

OPEN ACCESS

Molecular dynamics equation designed for realizing arbitrary density: Application to sampling method utilizing the Tsallis generalized distribution

To cite this article: Ikuo Fukuda and Haruki Nakamura 2010 *J. Phys.: Conf. Ser.* **201** 012011

View the [article online](#) for updates and enhancements.

You may also like

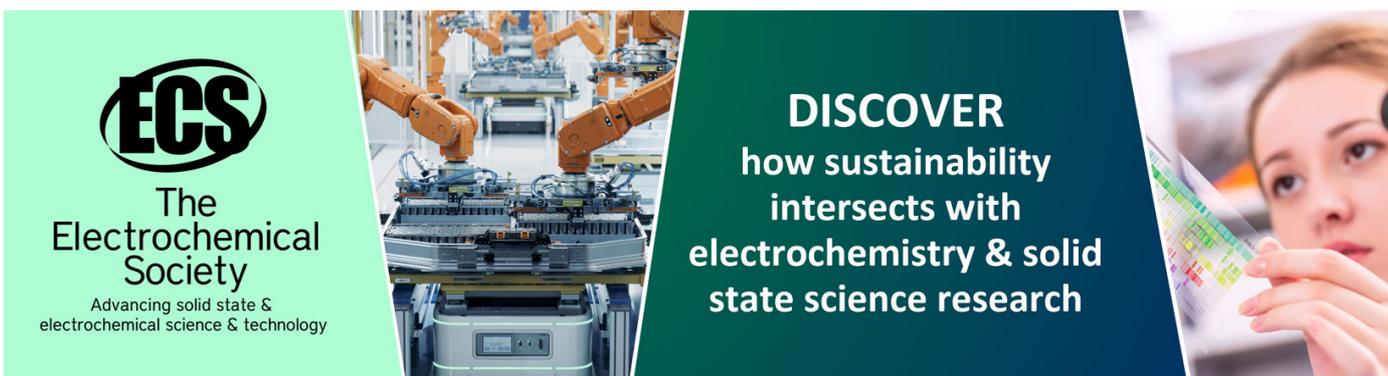
- [Transverse momentum distributions in proton–proton collisions at LHC energies and Tsallis thermodynamics](#)
M D Azmi and J Cleymans

- [Entropy per Rapidity in Pb-Pb Central Collisions using Thermal and Artificial Neural Network \(ANN\) Models at LHC Energies](#)

D. M. Habashy, Mahmoud Y. El-Bakry, Werner Scheinast et al.

- [SECTORS AND LARGE-SCALE MAGNETIC FIELD STRENGTH FLUCTUATIONS IN THE HELIOSHEATH NEAR 110 AU: VOYAGER 1, 2009](#)

L. F. Burlaga and N. F. Ness



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

Molecular dynamics equation designed for realizing arbitrary density: application to sampling method utilizing the Tsallis generalized distribution

Ikuro Fukuda¹ and Haruki Nakamura²

¹Computational Science Research Program, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

²Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-087, Japan

E-mail: ifukuda@riken.jp

Abstract. Several molecular dynamics techniques applying the Tsallis generalized distribution are presented. We have developed a deterministic dynamics to generate an arbitrary smooth density function ρ . It creates a measure-preserving flow with respect to the measure $\rho d\omega$ and realizes the density ρ under the assumption of the ergodicity. It can thus be used to investigate physical systems that obey such distribution density. Using this technique, the Tsallis distribution density based on a full energy function form along with the Tsallis index $q \geq 1$ can be created. From the fact that an effective support of the Tsallis distribution in the phase space is broad, compared with that of the conventional Boltzmann-Gibbs (BG) distribution, and the fact that the corresponding energy-surface deformation does not change energy minimum points, the dynamics enhances the physical state sampling, in particular for a rugged energy surface spanned by a complicated system. Other feature of the Tsallis distribution is that it provides more degree of the nonlinearity, compared with the case of the BG distribution, in the deterministic dynamics equation, which is very useful to effectively gain the ergodicity of the dynamical system constructed according to the scheme. Combining such methods with the reconstruction technique of the BG distribution, we can obtain the information consistent with the BG ensemble and create the corresponding free energy surface. We demonstrate several sampling results obtained from the systems typical for benchmark tests in MD and from biomolecular systems.

