

Automatic analysis of JET charge exchange spectra using neural networks

C M Bishop†, C M Roach‡ and M G von Hellermann§

† AEA Technology, Harwell Laboratory, Oxon OX11 0RA, UK

‡ AEA Technology, Culham Laboratory, Oxon OX14 3DB, UK

§ JET Joint Undertaking, Abingdon, Oxon OX14 3EA, UK

Received 22 March 1993

Abstract. The analysis of charge exchange recombination spectra represents a very challenging problem due to the presence of many overlapping spectral lines. Conventional approaches are based on iterative least-squares optimization and suffer from the two difficulties of low speed and the need for a good initial approximation to the solution. This latter problem necessitates considerable human supervision of the analysis procedure. In this paper we show how neural network techniques allow charge exchange data to be analysed very rapidly, to give an approximate solution without the need for supervision. The network approach is well suited to the fast intershot analysis of large volumes of data, and can readily be implemented in dedicated hardware for real-time applications. The neural network can also be used to provide the initial guess for the standard least-squares algorithm when high accuracy is required.

1. Introduction

Charge exchange recombination spectroscopy (CXRS or CHERS) provides a powerful technique for the analysis of ion temperatures, densities and rotation velocities in plasma physics experiments. Interaction of a beam of neutral particles with the plasma leads to transfer of electrons to the hot ions with subsequent radiative decay of the excited states. The corresponding spectral lines exhibit Doppler broadening due to the finite temperature, and wavelength shift due to bulk plasma rotation. A complete charge exchange spectrum can often be very complex due to the presence of many overlapping lines, only some of which arise from the charge exchange process. An example spectrum from the JET tokamak, together with its decomposition into component lines, is shown in figure 1.

The conventional approach to the analysis of CXRS spectra, outlined in section 2, involves the least-squares fitting of a parametrized functional form to the observed data. Such an approach yields good accuracy and has been in widespread use for several years. It suffers, however, from two principal drawbacks. The first arises from the fact that the least-squares fitting process involves the iterative minimization of an error function and is therefore computationally intensive, and hence relatively slow. This can result in a significant computational problem for diagnostics with high spatial, temporal and spectral resolution operating on long timescale plasmas, as is typical of current systems as well as next generation designs, and would preclude its use in real-time feedback applications such as burn control. The second drawback stems from the fact that the iterative approach requires an initial guess for the parameters of the functional form, and if this guess is not sufficiently accurate then the algorithm may fail to converge in a reasonable time, or may converge to a local minimum, or may suffer from problems due to the swamping of spectral

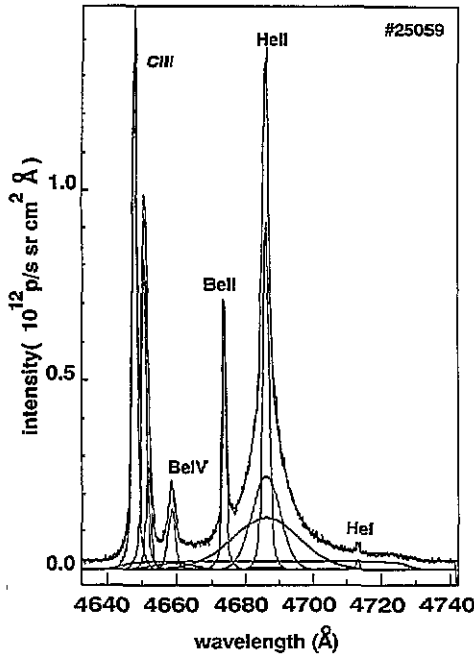


Figure 1. An example charge exchange spectrum from JET, together with its component lines obtained by the conventional least-squares technique.

lines (arising from convergence to the wrong global minimum). The provision of a suitable initial guess generally involves human intervention and this can severely limit the quantity of data which can be analysed as well as wasting valuable diagnosticians' time. It is not uncommon for a substantial fraction of the logged data to go unanalysed as a result of this problem.

