Fluctuation Relation beyond Linear Response Theory

A. Giuliani*,[1,2] F. Zamponi†,[1,3] and G. Gallavotti‡[1,2]

[1] *Dipartimento di Fisica, Università di Roma* La Sapienza, *P. A. Moro 2, 00185 Roma, Italy*
[2] *INFN, Università di Roma* La Sapienza, *P. A. Moro 2, 00185 Roma, Italy*
[3] *INFM – CRS Soft, Università di Roma* La Sapienza, *P. A. Moro 2, 00185 Roma, Italy*

(Dated: December 16, 2004)

The Fluctuation Relation (FR) is an asymptotic result on the distribution of certain observables averaged over time intervals τ as $\tau \to \infty$ and it is a generalization of the fluctuation–dissipation theorem to far from equilibrium systems in a steady state which reduces to the usual Green–Kubo (GK) relation in the limit of small external non conservative forces. FR is a theorem for smooth uniformly hyperbolic systems, and it is assumed to be true in all dissipative "chaotic enough" systems in a steady state. In this paper we develop a theory of finite time corrections to FR, needed to compare the asymptotic prediction of FR with numerical observations, which necessarily involve fluctuations of observables averaged over finite time intervals τ. We perform a numerical test of FR in two cases in which non Gaussian fluctuations are observable while GK does not apply and we get a non trivial verification of FR that is *independent of* and *different from* linear response theory. Our results are compatible with the theory of finite time corrections to FR, while FR would be *observably violated*, well within the precision of our experiments, if such corrections were neglected.

PACS numbers: 05.40.-a,05.45.-a,05.45.Pq

Keywords: entropy production rate, fluctuation theorem, non-Gaussian fluctuations, Green-Kubo relations

The Fluctuation Relation for the distribution $\pi_\tau(p)$ of fluctuations of the phase space contraction rate averaged over a time interval τ is studied numerically in a system of particles subject to an electric field and to an interparticle potential. The density of the system and the strength of the field are chosen so high that violations of linear response theory and non–Gaussianity of $\pi_\tau(p)$ are observable. In order to compare our results with the asymptotic ($\tau \to \infty$) prediction of the Fluctuation Theorem, a theory for the finite time corrections to the Fluctuation Relation becomes necessary as otherwise the Fluctuation Relation, expected to hold on the basis of the "Chaotic Hypothesis", would be apparently measurably violated in our rather extreme experimental situation. The finite time corrections theory is based on the thermodynamic formalism for "highly chaotic" systems which allows us to transform the analysis of the corrections into a rather familiar problem of "finite size effects" in Statistical Mechanics, where it becomes an application of a local form of the central limit theorem. Once the corrections to the Fluctuation Relation are taken into account our observed agreement with the theory provides an indirect confirmation of the validity of the Chaotic Hypothesis for our system.

I. INTRODUCTION

Anosov systems and the fluctuation theorem - The fluctuation theorem concerns fluctuations of phase space contraction in reversible hyperbolic (Anosov) systems. If time evolution is described by a differential equation on phase space M: $\dot{x} = X(x)$, $x \in M$, or by a map $S : x \to S(x)$ of M one defines the *phase space contraction* as, respectively, $\sigma(x) = -\operatorname{div} X(x)$ or $\sigma(x) = -\log|\det \partial_x S(x)|$. Reversibility means that there is a metric–preserving map I of M such that $IS = S^{-1}I$ if S is the time evolution over a certain time t (*e.g.* $t = 1$). *If the system is Anosov*, that is if M is compact and S is smooth and uniformly hyperbolic, see [1–5], the points x will have a well defined SRB distribution μ_{srb}, [5], *i.e.* almost all points w.r.t. the volume measure will evolve so that *all smooth observables* will have a well defined average equal to the integral over the SRB distribution. Hence, in particular, the time average of the function $\sigma(x)$ will be asymptotically given by the spatial average w.r.t. the SRB distribution. In the case of discrete time maps:

$$\sigma_+ \stackrel{def}{=} \lim_{\tau\to\infty} \frac{1}{\tau} \sum_{j=0}^{\tau-1} \sigma(S^j(x)) = \int_M \sigma \, d\mu_{srb} \stackrel{def}{=} \langle \, \sigma \, \rangle_{srb} \quad (1)$$

If $\sigma_+ > 0$, let:

$$p(x) = \frac{1}{\tau\sigma_+} \sum_{j=0}^{\tau-1} \sigma(S^j(x)) \quad (2)$$

Analogous definitions are given in the continuous time case. The function $p(x)$ will have average $\langle \, p \, \rangle_{srb} = 1$ and distribution $\pi_\tau(dp)$ such that

$$\pi_\tau(\{p \in \Delta\}) = e^{\tau \max_{p\in\Delta} \zeta_\infty(p) + o(\tau)} \, , \quad (3)$$

*e-mail: alessandro.giuliani@roma1.infn.it
†e-mail: francesco.zamponi@phys.uniroma1.it
‡e-mail: giovanni.gallavotti@roma1.infn.it

where the correction at the exponent is $o(\tau)$ w.r.t. τ as $\tau \to \infty$. The following *Fluctuation Relation*, discovered in a numerical simulation in [6] and formulated as a theorem for Anosov systems in [2], holds:

$$\zeta_\infty(p) = \zeta_\infty(-p) + p\sigma_+ \qquad \text{for all } |p| < p^* \tag{4}$$

where $\infty > p^* \geq 1$ is a suitable (model dependent) constant that, in general, is *different* from the maximum over τ and x of $p(x)$; note also that Eq. 4 is (strictly speaking) meaningless in the *equilibrium cases* in which the system is Hamiltonian and reversibility is the usual velocity sign change: because of the division by $\sigma_+ = 0$ in Eq. 2.