1. Introduction

Tsallis' statistics has been extensively investigated [1], and its studies include the fundamental aspect of statistical physics [2] and various applications to physical systems, such as granular matter and turbulence, and to data analysis for network flow and economics [3]. The explanation in view of the Tsallis' statistics are possible for certain physical phenomena that are intractable by the conventional Boltzmann-Gibbs (BG) extensive statistics [4]. Among the properties of this new statistics, we have paid much attention to the fact that the canonical ensemble theory derived from microcanonical ensemble is not unique and that the Tsallis statistics can be derived [5, 6, 7, 8]. Namely, the distribution representing a subsystem in a closed total system is not limited to the BG distribution, but can be realized by the Tsallis distribution. Utilizing the canonical ensemble for molecular dynamics (MD) is necessary in view of the comparison

between the MD simulation and the experiment. MD is a method that calculates interactions among microscopic particles composing the physical system and defines a time development of the system, in order to obtain thermal and dynamical properties of the physical system [9]. Thus, the construction of an MD method realizing the Tsallis distribution is very interesting and important.

In contrast to such a pure physical interest regarding the Tsallis distribution itself, there has been a practical approach such that the Tsallis distribution is used as a tool for effective sampling for the BG distribution, on the basis of the traditional statistical mechanics. BG distribution has been intensively investigated also for biological systems, and many studies have revealed that the properties of such systems are well described through the thermodynamical quantities, including the free energy, based on the BG distribution. For example, a stable three-dimensional protein structure is characterized by the lowest free energy state based on the BG distribution among a variety of putative structures of a polypeptide, which has the appropriate amino-acid sequence for an individual protein. However, such system has a complicated energy surface and it is thus hard to obtain an accurate BG distribution. One of the advantages of using the Tsallis distribution for the sampling of physical states is that we can deform the energy surface in a tractable manner. Specifically, an effective support of the Tsallis distribution in the phase space is broad, compared with that of the conventional BG distribution, since the Tsallis distribution obeys the power law with respect to the energy and decreases gradually in the case of the Tsallis index $q > 1$. Furthermore, the energy-surface deformation by the Tsallis distribution does not change the energy minimum points. For these reasons, This distribution enhances to traverse the energy well, without changing the stable states, and to effectively explore the wide region of the phase space. Another advantage of using the distribution is brought when applying it to the equations of motion of MD method. Specifically, we can effectively gain the ergodicity of the dynamical system, which is required to ensure generating the target distribution. This is because the Tsallis distribution provides a greater degree of the nonlinearity in the equations of motion. In addition, owing to the fact that the Tsallis distribution explicitly equips its functional form, we are free from the direct handling of the derivative of the energy density of states, which is required in the flat distribution method such as multicanonical MD.

For the purpose mentioned above, we require constructing an MD method that deterministically generates the Tsallis distribution. In this process, considering to the non-uniqueness of the canonical distribution suggested by the studies of the nonextensive statistics, it is preferable to have the ability to generate an arbitrary distribution, rather than a specific equilibrium distribution. Furthermore, in view of the sampling technique, it is flexible not to limit the distribution. Thus, it is useful to construct an equation that realizes an arbitrary given density function, although this is the opposite approach, where we find an underlying density function in the given equations of motion. Regarding such an attempt, we demonstrate the density dynamics in section 2. Applying this method to the Tsallis distribution, we introduce an MD method generating the Tsallis distribution. This MD method is an extension of the MD method that creates the BG distribution (section 3). Against certain physical systems intractable by the conventional methods due to the limited ability to generate the BG distributions, accurate and efficient results are given by the current method owing to the properties of the Tsallis distribution and the equations of motion (section 4).

