UCRL-JC-143205
Preprint

SPATIAL TREATMENT OF THE SLAB-GEOMETRY DISCRETE ORDINATES EQUATIONS USING ARTIFICIAL NEURAL NETWORKS

Patrick S. Brantley

This paper was prepared for submittal to
The 2001 ANS International Meeting on
Mathematical Methods for Nuclear Applications
Salt Lake City, Utah, USA
September 9-13, 2001

April 2001

Lawrence Livermore National Laboratory

DISCLAIMER

M&C 2001 Salt Lake City, Utah, USA, September 2001.

SPATIAL TREATMENT OF THE SLAB-GEOMETRY DISCRETE ORDINATES EQUATIONS USING ARTIFICIAL NEURAL NETWORKS

Patrick S. Brantley
Lawrence Livermore National Laboratory
P.O. Box 808, L-23
Livermore, CA 94551
brantley1@llnl.gov

Keywords: Neutron transport, discrete ordinates approximation, artificial neural networks, differential equations

ABSTRACT

An artificial neural network (ANN) method is developed for treating the spatial variable of the one-group slab-geometry discrete ordinates (S_N) equations in a homogeneous medium with linearly anisotropic scattering. This ANN method takes advantage of the function approximation capability of multilayer ANNs. The discrete ordinates angular flux is approximated by a multilayer ANN with a single input representing the spatial variable x and N outputs representing the angular flux in each of the discrete ordinates angular directions. A global objective function is formulated which measures how accurately the output of the ANN satisfies the discrete ordinates equations and boundary conditions at specified spatial points. Minimization of this objective function determines the appropriate values for the parameters of the ANN. Numerical results are presented demonstrating the accuracy of the method for both fixed source and incident angular flux problems.

1. INTRODUCTION

Artificial neural network (ANN) methods have been researched extensively within the nuclear community for applications in systems control, diagnostics, and signal processing. Multilayer ANNs are capable of representing continuous functions to an arbitrary accuracy (Fausett, 1994). Despite this ability, ANNs have not been previously considered as an approach for obtaining numerical solutions of neutron transport problems. In this paper, we investigate ANNs as a method for treating the spatial variable in the discrete ordinates (S_N) approximation to the slab geometry neutron transport equation.

We consider the use of multilayer perceptron ANNs (Fausett, 1994) as an alternative to standard spatial discretization methods for the slab-geometry discrete ordinates equations. The work presented in this paper is based on a method proposed by van Milligen et al. (1995) to obtain solutions of the differential equations arising in plasma physics applications. Brantley (2000) recently utilized the method to obtain solutions of slab geometry neutron diffusion problems with Marshak boundary conditions.

The proposed ANN method approximates the discrete ordinates angular flux as a multilayer ANN. The ANN has a single input representing the spatial variable x and N outputs representing the angular flux in each of the discrete ordinates angular directions. Since the ANN approximation to the angular flux is continuous and differentiable, the accuracy with which the ANN approximation to the angular flux satisfies the discrete ordinates equations and boundary conditions can be evaluated analytically. A global objective function is formulated which measures how accurately the output of the ANN satisfies the discrete ordinates equations and boundary conditions at specified spatial points. Minimization of this objective function determines the appropriate values for the parameters of the ANN. The ANN can then be used to compute the angular flux at any point in the spatial domain. Hence, any angular moments such as the scalar flux and current can also be readily computed. As opposed to traditional ANN applications, no a priori information regarding the discrete ordinates solution is required.

This paper is organized as follow. In Sec. 2, we describe the discrete ordinates transport problem to be solved, provide a brief description of the ANN architecture used, and present the ANN solution algorithm. In Sec. 3, we present numerical results from both a fixed source and an incident angular flux problem. We conclude with a brief discussion in Sec. 4.

