PREPARED FOR THE U.S. DEPARTMENT OF ENERGY,
UNDER CONTRACT DE-AC02-76CH03073

PPPL-3808
UC-70

PPPL-3808

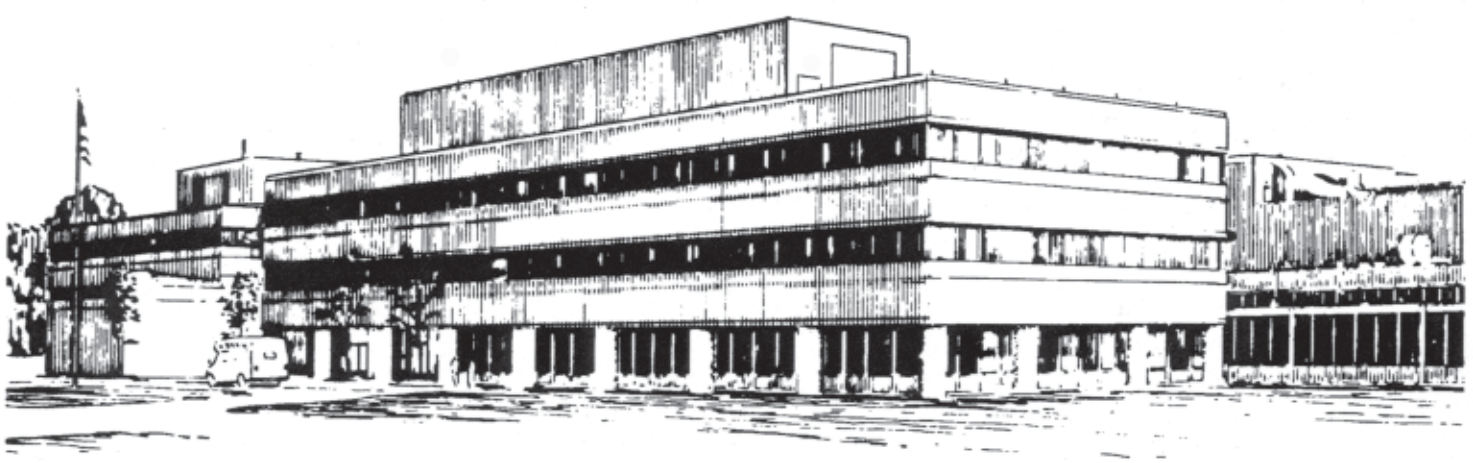
Neural Network Algorithm for Particle Loading

by

J.L.V. Lewandowski

April 2003

PPPL
PRINCETON PLASMA PHYSICS LABORATORY



PRINCETON PLASMA PHYSICS LABORATORY
PRINCETON UNIVERSITY, PRINCETON, NEW JERSEY

PPPL Reports Disclaimer

Availability

This report is posted on the U.S. Department of Energy's Princeton Plasma Physics Laboratory Publications and Reports web site in Fiscal Year 2003. The home page for PPPL Reports and Publications is: http://www.pppl.gov/pub_report/

DOE and DOE Contractors can obtain copies of this report from:

U.S. Department of Energy
Office of Scientific and Technical Information
DOE Technical Information Services (DTIS)
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401

Fax: (865) 576-5728

Email: reports@adonis.osti.gov

This report is available to the general public from:

National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road
Springfield, VA 22161

Telephone: 1-800-553-6847 or
(703) 605-6000

Fax: (703) 321-8547

Internet: http://www.ntis.gov/ordering.htm

Neural Network Algorithm for Particle Loading

J. L. V. Lewandowski
Princeton University
Plasma Physics Laboratory
Princeton, P.O. Box 451, NJ 08543
USA

Abstract

An artificial neural network algorithm for continuous minimization is developed and applied to the case of numerical particle loading. It is shown that higher-order moments of the probability distribution function can be efficiently renormalized using this technique. A general neural network for the renormalization of an arbitrary number of moments is given.

Pacs: 52.35.Py, 52.30.Jb, 52.55.Fa

1 Introduction

The collective behavior of a hot collisionless plasma is described in kinetic theory by the evolution of the distribution functions of its species $f_e(\mathbf{x}, \mathbf{z}, t)$ and $f_i(\mathbf{x}, \mathbf{z}, t)$ in continuous phase space. In the particle-in-cell (PIC) technique one returns to a discrete representation of the plasma in terms of a small sample of 10^4 to 10^8 "particles", or more appropriately, markers [1].

The most common (and natural) initialization of the simulation particles is to load the particle velocities, for example, with random numbers having the desired distribution. This method has obvious advantages; in particular it is in agreement with our physical intuition (that is the numerically-loaded probability distribution function (PDF) for the markers resembles the random distribution of actual particles). However, one disadvantage of random loading is that higher-order velocity moments are not well represented, and this may impact the long term behavior and the noise level of the simulated plasma.