2. Density dynamics

Let ρ be a density function, i.e., a smooth, positive, and integrable function defined on a domain Ω in \mathbb{R}^N . Consider an ordinary differential equation (ODE) having invariant density ρ ; i.e., given ρ becomes the density, with respect to Lebesgue measure $d\omega$ on \mathbb{R}^N , of an invariant measure for the flow $\{T_t : \Omega \rightarrow \Omega | t \in \mathbb{R}\}$ generated by an ODE

$$\dot{\omega} = X(\omega), \tag{1}$$

where $X : \mathbb{R}^N \supset \Omega \rightarrow \mathbb{R}^N$ is a smooth, complete vector field. From a generalized Liouville's theorem, it is sufficient that the vector field X satisfies the Liouville equation [10]

$$\operatorname{div} \rho X = 0. \tag{2}$$

It is shown that, by Birkhoff's individual ergodic theorem [11], a time average of any function f (a Borel measurable function on Ω with $\int_{\Omega} |f| \rho d\omega < +\infty$) has a long-time limit value for $\rho d\omega$ -almost every initial point ω . Further, if the flow is ergodic with respect to the measure space $(\Omega, \mathcal{L}_N \cap \Omega, \rho d\omega)$,

$$\lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^{\tau} f(T_t(\omega)) dt = \int_{\Omega} f \rho d\omega \bigg/ \int_{\Omega} \rho d\omega \tag{3}$$

holds almost everywhere. Substitution of $f \equiv \chi_A$, the characteristic function for any Borel set $A \subset \Omega$, yields

$$\lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^{\tau} \chi_A(T_t(\omega)) dt = \operatorname{const} \times \int_A \rho(\omega) d\omega \quad (\text{a.e.}), \tag{4}$$

indicating that the probability density for the realization of point ω is proportional to the given value $\rho(\omega)$. In the sense that an arbitrary density can be realized, we call (ρ, X) , a doublet of objective density ρ and a corresponding field X , density dynamics.

A specific form of objective X can be constructed by referring to the Nosé-Hoover (NH) equation [12, 13, 14]. The NH equation is an ODE that can realize the BG distribution, and has been extensively studied including the related ergodic property [15, 16, 17]. Now, the ODE [18] is

$$\begin{aligned} \dot{x}_i &= D_{p_i} \Theta(\omega), \quad i = 1, \dots, n, \\ \dot{p}_i &= -D_{x_i} \Theta(\omega) - D_{\zeta} \Theta(\omega) p_i, \quad i = 1, \dots, n, \\ \dot{\zeta} &= \sum_{j=1}^n D_{p_j} \Theta(\omega) p_j - n, \end{aligned} \tag{5}$$

with

$$\Theta = -\ln \rho : \Omega \rightarrow \mathbb{R}. \tag{6}$$

Here the phase space point is represented by $\omega \equiv (x, p, \zeta)$, where $x \equiv (x_1, \dots, x_n)$ and $p \equiv (p_1, \dots, p_n)$ represent coordinates and momenta, respectively, of the physical system, and ζ is an extended variable ($N \equiv 2n + 1$); $D_{x_i} \Theta(\omega)$, $D_{p_i} \Theta(\omega)$, and $D_{\zeta} \Theta(\omega)$ denote the partial derivative of Θ at ω with respect to x_i , p_i , and ζ , respectively. It can be directly checked that the Liouville equation (2) holds. A certain condition is required for the function Θ to ensure the completeness of X . We assume the fulfillment of this condition as well as the ergodicity.

Here, we have utilized the mathematical structure of invariant measure for the time development of the system and based our formalism on the ergodicity. But, it is, in general, difficult to judge exactly whether or not a given system is ergodic. Even if we can judge it, the system is not necessarily ergodic. When the system is not ergodic, however, ergodic decomposition theorem [19] ensures that on individual ergodic components, which uniquely decompose the phase space, the corresponding ergodicities are valid. If physically meaningful components are produced, it is natural to consider the ergodicity on each invariant set, provided that such a set and the induced invariant measure can be explicitly described.