2. ARTIFICIAL NEURAL NETWORK SOLUTION ALGORITHM

We consider the following discrete ordinates (S_N) problem in a homogeneous slab $x_L \le x \le x_R$ with linearly anisotropic scattering:

$$\mu_n \frac{d}{dx}\psi_n(x) + \sigma_t \psi_n(x) = \frac{1}{2}[\sigma_{s0}\phi_0(x) + 3\mu_n \sigma_{s1}\phi_1(x)] + \frac{1}{2}Q \ , \quad x_L \le x \le x_R \ , \ 1 \le n \le N \ , \tag{1a}$$

$$\psi_n(x_L) = \psi_n^{inc} \ , \ \mu_n > 0 \ , \tag{1b}$$

$$\psi_n(x_R) = \psi_n^{inc} \ , \ \mu_n < 0 \ . \tag{1c}$$

Our notation is standard (Lewis, 1993): $\psi_n(x)$ is the angular flux of particles traveling in the discrete ordinates direction μ_n, σ_t is the macroscopic total cross section, σ_{s0} and σ_{s1} are the isotropic and linearly anisotropic components of the macroscopic differential scattering cross section, respectively, Q is an interior source, and ψ_n^{inc} is a prescribed incident angular flux on the boundary. Here we use a standard even-order ($N \ge 2$) Gauss-Legendre quadrature set $\{(\mu_n, w_n) \mid 1 \le n \le N\}$ with the weights normalized such that $\sum_{n=1}^{N} w_n = 2$. The neutron scalar flux and current are computed from the discrete ordinates angular flux by

$$\phi_0(x) = \sum_{n=1}^{N} \psi_n(x) w_n \quad , \tag{2}$$

and

$$\phi_1(x) = \sum_{n=1}^{N} \mu_n \psi_n(x) w_n \quad , \tag{3}$$

respectively.

Standard approaches to the spatial discretization of Eq. (1a) involve exact integration over a spatial cell coupled with approximate auxiliary conditions relating the resulting cell-edge and cell-average angular fluxes. Instead, we propose to approximate the discrete ordinates angular flux by an ANN that has a single input for the spatial variable x and N outputs Y_n representing the angular flux in each of the discrete ordinates angular directions. We consider in this paper only multilayer, feedforward artificial neural networks, often called multilayer perceptrons (Fausett, 1994).

Consider an ANN with a single input node x, one hidden layer with J nodes, and N output nodes Y_n representing the discrete ordinates angular flux. This ANN architecture is shown schematically in Fig. 1, where connections between nodes explicitly represented are shown. At least one hidden layer is necessary so that the ANN can learn a continuous mapping to an arbitrary accuracy (Fausett, 1994). Each hidden and output layer node has a so-called bias term representing a base level of activation for the node. This bias term increases the flexibility and accuracy of the ANN. In addition, each hidden and output layer node has an activation function that determines the response of the node to the input it receives from nodes in the previous layer. We denote the connection weight from the input layer node to the j^{th} hidden layer node as α_j, the hidden layer bias as u_j, and the hidden layer activation function as $f_h(s)$. We denote the connection weight from the j^{th} hidden layer node to the n^{th} output layer node as β_{jn}, the output layer bias as v_n, and the output layer activation function as $f_o(t)$. The total number of parameters in the ANN (connection weights and bias terms) is then given by $2J + (J+1)N$.

The standard feedforward procedure for computing the output of the multilayer ANN is as follows (Fausett, 1994). For a specific input x value, the computed input to the j^{th} hidden layer node is given by

$$s_j = \alpha_j x + u_j \quad . \tag{4}$$

Given this input, the response or *activation* of the j^{th} hidden layer node is obtained by applying the hyperbolic tangent activation function given by