The chaotic hypothesis - Hyperbolicity is a paradigm for disordered systems similar to the small oscillations paradigm used for ordered motions: it does not hold exactly in essentially all the physically interesting systems. The *Chaotic Hypothesis* [1–3, 7, 8] is that nevertheless one can assume that chaotic motions (in the sense of motions with at least one positive Lyapunov exponent) exhibit some average properties of truly hyperbolic motions. This hypothesis is a natural generalization of the ergodic hypothesis, i.e. of the assumption that systems of many particles at equilibrium are well described on average by the microcanonical (or by the Gibbs) distribution, even if they are not really (or they are not proven to be) ergodic. A consequence of the Chaotic Hypothesis is that (dissipative) deterministic chaotic reversible motions should have fluctuations of phase space contraction satisfying Eq. 4.

One interesting example of such motions is given by a system of N interacting particles in d dimensions subjected to nonconservative forces and kept in a stationary state by a *reversible mechanical thermostat*. It will be defined by a differential equation $\dot{x} = X_E(x)$ where $x = (\underline{\dot{q}}, \underline{q}) \in R^{2dN} \equiv M$ (*phase space*) and

$$m\underline{\ddot{q}} = \underline{f}(\underline{q}) + \underline{g}_E(\underline{q}) - \underline{\theta}_E(\underline{\dot{q}}, \underline{q}) \tag{5}$$

where m is the mass of the particles, $\underline{f}(\underline{q})$ describes the internal (conservative) forces between the particles and $\underline{g}_E(\underline{q})$ represents the nonconservative "external" force acting on the system. Finally, $\underline{\theta}_E(\underline{\dot{q}}, \underline{q})$ is a mechanical force that prevents the system to acquire energy indefinitely: this is why we shall call it a *mechanical thermostat*. Systems belonging to this class are frequently used as microscopic models to describe nonequilibrium stationary states induced by the application of a driving force (temperature or velocity gradients, electric fields, etc.) on a fluid system in contact with a thermal bath, [8, 9]. In this context, the phase space contraction rate $\sigma(x)$ has been identified (setting $k_B = 1$) with the *entropy production rate* [1, 6–8], the variable $p(x)$ is defined as

$$p(x) = \frac{1}{\tau\sigma_+} \int_0^\tau dt\ \sigma(S_t x) \tag{6}$$

(where $x(t) \equiv S_t x$ is the solution of Eq. 5 with initial datum $x(0) = x$) and the fluctuation relation has been successfully tested in several numerical simulations [6, 10–15]. Having defined the notion of entropy production rate one can define a "duality" between fluxes $\underline{J}$ and forces $\underline{E}$ using $\sigma(x)$ as a "Lagrangian" [7]:

$$\underline{J}(\underline{E}, x) = \frac{\partial \sigma(x)}{\partial \underline{E}} \tag{7}$$

In the limit $\underline{E} \to \underline{0}$, *i.e.* close to equilibrium, the fluctuation relation leads to Onsager's reciprocity and to Green-Kubo's formulas for transport coefficients [16, 17]:

$$\mu_{ij} \equiv \lim_{\underline{E}\to\underline{0}} \frac{\langle J_i \rangle_{\underline{E}}}{E_j} = \int_0^\infty dt\ \langle J_i(t) J_j(0) \rangle_{\underline{E}=\underline{0}} \tag{8}$$

Gaussian distributions - If the distribution $\pi_\tau(p)$ is Gaussian, $\pi_\tau(p) \propto \exp\left[-\tau\frac{(p-1)^2}{2\delta_\tau^2}\right]$, from the fluctuation relation one can derive an *extension* of the Green-Kubo relation, i.e. of Eq. 8, to finite forces.

Indeed, the fluctuation relation for a Gaussian distribution implies that the dispersion δ_∞^2 of p around its average (equal to 1) is $\delta_\infty^2 = 2/\sigma_+$ which is, in such case, an extension of a Green–Kubo formula to non zero fields. One sees this by considering, for instance, cases in which $\sigma(x)$ is linear in E (as it will be in the cases that we study numerically below). Using time-translation invariance one can show that

$$\delta_\infty^2 = \frac{2}{\sigma_+^2} \int_0^\infty dt\ \langle (\sigma(t) - \sigma_+)(\sigma(0) - \sigma_+) \rangle_E \tag{9}$$

and from the fluctuation relation $\delta_\infty^2 \sigma_+ = 2$

$$\sigma_+ = \int_0^\infty dt\ \langle (\sigma(t) - \sigma_+)(\sigma(0) - \sigma_+) \rangle_E \tag{10}$$

Substituting $\sigma(t) = EJ(t)$ in the latter expression, one obtains the relation

$$\frac{\langle J \rangle_E}{E} = \int_0^\infty dt\ [\langle J(t) J(0) \rangle_E - \langle J \rangle_E^2] \tag{11}$$

valid, *subject to the Gaussian assumption*, also for $E \neq 0$.