3. Tsallis dynamics

The method introduced in the previous section is applied to the Tsallis distribution density of the following form [20],

$$\rho_{\text{Tsallis}}(x, p) = [1 - (1 - q)\beta E(x, p)]^{q/(1-q)}, \quad (7)$$

where $E(x, p)$ is a total energy, the sum of kinetic energy $K(p) \equiv \frac{1}{2} \|p\|^2$ and potential energy $U(x)$, which is defined on a domain $D \subset \mathbb{R}^n$. The parameter q is the Tsallis index, for which we treat $q \geq 1$. The limit of the density as $q \rightarrow 1$ is $\rho_{\text{BG}}(x, p) \equiv \exp[-\beta E(x, p)]$, the traditional BG density. In Eq. (7) we have considered the escort probabilities [4, 20], considering the normalized q -expectation value of E , and the renormalized temperature T' with $\beta = 1/k_{\text{B}}T'$. We often shift the origin of the energy value to handle the unknown energy minimum in the applications. It is noted that our method is not restricted to the form given in Eq. (7), but available to utilize the other forms, including the q -exponential function.

To apply the density, Eq. (7), to Eq. (5), we set

$$\rho(x, p, \zeta) \equiv \rho_{\text{Tsallis}}(x, p) \rho_z(\zeta). \quad (8)$$

Here we have put a general form of the density component with respect to ζ , as $\rho_z(\zeta)$, which should contribute to meet the condition for ρ on $\Omega \equiv D \times \mathbb{R}^n \times \mathbb{R}$. Then Eq. (5) yields the Tsallis dynamics (TD) [18, 21]:

$$\begin{aligned} \dot{x}_i &= g(x, p) p_i, \quad i = 1, \dots, n, \\ \dot{p}_i &= -g(x, p) DU(x) - \tau(\zeta) p_i, \quad i = 1, \dots, n, \\ \dot{\zeta} &= g(x, p) \|p\|^2 - nk_{\text{B}}T', \end{aligned} \quad (9)$$

where

$$\begin{aligned} g(x, p) &\equiv \frac{q}{1 - (1 - q)E(x, p)/k_{\text{B}}T'}, \\ \tau(\zeta) &\equiv -k_{\text{B}}T' D \ln \rho_z(\zeta). \end{aligned}$$

Here we have multiplied the right-hand side of Eq. (9) by constant factor $k_{\text{B}}T'$, so that we get variables x and p with the ordinary physical dimensions (similar procedure is clearly possible in Eq. (5)). Corresponding to the fact that the $q \rightarrow 1$ limit for Tsallis density (7) becomes the BG density, substitution of $q = 1$ [and $\rho_z(\zeta) = \exp[-(\beta/2Q)\zeta^2]$] into TD equation (9), which generates the Tsallis density, yields the NH equation, which generates the BG density. Namely, the extensive limit of the TD is the (generalized) NH equation. As a similar method, that [22] using only the potential energy $U(x)$ (instead of the total energy) for defining the distribution or that [23] targeting the superextensive region ($0 < q < 1$) has been proposed. The current method first succeeded in generating the Tsallis distribution of the full energetic form in the subextensive region ($q > 1$). Note that also a Hamiltonian approach to achieve the Tsallis distribution, in conjunction with the Winkler's method [24], was proposed [25].

Under the conditions stated above and the ergodicity, a long-time average value of physical variable O represented by a function of x and p exists for almost everywhere and equals to the

space average in the Tsallis distribution:

$$\begin{aligned}
 & \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau O(x(t), p(t)) dt \\
 &= \int_{\Omega} O(x, p) \rho(\omega) d\omega \bigg/ \int_{\Omega} \rho(\omega) d\omega \\
 &= \int_{D \times \mathbb{R}^n} O(x, p) \rho_{\text{Tsallis}}(x, p) dx dp \bigg/ \int_{D \times \mathbb{R}^n} \rho_{\text{Tsallis}}(x, p) dx dp .
 \end{aligned} \tag{10}$$