$$f_h(s_j) = \tanh(s_j) = \frac{1-\exp(-2s_j)}{1+\exp(-2s_j)} \quad . \tag{5}$$

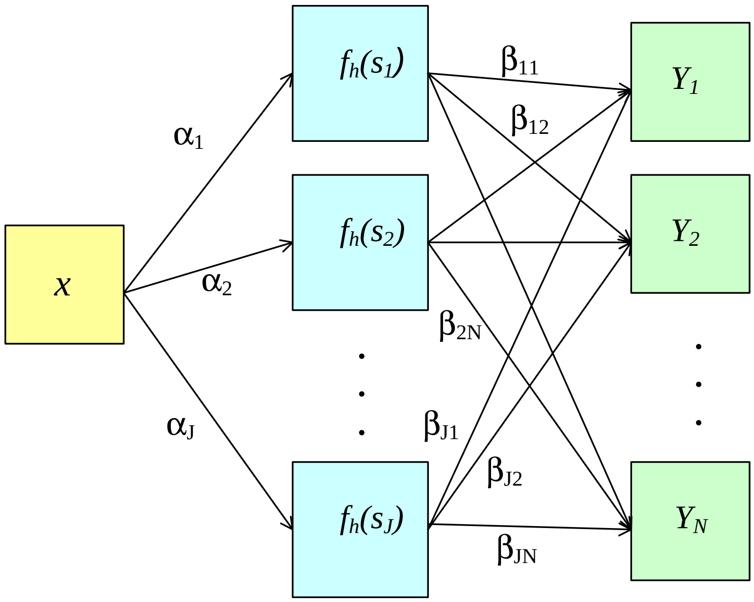

Fig. 1 Multilayer ANN architecture.

This function is a continuous analog to the hidden layer node being "on" for positive argument values and "off" for negative argument values. We note that other forms of the activation function, such as binary or bipolar sigmoid, can also be used (Fausett, 1994). Given the activation of each of the hidden layer nodes, the input to the n^{th} output layer node is computed as

$$t_n = \sum_{j=1}^{J} \beta_{jn} f_h(s_j) + v_n \quad . \tag{6}$$

The ANN approximation of the angular flux in the discrete ordinates direction μ_n is obtained by applying the output layer activation function:

$$\begin{aligned} \psi_n \cong Y_n &= f_o(t_n) \\ &= t_n \quad . \end{aligned} \tag{7}$$

As implied by Eq. (7), the identity function was used as the output layer activation function for all of the work presented in this paper. We discuss a possible alternate form for the output layer activation function in Sec. 4.

The connection weights and bias terms are typically initialized to small random values and are systematically adjusted during the *training* of the ANN. The ANN is trained to approximately satisfy the discrete ordinates problem Eqs. (1) by minimizing an objective function E given by

$$\begin{aligned} E = & \frac{1}{2}\sum_{i=1}^{I}\gamma_i \sum_{\mu_n}\left\{\mu_n \frac{d}{dx}Y_n(x_i) + \sigma_t Y_n(x_i) - \frac{1}{2}\left[\sigma_{s0}\phi_0^{ann}(x_i) + 3\mu_n\sigma_{s1}\phi_1^{ann}(x_i)\right] - \frac{1}{2}Q\right\}^2 \\ & + \frac{1}{2}\gamma_L \sum_{\mu_n>0}\left[Y_n(x_L) - \psi_n^{inc}\right]^2 \\ & + \frac{1}{2}\gamma_R \sum_{\mu_n<0}\left[Y_n(x_R) - \psi_n^{inc}\right]^2 \quad , \end{aligned} \tag{8}$$

where I is the number of spatial training points (including two points on the boundaries of the slab) and $\phi_0^{ann}(x)$ and $\phi_1^{ann}(x)$ are the ANN approximations of the neutron scalar flux and current given by

$$\phi_0^{ann}(x) = \sum_{n=1}^{N} Y_n(x) w_n \quad , \tag{9}$$

and

$$\phi_1^{ann}(x) = \sum_{n=1}^{N} \mu_n Y_n(x) w_n \quad , \tag{10}$$

respectively. The sum over μ_n in the first term of Eq. (8) represents a sum over all discrete ordinates directions in the interior of the slab and a sum over exiting directions on the boundaries of the slab. The incident directions on the boundaries of the slab are accounted for in the second and third (boundary condition) terms of the objective function. The γ weights in Eq. (8) serve both to increase the accuracy and efficiency of the ANN method, as described below. We note that, as opposed to typical ANN training methods, no a priori information regarding the discrete ordinates solution is required.