In this paper we report on a novel approach to the analysis of charge exchange data based on neural network techniques. It has already been shown that neural networks provide a practical approach to the analysis of simple spectra (Bishop and Roach 1992). Here we apply neural networks to the much more complex problem of the analysis of charge exchange spectra. We show that this technique can analyse a charge exchange spectrum without the need for an initial guess, and can therefore operate automatically without requiring human supervision. Furthermore, the network algorithm does not involve iterative procedures and is therefore intrinsically very fast. The improvement in speed over the conventional approach is typically two orders of magnitude. A detailed account of our approach is given in section 3, and some preliminary results are given in section 4.

Finally, the advantages and limitations of the neural network technique are discussed in detail in section 5.

2. The conventional approach

In order to establish notation, and to provide the necessary background, we give here a brief overview of the conventional approach to the analysis of charge exchange recombination spectra. This is necessarily a somewhat superficial treatment but will serve to establish the basic principles and features. A more complete discussion can be found in von Hellermann and Summers (1992).

The raw charge exchange spectrum is composed of a number of spectral lines which have finite widths and which typically are strongly overlapping. In addition, there may be a significant background contribution arising from bremsstrahlung. The data are collected as a set of photon counts $I(\lambda_k)$ in bins at successive wavelengths λ_k , and are subject to random errors arising from the photon statistics. In most cases the principal goal is to determine the amplitude, location and width of one or more charge exchange lines, thereby allowing the ion density, toroidal velocity and temperature respectively to be calculated. If the lines were isolated this would be a straightforward problem to solve, but the presence of many overlapping lines leads to a very complex and difficult data analysis problem.

In the conventional analysis technique, a particular functional representation is chosen for the charge exchange spectrum. This would typically take the form of a sum of functions, one for each spectral line. Thus, if it is believed that the spectrum can be adequately represented in terms of N lines, the intensity of the spectrum at wavelength λ would be written as

$$I_{\text{fit}}(\lambda) = I_0 + \sum_{i=1}^N g_i(\lambda) \quad (1)$$

where I_0 represents the background of continuum radiation. Each spectral line is described by a function $g_i(\lambda)$ which is characterized by a small number of parameters whose values must be determined from the data. In the simplest case, each line would be represented by a Gaussian function

$$g_i(\lambda) = A^i \exp \left\{ -\frac{(\lambda - \lambda_0^i)^2}{2(\sigma^i)^2} \right\} \quad (2)$$

where A^i , λ_0^i , and σ^i represent respectively the amplitude, location and width of the line. In practice, the effects of instrument functions and calibration factors must also be taken into account in defining $I_{\text{fit}}(\lambda)$. For each new spectrum, the values of these parameters are determined by seeking the best fit between the parametrized form $I_{\text{fit}}(\lambda)$ and the measured spectrum $I_{\text{exp}}(\lambda)$. Closeness of fit is generally measured in terms of a sum of squares error, or χ^2 , given by

$$\chi^2 = \sum_k \frac{\{I_{\text{exp}}(\lambda_k) - I_{\text{fit}}(\lambda_k)\}^2}{\epsilon_k^2} \quad (3)$$

where λ_k is the wavelength of the k th bin, and ϵ_k is the experimental uncertainty associated with the data in bin k .

We can regard χ^2 as a function of the spectral parameters A^i , λ_0^i , and σ^i whose values can be determined by seeking a minimum of χ^2 . The dependence of χ^2 on the parameters is typically highly non-linear, and iterative algorithms must be employed. For lineshape functions which have a simple analytic form, such as the Gaussians in (2), the derivatives of χ^2 with respect to the parameters may be obtained as analytic expressions whose values can be determined numerically. It then becomes possible to use conventional gradient based optimization techniques, such as conjugate gradients, to minimize χ^2 . These make use of an initial guess for the values of the parameters, and then iteratively adjust their values using the gradient information. The standard iterative approach to the analysis of charge exchange spectra is summarized in schematic form in figure 2. An example of the analysis of a spectrum by this method is shown in figure 1.