We can reconstruct the BG distribution using ODE (9), by applying the re-weighting technique [26]. Under the same assumptions, from Eq. (10) we see

$$\overline{O \rho_{\text{BG}} / \rho_{\text{Tsallis}}} \bigg/ \overline{\rho_{\text{BG}} / \rho_{\text{Tsallis}}} = \int_{D \times \mathbb{R}^n} O(x, p) \rho_{\text{BG}}(x, p) dx dp \bigg/ \int_{D \times \mathbb{R}^n} \rho_{\text{BG}} dx dp \quad (\text{a.e.}) \tag{11}$$

provided that the space integrals are finite. Therefore, an expectation of O in the BG distribution is obtained, when we evaluate the long-time average of O scaled by the ratio $\rho_{\text{BG}} / \rho_{\text{Tsallis}}$ and the long-time average of the ratio. In Eq. (11), the temperature T in $\rho_{\text{BG}}(x, p) = \exp[-E(x, p)/T]$ is not necessary to accord with T' defined in ρ_{Tsallis} , but can be given by any positive value. In addition, during a single process of integrating ODE (9) with a fixed T' , we can calculate the expectation values for many temperatures $T = T_1, T_2, \dots$. This means that many BG distributions with individual temperatures can be realized in a single integration process of the equation of motion. A suitable setting of the parameters in the distribution is not trivial matter. An attempt of the setting for the purpose of the sampling of states is described in Ref. [27].

4. Application of the Tsallis dynamics

As an illustration of the application of TD method to MD, we discuss how the problems in MD can be solved, regarding the 1-dimensional harmonic oscillator (1HO) and 1-dimensional double-well (1DW) systems. Harmonic oscillation becomes an approximation of the small oscillation around the energy minima. However, even for such a simple motion, the correct BG distribution cannot be successfully created, which is a peculiar problem in MD such as in the NH equation. Most evident example is 1HO, in which the trajectory does not fully explore the phase space and show nonergodic behaviours, leading to the failure of generating the BG distribution. This failure would come from an ‘‘insufficient complexity’’ in the equations of motion, due to the fact that the original 1HO Newton equation is a simple (integrable) and small system (‘‘complexity’’ of the original physical system is one of the factors (but an important factor) to characterize the ‘‘complexity’’ in the equations of motion of MD). To solve this problem, several methods have been proposed, where the number of the extended variables was increased. For example, by increasing arbitrarily the number of thermostats and enforcing their coupling, the Nosé-Hoover chain method [28] gains the complexity in the equations of motions and has improved the ability to obtain ergodicity, even for a simple given system such as 1HO. Another route to gain the complexity is found by the TD method, which uses only one extended variable. That is, by increasing the nonlinearity in the equation of motion by using the ‘‘ q -deforming’’ [18] of the NH equation together with a suitable choice of $\rho_z(\zeta)$, TD can generate the BG distribution for 1HO when we recover the BG distribution by the re-weighting [29].

To explore the energy landscape it is necessary to visit other energy minima, by escaping the harmonic-like motion in one energy well, through an unharmonic motion. The most fundamental

model of such a motion is that in 1DW. As the energy well is deep, the trajectory is confined over a long time, so that it needs much time to sufficiently sample the phase space, even if the system is ergodic. Many BG dynamics (BGD), i.e., deterministic equation that can generate the BG distribution in a direct way, have been proposed [30, 31, 32]. However, as the accuracy of BGD is increased, the accuracy of the probability $\propto \exp(-\beta E)$ even at a high energy barrier will be increased, so that the simulation requires a long time to escape beyond such barriers. This second problem in MD is emphasized as “increasing the complexity” of the system, in contrast to the first problem stated above. TD solves this problem, using the slow decay feature of the Tsallis distribution with respect to the energy, which means that high energy states can be realized more frequently than those in the BG distribution and so overcoming the energy barriers is enhanced. In systems having such energy surfaces, it is shown [21, 29] that much effective sampling is performed in TD, compared with that in BGD.