The objective function given by Eq. (8) is a measure of how accurately the ANN approximation to the angular flux satisfies the discrete ordinates problem at the spatial training points. This objective function is minimized by systematically adjusting the weights of the ANN using a minimization algorithm. We utilize the conjugate gradient and BFGS variable metric methods as implemented in the CONMIN algorithm (Shanno, 1980). The derivatives arising from the action of the discrete ordinates transport operator on the ANN can be calculated analytically. In addition, the gradient of the objective function with respect to the ANN parameters, utilized by the minimization algorithm, can be calculated analytically. The minimization is deemed to have converged when $\left\| \underline{\nabla}_p E \right\|_2 \le \varepsilon \max\left(1, \left\| \underline{p} \right\|_2\right)$, i.e. the L_2 norm of the gradient of the objective function with respect to the ANN parameters is less than or equal to the convergence criterion ε times the maximum of one and the L_2 norm of the ANN parameter vector $\underline{p}$.

The training of the ANN requires minimization of a nonlinear objective function, which can be susceptible to trapping in local minima. The γ weights in Eq. (8) are adapted during (but independent from) the minimization procedure without user intervention and are set to larger values at spatial training points with larger residuals. These weights help to alleviate the local minima problem by focusing the effort of the minimization on the spatial points with larger residuals. In addition, the γ terms incorporate an inverse scalar flux weighting such that all spatial training points contribute roughly equally to the objective function regardless of the magnitude of the scalar flux at the point. This was found to be important for obtaining consistently accurate results, particularly in deep penetration problems.

After the ANN is trained, the angular flux can be efficiently computed at any spatial point in the slab. In addition, any angular moments of the angular flux (such as the scalar flux or current) can be readily computed from this angular flux.

3. NUMERICAL RESULTS

To demonstrate the accuracy of the ANN method, we consider two numerical problems solved in a homogeneous slab $0 \le x \le 1$ with total cross section $\sigma_t = 10$ (i.e. a ten mfp slab). Problem I has isotropic scattering with $\sigma_{s0} = 9.5$, a constant interior source $Q = 0.5$, and vacuum boundaries. Problem II has linearly anisotropic scattering with $\sigma_{s0} = 9.5$ and $\sigma_{s1} = 5.0$, an isotropic incident angular flux on the left edge of the slab of $\psi_n^{inc} = 1$, $\mu_n > 0$, no interior source, and a vacuum boundary on the right edge of the slab. We solve these problems using quadrature sets of order two and four and compare to fine-mesh discrete ordinates reference solutions of the same quadrature order obtained using a standard linear discontinuous spatial discretization (Lewis, 1993). For the S_2 case, the ANN used five hidden layer nodes and was trained using eleven spatial training points spaced one mean free path apart (one point on each of the two boundaries and nine interior points). For the S_4 case, the ANN used ten hidden layer nodes and was trained using twenty-one spatial training points spaced one-half mean free paths apart (one point on each of the two boundaries and nineteen interior points). The trained ANN was then used to compute the scalar flux at 1000 equally-spaced intervals within the slab. All calculations utilized a minimization convergence criterion of $\varepsilon = 10^{-6}$.

Both the discrete ordinates angular flux and the scalar flux for Problem I are concave-down. Fig. 2 plots the S_2 scalar flux and the percent relative error (PRE) in the scalar flux obtained using the ANN method for this problem. The ANN scalar flux is virtually indistinguishable from the standard S_2 results. The magnitude of the PRE in the ANN solution is less than 0.3% at all points in the slab. The ANN solution is extremely accurate in the interior of the slab, with a small degradation in accuracy within a couple of mean free paths of the boundary. The root-mean-squared (RMS) error in the ANN scalar flux is shown in Table 1 for both the S_2 and the S_4 calculations. This RMS error represents an average error in the scalar flux across the slab. Both of the average errors in the scalar flux are reasonably small, with the S_4 error being extremely small since a larger number of hidden layer nodes and spatial training points were used. This problem illustrates that the ANN method is able to accurately compute a concave-down cosine-shaped solution. Although the true solution to the problem is symmetric, the ANN solution is not rigorously symmetric.

