4.a), the Phillips-Tikhonov regularization technique was applied (Tikhonov and Arsenin, 1977). Since \( \mathbf{C}_1 \) is an ill-conditioned matrix, then a regularized pseudo-inverse of \( \mathbf{C}_1 \) may be computed through:

\[
\mathbf{C}_1^{-[1]} = (\mathbf{C}_1^T \mathbf{C}_1 + \alpha \mathbf{H})^{-1} \mathbf{C}_1^T
\]

where \( \mathbf{H} \) is a regularized matrix (Tikhonov and Arsenin, 1977), and \( \alpha \) is an adjustable parameter. Thus, the vector form of the estimated PSD, \( \mathbf{f} \), can be obtained by inverting Eq. (9) as follows:

\[
\mathbf{f} = \mathbf{C}_1^{-[1]} \mathbf{i} = (\mathbf{C}_1^T \mathbf{C}_1 + \alpha \mathbf{H})^{-1} \mathbf{C}_1^T \mathbf{i}
\]

(10)

The inversion method of Eq. (10) was applied onto the measured \( i(\theta) \) of Table 1; and the regularization parameter \( \alpha \) was adjusted by trial-and-error, in order to obtain the better possible solutions (i.e., the lower \( \alpha \) value that produces positive estimates of the PSD ordinates). The estimated PSD is indicated as “ELS” in Fig. 2a.

A visual comparison of the four PSDs indicates that the NN solution is a good approximation to the “better” estimation obtained by TEM. In contrast, CHDF and ELS estimates are clearly wider than TEM; and additionally, the ELS estimate exhibits a secondary population (of a high mean diameter) not revealed by TEM. Note that the numerical inversion of Eq. (10) only included information on the ELS measurement; and in such sense the information content is lower than in the case of the GRNN method.

From the PSDs of Fig. 2a), the corresponding average diameters \( D_{DLS}(\theta) \) were calculated through Eq. (4.b); and are represented in Fig. 2b). Additionally to \( D_{DLS}(90^\circ) \), Table 2 summarizes some characteristic mean diameters of a PSD, such as the standard number-mean diameter, \( D_{1.0} \), and the \( D_{6.5} \) diameter, defined by:

\[
D_{1.0} = \frac{\sum D_i f_i}{\sum f_i}; \quad D_{6.5} = \frac{\sum D_i^6 f_i}{\sum D_i^5 f_i}
\]

(11)

The \( D_{6.5} \) diameter approaches \( D_{DLS}(90^\circ) \) for particles in the Rayleigh scattering regime, typically observed for particles with \( D < 100 \) nm. For this reason, \( D_{6.5} \approx D_{DLS}(90^\circ) \) for the PSDs obtained by NN and TEM (see Table 2).

The prediction of \( D_{DLS}(\theta) \) from the PSD estimated by the NN is close to the experimental points determined by DLS (symbols in Fig. 2b). This result also supports the idea of the quality of the estimation of the PSD through the method here proposed.

### IV. CONCLUSIONS

A NN for estimating the PSD of a PS latex from combined ELS and DLS measurements was developed. The proposed model uses a GRNN that was trained on the basis of simulated log-normal PSDs, with particles in the diameter range [50-1100] nm.

The method was successfully evaluated on the basis of both synthetic and experimental examples. It was observed that the trained network was able of: a) adequately recovering the log-normal and EMG PSDs; and b) estimating narrow PSDs of commercial PS standards, with improved results with respect to CHDF measurements. The NN also provided better results than those obtained by solving the inverse problem related to the ELS measurements.

The trained GRNN is a fast and robust tool. With respect to the standard inversion techniques, the network presents the advantages of not requiring any numerical inversion or parameter adjustment; and it is relatively insensitive to standard noisy measurements. The success of the model is mainly due to the increased information content gained through the measurement combination. Although not shown, the performance of a GRNN developed only on the basis of ELS measurements was rather poor.

The main limitation of the model is the imposed restriction to unimodal PSDs in the selected diameter range. However, the network performance can be significantly improved by increasing the number of simulated experiments used during the training stage. For example, the accuracy of the PSD estimate could be improved by including a large spectrum of PSD shapes (e.g., bimodals PSDs), or by increasing the PSD resolu-
tion by training the network with lower discretization intervals (either $\Delta D$ or $\Delta \theta$).

ACKNOWLEDGMENTS

To CONICET, MinCyT, Universidad Tecnológica Nacional (Facultad Regional Santa Fe), and Universidad Nacional del Litoral, for the financial support.

REFERENCES


