Available online at www.sciencedirect.com
ScienceDirect
Journal of Magnetic Resonance 182 (2006) 22–28
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

Diffusion coefficient distribution from NMR-DOSY experiments using Hopfield neural network

Rita C.O. Sebastião [a,*], Carlos N. Pacheco [b], J.P. Braga [a], Dorila Piló-Veloso [a]

[a] *Depto. de Química—ICEx, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil*
[b] *Princeton University, Department of Chemistry, Princeton, NJ, USA*

Received 27 March 2006; revised 29 May 2006
Available online 27 June 2006

Abstract

Diffusion ordered spectroscopy (DOSY) is a powerful two-dimensional NMR method to study molecular translation in various systems. The diffusion coefficients are usually retrieved, at each frequency, from a fit procedure on the experimental data, considering a unique coefficient for each molecule or mixture. However, the fit can be improved if one regards the decaying curve as a multiexponential function and the diffusion coefficient as a distribution. This work presents a computer code based on the Hopfield neural network to invert the data. One small-molecule binary mixture with close diffusion coefficients is treated with this approach, demonstrating the effectiveness of the method.

Keywords: Diffusion coefficient; Inverse problem; Hopfield neural network

1. Introduction

The diffusion coefficient is an important property in studies of mixtures, providing the size and structural information of the particles [1]. In nuclear magnetic resonance, the diffusion ordered spectroscopy, DOSY experiments [2], measure this property of compounds. This technique is based on the application of the gradient field, which encodes and decodes the translational diffusion motion of the components in samples. The signal attenuation is detected and it depends on the gradient strength, the waiting time between the gradients and the diffusion coefficient of the molecules.

In general, the NMR spectrometers work with a fit procedure of the intensity decaying function, providing the diffusion coefficient as a parameter. The distribution function can also be obtained using the inverse Laplace transform. If there is more than one component in the sample with closed diffusion coefficient, the decaying intensity is composed of these contributions and the distribution function concept is more adequate to deal with these problems.

To obtain the diffusion coefficient distribution from these experiments, one has to solve a linear inverse problem [3]. Generally, these inverse problems are classified as an ill-conditioned problem with a decreasing character of the singular values and special methods, e.g., Tikhonov regularization [4,5], truncated singular value decomposition [5,6] and inverse Laplace transform [7], have to be employed.

In this work, an alternative method based on the Hopfield neural network is proposed. Experimental data of an equimolar mixture of Brucine and Isopinocampheol was used to emphasize the applicability of the method. This approach has been also successfully tested to recover the transverse relaxation time distribution from spin-echo experimental data [8], macromolecular properties from light scattering experimental data [9], and the probability density function from experimental positron annihilation lifetime spectra [7].

* Corresponding author.
E-mail address: ritacos@gmail.com (R.C.O. Sebastião).

1090-7807/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2006.06.005

2. DOSY—theoretical background

In nuclear magnetic resonance the diffusion motion is studied by the diffusion ordered spectroscopy—DOSY. This technique is based on the application of pulsed field gradient and each infinitesimal volume in the sample is encoded. The angular frequency of nuclear spins depends on its position and the magnetogyric ratio, γ, [10]

$$\omega(r) = -\gamma B(r). \tag{1}$$

The magnetic field gradient is aligned with the principal magnetic field, conventionally in the z direction. The application of the gradient induces a change in the phase angle for each spin as,

$$\phi(z) = \gamma g_z z t, \tag{2}$$

with g_z being the gradient strength and z the coordinate of spin at time t. The total change in the phase angle has another important contribution, related with the principal magnetic field, $\phi(z) = \gamma B_0 t + \gamma g_z z t$.

The pulse sequence in DOSY is similar with the spin-echo experiments. One of the most popular DOSY sequences, the BPPSTE (bipolar pulses with stimulated echo), may be described as follows: two pulsed field gradients with equivalent intensity and opposite signs are applied before and after a hard 180° pulse. This 180° pulse is applied after the first 90° and before the second 90° pulse. This second hard 90° pulse tips the magnetization towards the z-axis, and the molecule then diffuses aligned with main field. After the diffusion delay, a third 90° pulse is issued, followed by the second 180° pulse, flanked by two pulsed field gradients, also with equivalent intensity and opposite signs, aiming at regrouping the magnetization in the transversal plane, resulting in a Hahn-echo [11]. To note a change in the phase angle, the second gradient has to be applied after the spins moved, in a Δ interval. This interval is chosen depending on the diffusion coefficient of the particles.