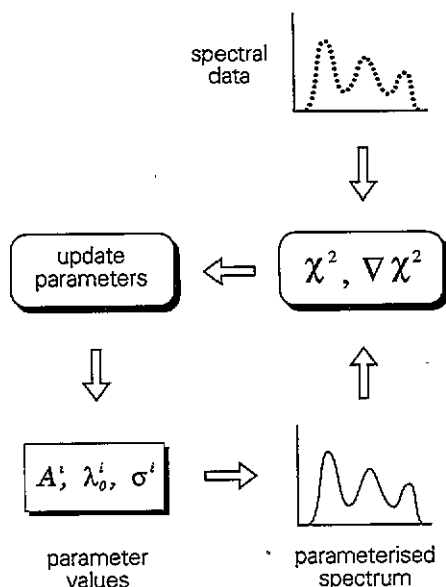


Figure 2. A schematic illustration of the conventional approach to the analysis of a charge exchange spectrum. The algorithm must iterate around the loop many times to generate a converged solution.

There will generally be several sets of values for the parameters all of which correspond to minimum χ^2 solutions. These represent solutions in which the roles of different lines have been interchanged. This represents a potential problem since it will lead to the parameter values being interpreted in the wrong way. The particular solution which is obtained will depend on the choice of initial values for the parameters, and the algorithm will only converge to the correct minimum if the initial guess is already sufficiently close to that minimum. In practice, a more severe problem is that the iteration fails to converge to a satisfactory solution within a reasonable period of time. Again, this can be attributed to the initial guess being too far from the desired solution.

A further complication arises from atomic physics considerations which result in constraints relating various spectral line parameters. These constraints must be taken into account in formulating the iterative algorithm in order to achieve satisfactory results. Their inclusion can, however, increase the robustness of the algorithm since information to determine a given parameter may be available from several parts of the spectrum. This can help to reduce the problem of parameter swap discussed above, but is generally still insufficient to allow for fully automatic analysis of charge exchange spectra.

3. The neural network approach

As we have seen, the conventional approach to the analysis of charge exchange spectra suffers from two principal difficulties. First, it corresponds to the solution of a nonlinear multivariate optimization problem and is therefore computationally intensive and hence slow. Second, convergence to the correct solution requires a sufficiently accurate initial guess for the parameters, which in turn requires significant levels of human supervision. Here we introduce a complementary approach to the analysis of such data which is able to provide a practical solution to these two problems. The ultimate goal is a fully automatic system for the routine analysis of charge exchange spectra.

Neural network algorithms draw their inspiration from studies of the detailed structure of the brain. However, they can also be regarded as a natural extension of conventional

pattern recognition techniques. An introductory review of neural networks, and a survey of some typical applications in physics and engineering, can be found in Bishop (1993). They have already been used with considerable success in several applications in fusion research (Bishop and Roach 1992, Allen and Bishop 1992, Bishop *et al* 1993, Lister and Schnurrenberger 1991, Bishop *et al* 1992).

In this paper we show how a particular type of network, known as the multilayer perceptron (MLP), can be applied to the automatic analysis of charge exchange spectra. The MLP can be regarded as a class of nonlinear functions which map a set of input variables to a set of output variables. The number of input and output variables is generally determined by the problem being solved. The particular function which the MLP implements is determined both by the architecture of the network and by the values of a set of parameters known as weights. It can be shown that the MLP provides a universal procedure for generating nonlinear mappings between multi-dimensional spaces in the sense that, for a sufficiently large network, there will exist a choice of weight values such that the network mapping approximates any given functional mapping to arbitrary accuracy. Procedures for determining the appropriate architecture for a network are reviewed in Bishop (1993).

The determination of the optimum values for the weights in a neural network is known as network training and depends on the provision of a suitable set of training data. This data comprises a large number of examples of the desired mapping, i.e. examples of input vectors and their corresponding output vectors. Network training is a computationally intensive process. Once the network is trained, however, it can process new data very rapidly. Detailed accounts of the structure of the MLP, and of various techniques used for MLP training, can be found in Bishop and Roach (1992) and Bishop (1993).