In this paper, we present a neural network algorithm to renormalize a numerically-loaded PDF; we specifically consider the case of a Maxwellian PDF (because of its applications to plasma particle-in-cell simulations) but the method can be easily generalized to arbitrary distribution functions. A numerically-loaded PDF with accurate higher-order moments allows for noise-free simulations to be run for longer times; this is especially important for the modelling of the fast-moving electrons.

The paper is organized as followss; in section 2, we discuss the main features of a continuous neural minimization algorithm. This algorithm is tested numerically for the case of a Maxwellian distribution in section 3; the issues related to the parallelization of the algorithm are discussed in the same section. Concluding remarks are given in section 4.

2 Neural Minimization Algorithm

In this section, we present the general features of a neural network algorithm for continuous minimization. The specific case of a neural algorithm for particle loading is specifically discussed, and numerical results are given.

2.1 General Aspects of Neural Networks

Over the past decade highly parallel neural networks have been investigated extensively to solve complicated problems such as robotic control [2], vision and image processing [3] and speech recognition [4]. Artificial neural networks (ANNs) is an abstract simulation of a real nervous system that contains a collection of neuron units communicating with each other via axon connections (such a model bears strong resemblance to axons and dendrites in a nervous system). ANNs consist of individual processors and interconnections between the processors, or neurons (Figure 1). Each neuron can have two different states, i.e. on and off which are represented by binary numbers 1 and 0; each neuron sums all the signals coming from all the other neurons through the weighted interconnections, thresholds the summed signal to 0 or 1, and changes its state according to the thresholded output. This method of computation is based on the collective interaction between the neurons and it exhibits a high degree of parallelism.

In this paper, we focus out attention on the Hopfield model, which is based on the nonlinear dynamics between globally interconnected neurons [5, 6]. The states of individual neurons are specified by their outputs θ_i which range between 0 and 1. In the (continuous) Hopfield model,

neurons change their states according to the following equations

$$\frac{d\zeta_i}{dt} = \sum_j T_{ij}\theta_j \ , \tag{2.1}$$

$$\theta_i = g\left(\zeta_i\right) \ , \tag{2.2}$$

where t is the continuous time which corresponds to the updating parameter, T_{ij} is the interconnection strength, and $g(x)$ is a monotonically increasing nonlinear function bounded between 0 and 1, and ζ_i is an intermediate variable. The nonlinear function $g(x)$ is typically of the form $g(x) = \frac{1}{2}\ [1 + \tanh(x/x_0)]$, where x_0 is a constant. Hopfield [6] has shown that if the interconnection strengths are symmetric, $T_{ij} = T_{ji}$, neurons in the continuous model always their states in such a way as to minimize an energy functional defined by

$$E = -\frac{1}{2}\ \sum_i \sum_j T_{ij}\theta_i\theta_j \ , \tag{2.3}$$

and stop at the minima of this function. Clearly, the updating rules (2.1,2.2) for minimizing the energy function (2.3) are highly parallel The Hopfield model has been generalized to arbitrary energy function by Koch *et al* [7] as follows; write the positive-definite energy function as

$$E = F\left(\theta_1, \theta_2, \cdots, \theta_M\right) \equiv F\left(\boldsymbol{\theta}\right) \ , \tag{2.4}$$

where F is a non-singular, bounded function of the variable θ_i. The updating rules given by

$$\frac{d\zeta_i}{dt} = -\frac{\partial F}{\partial \theta_i} \ , \tag{2.5}$$

$$\theta_i = g\left(\zeta_i\right) \ , \tag{2.6}$$

minimize the energy function (2.4) [7].

2.2 Neural Network Algorithm for Particle Loading

Having introduced the principal features of the neural network algorithm for continuous minimization, we can proceed with the renormalization problem discussed in the Introduction. Let us assume that a Maxwellian PDF, based on a set of N samples (markers), has been numerically generated using some arbitrary random technique; the method used can be straightforward methods (for example, Neumann's rejection technique) or more sophisticated methods, such as those based on number theory [8,9]. Some methods based on random number generators tend to introduce a 'background noise' that could be detrimental for the observation of low-amplitude instabilities. Apart from the noise properties associated with the initial PDF, the accuracy of higher-order velocity moments can be also important in some applications. This paper attempts to address some of the issues associated with higher-order velocity moments (irrespective of the loading method of the initial PDF). The problematic velocity moments are the odd moments because of an obvious cancellation problem between markers with positive and negative velocities.