Within the Δ interval, the change of the phase angle is proportional to the strength of the field gradient. The coherence of the signal is lost and this provides an attenuation of the registered signal. To determine the diffusion coefficient of the particles, a set of experiments has to be realized varying the gradient amplitude. This relation can be expressed by [10,11],

$$I(q) = I_0 \exp(-Dq^2\Delta'), \tag{3}$$

being I the intensity of the signal, D the diffusion coefficient, $\Delta' = \Delta - \delta/3$, with Δ the diffusion time and δ the gradient duration; $q = \gamma g \delta$, with γ being the magnetogyric ratio and g the gradient amplitude.

Generally, the diffusion coefficient is recovered in a specific chemical shift by a fitting procedure of the Eq. (3) and this methodology provides a diffusion coefficient as an appropriate parameter. Nevertheless, if the signal analyzed in the experiment is overlapped, commonly in mixtures, the diffusion coefficients for each component cannot be retrieved.

In a multi-component solution, one has to consider the signals as a sum of several decaying functions [11,12],

$$I(q, v_m) = \sum A_n(v_m) \exp(-Dq^2\Delta'), \tag{4}$$

with $A_n(v_m)$ being the intensity of the signal in NMR experiment at the frequency v_m.

With a continuous distribution of the diffusion coefficients at a particular v frequency, the Eq. (4) can be expressed as,

$$I(s) = \int \exp(-Ds) f(D) \mathrm{d}D, \tag{5}$$

with $s = q^2\Delta'$ and $f(D)$ the distribution function of the diffusion coefficients. The diffusion coefficient distribution has been obtained in NMR spectrometers by the inverse Laplace transform in the CONTIN program [8,10,12]. This code, developed by S.W. Provencher, solves the problem by numerical methods similar to the Tikhonov regularization approach, with the regularization parameter being chosen in a statistic set of data. In this work, the Hopfield neural network was proposed.

3. Inverse problems and Hopfield neural network

In a variety of problems in science, some microscopic properties only can be obtained by the solution of model functions that describes macroscopic experimental data. These model functions, as Eq. (5), are commonly known as Fredholm integral equation. The retrieval of the microscopic property, $f(D)$, from the experimental data, $I(s)$, is an ill-conditioned inverse problem and requires some appropriate techniques to be solved [13].

The standard procedure was adopted to calculate the function $\mathbf{f}$ in Eq. (5) from data $I(s)$ and $A = \exp(-sD)$. It consist in discretize the variable s in an appropriate interval and convert the integration in a convenient weighted sum to get $\mathbf{Kf} = \mathbf{I}$, being $\mathbf{I} = [I(s_1), I(s_2), \ldots, I(s_m)]^{\mathrm{T}}$, $\mathbf{f} = [f(D_1), f(D_2), \ldots, f(D_n)]^{\mathrm{T}}$ and $\mathbf{K}$ the $(m \times n)$ matrix, defined as

$$\mathbf{K} = \begin{pmatrix} w_1 A(D_1, s_1) & w_2 A(D_2, s_1) & \ldots & w_n A(D_n, s_1) \\ w_1 A(D_1, s_2) & w_2 A(D_2, s_2) & \cdots & w_n A(D_n, s_2) \\ \vdots & \vdots & \ddots & \vdots \\ w_1 A(D_1, s_m) & w_2 A(D_2, s_1) & \ldots & w_n A(D_n, s_m) \end{pmatrix},$$

in which the $w_1, w_2, \ldots, w_n$ are the appropriate coefficients determined by the quadrature to represent the integration equation.

In the Hadamard sense [14], an ill-conditioned inverse problem is established if the solution does not exist, is not unique or continuous in R^n. The decreasing character of the singular values in the $\mathbf{K}$ matrix induces an inverse matrix presenting values bigger than the original one. Therefore, in a problem with $m = n$, if the solution $\mathbf{f} = \mathbf{K}^{-1}\,\mathbf{I}$ is tried, the experimental error in $\mathbf{I}$ matrix is magnified and a wrong answer to the inverse problem is obtained.