The leading order in E of the latter relation (*linear response*) is the Green-Kubo formula for the equilibrium transport coefficient, Eq. 8.

Numerical verification of the chaotic hypothesis - The simplest check of the applicability of the Chaotic Hypothesis is a check of the fluctuation relation: of course even if the check has a positive result this will not "prove" the hypothesis but it will at least add confidence to it. A rather stringent test of the fluctuation relation would be a check which *cannot be reduced to a kind of Green-Kubo relation*; this requires at least one of the two following conditions to be satisfied:

1. the distribution $\pi_\tau(p)$ is distinguishable from a Gaussian, or

2. deviations from the leading order in E in Eq. 11, *i.e.*, deviations from the Green-Kubo relation, are observed.

This is very hard to obtain in numerical simulations of Eq. 5 for the following reasons:

1. to observe deviations from linearity in Eq. 11 one has to apply very large forces E, then σ_+ is very large and it becomes very difficult to observe the negative values of $p(x)$ that are needed to compute $\zeta_\infty(-p)$ in Eq. 4;

2. deviations from Gaussianity in $\pi_\tau(p)$ are observed only for values of p that differ significantly (of the order of $2\delta_\infty$) from 1 and, again, it is very difficult to observe such values of p.

Due to the limited computational resources available in the past decade, all numerical computations that can be found in the literature on systems described by Eq. 5 found that the measured distribution $\pi_\tau(p)$ could not be distinguished from a Gaussian distribution in the interval of p accessible to the numerical experiment [6, 10, 11, 15].

The purpose of the present paper is to test the fluctuation relation, in a numerical simulation of a system described by Eq. 5, for large applied force when deviations from linearity can be observed, and the distribution $\pi_\tau(p)$ is appreciably non-Gaussian. This has become possible thanks to the fast increase of computational power in the last decade. However, it is still very difficult to reach values of τ which can be confidently regarded as "close" to the asymptotic limit $\tau \to \infty$; thus to interpret our results we develop a theory of the $o(1)$ corrections to the function $\zeta_\infty(p)$ in order to extract the limiting function $\zeta_\infty(p)$ from the numerical data. Taking into account the latter finite time corrections, we successfully test the fluctuation relation for non–Gaussian distributions and beyond the linear response theory.

The paper is organized as follows: section II is devoted to the theory of finite τ corrections to the large deviation function $\zeta_\infty(p)$ that is needed in the analysis of the data; in section III we present the model and the details of the numerical simulation; in section IV we present the details of the data analysis; finally, in section V and VI we report the result of our simulations.

II. FINITE TIME CORRECTIONS TO THE FLUCTUATION RELATION

In the present section we describe a strategy to study (in principle constructively) the $O(1)$ corrections in the exponent of Eq. 3. The theory we propose will hold *assuming that the time evolution is hyperbolic* so that it can be applied to physical systems only if the chaotic hypothesis is accepted. For simplicity we consider only the case of discrete time evolution via a map S.

A. SRB measure, symbolic dynamics and statistical mechanics

We study the distribution of p at fixed τ via its Laplace transform (*characteristic function*) $z_\tau(\lambda)$:

$$z_\tau(\lambda) = -\frac{1}{\tau}\log\langle e^{-\lambda\sum_{j=0}^{\tau-1}\sigma(S^j x)}\rangle_{srb} \tag{12}$$

The main consequence of the hyperbolicity is, [2–4, 17–19], that one can find a symbolic representation of the points of M in terms of sequences $\underline{\varepsilon} = (\varepsilon_i)_{i=-\infty}^{\infty}$ of finitely many digits $\varepsilon = 1, \dots, k$ subject only to a simple *hard core* restriction, namely $T_{\varepsilon_j,\varepsilon_{j+1}} \equiv 1$ if T is a matrix (*compatibility matrix*) with entries 0 or 1 and such that $T^N_{\varepsilon,\varepsilon'} > 0$ for some $N > 0$ and all $\varepsilon, \varepsilon'$ (*mixing* condition). Moreover in such a representation the dynamics becomes simply the left shift, *i.e.* if $\underline{\varepsilon}(x)$ represents x then $S(x)$ is represented by the sequence $\underline{\varepsilon}$ shifted to the left by one unit.