The odd velocity moments of the continuous (exact) Maxwellian distribution function $F_M(v) = (2\pi)^{-1/2} \exp(-v^2/2)$ vanish

$$\mathcal{M}_{2n+1} \equiv \int_{-\infty}^{+\infty} v^{2n+1} F_M(v) dv = 0 \ , \tag{2.7}$$

for $n = 0, 1, 2, ...$ Now consider the moments of the numerically-loaded PDF. It is sufficient to carry out the velocity space integration using Riemman sums; the velocity moment of order n, based on a set of N markers, is then

$$M_n \equiv \frac{1}{N} \sum_{k=1}^{N} {V_k}^n \ . \tag{2.8}$$

In general, however, the odd velocity moments M_{2n+1} do not vanish; this break of symmetry in the PDF can be problematic for long timescale simulations. Let us consider a neural network algorithm to renormalize the first-order ("particle flux") and the third-order ("heat flux") velocity moments; the energy functional is written as

$$E = C_1 M_1^2 + M_3^2 \equiv F(V_1, V_2, \cdots, V_N) \equiv F(\mathbf{V}) > 0 \ , \tag{2.9}$$

where C_1 is a positive constant. Without thresholding the evolution equation for the k^{th} marker is

$$\begin{aligned} \frac{dV_k}{dt} &= -\frac{\partial F}{\partial V_k} \\ &= -\frac{2}{N} C_1 M_1 \sum_{\ell=1}^{N} \frac{\partial V_\ell}{\partial V_k} - \frac{6}{N} M_3 \sum_{\ell=1}^{N} V_\ell^2 \frac{\partial V_\ell}{\partial V_k} \\ &= -\frac{2}{N} \left(C_1 M_1(t) + 3 M_3(t) V_k^2 \right) \ , \end{aligned} \tag{2.10}$$

where we have used the definition (2.8) and the relation $\partial V_\ell / \partial V_k = \delta_{k,l}$, where $\delta_{k,l}$ is the Kronecker symbol for integers k and ℓ. Before the renormalization ($t = 0$), we usually have $|M_3(0)| \gg |M_1(0)|$, since the higher-order velocity moments are more "noisy". If C_1 is chosen to be unity in Eq(2.10) then the neural network will be biased towards the highest velocity moment (here M_3). In order to remedy to this situation, we have chosen $C_1 = \left(M_3(0)/M_1(0)\right)^2 \gg 1$; numerical simulations confirm the validity of this choice. The generalization of the above algorithm for the renormalization of the first Q velocity moments is straightforward; define the energy functional as

$$E = \sum_{q=1}^{Q} C_q \left(M_q - I_q\right)^2 = F(\mathbf{V}) \ , \tag{2.11}$$

where the constants C_q are conveniently chosen as $C_q = \left(M_Q(0)/M_q(0)\right)^2$, for $q = 1, 2, \cdots, Q$; and

$$I_q \equiv \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} x^q \exp\left(-x^2\right) dx$$

for positive integer q is related to the exact moment of the Maxwellian PDF (note that $I_q = 0$ for q odd; the first nonvanishing moments are $I_2 = \frac{1}{2}$, $I_4 = \frac{3}{4}$, $I_6 = \frac{15}{8}$, $I_8 = \frac{105}{16}$, $I_{10} = \frac{945}{32}$, $\cdots$). Using the chain rule as before, the evolution equation for the k^{th} marker then becomes

$$\frac{dV_k}{dt} = -\frac{2}{N} \sum_{q=1}^{Q} q C_q \left(M_q - I_q\right) V_k^{q-1} \ . \tag{2.12}$$

2.3 Numerical Results

For the simulations presented in this paper, we have a second-order Runge-Kutta algorithm [10] to integrate the N equations (2.10). The algorithm has been parallelized using the Message Passing Interface (MPI); the parallel version of the algorithm is based on distributing the N equations (2.10) across different processors; the computation of the velocity moments [M_q in Eq.(2.12] is the operation that requires synchronization and global communication across processors.

In this paper, we compare the renormalization of a Maxwellian PDF for the case of random loading and the so-called "quiet start loading" of Denavit and Walsh [9] which is based on number theory [8]. The original PDF (i.e. before the renormalization procedure) are shown in Figure 1 (random loading) and Figure 2 (quiet start loading based on Fibonacci numbers [9]) for the case on $N = 10946$ markers. Figure 2 shows the regular array in $X - V$ space; note that the velocities are staggered to prevent a pattern of velocity beams with their associated unphysical recurrences and instabilities [11].

Figure 3 shows the time evolution of the absolute values of the odd velocity moments M_1 (dotted line) and M_3 (plain line) computed using the neural network algorithm with a time step $\Delta t = 10$. The time step can be chosen by considering the order of magnitude of the right-hand side of Eq.(2.10)

$$\frac{1}{\tau} \equiv \frac{2}{N} \left\{ C_1 \left| M_1(0) \right| + 3 M_2(0) \left| M_3(0) \right| \right\}$$

The above estimate can be extended to the general neural network (2.11,2.12). For reasonable accuracy (and possibly stability), we must chose a time step such that $\Delta t / \tau \ll 1$. Figure 3 shows that the renormalized particle flux, M_1, has decreased by 4 orders of magnitude, whereas the heat flux, M_3, has decreased by 3 orders of magnitude.