To avoid inappropriate solutions in Eq. (5), the Hopfield neural network [3] was adapted to solve the inverse problem. This neural network is composed by a single recurrent layer with logic units fully connected. The connections between the neurons i and j are weighted by a factor T_{ij} and external contributions are considered as a $O_i(t)$ term [3,13]. Therefore, the state of the neurons, u_i, is calculated by a weighted sum of all its inputs,

$$\frac{\mathrm{d}u_i(t)}{\mathrm{d}t} = -u_i(t) + \left[\sum_{j=1}^{n} T_{ij} f(u_j(t)) + O_i(t)\right], \tag{6}$$

with $f(u_j(t)) = f_j = \tanh(u_i)$ being the activated neurons connected to the neuron i. The activation function is chosen as an increasing function, i.e., $\partial f/\partial u > 0$, to satisfy the convergence criteria in the neural network.

The neural network has an associated energy function described by,

$$E = \frac{1}{2}\left(\sum_{j=1}^{n} K_{1j} f_j - I_1\right)^2 + \frac{1}{2}\left(\sum_{j=1}^{n} K_{2j} f_j - I_2\right)^2 + \ldots + \frac{1}{2}\left(\sum_{j=1}^{n} K_{mj} f_j - I_m\right)^2, \tag{7}$$

with n being the number of points used to represent Eq. (5) and m the number of available experimental data.

The convergence criteria, $\frac{\mathrm{d}E}{\mathrm{d}t} < 0$ is satisfied if the Hamiltonian relation [3,13,15],

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} = -\frac{\partial E}{\partial f} \tag{8}$$

is established. Therefore, the Eq. (8) is developed as [3],

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} = -\left(\sum_{j=1}^{n} K_{1j} f_j - I_1\right) K_{1i} - \left(\sum_{j=1}^{n} K_{2j} f_j - I_2\right) K_{2i} - \cdots - \left(\sum_{j=1}^{n} K_{nj} f_j - I_n\right) K_{ni} \tag{9}$$

or,

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} = \sum_{j=1}^{n} T_{ij} f_j + O_i \tag{10}$$

with,

$$T_{ij} = \sum_{l=1}^{n} K_{li} K_{lj} = T_{ji} \quad \text{and} \quad O_i = \sum_{j=1}^{n} K_{ji} I_j.$$