The key remarks are

1. Smooth observables on phase space can be represented by *short range potentials*: in the case of the observable $\sigma(x)$ this means that there are functions $s_X(\underline{\varepsilon}_X)$ defined for all intervals $X = (a, \dots, a+2n+1)$ and $\underline{\varepsilon}_X = (\varepsilon_a, \dots, \varepsilon_{a+2n+1})$, *translationally invariant* $s_X = s_{X+b}$ and *exponentially decaying* on time scale κ^{-1} (*i.e.* $|s_X(\underline{\varepsilon}_X)| < Ce^{-\kappa n}$ for some $C, \kappa > 0$), such that

$$\sigma(x) = s(\underline{\varepsilon}(x)), \qquad s(\underline{\varepsilon}) = \sum_{X\circ 0} s_X(\underline{\varepsilon}_X) \tag{13}$$

where the sum is over the intervals X centered at the origin (noted by $X \circ 0$). Another important smooth observable is the *expansion rate* $L(x)$ defined as the logarithm of the determinant of the linearization matrix $\partial S(x)$ (*i.e.* the Jacobian matrix of the map) restricted to the unstable manifold: $L(x) = \log\det\partial S(x)_u$. This is also expressible via an exponentially decaying potential Φ:

$$L(x) = \ell(\underline{\varepsilon}(x)), \qquad \ell(\underline{\varepsilon}) = \sum_{X\circ 0} \Phi_X(\underline{\varepsilon}_X) \tag{14}$$

2. The SRB distribution, represented as a distribution over the (compatible) symbolic sequences $\underline{\varepsilon}$, is a Gibbs state for the short range potential $\Phi = (\Phi_X(\underline{\varepsilon}_X))$ defined in Eq. 14, *i.e.*

$$\langle F\rangle_{srb} = \lim_{R\to\infty}\frac{\sum_{\underline{\varepsilon}} e^{-\sum_{X\subset\Lambda_R}\Phi_X(\underline{\varepsilon}_X)}F(\underline{\varepsilon}')}{\sum_{\underline{\varepsilon}} e^{-\sum_{X\subset\Lambda_R}\Phi_X(\underline{\varepsilon}_X)}} \tag{15}$$

where $\Lambda_R = (-R, \dots, R) \subset \mathbb{Z}$, the sums extend over compatible configurations $\underline{\varepsilon} = (\varepsilon_{-R}, \dots, \varepsilon_R)$ (*i.e.* with $T_{\varepsilon_j,\varepsilon_{j+1}} = 1$ for $j = -R, \dots, R-1$), and $F(\underline{\varepsilon}')$ is an arbitrary smooth observable defined on phase space regarded as a function on the symbolic sequences and evaluated at a sequence $\underline{\varepsilon}'$ which extends (rather arbitrarily) $\underline{\varepsilon}$ to an infinite compatible sequence by continuing $\underline{\varepsilon}$ to the right with a sequence $\underline{\varepsilon}_>$ and to the left with a sequence $\underline{\varepsilon}_<$ into $\underline{\varepsilon}' = (\underline{\varepsilon}_<, \underline{\varepsilon}, \underline{\varepsilon}_>)$ so that $\underline{\varepsilon}_<$ depends only on the symbol ε_{-R} and $\underline{\varepsilon}_>$ depends only on the symbol ε_R: see [18, 20, 21].

The surprising reduction of the problem of studying the SRB distribution to that of a *Gibbs distribution* for a one dimensional chain with short range interaction (this is the physical interpretation of Eq. 15) generated the possibility of studying more quantitatively at least some of the problems of nonequilibrium statistical mechanics outside the domain of "nonequilibrium thermodynamics", [22].

B. Finite time corrections to the characteristic function

The characteristic function $z_\tau(\lambda)$ of p, see Eq. 12, can therefore be computed as

$$e^{-\tau z_\tau(\lambda)} = \lim_{R\to\infty} \frac{\sum_{\underline{\varepsilon}} e^{-\sum_{X\subset\Lambda_R}\Phi_X(\underline{\varepsilon}_X) - \lambda\sum_{X\circ[0,\tau-1]} s_X(\underline{\varepsilon}_X)}}{\sum_{\underline{\varepsilon}} e^{-\sum_{X\subset\Lambda_R}\Phi_X(\underline{\varepsilon}_X)}} \tag{16}$$

This means that it is the (limit as $R\to\infty$ of the) ratio between the partition functions $Z_R(\Phi)$ of a Gibbs distribution in Λ_R with potential Φ (the denominator) and the partition function $Z_R(\Phi,\lambda s)$ with the same potential *modified* in the *finite* region $[0,\tau-1]\subset\mathbb{Z}$ by the addition of a potential $\lambda s_X(\underline{\varepsilon}_X)$.

The one dimensional systems are very well understood and the above is a well studied problem in statistical mechanics, known as *a finite size corrections* calculation. For instance it can be attacked by *cluster expansion*, [21]; this is a technique to deal with the average of the exponential of a spin Hamiltonian which is defined in terms of potentials ϕ_X exponentially decaying with rate κ, such as those appearing in the numerator and in the denominator of Eq. 16. It allows us to represent them as:

$$\sum_{\underline{\varepsilon}} e^{-\sum_{X\subset\Lambda_R}\phi_X(\underline{\varepsilon}_X)} = e^{-\sum_{X\subset\Lambda_R}\widetilde{\phi}_X}\ , \tag{17}$$

where $\widetilde{\phi}_X$ are new *effective* potentials, explicitly computable in terms of suitable averages of products of $\phi_X(\underline{\varepsilon}_X)$'s, and which can be proven to be still exponentially decaying with the diameter of X with a rate $0<\kappa'\le\kappa$.