In order to compare the random and quiet start loading procedures, one can consider the probability spectrum, $P(\Delta V)$, associated with the absolute value of the velocity change, $\Delta V \equiv |V(0) - V(T)|$, where T is the total simulation time. Figures 5 and 6 show $P(\Delta V)$ for the case of random loading and quiet start loading, respectively. The probability spectrum for the case of random loading displays a maxima at relatively large ΔV; on the other hand, the probability spectrum for the quiet start loading (Figure 6) follows a power law $P(\Delta v) \propto (\Delta v)^{-\alpha}$. This suggests that the renormalization of the quiet start PDF using the neural network algorithm will be more effective as compared to the random loading case; this is confirmed by the simulations.

Figure 7 shows the renormalization of the odd velocity moments M_1, M_3 and and M_5. All velocity moments decrease by 3 to 4 orders of magnitude. Similar performance is observed in the general case (2.12).

3 Concluding Remarks

We have presented a neural network algorithm for particle loading based on continuous minimization. It has been shown that such renormalization procedure is efficient and be can easily generalized. The success of the method presented in this paper can be enhanced further when coupled with loading techniques based on number theory [8,9]. Numerically-loaded distribution functions with accurate higher-order moments are useful for long-time, noise-free simulations of turbulent plasmas.

Acknowledgments This research was supported by Contract No. DE-AC02-CH0-3073 and the Scientific Discovery through Advanced Computing (SciDAC) initiative (U.S. Department of Energy).

References

[1] G. Hu and J.A. Krommes, *Phys. Plasmas*, **1**, 863 (1994).

[2] A.M.S. Zalala and A.S. Morris, *Neural Networks for Robotic Control: Theory and Applications*, (Elis Horwood Press, London, 1996).

[3] G.A. Carpenter and S. Grossberg, *Neural Networks for Vision and Image Processing*, (MIT Press, Cambridge, 1992).

[4] Y. Bengio, *Neural Networks for Speech and Sequence Recognition*, (International Thomson Computer Press, London, 1996).

[5] J.J. Hopfield and D.W. Tank, *Biol. Cybernet.*, **52**, 141 (1985).

[6] J.J. Hopfield, *Proc. Nat. Acad. Sci.*, **79**, 2554 (1982).

[7] C. Koch, J. Marroquin and A. Yulle, *Proc. Nat. Acad. Sci.*, **83**, 4263 (1986).

[8] J.H. Halton, *Numer. Math.*, **2**, 84 (1960).

[9] J. Denavit and J.M. Walsh, *Comments Plasma Phys. Cont. Fusion*, **6**, 209 (1981).

[10] W.S. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery, *Numerical Recipes in Fortran* (Cambridge University Press, New York, 1992).

[11] J.M. Dawson, *Phys. Rev.*, **118**, 1320 (1961).

Figure 1 Schematic diagram of a simple neural network; T_{ij}, T_{jk} and T_{ik} are the interconnections strengths between the pairs of neurons (i,j), (j,k) and (i,k), respectively.

Figure 2 Random loading of a Maxwellian distribution function for a set of $N = 10946$ markers.

Figure 3 Quiet start loading (based on Fibonacci numbers [9]) of a Maxwellian distribution function for a set of $N = 10946$ markers.

Figure 4 Renormalization of the first two odd moments [M_1 (dotted line) and M_3 (plain line)] for a quiet start loading of a Maxwellian distribution based on a set of $N = 196418$ markers. The time step used is $\Delta t = 10$.

Figure 5 Probabibility spectrum for the velocity change $\Delta V = |V(T) - V(0)|$ for the case of a random loading based on a set of $N = 196418$ markers.

Figure 6 Probabibility spectrum for the velocity change $\Delta V = |V(T) - V(0)|$ for the case of the quiet start loading for the same parameters as in Figure 5. The probability spectrum follows a power law distribution, $P(\Delta V) \propto (\Delta v)^{-\alpha}$.

Figure 7 Renormalization of the first three odd moments [M_1 (dotted line), M_3 (plain line) and M_5 (dashed line)] for the same parameters as in Figure 4.