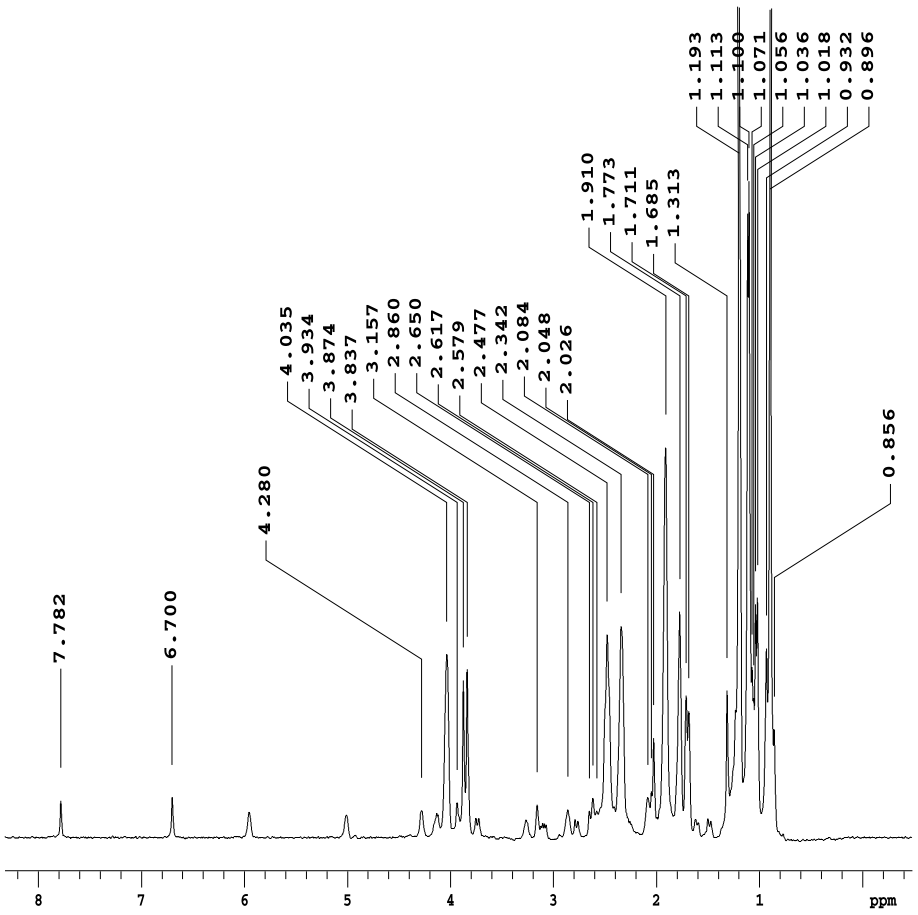

Fig. 1. Brucine and Isopinocampheol mixture spectrum.

The state of the neurons is calculated by integration of Eq. (10). A stable state, $\mathbf{f} = [f_1 f_2 \ldots f_n]$, that minimizes $\|\mathbf{Kf} - \mathbf{I}\|_2^2$ is reached and the multiple solution character of the problem can be observed along the integration procedure [16,17]. The appropriate solution is chosen based on the experimental error and chemical coherence.

4. Results and discussion

Experimental data of signal intensity, **I**, as a function of gradient strength for Brucine and Isopinocampheol 50:50 molar mixture were performed in a 500 MHz Varian spectrometer. The pulse sequence used was the Stimulated Echo Sequence with self-compensating gradient using a 2-ms purge pulse before acquisition. This sequence employs bipolar gradient pulses flanking the second 90°-pulse, which tips the magnetization along the z-direction, where it is stored during diffusion delay before the read-out. Chemical shifts, used for decaying curves measurements, were predicted by the ^{1}H NMR experiments of the mixture. The Figs. 1 and 2 present the ^{1}H NMR experiments of the mixture and substances, respectively. The spectrum of the substances was performed just to confirm the overlapping areas.

In Fig. 3 is shown the experimental decaying curve at $\delta = 1.036$ ppm. Simulated data in each chemical shift, was proposed as a bi-exponential function,

$$I(s) = \sum_{i=1}^{2} A_i \exp(-sD). \tag{11}$$

The amplitudes A_i and the s terms were determined by a fit procedure of experimental data. These simulated data were in fair agreement whit the experimental ones, as is also shown in Fig. 3.

The Eq. (5) stands for the Laplace transform of the diffusion coefficient distribution and was represented in a rectangular base as $\mathbf{Kf} = \mathbf{I}$, being 64 experimental data and $n = 264$ points. The base size and the kernel calculation were tested in the direct problem. In this case, the analytical inverse Laplace transform (ILT) was used as the solution [18],

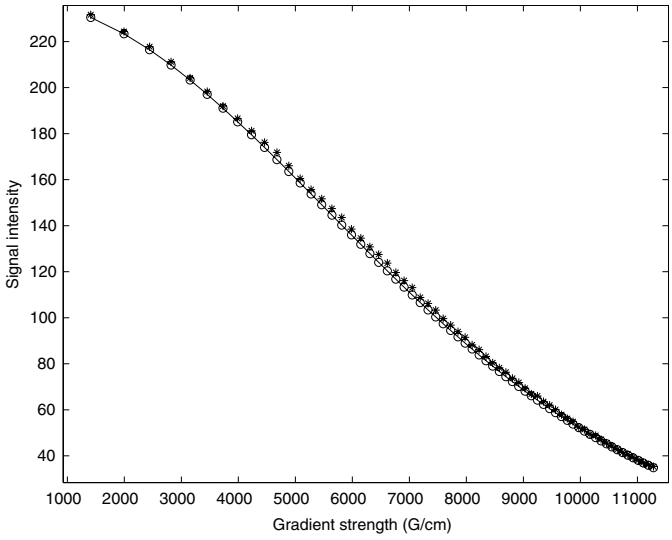

Fig. 3. Experimental signal intensity decay (*), simulated data using the Eq. (11) (full line) and the recovered data (bullets) in the direct problem, $\mathbf{Kf} = \mathbf{I}$, using the ILT distribution function, **f**, at $\delta = 1.036$ ppm.

