Solving N-Body Problems with Neural Networks

Marcelino Quito, Jr., Christopher Monterola, and Caesar Saloma*
National Institute of Physics, University of the Philippines, Diliman, Quezon City 1101, The Philippines
(Received 9 December 2000)

We show a new approach for solving the N-body problems based on neural networks. Without loss of generality, we derived a network solution for the time-dependent positions of N bodies in self-gravitating systems. The simulation is limited to a system of collisionless disks — a case for determining the spatial distributions of dark matter and in reproducing global effects such as formation of spiral galaxies. Our approach yields a solution that is analytic with time-reversed path-tracing capabilities that could lead to new findings in the study of the collective behavior of N-body systems.

DOI: 10.1103/PhysRevLett.86.4741 PACS numbers: 07.05.Mh, 45.50.Pk, 84.35.+i, 95.10.Ce

The N-body problem occurs in almost all branches of physics from studies on submicroscopic systems to macroscopic ones (Bose-Einstein condensation, molecular oscillations, protein-folding, granular dynamics, swarming, multilane traffic flow, galaxy formation, etc). The problem is characterized by a set of coupled differential (or difference) equations (DE's) and progress in our understanding of complex system dynamics has depended in our ability to solve accurately the set which rapidly becomes complicated with N [1–3]. Investigations of N-body systems have resulted in the discovery of new interesting collective phenomena such as equipartition, mass segregation, jamming, and phase synchronization [4,5].

N-body problems are difficult to solve. In fact, the exact solution for the relativistic 2-body self-gravitating systems has just been found [6]. The N-body system ($N \geq 3$) behaves chaotically with no known analytic solution even in the nonrelativistic regime [3]. As an ordinary differential equation (DE) the system is considered an "initial value problem" that satisfies the Lipschitz condition [7] of only one possible outcome for a particular initial state.

Large N-body systems are investigated numerically using standard algorithms whose efficiency depends on [8–11]: (i) accuracy improvement with step-size reduction, (ii) adaptability of time step in response to prevailing boundary conditions, and (iii) growth of computational complexity C with N. Iterative methods have two serious limitations which affect the accuracy and information content of their solutions: (1) rapid propagation of rounding-off errors with C which increases with N and time-step resolution [11], and (2) availability of solution values only at discrete instants of time defined by the step size.

Here, we propose a new strategy for solving N-body problems without the limitations of the standard methods. It is based on artificial neural networks (NN) which has been previously used to solve noncoupled DE's [12,13]. The NN solution is analytic with time-reversed path-tracing capabilities — a solution value is obtained at any instant of time. While the NN solution is still affected by floating-point errors, its analytic form allows for a systematic truncation error throughout the temporal evolution of motion instead of cumulative as in the standard methods. Hence, the errors propagate predictably and their corrections are easy to implement. The complexity C of the solution search is also independent of time step.

The importance of error propagation analysis traces back to Lecar *et al.* [14] who showed that 11 different standard methods yield solutions which differed by as much as 100% for $N = 25$. This basic failure is argued to be solely a product of encounters [15], that induces the solution to diverge exponentially with increasing time steps. Surprisingly, as far as we know, no single experiment has been utilized to test the divergence of collisionless-disks solution and perhaps settle the issue. Here we demonstrate using Hamiltonian conservation and time-reversal tracing that the NN solution does not diverge exponentially with time for an N-body system. While the NN architecture is tested only for a self-gravitating system of collisionless disks, our proposed method is general and applies to other N-body systems as well.

We consider N particles of mass m_i with positions $\{r_i\}$ which are each moving under an attractive force from all the other bodies where particle index $i = 1, \ldots, N$. The system is described by

$$F_i(r_i, t) = \frac{\partial^2 r_i}{\partial t^2} + \sum_{k,k\neq i}^{N} \frac{G m_k \hat{r}}{|r_i - r_k|^2} = 0 . \qquad (1)$$

A feedforward three-layer perceptron architecture with N outputs is used to determine the N particle positions $r_i(t)$ at time t. Its $(2N + 1)$ inputs $\{z_i\}$ represent the specific t value, and the initial particle positions $\{r_{io}\}$ and velocities $\{v_{io}\}$. The NN maps to $r_i(t)$ according to $r_i(t) = \sum_{j=1}^{q}(w_{ji} Y_j)$ and $Y_j = \sum_{i=1}^{2N+1} \tanh(w_{ij} z_i)$ where w_{mn} represents the interconnection weight of the mth and the nth layer and q is the number of hidden nodes [11,12]. The activation functions for the hidden and output nodes are $\tanh(x)$ and linear, respectively.