FIG.1 Lewandowski

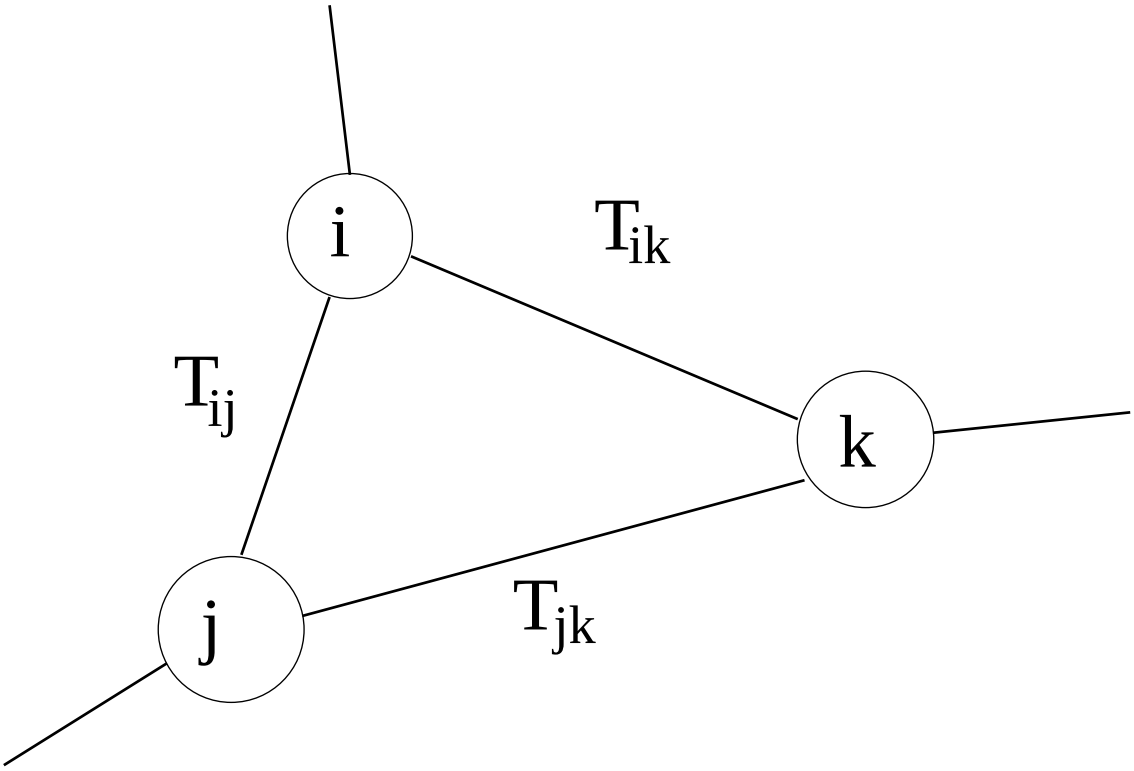


FIG.2 Lewandowski

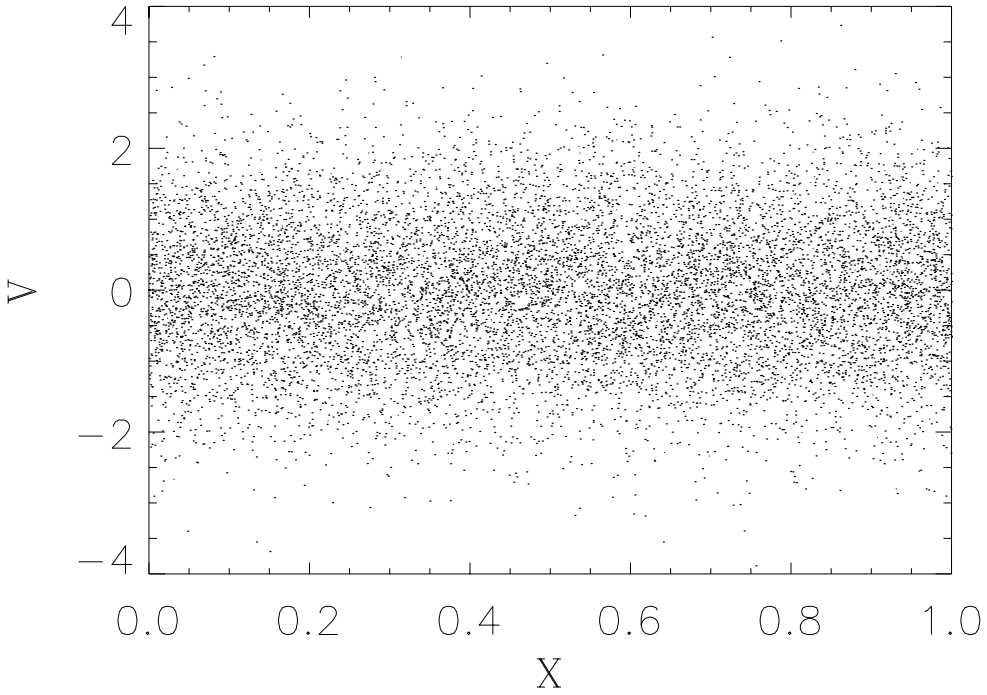


FIG.3 Lewandowski