$$f(D) = \lim_{k\to\infty} \frac{(-1)^k}{k!} \left[I^{(k)}\left(\frac{k}{D}\right) \right] \left(\frac{k}{D}\right)^{k+1}, \tag{12}$$

in which $I^{(k)}(k/D)$ is the order k derivate of $I(s)$ at the (k/D) point and D is the diffusion coefficient. Since the Eq. (11) describes the $I(s)$ function, this derivative can be expressed by

$$I^{(k)}(k/D) = (-1)^k \sum A_i D_i^k \exp\left(-D_i \frac{k}{D}\right) \left(\frac{k}{D}\right)^{k+1}, \tag{13}$$

in which the k parameter is chosen according to the diffusion coefficient magnitude, since the (k/D) term becomes infinite with k. The $k = 20$ was chosen for this problem and the Fig. 3 also presents the signal intensity recovered in the direct problem.

For each chemical shift, one inverse problem has to be performed and the distribution function is obtained by the Hopfield neural network. An initial condition to the

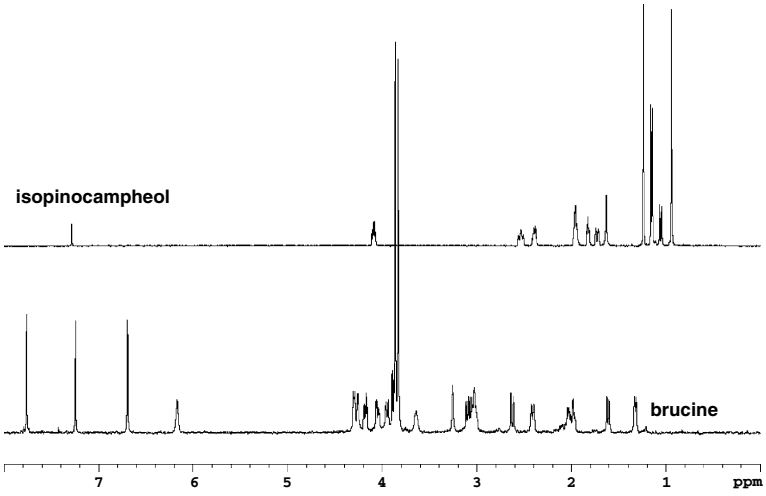

Fig. 2. Brucine and Isopinocampheol spectrum.

neurons is necessary for integration procedure of Eq. (10) and the computational time depends on the feature of this initial state of the neurons. In this sense, the analytical inverse Laplace transform, Eq. (12), can be used.

At this point is important to emphasize the decreasing energy character of the Hopfield neural network. Since the temporal derivation of energy is always negative, the initial condition given to the network is improved and the experimental data recovered in the direct problem has a smaller residual error. Another important characteristic of the network is related to the initial guess of the problem. If the initial condition corresponds to a solution of the problem, $dE/d\mathbf{f} = 0$ and the integration in neurons states is not performed.

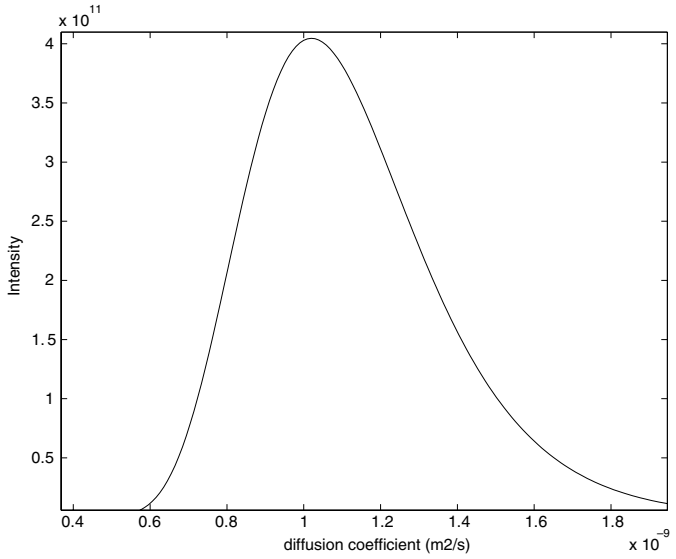

Fig. 4. Probability density function obtained by the Hopfield neural network at $\delta = 1.036$ ppm.