Table 1 shows the computed absorption and leakage rates obtained from the reference calculation and the PRE of the computed ANN quantities. The ANN absorption rates were obtained by computing the ANN scalar flux at 1000 intervals in the slab and using the trapezoidal rule to perform the integration across the slab. The ANN method yields very accurate results for the integral absorption rate and slightly less accurate results for the leakage rates. In all cases, the ANN method yields values with a PRE less than 0.2%. It is evident from the error in the S_4 leakage rates that the ANN solution is not rigorously symmetric.

Table 1 Computed Errors for Problem I

				Leakage Rate	
Quadrature Order		Scalar Flux RMS Error	Absorption Rate	Left	Right
2	Ref.	–	0.296065	0.101968	0.101968
	ANN	0.112%	-0.095%	-0.182%	-0.182%
4	Ref.	–	0.303815	0.098093	0.098093
	ANN	0.015%	0.001%	-0.029%	0.003%

Both the discrete ordinates angular flux and the scalar flux for Problem II are concave-up. Problem II is a reasonably challenging problem, with the scalar flux decreasing by a factor of over thirty across the slab. Fig. 3 plots the S_4 scalar flux and the percent relative error (PRE) in the scalar flux obtained using the ANN method for this problem. Again, the ANN scalar flux is virtually indistinguishable from the standard S_4 results. The magnitude of the error in the ANN solution is less than 0.2% at all points in the slab, with the largest error occurring deep in the slab at the right boundary. The RMS errors in the ANN scalar flux are shown in Table 2 for the S_2 and the S_4 calculations. As for Problem I, these RMS scalar flux errors are small. This problem illustrates that the ANN method is able to accurately compute a concave-up exponential-shaped solution. This problem also demonstrates the ability of the ANN method to compute accurate solutions for problems with linearly anisotropic scattering.

Table 2 shows the computed absorption and leakage rates obtained from the reference calculation and the PRE of the computed ANN quantities. As for Problem I, the ANN method yields very accurate results for the integral absorption rate. In addition, the leakage rates at both the left and right edges of the slab are in close agreement with the reference solution.

Table 2 Computed Errors for Problem II

				Leakage Rate	
Quadrature Order		Scalar Flux RMS Error	Absorption Rate	Left	Right
2	Ref.	–	0.251049	0.299013	0.027288
	ANN	0.015%	0.010%	0.017%	-0.025%
4	Ref.	–	0.239014	0.256208	0.026045
	ANN	0.034%	-0.002%	0.020%	-0.061%

4. CONCLUSIONS

We have examined the use of artificial neural networks for the treatment of the spatial variable in the discrete ordinates approximation of the slab-geometry neutron transport equation. In this paper, we considered homogeneous slabs with linearly anisotropic scattering. The proposed ANN method approximates the discrete ordinates angular flux as a multilayer ANN with a single input representing the spatial variable x and N outputs representing the angular flux in each of the discrete ordinates angular directions. A global objective function is formulated which measures how accurately the output of the ANN satisfies the discrete ordinates equations and boundary conditions at specified spatial points. Minimization of the objective function determines the appropriate values for the parameters of the ANN. The ANN can then be used to efficiently compute the angular flux at any point in the spatial domain. Hence, any angular moments such as the scalar flux and current can also be readily computed.

Our numerical results demonstrate that the ANN method can provide accurate results for the scalar flux using a small number of spatial points for problems whose solutions are either concave-up or concave-down. The solutions obtained by the method are also reasonably accurate near the boundaries of the problem. In addition, the computed integral absorption rate and leakage rates at the boundaries of the slab show good agreement with values obtained from the reference solution. Although not examined in detail in this paper, our numerical experiments also demonstrate that, much like standard discretization methods (Lewis, 1993), the number of spatial training points used in the ANN method must be increased as the quadrature order is increased to maintain the accuracy of the solutions. In addition, for a given number of spatial training points, increasing the number of hidden layer nodes increases the accuracy of the solutions with a commensurate increase in computational effort.