In particular, a representation like Eq. 17 allows us to rewrite the partition function in the denominator of Eq. 16 as:

$$Z_R(\Phi) = \exp\Big[(2R+1)f_\infty(\Phi) - c_\infty(\Phi) + O(e^{-\kappa' R})\Big] \tag{18}$$

and the one in the numerator as

$$\begin{aligned}Z_R(\Phi,\lambda s) = \exp\Big[&(2R+1-\tau)f_\infty(\Phi) + \tau f_\infty(\Phi+\lambda s)\\ &- c_\infty(\Phi) - g_\infty(\lambda) + O(e^{-\kappa' R}+e^{-\kappa'\tau})\Big]\end{aligned} \tag{19}$$

Therefore

$$\begin{aligned}z_\tau(\lambda) &= f_\infty(\Phi) - f_\infty(\Phi+\lambda s) + \frac{g_\infty(\lambda)}{\tau} + O(e^{-\kappa'\tau})\\ &\stackrel{def}{=} z_\infty(\lambda) + \frac{g_\infty(\lambda)}{\tau} + O(e^{-\kappa'\tau})\end{aligned} \tag{20}$$

The function $z_\infty(\lambda)$ is convex in λ and the functions $g_\infty(\lambda)$ and $z_\tau(\lambda)$ are analytic in λ (a consequence of the 1-dimensionality and of the short range nature of the SRB distribution): namely $g_\infty(\lambda) = g_\infty^{(1)}\lambda + \frac{1}{2}g_\infty^{(2)}\lambda^2 + \ldots$ and $z_\tau(\lambda) = z_\tau^{(1)}\lambda + \frac{1}{2}z_\tau^{(2)}\lambda^2 + \ldots$ and the coefficients of their expansion in a power series of λ can be expressed in terms of correlation functions of $\sigma(x)$. For instance, from Eq. 12 and using the translational invariance of the SRB measure,

$$\begin{aligned}z_\tau^{(1)} &= \tau^{-1}\langle\sum_{j=0}^{\tau-1}\sigma(S^jx)\rangle_{srb} = \sigma_+\\ z_\tau^{(2)} &= \tau^{-1}\left[\langle\sum_{j=0}^{\tau-1}\sigma(S^jx)\rangle_{srb}^2 - \langle\sum_{j=0}^{\tau-1}\sigma(S^jx)\sum_{k=0}^{\tau-1}\sigma(S^kx)\rangle_{srb}\right]\\ &= -\sum_{k=-\tau+1}^{\tau-1}\left[1-\frac{|k|}{\tau}\right]\langle\sigma(S^kx)\sigma(x)\rangle_c\end{aligned} \tag{21}$$

where $\langle\sigma(S^kx)\sigma(x)\rangle_c = \langle\sigma(S^kx)\sigma(x)\rangle_{srb} - \sigma_+^2$. Using Eq. 20, $g_\infty(\lambda) = \lim_{\tau\to\infty}\tau[z_\tau(\lambda)-z_\infty(\lambda)]$, and the analyticity of $z_\tau(\lambda)$, we have $g_\infty^{(j)} = \lim_{\tau\to\infty}\tau[z_\tau^{(j)} - z_\infty^{(j)}]$. Since the connected correlation function $\langle\sigma(S^kx)\sigma(x)\rangle_c$ *decays exponentially* for $k\to\infty$, we obtain

$$\begin{aligned}g_\infty^{(1)} &= 0\\ g_\infty^{(2)} &= \sum_{k=-\infty}^{\infty}|k|\langle\sigma(S^kx)\sigma(x)\rangle_c\end{aligned} \tag{22}$$

C. Finite time corrections to $\zeta_\infty(p)$

A direct measurement of $z_\tau(\lambda)$ from the numerical data is difficult. What is really accessible to numerical observation are the quantities $\frac{1}{\tau}\log\pi_\tau(\{p\in\Delta\})$ in Eq. 3 because the measured values of p are used to build an histogram obtained by dividing the p–axis into sufficiently small bins Δ and counting how many values of p fall in the various bins. Let us choose the size of the bins Δ as $|\Delta| = O(\varepsilon_\tau/\tau)$, with ε_τ a small parameter which will be eventually chosen $\varepsilon_\tau = o(1)$, see Appendix A for a discussion of this point. Let also p_Δ be the center of the bin Δ. An application of a local form of central limit theorem, discussed in Appendix A, shows that the following asymptotic representation of $\pi_\tau(\{p \in \Delta\})$ holds:

$$\pi_\tau(\{p \in \Delta\}) = e^{\tau\zeta_\tau(p_\Delta)}\Big(1 + o(1)\Big) \tag{23}$$

where $\zeta_\tau(p_\Delta)$ can be interpolated by an analytic function of p, satisfying the equation

$$\zeta_\tau(p) = -z_\tau(\lambda_p) + \lambda_p p\sigma_+ - \frac{1}{2\tau}\log\left[\frac{2\pi}{\tau}\Big(-\frac{z''_\tau(\lambda_p)}{\sigma_+^2}\Big)\right] \tag{24}$$

and λ_p is the inverse of $p(\lambda) = z'_\tau(\lambda)/\sigma_+$.