The Fig. 4 presents the diffusion coefficient distribution obtained from the inversion procedure at $\delta = 1.036$ ppm. At this chemical shift, there are not overlapped signals, as can be verified by the substances spectrum in Fig. 2 and the distribution function with only one peak is expected. Integration of Eq. (10) was performed until the residual error $\|\mathbf{KV} - \mathbf{I}\|_2^2$, in which $\mathbf{V}$ is the network response, reaches a desired tolerance.

To test the potentiality of this approach, in the overlapping regions a random distribution function is also given to the network as initial condition and the relative errors expressed in percent is presented in Table 1. The results obtained by the network when the ILT initial condition is considered are also presented in Table 1. One can note the excellent agreement with the experimental and the recovered data by the neural network in all chemical shifts. The diffusion coefficients, given by the center of the probability density function, are also presented.

In the overlapping regions, at $\delta = 1.685$, 1.910, 2.342 and 2.477 ppm one can observe the neural network

Table 1
Diffusion coefficients and relative error in the Hopfield neural network approach (HNN) and inverse Laplace transform method (ILT)

Chemical shift	ILT error (%) $\mathbf{V}$ = ILT response	HNN error (%) $\mathbf{V}$ = HNN distribution with 1 peak	HNN error (%) $\mathbf{V}$ = HNN distribution with 2 peaks	Random distribution function error (%) $\mathbf{V}$ = initial condition	Diffusion coefficient obtained by the HNN/(m^2/s)
0.896	1.931	0.4710			10.38 (−10)
0.932	1.909	0.5277			7.827 (−10)
1.018	1.895	0.5017			10.48 (−10)
1.036	1.797	0.4632			10.21 (−10)
1.100	1.8177	0.4794			10.73 (−10)
1.113	1.8718	0.4952			10.21 (−10)
1.193	2.060	0.5077			10.19 (−10)
1.313	1.9253	0.4997			9.287 (−10)
1.685	1.8284	0.4609	0.8035	41.30	6.254 (−10) and 10.98 (−10)
1.711	1.8708	0.4574			10.54 (−10)
1.773	1.8566	0.4654			10.44 (−10)
1.910	1.794	0.4509	0.8286	38.86	6.254 (−10) and 10.88 (−10)
2.026	1.995	0.5410	0.5409	39.13	6.254 (−10) and 10.78 (−10)
2.342	1.956	0.3614	0.6358	37.15	6.254 (−10) and 10.09 (−10)
2.477	1.72	0.4761	0.8410	47.8	6.254 (−10) and 11.28 (−10)
2.860	2.271	1.308			6.70 (−10)
3.157	2.314	1.485			5.75 (−10)
3.837	1.888	0.4389			5.91 (−10)
3.874	1.8760	0.4250			5.81 (−10)
3.934	2.347	1.509			6.00 (−10)
4.280	2.305	1.411	1.4023	55.66	6.254 (−10) and 10.88 (−10)
5.00	2.2348	1.116			7.56 (−10)
6.700	2.0488	0.9412			5.69 (−10)
7.782	2.171	1.213			5.85 (−10)

The relative error was calculated as $E\% = \sqrt{S_x^2}/\mathbf{I} \times 100$, were $\sqrt{S_x^2} = \sqrt{\sum(\mathbf{KV} - \mathbf{I})^2/N - 1}$, being $\mathbf{V}$ the neural network response or ILT response, N the number of available data and $\mathbf{I}$ the experimental data. The parentheses are for power of 10.

provides a distribution function with one peak and relative errors smaller than the ILT ones and provides a distribution function with two peaks when a random distribution function is given as initial condition. Nevertheless, in the both situations, the neural network response presents smaller relative errors than the ILT solutions that have a distribution function with only one peak. This characteristic of the inverse problems is essential to confirm the required contribution of the experimental research. Despite the relative errors, the distribution functions with two peaks have to be designed as the chemical solution of the problem.

At the chemical shifts $\delta = 2.026$ and 4.280 ppm, one can note the relative errors is smaller when the distribution function with two peaks is considered. In this case the chemical solution and the mathematic criteria are coincident. The Fig. 5 presents the distribution function obtained by the neural network at $\delta = 2.026$ ppm.