The unsupervised NN is trained via a modified gradient descent rule where we minimize a positive-definite cost function $E(r_i)$ that is derived from the dynamics and boundary conditions of the system [11,12]:

$$E(r_i) = \sum_i |F_i(r_i^{(p)}, t)|^2 + \sum_i |r_i^{(p)}(t=0) - r_{i0}|^2 + \sum_i \left| \frac{\partial r_i^{(p)}(t=0)}{\partial t} - v_{i0} \right|^2 . \quad (2)$$

$E(r_i)$ is chosen such that as $E(r_i) \to 0$ the exact solution $\{r_i(t)\}$ is approached by the NN output. $E(r_i)$ represents the degree of instability of the NN environment and the unsupervised NN training is aimed at reaching its minimum state from an initially random state. The weights are updated at a fixed learning rate $\sigma = 0.001$, from w_{mn}^p to $w_{mn}^{p+1} = w_{mn}^p - \sigma \frac{\partial E}{\partial w_{mn}}$, where index m is the mth node in the previous layer, while n is the nth node in the next layer [16]. Because the computational load increases with N and could exceed the processing capacity of the computer (DEC Alpha, 667 MHz, GCC 2.7.2 compiler, double precision), we introduce a normalization factor based on N before the signals are evaluated in the NN nodes.

In Fig. 1 are two-dimensional (2D) time traces of N point-particle systems [$N = 2(1a)$, $3(1b)$, $4(1c)$, and $10(1d)$] from its initial state [$m_i = 1.0$, gravitational constant $G = 0.02$, ($|ri| \geq 3500.0$)]. Two-body motion is known to trace those paths that correspond to one of the conic sections. On the other hand, Figs. 1b–1d depict the collapse of an N-body system ($N \geq 3$) from an initial state of rest. The 2D presentations achieve a sense of connection with real galactic systems that a 1D self-gravitating system cannot while avoiding the complications encountered (e.g., singularities, evaporation [4]) in a 3D analysis. The examples are outputs of a trained NN with its $E(r_i)$ reduced to 10^{-3}. The trained NN performs robustly against slight changes in the initial conditions which is crucial for maintaining the accuracy of $r_i(t)$ with increasing t considering that the gravitational N-body system is quite sensitive to changes in the initial state for $N \geq 3$ [13].

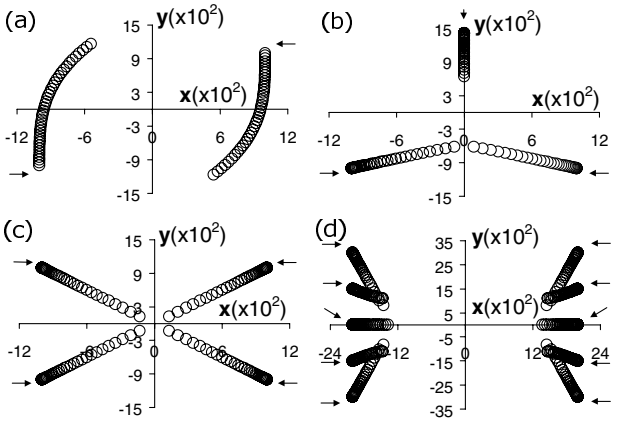

FIG. 1. Temporal evolution of N-body gravitational system on a 2D plane: (a) $N = 2$, (b) $N = 3$, (c) $N = 4$, and (d) $N = 10$. Initial particle positions are designated by arrows. Particles are initially at rest except for the two-body system where bodies are given a slight initial vertical velocity.