The ANN method presented in this paper possesses some negative characteristics as well. From a theoretical viewpoint, the method does not rigorously preserve particle balance or any symmetry present in the solution. However, the numerical experiments examined to date show that the particle conservation errors are small (typically less than 0.1%) for globally converged solutions. The method is also not guaranteed to produce

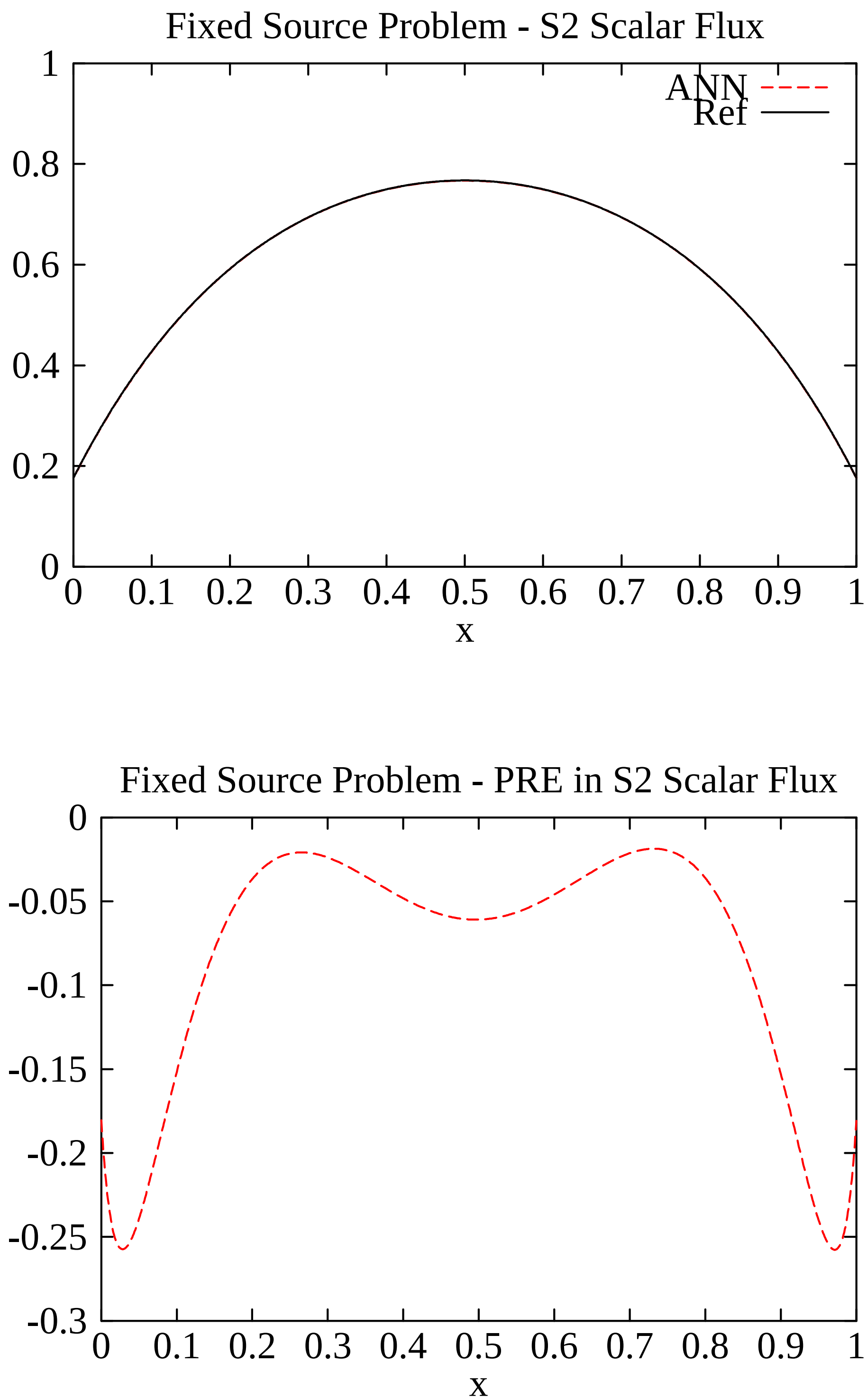

Fig. 2 ANN S_2 scalar flux and percent relative error for Problem I.