The diffusion coefficients as a function of the chemical shift were plotted in Fig. 6a. From this figure and Table 1, one can note two principal diffusion coefficients, one averaged about $6.308 \times 10^{-10}\ m^2/s$ and another in $10.47 \times 10^{-10}\ m^2/s$, characterizing the two molecules in the mixture. This DOSY-plot can also be used as an important tool to help the assignment of peaks in NMR experiments. In Fig. 6b, is shown the DOSY-plot obtained in the ILT code of the Varian spectrometer. This code is commonly used in all available spectrometers.

A close inspection of Figs. 6a and b show that the striking difference between the ILT and the HNN processing is that the latter presents a less scattered diffusion plot, which posits a higher accuracy for the HNN method. Another positive character of this method is the HNN capability of retrieve the distribution function with more than one diffusion peak per frequency, which is extremely useful in overlapped regions.

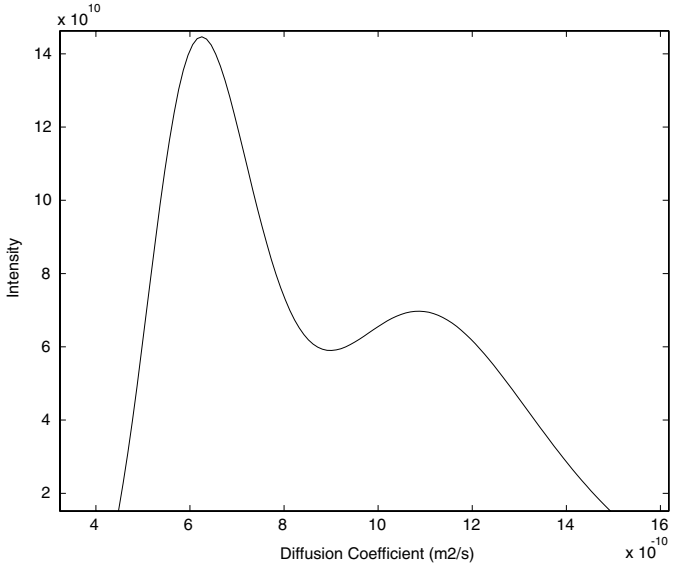

Fig. 5. Distribution function obtained by the Hopfield neural network at $\delta = 2.026$ ppm.

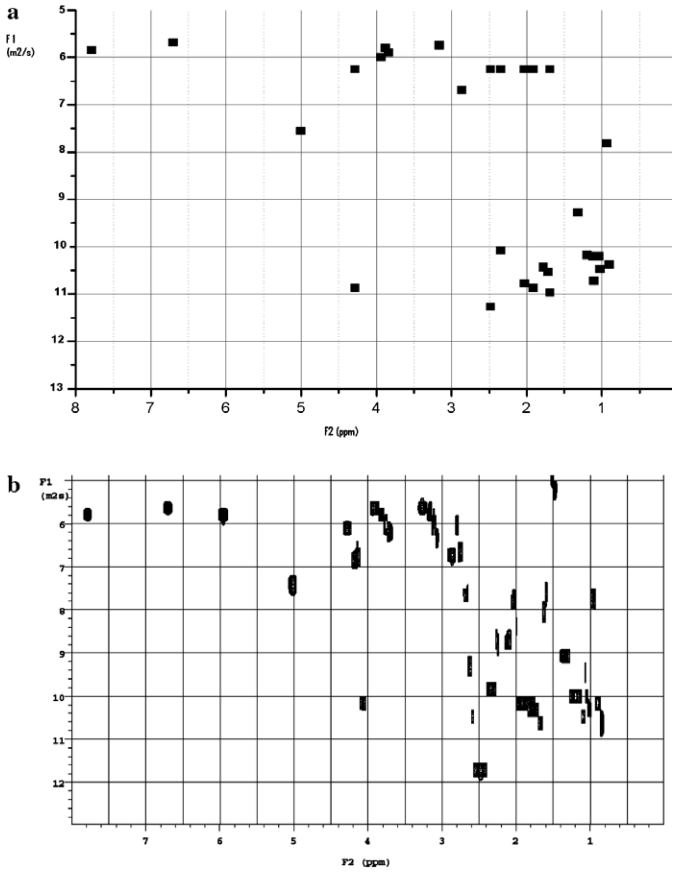

Fig. 6. Comparison between DOSY plots: (a) diffusion coefficients obtained by HNN and (b) diffusion coefficients obtained by ILT. DOSY spectrum obtained with DgcsteSL sequence, 64 gradient strengths. F1 scale is in $10^{-10}\ m^2/s$.