Accuracy.—The functional dependence between the NN output error ϵ_r and $E(r_i)$ could not be established because the true solution of Eq. (1) is unknown *a priori*. Hence, ϵ_r is not directly computable from the $E(r_i)$ value. However, strategies exist that permit the minimization of errors such as generalization errors, training set errors, etc. [17]. Standard methods also suffer from the same fate and the capability to handle errors normally rates the superiority of a particular method over others [11]. Accuracy is often measured by comparing the solution with those obtained by other methods [8,18,19].

For the gravitational N-body problem, a possible accuracy measure is the ability of the solution to maintain the value of the system Hamiltonian H through time [20]. Deviations (e.g., distortion, perihelion shift, and overstep phenomenon) from ideal star movements have occurred due to discretization and rounding-off errors. Mostly, they are due to variations in the total energy of the system. For example, Euler integration which is a basic numerical method for N-body problems, is known to gain energy over time.

Figure 2 plots the deviation δH of H as a function of t for $N = 2$, 3, 4, 10, and 25 for a solution of a trained NN with $E(r_i) \sim 10^{-3}$ ($p = 10\,000$) and δH must be 0 at $t = 0$. The results show that while the NN solution could not describe the ideal initial state of the system [$\delta H(t = 0) \neq 0$], it maintains a systematic linear decay for H which approaches the theoretical energy conservation law with increasing p. For systems with more particles, longer

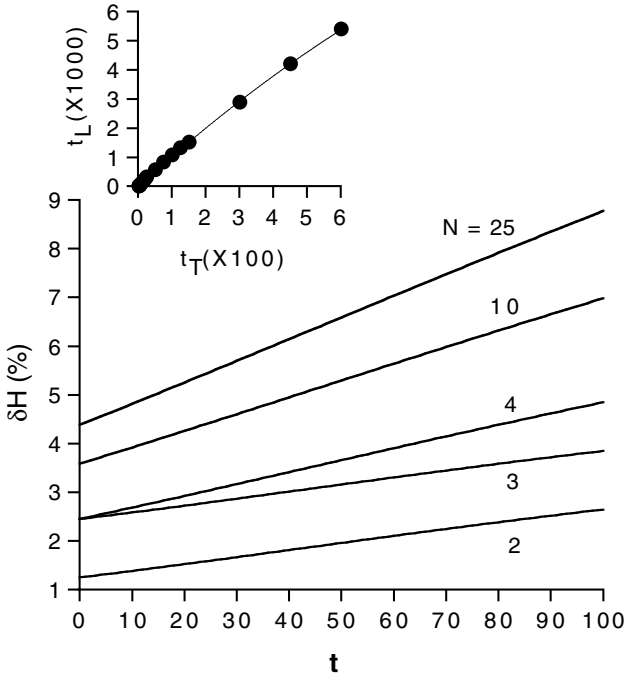

FIG. 2. Percentage deviation of the Hamiltonian as a function of t for $N = 2$, 3, 4, 10, and 25. The inset plot shows the dependence of t_L with t_T where the solid is described by $t_L = -0.003t_T^2 + 10.653t_T + 7.590$.

training times (larger p's) are needed to reach a desired $E(r_i)$ value.

A limit t_L exists for the duration that the NN solution could maintain H. Figure 2 (inset plot) reveals a nonlinear dependence of t_L with the time-range t_T used in the training set that implies an upper limit for the effective training set size. While the NN solution is not exact, the systematic propagation of the error allows for an accurate observation of the temporal evolution of the motion of the individual particles. Our analysis revealed that for N collisionless disks the NN solution did not diverge exponentially. The Hamiltonian has remained accurately conserved (deviation of 0.0132%) even after 10^4 time steps with size $\Delta t = 1$.

The NN solution is analytic and its information content is not limited by sampling conditions (e.g., magnitudes of Δt and sampling period). Its analyticity leads to a more efficient way of storing data concerning the temporal evolution of the system.

Negative time.—The dynamics of retracing the constrained motion of a system of N bodies from their present states is not covered by the Lipschitz condition for initial value problems. The iterative nature of standard solution searches allows only for a forward (temporal) progression of the system from its initial state. They could not address accurately the uniqueness of backtracking the paths taken by the particles [20,21]. Temporal retracing is crucial in issues about the history of a particular set of heavenly bodies from their current states.