Using the previous equations, we now compute the lowest order finite time correction to $\zeta_\infty(p)$ around the maximum.

We rewrite $\zeta_\tau(p)$ as $\zeta_\tau(p) = \zeta_\infty(p) + \frac{\gamma_\infty(p)}{\tau} + O(\frac{1}{\tau^2})$. By the analyticity of $\zeta_\tau(p)$, we can write $\zeta_\infty(p), \gamma_\infty(p)$ around $p = 1$ in the form: $\zeta_\infty(p) = \frac{1}{2}\zeta_\infty^{(2)}(p-1)^2 + \frac{1}{3!}\zeta_\infty^{(3)}(p-1)^3 + \ldots$ and $\gamma_\infty(p) = \gamma_\infty^{(0)} + \gamma_\infty^{(1)}(p-1) + \ldots$.

Up to terms of order $(p-1)^2$ and higher in the series for $\gamma_\infty(p)$ we can rewrite:

$$\begin{aligned}
&\zeta_\tau(p) = \\
&= \zeta_\infty(p) + \frac{\gamma_\infty^{(0)}}{\tau} + \frac{\gamma_\infty^{(1)}}{\tau}(p-1) + O\left(\frac{(p-1)^2}{\tau}\right) + o\left(\frac{1}{\tau}\right) = \\
&= \zeta_\infty\left(p + \frac{\gamma_\infty^{(1)}}{\tau\zeta_\infty^{(2)}}\right) + \frac{\gamma_\infty^{(0)}}{\tau} + O\left(\frac{(p-1)^2}{\tau}\right) + o\left(\frac{1}{\tau}\right) .
\end{aligned} \tag{25}$$

Thus, the finite time corrections to $\zeta_\infty(p)$ around its maximum begin with a shift of the maximum at

$$p_0 = 1 - \frac{\gamma_\infty^{(1)}}{\tau\zeta_\infty^{(2)}} + o\left(\frac{1}{\tau}\right) \tag{26}$$

To apply the latter result we need to compute $\gamma_\infty^{(1)}$ in terms of observable quantities. And, in order to compute $\gamma_\infty^{(1)}$ we apply Eq. 24. First of all, we note that λ_p is determined by the condition

$$p\sigma_+ = z'_\tau(\lambda_p) = \sigma_+ + z''_\tau(0)\lambda_p + O(\lambda_p^2) \tag{27}$$

where we used Eq. 21 and Eq. 22. Then, $\lambda_p = \frac{\sigma_+(p-1)}{z''_\tau(0)} + O\big((p-1)^2\big)$. Substituting this result into Eq. 24 and equating the terms of order $O(\frac{p-1}{\tau})$ at both sides we find:

$$\gamma_\infty^{(1)} = -\frac{1}{2}\frac{z_\infty^{(3)}\sigma_+}{(z_\infty^{(2)})^2} . \tag{28}$$

The last equation can also be rewritten as:

$$\gamma_\infty^{(1)} = \frac{\zeta_\infty^{(3)}}{2\zeta_\infty^{(2)}} \tag{29}$$

This can be proven recalling that $\zeta_\infty^{(2)}$ and $\zeta_\infty^{(3)}$ are derivatives of $\zeta_\infty(p)$ in $p = 1$, that can be obtained by differentiating w.r.t. λ (two or three times, respectively) the definition $\zeta_\infty\big(z'_\infty(\lambda)/\sigma_+\big) = -z_\infty(\lambda) + \lambda z'_\infty(\lambda)$ and computing the derivatives in $\lambda = 0$ recalling that $z'_\infty(0)/\sigma_+ = 1$. Plugging Eq. 29 into Eq. 26 we finally get

$$p_0 = 1 - \frac{\zeta_\infty^{(3)}}{2\tau(\zeta_\infty^{(2)})^2} + o\left(\frac{1}{\tau}\right) \tag{30}$$

that is the main result of this section. The key point is that the moments $\zeta_\infty^{(2)}$ and $\zeta_\infty^{(3)}$ in Eq. 30 are quantities that can be measured from our empirical data (within an $O(\tau^{-1})$ error). We then have a verifiable prediction on the expected shift of the maximum at finite τ. Our data agree very well with this prediction, see Fig. 2 and corresponding discussion in sec. IV below.

Substituting Eq. 30 in Eq. 25, we finally find:

$$\zeta_\infty(p) = \eta_\tau(p) + O\Big(\frac{(p-1)^2}{\tau}\Big) + o(\tau^{-1}) , \tag{31}$$

where $\eta_\tau(p)$ is defined as

$$\eta_\tau(p) \stackrel{def}{=} -\frac{\gamma_\infty^{(0)}}{\tau} + \zeta_\tau\left(p - \frac{\zeta_\infty^{(3)}}{2\tau(\zeta_\infty^{(2)})^2}\right) . \tag{32}$$

The key point of the above discussion was the validity of Eq. 23–24; see Appendix A for their derivation.