5. Conclusion

The diffusion ordered spectroscopy is a useful technique to study the diffusion motion in a wide variety of systems. In this work, the inversion of experimental data of Brucine and Isopinocampheol equimolar mixture was proposed using Hopfield neural network. The inversion procedure is an ill-conditioned problem and requires a special technique for its numerical treatment.

In NMR, is usual to invert problems described by the Laplace transform, as in Eq. (5), using the CONTIN program. However, this program is based on the Tikhonov regularization and has about 5000 lines with 66 subroutines. The computer code developed in the present work is composed by approximately 100 lines with the principal attributes of the Hopfield neural network combined with the analytical inverse Laplace transform.

The usage of the inverse Laplace transform as initial condition of the neurons proved to be appropriate. The recurrent neural network, due to its property of decreasing energy, will improve the initial condition, resulting in a distribution function with lower relative error. Based on the results obtained in this work, it can be concluded the

HNN can be used for DOSY processing, providing more accurate results in comparison with the ILT approach. The methodology used here is not restricted and can be extended for several problems in science.

Acknowledgments

The authors would like to thanks the CNPq and Fapemig for financial support.

References

[1] P. Stilbs, Molecular self-diffusion coefficients in Fourier transform nuclear magnetic resonance spectrometric analysis of complex mixtures, Anal. Chem. 53 (1981) 2135–2137.
[2] C.S. Johnson Jr., Diffusion ordered nuclear magnetic resonance spectroscopy: principles, applications, Prog. NMR Spectrosc. 34 (1999) 203–256.
[3] J.J. Hopfield, Neural networks and physical systems with collective computational abilities, Proc. Natl. Acad. Sci. USA 79 (1982) 2554–2558.
[4] A.N. Tikhonov, V. Arsénine, Solutions of Ill-posed Problems, Mir, Moscow, 1974.
[5] J.P. Braga, Numerical comparison between Tikhonov regularization and singular value decomposition methods using the L curve criterion, J. Math. Chem. 29 (2) (2001) 151–161.
[6] R.C.O. Sebastião, N.H.T. Lemes, L.S. Virtuoso, J.P. Braga, Nonlinear global inversion of potential energy surfaces from the experimentally determined second viral coefficients, Chem. Phys. Lett. 378 (2003) 406–409.
[7] V.C. Viterbo, R.C.O. Sebastião, R.P.G. Monteiro, W.F. Magalhães, J.P. Braga, Probability density function from experimental positron annihilation lifetime spectra, J. Braz. Chem. Soc. 16 (1) (2005) 93–97.
[8] R.C.O. Sebastião, Transverse relaxation time distribution from spin-echo experiments using Hopfield neural network, J. Magn. Reson. 177 (1) (2005) 146–151.
[9] S.F. Sun, Physical chemistry of macromolecules, Basic Principles and Issues, Wiley Interscience, New York, 1994.
[10] C.S. Johnson Jr., Diffusion ordered nuclear magnetic resonance spectroscopy: principles and applications, Prog. NMR Spectrosc. 34 (1999) 203–256.
[11] J.E. Tanner, Use of the simulated echo in NMR diffusion studies, J. Chem. Phys. 52 (5) (1970) 2523–2526.
[12] E.O. Stejskal, J.E. Tanner, Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient, J. Chem. Phys. 42 (1) (1965) 288–292.
[13] G.M. Wing, J.D. Zahrt, A Primer on Integral Equations of First Kind, SIAM, Philadelphia, 1991.
[14] J. Hadamard, Le probléme de Cauchy et les équations aux derive partielle linéaires hyperboliques, Herman, Paris, 1932.
[15] S.J. Leon, Linear Algebra with Applications, sixth ed., Prentice Hall, New Jersey, 2002.
[16] G.E. Forsythe, M.A. Malcolm, C.B. Moler, Computer Methods for Mathematical Computations, Prentice-Hall, New Jersey, 1977.
[17] G.H. Golub, C.F. Van Loan, Matrix Computations, Johns Hopkins University Press, Baltimore, 1989.
[18] D.V. Widder, Advanced Calculus, second ed., Prentice-Hall, Englewood Cliffs, NJ, 1961.