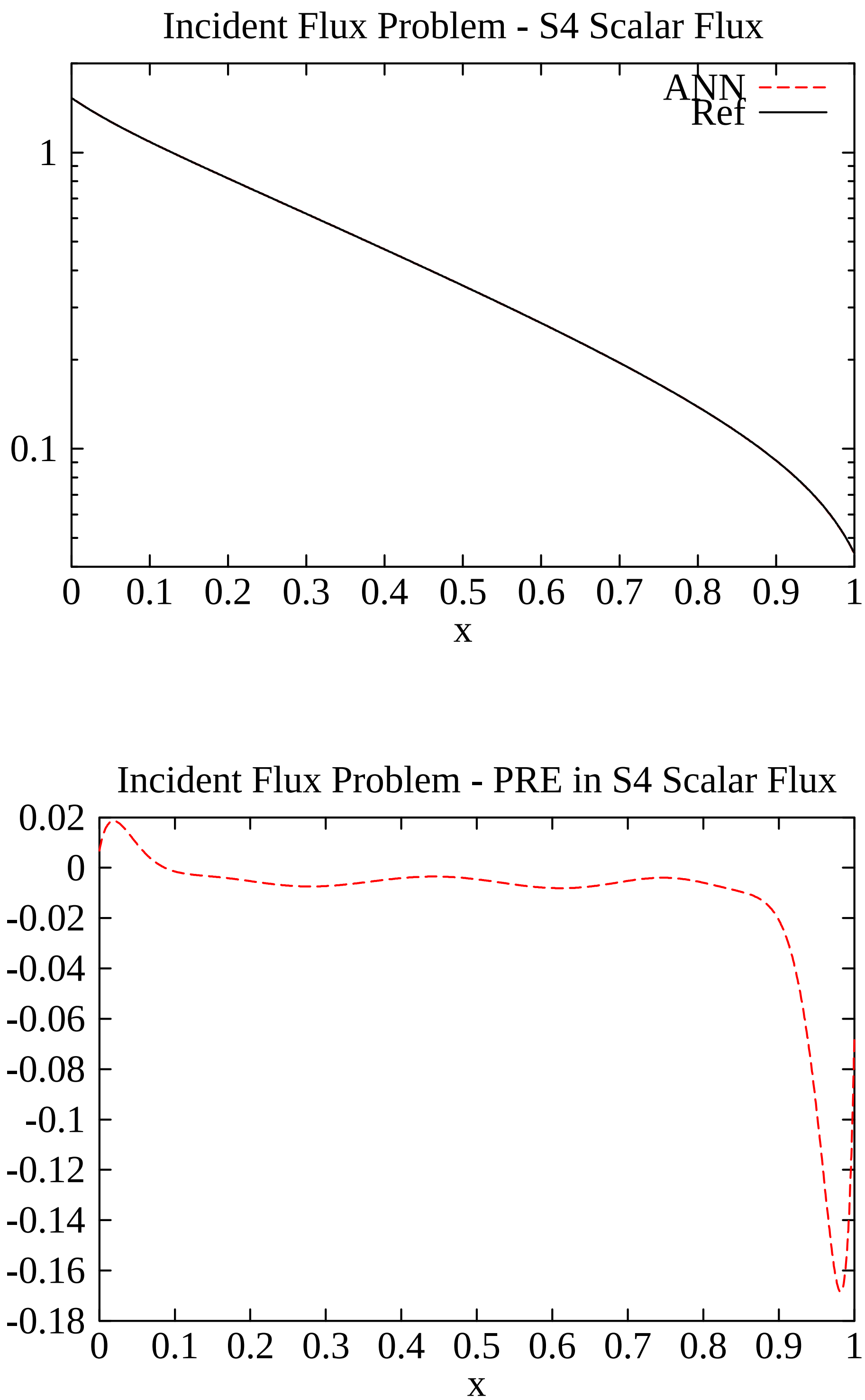

Fig. 3 ANN S_4 scalar flux and percent relative error for Problem II.