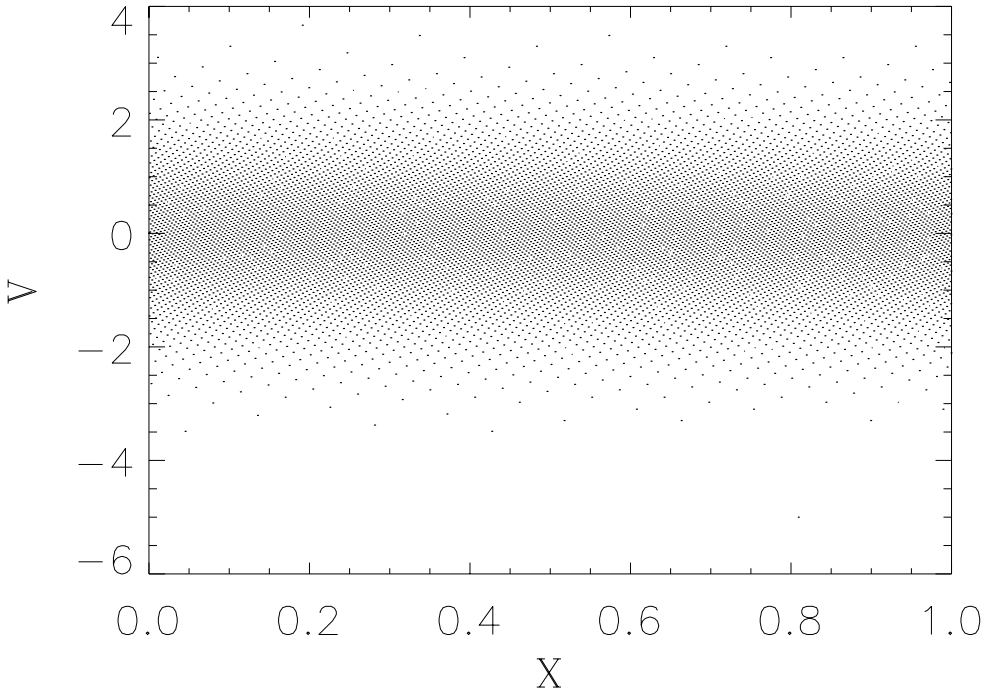


FIG.4 Lewandowski

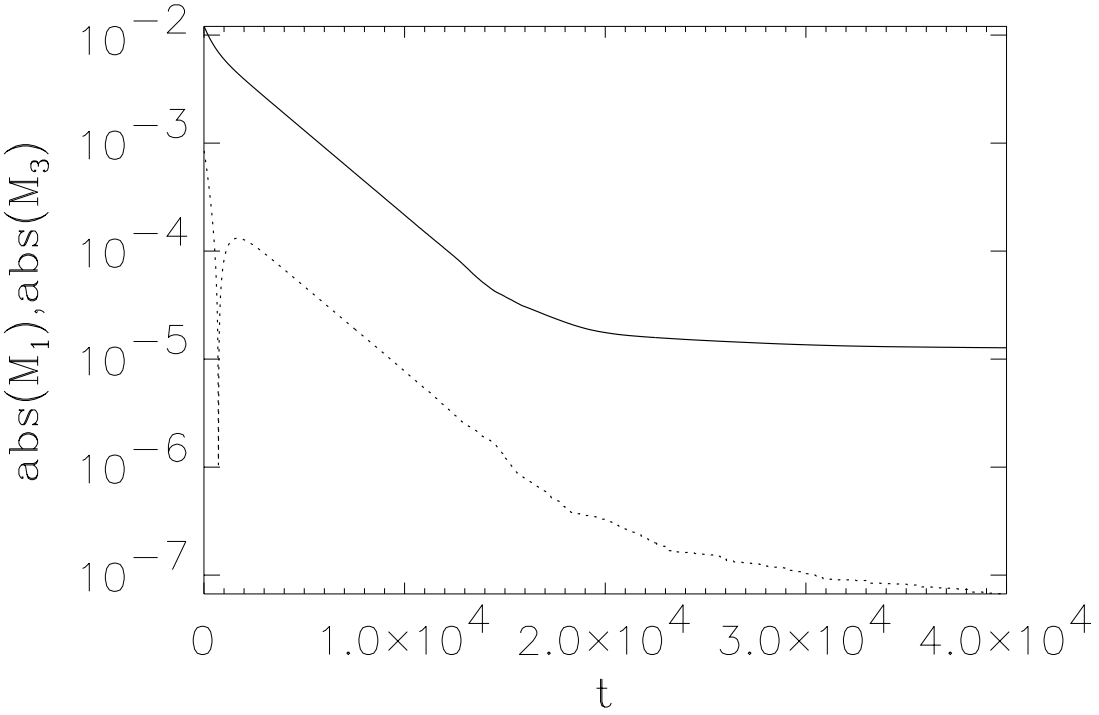


FIG.5 Lewandowski