D. Remarks

(1) The shift away from 1 of the maximum of the function $\zeta_\tau(p)$ at finite τ, expressed by the second term in Eq. 32, is due to the asymmetry of the distribution $\pi_\tau(p)$ around the average value $p = 1$; consequently, it is proportional, at leading order in τ^{-1}, to $\zeta_\infty^{(3)}$ which is indeed a measure of the asymmetry of $\zeta_\infty(p)$ around $p = 1$. This shift would be absent in the case of a symmetric distribution (*e.g.*, a Gaussian) and for this reason it was not observed in previous experiments [6, 10, 11, 15].

(2) The error term in the r.h.s. of Eq. 23 is $o(1)$ w.r.t. τ and it does not affect the computation of $\gamma_\infty(p)$. It is then clear that with a calculation similar to that we performed, one can get equations for the coefficients $O(\lambda^k)$ in the exponents of Eq. 23; in this way one can iteratively construct the whole sequence of coefficients $\gamma_\infty^{(k)}$ defining the power series expansion of $\gamma_\infty(p)$.

(3) In models with continuous time evolution the quantity σ_+ is not dimensionless but it has dimensions of inverse time: in such cases one can imagine that one is still studying a map which maps a system configuration at a time when some prefixed event happens in the system (typically a "collision") into the next one in which a similar event takes place. If τ_0 is the average time interval between such events then $\tau_0\sigma_+$ will play the role

played by σ_+ in the discrete time case: it will be the adimensional parameter entering the estimates of the error terms.

Note that the coefficients $g_\infty^{(k)}$ are of order σ_+^k, and their size is necessarily estimated by (the adimensional) entropy production to the k–th power. Then, in the continuous time case, the choice of τ_0 affects the estimates of the remainders, because it affects the size of the adimensional parameter $\tau_0\sigma_+$; and the size of the mixing time (that is connected with the estimated range of decay of the potentials, see [21]). The natural (and physical) choice for τ_0 is the mixing time. Consistently with this remark, at the moment of constructing numerically the distribution function for the entropy production rate averaged over a time τ, we will always consider time intervals of the form $\tau = \tau_0 n$, $n \geq 1$, see section III C below.

III. THE MODEL

We consider a system of N classical particles of equal mass m in dimension d; they are described by their position q_i and momenta $p_i = m\dot{q}_i$, $(p_i, q_i) \in R^{2d}$, $i = 1, \dots, N$. The particles are confined in a cubic box of side L with periodic boundary conditions. Each particle is subject to a *conservative force*, $f_i(\underline{q}) = -\partial_{q_i} V(\underline{q})$, and to a *nonconservative force* E_i that does not depend on the phase space variables. The force E_i is locally conservative but not globally such due to periodic boundary conditions. The *mechanical thermostat* is a Gaussian thermostat [9], $\theta_i(\underline{p}, \underline{q}) = -\alpha(\underline{p}, \underline{q})\, p_i$, and the function $\alpha(\underline{p}, \underline{q})$ is defined by the condition that the total kinetic energy $K(\underline{p}) \equiv \frac{1}{2m}|\underline{p}|^2 = \frac{1}{2m}\sum_i p_i^2$ should be a constant (*isokinetic ensemble*). The equations of motion are:

$$\begin{cases} \dot{q}_i = \frac{p_i}{m} \ , \\ \dot{p}_i = f_i(\underline{q}) + E_i - \alpha(\underline{p}, \underline{q})\ p_i \ , \end{cases} \tag{33}$$

From the constraint $\frac{dK}{dt} = 0$ one obtains

$$\alpha(\underline{p}, \underline{q}) = \frac{\sum_i E_i\ p_i + \sum_i f_i(\underline{q})\ p_i}{\sum_i p_i^2} \ . \tag{34}$$

A. Entropy production rate

The total phase space volume contraction rate for this system is given by:

$$\begin{aligned} \sigma(\underline{p}, \underline{q}) &= -\sum_i \left(\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \right) \\ &= dN\alpha(\underline{p}, \underline{q}) + \sum_i \frac{\partial \alpha}{\partial p_i} p_i = (dN-1)\ \alpha(\underline{p}, \underline{q}) \ . \end{aligned} \tag{35}$$

Defining the *kinetic temperature*, $T \equiv 2K(\underline{p})/(dN-1)$, [9], the phase space contraction rate can be rewritten as

$$\sigma(\underline{p}, \underline{q}) = \frac{\sum_i E_i\ \dot{q}_i - \dot{V}}{T} \ . \tag{36}$$