strictly positive angular fluxes. However, the production of negative angular fluxes has not been problematic in the numerical problems examined to date with the exception of strongly absorbing deep-penetration problems in which the scalar flux varies by many orders of magnitude. We are currently investigating the use of an exponential activation function for the output layer nodes to eliminate the production of negative angular fluxes for these types of problems. This approach has the potential added benefit of allowing the ANN to accurately compute a much broader range of angular flux values. From an operational viewpoint, the proposed ANN method converts the solution of a linear transport problem into a nonlinear minimization problem. Nonlinear minimization can be numerically intensive and comes with pitfalls of its own, such as trapping in local minima.

We have not discussed the numerical efficiency of the ANN method in this paper. The number of ANN parameters, and hence the computational expense of the ANN method, grows as the quadrature order and the number of hidden layer nodes is increased. In addition, the evaluation of the global objective function and its gradient during the minimization becomes computationally more expensive as the number of spatial training points is increased. Our current implementation combines an object-oriented implementation of the ANN coupled with a procedural implementation of the minimization algorithm. This coupling possesses inherent inefficiencies, but we believe that it can be improved. Future efforts will be directed at improving the efficiency of the ANN method through both implementation and algorithmic improvements.

We considered the case of linearly anisotropic scattering in this paper, and the extension of the ANN method to general anisotropic scattering is straightforward. The extension to inhomogeneous cross sections and interior sources is similarly straightforward, although the accuracy of the method for inhomogeneous slabs remains to be seen. Furthermore, there are no conceptual difficulties in extending the method to solve the multidimensional discrete ordinates equations. However, the accuracy and efficiency of the ANN method for these multidimensional problems remains an open question.

We have not addressed in this paper the application of the ANN method to optically-thick, highly-scattering problems. An appealing aspect of the application of this method to these thick diffusive problems is that source iteration, which can converge arbitrarily slowly for these problems, is avoided. We plan to investigate this topic in future work.

NOMENCLATURE

μ_n, w_n	Gauss-Legendre quadrature directions and weights
x	spatial variable
$\psi_n(x)$	discrete ordinates angular flux at position x in direction μ_n
$\phi_0(x), \phi_1(x)$	scalar flux, current
σ_t	macroscopic total cross section

σ_{s0}, σ_{s1}	isotropic, linearly anisotropic components of macroscopic differential scattering cross section
Q	interior source
ψ_n^{inc}	incident angular flux
$Y_n(x)$	ANN approximation to discrete ordinates angular flux at position x in direction μ_n
$\phi_0^{ann}(x)$, $\phi_1^{ann}(x)$	ANN approximations to the scalar flux, current

Subscripts

n	discrete ordinates angular direction

ACKNOWLEDGMENTS

This work was performed under the auspices of the U.S. Department of Energy by the University of California Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48.

REFERENCES

Brantley, P.S., 2000. Artificial Neural Network Solutions of Slab-Geometry Neutron Diffusion Problems. *Trans. Am. Nuc. Soc.* **83**, 251-252.

Fausett, L., 1994. *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications*, Prentice-Hall, Upper Saddle River, New Jersey, Chapters 1 and 6.

Lewis, E.E, Miller Jr., W.F., 1993. *Computational Methods of Neutron Transport*, American Nuclear Society, Lagrange Park, Illinois, Chapter 3.

Shanno, D.F., Phua, K.H., 1980. Remark on Algorithm 500. *ACM Trans. Math. Software,* **6**, 618-622.

Van Milligen, B.Ph., Tribaldos. V., Jiménez, J.A., 1995. Neural Network Differential Equation and Plasma Equilibrium Solver. *Phys. Rev. Lett.* **75**, 3594-3597.