The NN solution has information about the "negative" time evolution because the NN assigns $t = 0$ as "present" time when the system is its initial state. Because the NN solution is analytic, values for $\{r_i\}$ at $t < 0$ could be known even if the said t range was not taken during training.

We investigated if a trained NN could retrace a known forward temporal evolution of an N-body system using the concept of negative time (see Fig. 3). A path is first predicted by the trained NN for a two-body system from $t = 0$ to t_{final} then $\{r_i(t = t_{\text{final}/2})\}$ and $\{v_i(t = t_{\text{final}/2})\}$ were chosen and used as initial conditions ($t = 0$) for the new (retraced) path. Two time ranges were predicted: $t = 0$ to $t_{\text{final}/2}$ (forward tracing) and $t = 0$ to $-t_{\text{final}/2}$ (back tracing). We found that the trained NN could retrace the path accurately, which proves that the solution for collisionless N-body system does not diverge exponentially.

Complexity.—The accuracy and efficiency of an N-body simulator depend on N [21,22] and is limited by: (i) computer memory size, (ii) sampling period, and (iii) statistical distributions of real N-body system [6–9,22]. In standard methods, large N's may be accommodated by performing s separate simulations with N/s particles instead of one simulation with N particles. The former is computationally less expensive and easy to parallelize [21]. However, this scaling procedure works only for clusters whose dynamics scales linearly with N. For bodies that interact diversely in forming clusters, the best way is to modify or reformulate the algorithm, a strategy

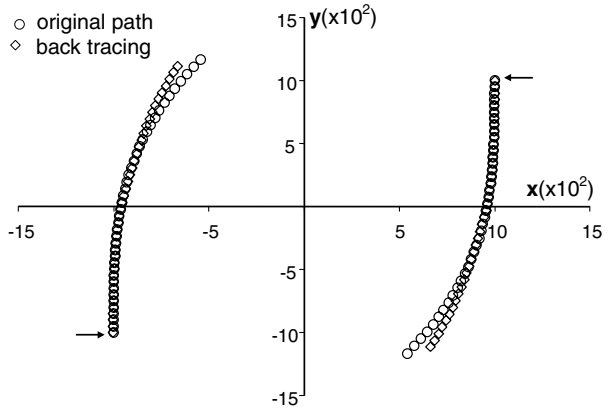

FIG. 3. Retracing the motion of a two-body system with a trained NN. Retraced path exhibits an average error of 0.21% per time step. Initial particle positions are designated by arrows. From the original predicted path, points $\{x = \pm 9.5614 \times 10^2, y = \mp 0.3041 \times 10^2\}$ at $t = t_{\text{final}/2}$ were used as initial states for retracing.

the leads to the tree- and particle-mesh methods whose computational complexity C per time-step scale with $O(N \log N)$ and $O(N + N_g \log N_g)$ respectively, where N_g is the number of grids that statistically confine the N bodies in the particle-mesh method. A large C reduction is achieved relative to the brute force particle-particle method where $C \sim O(N^2)$. Note that both the tree- and particle-mesh algorithms are also implementable via the NN approach.

The analyticity of the NN solution implies that for an optimized NN, the total C scales with $O(1.91N + 3)$ independent of Δt. Optimization of the NN architecture for a given N is obtained by searching for that hidden node number q with the smallest possible $E(r_i)$ for a fixed