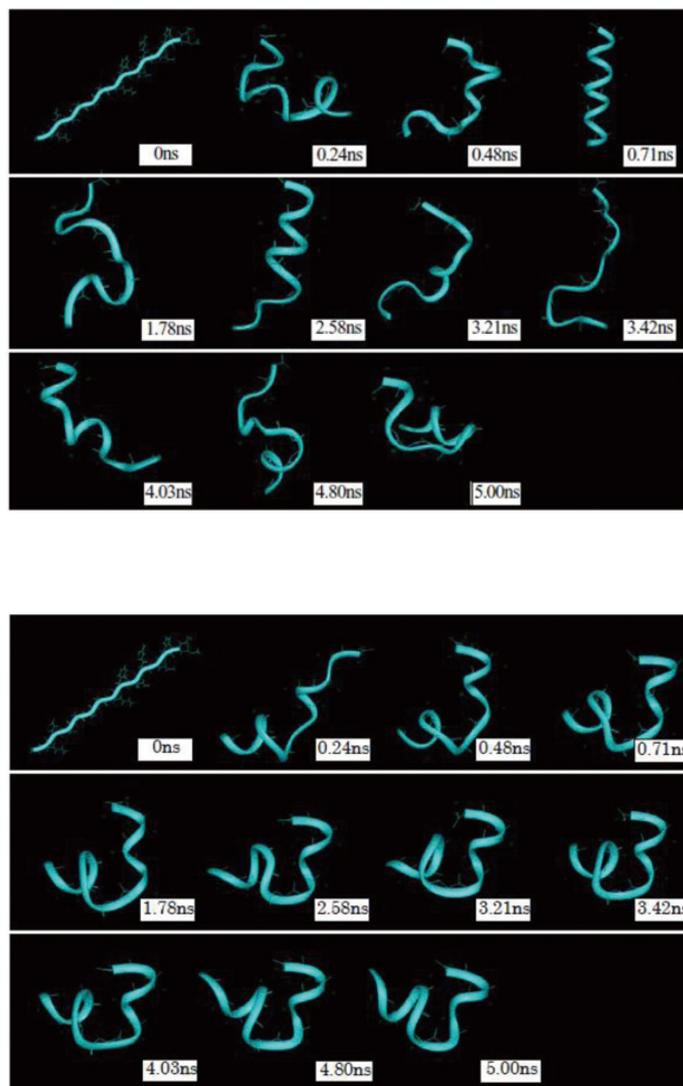


Figure 1. Structural change of main chain of C-peptide obtained in (upper panel) TD simulation and (lower panel) a conventional MD simulation.

As an application of TD to biomolecular system, we show the results of C-peptide, which is a fragment of bovine pancreatic ribonuclease A (PDB id: 1a2w). Figure 1 shows the results of a numerical simulation of 5 ns of the peptide. An extended conformation of the molecule was used as an initial condition. We observe that much effective conformational sampling was performed by the TD (the upper panel of the figure), compared with the results of the conventional MD (BGD at the temperature of 300 K; shown in the lower panel). In the BGD, getting trapped into a local backbone structure is observed, while in the TD a variety of the structures, including the α -helical structure (see e.g., 2.58 ns), which is considered to be the native structure, were sampled much faster. Similar results were confirmed for different initial conditions.

For a development, a molecular-dynamics sampling scheme has been proposed [33] in which any summation of multiple arbitrarily given distributions can be realized deterministically. This method is also based on the density dynamics demonstrated in section 2.

5. Concluding remarks

We have constituted the density dynamics, which is a deterministic method to realize an arbitrary density function. Applying this method to the Tsallis distribution density, we constructed an ODE generating the Tsallis distribution that is described by both the coordinates and the momenta and by the Tsallis index $q \geq 1$, whereas such a construction was impossible in conventional approaches. We demonstrated, via the investigations of 1HO and 1DW systems, that the typical problems in MD were solved by the characteristic features of this dynamics. Application to the peptide system shows that the current method provides efficient results against the case in which the conventional deterministic method unsuccessfully sample the phase space. Other related numerical techniques are shown in Ref. [34], and a recent development relevant to the generalized ensemble method will be discussed in detail elsewhere.