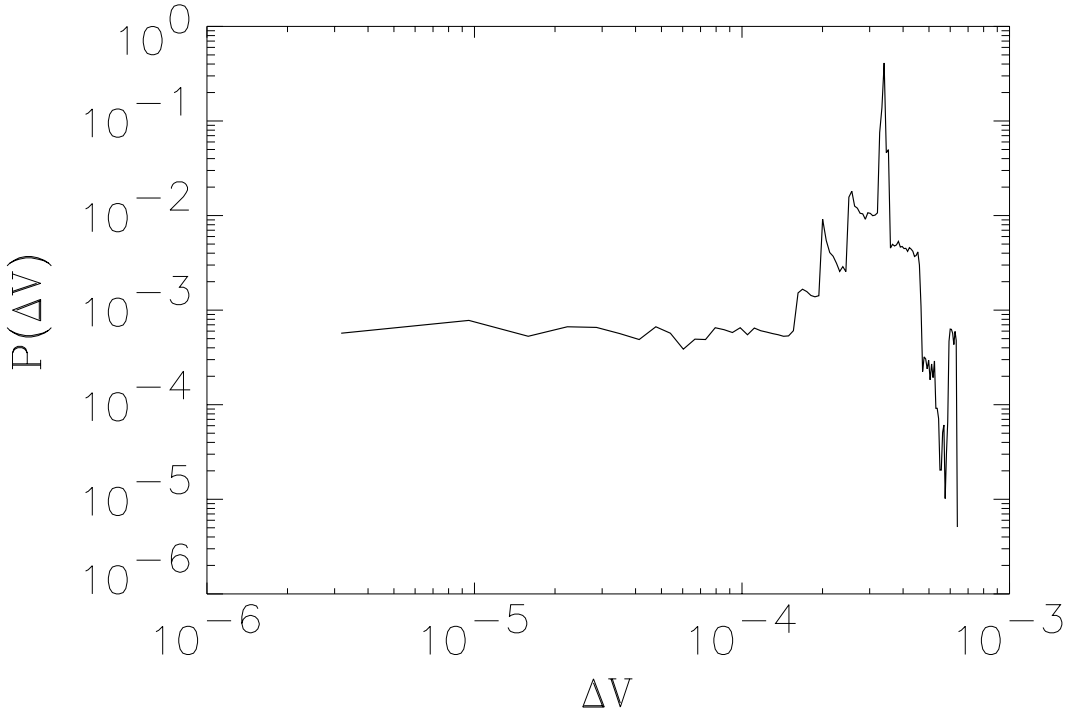


FIG.6 Lewandowski

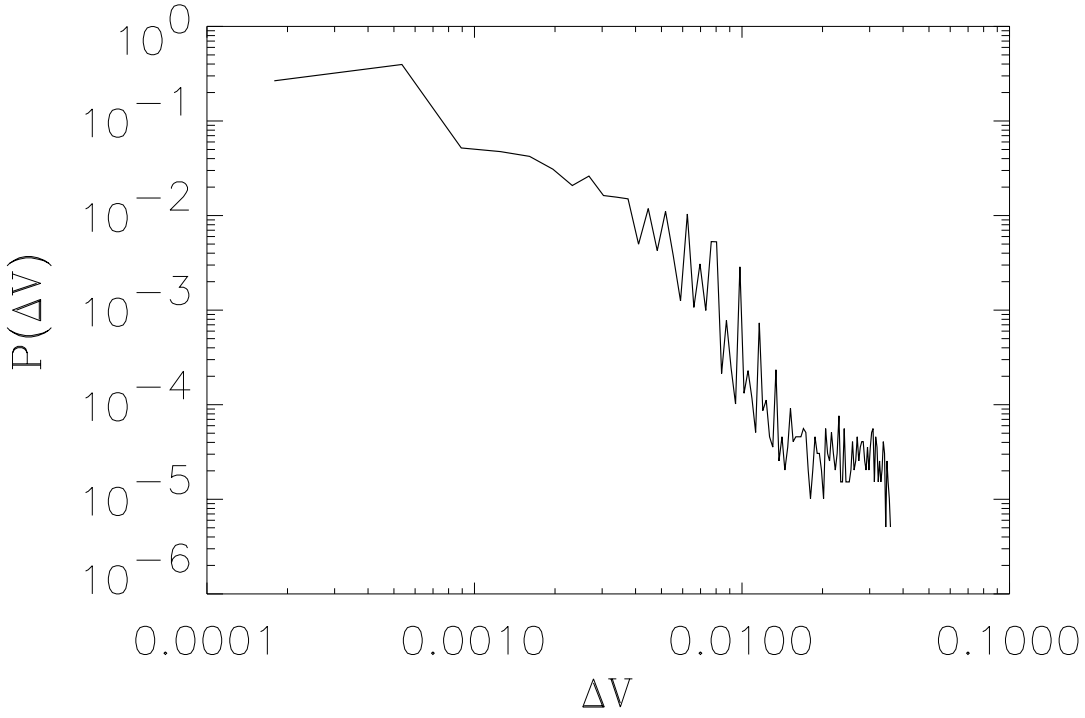
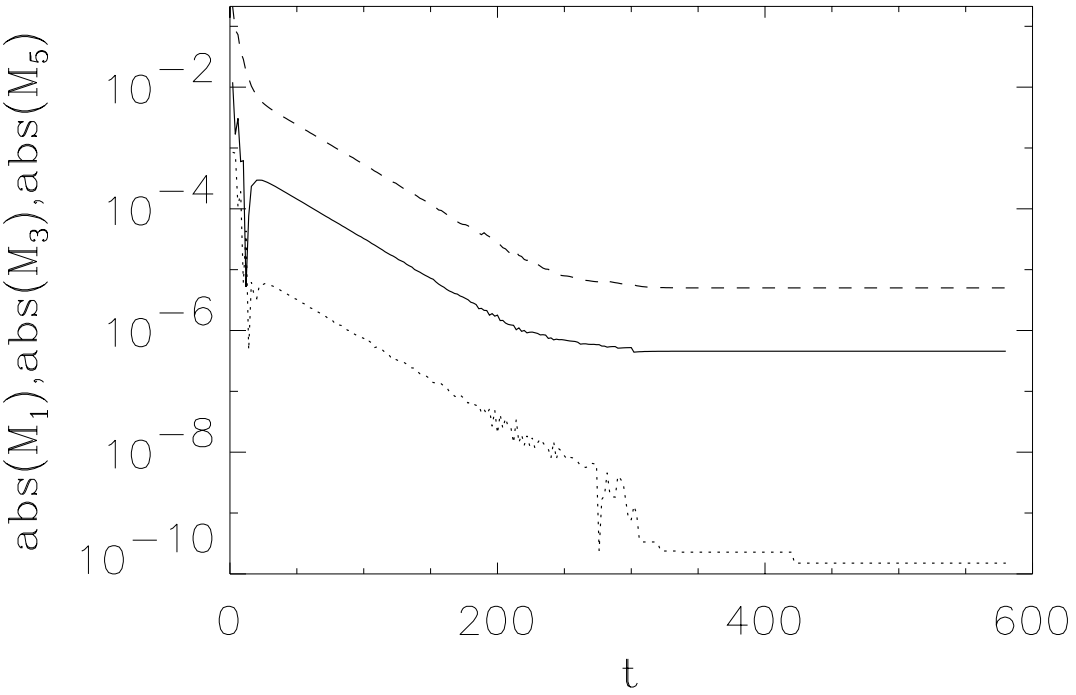