The analysis of a charge exchange spectrum can be reformulated as a nonlinear mapping problem. Inputs consist of the spectral data expressed as the amplitude of the spectrum in each wavelength bin. Generally, some initial pre-processing is applied to the data, as discussed in section 4. This input data must be transformed to a set of output variables which comprise the parameters of the constituent spectral lines. The required mapping can be performed using an MLP neural network. Training of the network requires a data set of charge exchange spectra for which the corresponding spectral line parameter values are known. This data set can be provided with the aid of the standard least-squares technique. We can therefore consider the neural network approach as three separate stages:

- (1) Collect a suitable set of charge exchange spectra and determine the corresponding values of the spectral line parameters using the standard iterative least-squares approach. This may necessitate some degree of human supervision, both to ensure satisfactory convergence of the algorithm, and for careful validation of the results.

- (2) Use the data set from step 1 to train a network. In practice the data set is split into training and test sets which allows an optimization of the network architecture to be performed. This is achieved by training several networks with different topologies and selecting the one with the smallest test error procedure. A more systematic version of this is known as cross-validation (Bishop 1993).

- (3) The trained network can now be used to process new spectra, and thereby determine the spectral parameter values. The network is said to be operating in feedforward mode since the outputs are calculated as an explicit function of the inputs. This is illustrated schematically in figure 3. There is no iteration involved in this process which is therefore typically very fast.

The data set of charge exchange spectra needs to be sufficiently large in two senses. First, it should span the range of parameter values which are likely to be encountered

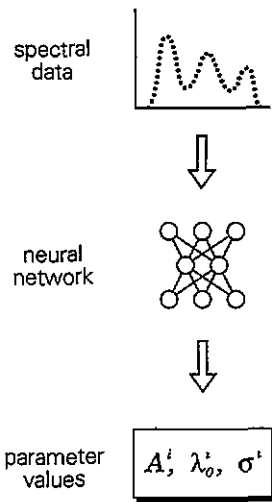


Figure 3. A schematic illustration of the neural network analysis of a charge exchange spectrum. The processing is performed in a single step and is therefore very fast.

during stage 3. This is necessary because the network mapping effectively 'interpolates' between training points and cannot be expected to give satisfactory results if presented with substantially novel data. This point is discussed further in section 5. Second, the network training procedure must be 'overdetermined' in that the number of training examples should be sufficiently large in relation to the number of degrees of freedom in the network. Additionally, the data set should be large enough to allow division into training and test sets while still satisfying these two criteria.

4. Preliminary results

In order to illustrate the neural network approach to charge exchange analysis we present some preliminary results from the analysis of spectra from the JET tokamak. A data set of around 3000 fully analysed Be IV/C III spectra has been assembled (see figure 1 for an example). Each spectrum contains up to 11 significant transitions together with an approximately uniform background due to bremsstrahlung. Since each spectral line is characterized by its amplitude, location and width, this represents a total of 34 parameters (including the amplitude of the background) to be determined for each spectrum. Atomic physics constraints reduces the number of independent parameters to 18.

The raw spectra consist of amplitude measurements at 600 equally spaced wavelengths. This itself raises a difficulty since the input data effectively lives in a 600 dimensional space. A simplistic approach of presenting all of this data to a network with 600 inputs would give poor results for three reasons. First, the network would be very large and would therefore need long training times. Second, very large data sets would be needed to ensure that the training represented an overdetermined problem as discussed in section 3. Third, the accuracy of the predictions made by the network would tend to deteriorate severely as a result of the 'curse of dimensionality', a well known problem in pattern recognition (Bishop 1993). As the dimensionality of the input space grows, the difficulty of building an effective pattern recognition system, or in this case a nonlinear mapping system, increases dramatically (Duda and Hart 1973). We have addressed this problem initially by techniques such as the selection of a subset of the bins as inputs, and by using several networks each dealing with subsections of the spectrum. This latter approach is not ideal, since the