Acknowledgments

This research was supported by Research and Development of the Next-Generation Integrated Simulation of Living Matter, a part of the Development and Use of the Next-Generation Supercomputer Project of the Ministry of Education, Culture, Sports, Science and Technology (MEXT). We thank Masaru Horie for discussions and his valuable help in numerical simulations.

References

- [1] Tsallis C 1988 *J. Stat. Phys.* **52** 479
- [2] Suyari H 2006 *Physica A* **368** 63
- [3] Tsallis C 2009 *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World* (Berlin: Springer)
- [4] Tsallis C 1999 *Braz. J. Phys.* **29** 1; for an updated bibliography, cf. <http://tsallis.cat.cbpf.br/biblio.htm>.
- [5] Plastino A R and Plastino A 1994 *Phys. Lett. A* **193** 140
- [6] Abe S and Rajagopal A K 2000 *Phys. Lett. A* **272** 341
- [7] Abe S and Rajagopal A K 2000 *J. Phys. A* **33** 8733
- [8] Abe S and Rajagopal A K 2001 *Europhys. Lett.* **55** 6
- [9] Allen M P and Tildesley D J 1989 *Computer Simulation of Liquids* (Oxford: Oxford Univ. press)
- [10] Mañé R 1987 *Ergodic Theory and Differentiable Dynamics* (Berlin: Springer) p 35
- [11] Inoue K 1989 *Seminar on Probability* **57** 36
- [12] Nosé S 1984 *J. Chem. Phys.* **81** 511
- [13] Hoover W G 1985 *Phys. Rev. A* **31** 1695
- [14] Nosé S 1991 *Prog. Theor. Phys. Suppl.* **103** 1
- [15] Posch H A, Hoover W G and Vesely F J 1986 *Phys. Rev. A* **33** 4253
- [16] Cho K and Joannopoulos J D 1992 *Phys. Rev. A* **45** 7089
- [17] Di Tolla F D and Ronchetti M 1993 *Phys. Rev. E* **48** 1726
- [18] Fukuda I and Nakamura H 2002 *Phys. Rev. E* **65** 026105
- [19] Totoki H 1971 *Introduction to ergodic theory* (Tokyo: Kyoritsu)
- [20] Tsallis C, Mendes R S and Plastino A R 1998 *Physica A* **261** 534

- [21] Fukuda I and Nakamura H 2003 *Chem. Phys. Lett.* **382** 367
- [22] Andricioaei I and Straub J E 1997 *J. Chem. Phys.* **107** 9117
- [23] Plastino A R and Anteneodo C 1997 *Ann. Phys.* **255** 250
- [24] Winkler R G 1992 *Phys. Rev. A* **45** 2250
- [25] Andrade J S Jr., Almeida M P, Moreira A A and Farias G A 2002 *Phys. Rev. E* **65** 036121
- [26] Ferrenberg A M and Swendsen R H 1988 *Phys. Rev. Lett.* **61** 2635
- [27] Fukuda I, Horie M and Nakamura H 2005 *Chem. Phys. Lett.* **405** 364
- [28] Martyna G J, Klein M L and Tuckerman M 1992 *J. Chem. Phys.* **97** 2635
- [29] Fukuda I and Nakamura H 2004 *J. Phys. Chem. B* **108** 4162
- [30] Hoover W G and Holian B L 1996 *Phys. Lett. A* **211** 253; and the references therein
- [31] Liu Y and Tuckerman M E 2000 *J. Chem. Phys.* **112** 1685
- [32] Sergi A and Ferrario M 2001 *Phys. Rev. E* **64** 056125
- [33] Fukuda I and Nakamura H 2005 *Phys. Rev. E* **71** 046708
- [34] Fukuda I and Nakamura H 2006 *Phys. Rev. E* **73** 026703