FIG.7 Lewandowski



External Distribution

Plasma Research Laboratory, Australian National University, Australia
Professor I.R. Jones, Flinders University, Australia
Professor João Canalle, Instituto de Fisica DEQ/IF - UERJ, Brazil
Mr. Gerson O. Ludwig, Instituto Nacional de Pesquisas, Brazil
Dr. P.H. Sakanaka, Instituto Fisica, Brazil
The Librarian, Culham Laboratory, England
Mrs. S.A. Hutchinson, JET Library, England
Professor M.N. Bussac, Ecole Polytechnique, France
Librarian, Max-Planck-Institut für Plasmaphysik, Germany
Jolan Moldvai, Reports Library, MTA KFKI-ATKI, Hungary
Dr. P. Kaw, Institute for Plasma Research, India
Ms. P.J. Pathak, Librarian, Insitute for Plasma Research, India
Ms. Clelia De Palo, Associazione EURATOM-ENEA, Italy
Dr. G. Grosso, Instituto di Fisica del Plasma, Italy
Librarian, Naka Fusion Research Establishment, JAERI, Japan
Library, Plasma Physics Laboratory, Kyoto University, Japan
Research Information Center, National Institute for Fusion Science, Japan
Dr. O. Mitarai, Kyushu Tokai University, Japan
Dr. Jiangang Li, Institute of Plasma Physics, Chinese Academy of Sciences, People's Republic of China
Professor Yuping Huo, School of Physical Science and Technology, People's Republic of China
Library, Academia Sinica, Institute of Plasma Physics, People's Republic of China
Librarian, Institute of Physics, Chinese Academy of Sciences, People's Republic of China
Dr. S. Mirnov, TRINITI, Troitsk, Russian Federation, Russia
Dr. V.S. Strelkov, Kurchatov Institute, Russian Federation, Russia
Professor Peter Lukac, Katedra Fyziky Plazmy MFF UK, Mlynska dolina F-2, Komenskeho Univerzita, SK-842 15 Bratislava, Slovakia
Dr. G.S. Lee, Korea Basic Science Institute, South Korea
Institute for Plasma Research, University of Maryland, USA
Librarian, Fusion Energy Division, Oak Ridge National Laboratory, USA
Librarian, Institute of Fusion Studies, University of Texas, USA
Librarian, Magnetic Fusion Program, Lawrence Livermore National Laboratory, USA
Library, General Atomics, USA
Plasma Physics Group, Fusion Energy Research Program, University of California at San Diego, USA
Plasma Physics Library, Columbia University, USA
Alkesh Punjabi, Center for Fusion Research and Training, Hampton University, USA
Dr. W.M. Stacey, Fusion Research Center, Georgia Institute of Technology, USA
Dr. John Willis, U.S. Department of Energy, Office of Fusion Energy Sciences, USA
Mr. Paul H. Wright, Indianapolis, Indiana, USA

The Princeton Plasma Physics Laboratory is operated
by Princeton University under contract
with the U.S. Department of Energy.

Information Services
Princeton Plasma Physics Laboratory
P.O. Box 451
Princeton, NJ 08543

Phone: 609-243-2750
Fax: 609-243-2751
e-mail: pppl_info@pppl.gov
Internet Address: http://www.pppl.gov