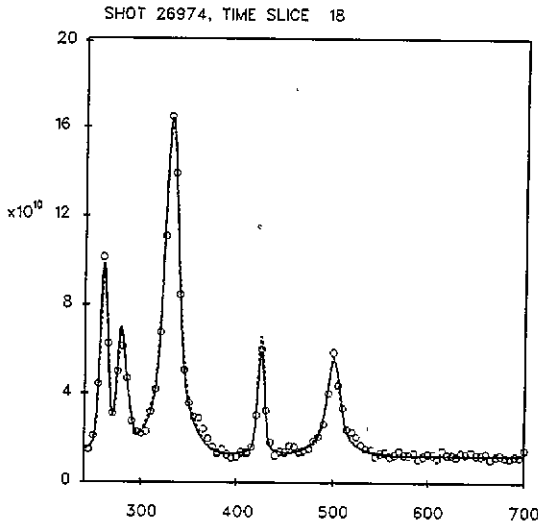


Figure 4. Plot of a conventional (solid) and neural network (dashed) fit to a JET charge exchange spectrum. The circles indicate every fifth spectral data point.

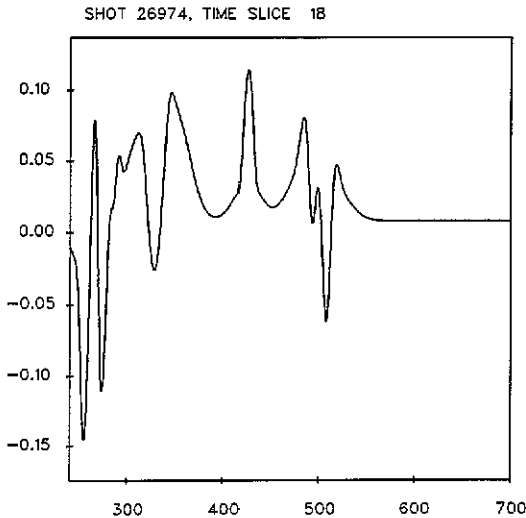


Figure 5. Plot of the residual error between the neural network and conventional spectral reconstructions, expressed as a fraction of the conventional spectrum amplitude, for the spectrum shown in figure 4.

presence of the atomic physics constraints implies that information to determine a particular parameter may be available from more than one part of the spectrum and so a given network may not have access to all relevant information. The largest network trained had 50 inputs, 15 hidden units and 6 outputs. More sophisticated approaches to this problem are under development, as will be discussed in the next section.

Initial results are very encouraging. The neural network outputs generally show good agreement with the corresponding conventional results, with ion temperatures being determined to around 30% accuracy. It is anticipated that there is still significant scope for improvements in the results as a consequence of the adoption of more sophisticated preprocessing techniques. Figure 4 shows an example of the neural network and conventional fits to a complete spectrum. The circles show every fifth data point from the experimental spectrum. The reconstructed spectrum, obtained from the neural network predictions for the spectral line parameters, is shown by the dashed curve, and the

corresponding spectrum obtained by the conventional approach is shown by the solid curve. Since the two curves are nearly coincident, we have plotted the residual difference between them, expressed as a fraction of the conventional spectrum amplitude, in figure 5. These results were obtained using four networks to extract different sets of parameters. The network approach (without optimization) requires 15 ms to process a new spectrum as compared with an average of 500 ms for the conventional algorithm (both running on the JET IBM mainframe).

5. Discussion

We have shown how neural network techniques can be applied to the problem of analysing charge exchange spectra. Such an approach allows a spectrum to be analysed without human supervision, and also greatly improves the speed of analysis. Input spectral data is mapped, via an explicit functional transformation, directly onto the desired spectral line parameters. The network could equally well be trained to generate physical quantities such as temperatures, densities and rotation velocities directly at the outputs. Note, however, that it is generally helpful to provide at least some simple normalization of the data prior to its use in network training (Bishop 1993).