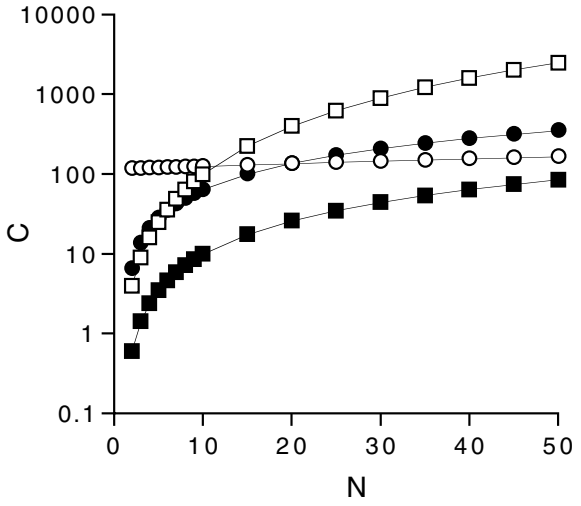

FIG. 4. Comparison of the complexity C as a function of N for particle-particle (squares), tree- (solid squares), particle-mesh (circles), and NN methods (solid circles) with $t = 1$. The first three methods are implemented via the standard integration methods.

iteration length ($p = 10^4$). In the standard methods, the number of operations needed to bring the system from $t = 0$ to $\eta\Delta t$ varies as ηC. The NN method does not depend on η and the trained NN could forecast faster farther into time. Figure 4 compares the dependence of C with N for the various methods.

We have presented a new approach for solving N-body problems. The distinct advantages (analyticity of solution, sytematic propagation of errors, time-step independent C) of the approach over standard methods have been discussed thoroughly.

This work is partially supported by a grant from the Commission on Higher Education of the Philippines.

*Email address: csaloma@nip.upd.edu.ph

[1] A. Muriel, Phys. Rev. Lett. **75**, 2255 (1995).
[2] J-P. Connerade, Nature (London) **391**, 439 (1998).
[3] G. Buck, Nature (London) **395**, 51–53 (1998).
[4] K. R. Yawn and B. N. Miller, Phys. Rev. Lett. **79**, 3561 (1997).
[5] Z. Zheng, G. Hu, and B. Hu, Phys. Rev. Lett. **81**, 5318 (1998).
[6] R. B. Mann, D. Robbins, and T. Ohta, Phys. Rev. Lett. **82**, 3738 (1999).
[7] A. Marciniak, *Numerical Solutions of the N-Body Problem* (Dordrecht, Boston, 1985).
[8] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C* (Cambridge University Press, New York, 1992), 2nd ed.; R. Crandall, *Topics in Advanced Scientific Computation* (TELOS, New York, 1996).
[9] W. Hayes, Ph.D. thesis, University of Toronto, Toronto, Canada, 1995.
[10] D. F. Hevia and A. Ranada, Eur. J. Phys. **17**, 295 (1996).
[11] C. Saloma, J. Appl. Phys. **74**, 5314 (1993); R. Corles, Contemp. Math. **172**, 31 (1994).
[12] B. Ph. van Milligen *et al.,* Phys. Rev. Lett. **75**, 3594 (1995).
[13] C. Monterola and C. Saloma, Phys. Rev. E **57**, 1247R (1998).
[14] M. Lecar, Bull. Astron. **3**, 91 (1968).
[15] J. Goodman, D. Heggie, and P. Hut, Astrophys. J. **415**, 715 (1993); H. Kandrup and H. Smith, Astrophys. J. **374**, 255 (1991); **386**, 635 (1992).
[16] S. Haykin, *Neural Networks: A Comprehensive Foundation* (Prentice Hall, New York, 1999), 2nd ed.
[17] V. Vapnik, E. Levin, and Y. Le Cun, Neural Comput. **5**, 2 (2000).
[18] G. Quinlan and S. Tremaine, in *Proceedings of the Conference on Dynamics of Disc Galaxies,* edited by B. Sundelius (Göteborg University and Chalmers University of Technology, Göteborg, Sweden, 1991), p. 143.
[19] J. M. Sanz-Serna and S. Larsson, Appl. Numer. Math **13**, 181 (1993); G. Quinlan and S. Tremaine, Mon. Not. R. Astron. Soc. **259**, 505 (1992).
[20] R. Skeel and C. Gear, Physica (Amsterdam) **60D**, 311 (1992).
[21] *Predictability, Stability and Chaos in N-Body Dynamical Systems,* edited by D. Heggie and A. Roy (Plenum Press, New York, 1991), pp. 47–62.
[22] H. Smith, Jr., Astron. Astrophys. **76**, 192 (1979).