In general, we do not expect the network to achieve as high an accuracy as the conventional least-squares approach. Nevertheless, the speed of the network makes it ideally suited to processing large volumes of data very rapidly for intershot analysis. Furthermore, the solution generated by the network can be used as an initial guess for the iterative code. This would allow the iterative code to run essentially without human guidance since the network provides a reasonably accurate starting condition. Such a hybrid technique would avoid the principal difficulty of the standard approach, namely the need for human intervention to provide the initial guess. In addition, by starting from a good initial set of parameter values we would expect the iterative algorithm to converge in fewer steps, leading to improved speed. (For situations where a succession of very similar spectra are to be analysed, the result of fitting one spectrum can be used to provide the initial guess for the next spectrum). This hybrid method provides an automatic analysis procedure which also achieves high accuracy.

Since the basic elements of the neural network algorithm involve vector-matrix operations they can be implemented very efficiently on vector and pipelined processors thereby giving very fast data throughput. For extremely high processing speeds, as would be required for real-time applications such as burn control, the neural network is ideally suited to implementation in special purpose highly parallel analogue hardware. Such hardware has already been developed for the neural network approach to equilibrium feedback control on the COMPASS tokamak (Bishop *et al* 1992).

As we have already indicated, there are two key issues associated with the neural network approach to the analysis of charge exchange spectra. The first of these stems from the high dimensionality of the raw data and manifests itself in the 'curse of dimensionality' familiar in pattern recognition problems. The solution is to preprocess the data to reduce its dimensionality, a technique known as feature extraction. The values of the features can be taken as inputs to the network which is then trained in the usual way. Selection of the most appropriate features to extract is an important problem which has a direct bearing on the performance of the complete system, and work on this is underway.

The second key issue regarding the neural network technique concerns the capability of the network to deal adequately with data which is significantly different from that on which

it was trained. In general we expect the network to perform well only when presented with data which is similar to that used during training. For a system which is to run automatically it is necessary to provide some form of validation to ensure that the network outputs are satisfactory. One approach is to try to detect the presence of novel input data, and this has been demonstrated in a different context in Bishop and James (1992). In the present application, however, we can adopt an alternative procedure. The determination of the spectral line parameters from the spectral data can be regarded as an 'inverse' problem. The 'forward' problem consists of the evaluation of the spectrum for a given set of parameters, and is a well defined calculation with a unique solution. By contrast the inverse problem is highly nonlinear and, as we have seen, may exhibit many potential solutions, only one of which is physically meaningful. We can check the neural network solution to the inverse problem by running it through the forward problem, that is by evaluating the spectrum predicted by the parameter values generated by the network, and computing χ^2 . If the reconstructed spectrum corresponds closely to the original spectrum we can be reasonably sure that the network outputs are satisfactory.

As we have seen, the neural network approach to charge exchange is complementary to conventional iterative techniques. It can be used alone for high speed analysis of the data, or combined with standard methods to give a hybrid approach, for situations in which very high accuracy is desired.

Acknowledgments

We would like to thank Dr P R Thomas of JET for the suggestion that charge exchange spectra would make an interesting and valuable subject for neural network analysis. Other JET staff have also provided useful contributions and we particularly wish to thank J Ellis, J Frieling and M Wallwork.

References

- Allen L and Bishop C M 1992 *Plasma Phys. Control. Fusion* **34** 1291
- Bishop C M 1993 Neural networks and their applications *Rev. Sci. Instrum.* to appear
- Bishop C M, Cox P M, Haynes P S, Roach C M, Smith M E U, Todd T N and Trotman D L 1992 A neural network approach to tokamak equilibrium control *Applications of Neural Networks* ed J G Taylor (Berlin: Springer) p 114
- Bishop C M and James G D 1993 Analysis of multi-phase flows using dual-energy gamma densitometry and neural networks *Nucl. Instrum. Methods A* **327** 580
- Bishop C M and Roach C M 1992 *Rev. Sci. Instrum.* **63** 4450
- Bishop C M, Strachan I G D, O'Rourke J, Maddison G P and Thomas P R 1993 *Neural Comput. Appl.* **1** 4
- Duda R and Hart P 1973 *Pattern Classification and Scene Analysis* (New York: Wiley)
- Lister J B and Schnurrenberger H 1991 *Nucl. Fusion* **31** 1291
- von Hellerman M and Summers H P 1992 *Rev. Sci. Instrum.* **63** 5312