The first term is the power dissipated by the external force divided by the kinetic temperature, and can be identified with the entropy production rate [6, 7, 9]. The second term is the total derivative w.r.t. time of the potential energy divided by the temperature: this term does not affect the validity of the Fluctuation Relation in the asymptotic limit $\tau \to \infty$, as total derivatives give a contribution $O(\tau^{-1})$ in $p(x)$ [7, 23], hence they do not contribute to ζ_∞; however it has effect on the distribution of fluctuations over a finite time τ and its influence on the numerical computations has been recently discussed in detail [15]. The most convenient thing to do, in order to have a finite time distribution that approximates in the best possible way the asymptotic distribution of fluctuations, is to study the distribution of fluctuations for the *entropy production rate* $\dot{s}$, where $\dot{s}$ is identified with σ *minus* the total derivative term $-\dot{V}/T$ in Eq. 36:

$$\dot{s}(\underline{p}, \underline{q}) = \frac{\sum_i E_i\ \dot{q}_i}{T} \tag{37}$$

From now on we will call $\zeta_\infty(p)$ and $\zeta_\tau(p)$ the distributions for the fluctuations of the entropy production rate $\dot{s}$ averaged over infinite or finite time, respectively. These will be the objects we will measure and use from now on.

In order to define the *current* $J(x, E)$, let us rewrite $E_i = E\, u_i$, where u_i is a (constant) unit vector that specifies the direction of the force acting on the i-th particle. Then, according to Eq. 7,

$$J(\underline{p}, \underline{q}) = \frac{\partial \sigma}{\partial E} = \frac{\sum_i u_i\ \dot{q}_i}{T} \tag{38}$$

B. Discretization of the equations of motion

To perform the numerical simulation, one has to write the equations of motion in a discrete form. One possibility is to use the *Verlet algorithm* [24]; for Hamiltonian equations of motion (*i.e.*, $\underline{E} = \underline{0}$ and $\alpha = 0$)

$$\begin{cases} \dot{q}_i = \frac{p_i}{m} \ , \\ \dot{p}_i = f_i(\underline{q}) \ , \end{cases} \tag{39}$$

the Verlet discretization has the form

$$\begin{cases} q_i(t+dt) = q_i(t) + \frac{p_i(t)}{m} dt + \frac{1}{2} f_i(t) dt^2 \ , \\ p_i(t+dt) = p_i(t) + \frac{1}{2}\big[f_i(t) + f_i(t+dt) \big] dt \ , \end{cases} \tag{40}$$

where dt is the *time step size.* This discretization ensures that the error is $O(dt^4)$ on the positions $q_i(t)$ in a single time step. The implementation of this algorithm on a computer is discussed in detail in [24].

However, this method requires the forces $f_i(t)$ to depend only on the positions and not on the velocities: hence, it has to be adapted to Eq.s 33. This has been done in the following way. We write the discretized equations as

$$\begin{cases} q_i(t+dt) & = q_i(t) + \frac{p_i(t)}{m} dt \\ & + \frac{1}{2}\big[f_i(t) + E_i - \alpha(t) p_i(t)\big] dt^2 \ , \\ p_i(t+dt) & = p_i(t) + E_i + \frac{1}{2}\big[f_i(t) + f_i(t+dt) \\ & -\alpha(t)p_i(t) - \alpha(t+dt)p_i(t+dt)\big] dt \ , \end{cases} \tag{41}$$

with the same error as in the standard Verlet discretization. We store in the computer, at time t, the positions $q_i(t)$, the momenta $p_i(t)$, the forces $f_i(t)$, and the Gaussian multiplier $\alpha(t)$. Then, we perform the following operations:

1. we calculate the new positions $q_i(t+dt)$ using the first equation;

2. using the new positions we calculate the new forces $f_i(t+dt)$ (the conservative forces depend only on the positions);

3. we calculate the quantity $\xi_i = p_i(t) + E_i + \frac{1}{2}\big[f_i(t) + f_i(t+dt) - \alpha(t)p_i(t)\big] dt$ and we observe that $p_i(t+dt)$ can be expressed in terms of the (known) ξ_i and the (unknown) $\alpha(t+dt)$ as
$$p_i(t+dt) = \frac{\xi_i}{1-\alpha(t+dt)dt/2} \ ; \tag{42}$$

4. substituting Eq. 42 in the definition of $\alpha(t+dt)$, Eq. 34, we get a self-consistency equation for $\alpha(t+dt)$, whose solution is
$$\begin{aligned} \alpha(t+dt) &= \frac{\alpha_0}{1-\alpha_0 dt/2} \ , \\ \alpha_0 &= \frac{\sum_i E_i \ \xi_i + \sum_i f_i(t+dt) \ \xi_i}{\sum_i \xi_i^2} \ ; \end{aligned} \tag{43}$$

5. substituting Eq. 43 in Eq. 42 we calculate $p_i(t+dt)$.

This procedure allows us to calculate the new positions, momenta, forces, and α, at time $t+dt$ according to Eq.s 41 *without approximations*, defining a map S such that $(\underline{p}(t+dt), \underline{q}(t+dt)) = S(\underline{p}(t), \underline{q}(t))$.

Our (*discrete*) dynamical system will be defined by the map $S(\underline{p}, \underline{q})$ and will approximate the differential equations of motion, Eq. 33, with error $O(dt^4)$ for the positions and $O(dt^3)$ for the velocities.
The map S satisfies the following properties:

1. it is *reversible*, *i.e.* it exists a map $I(\underline{p}, \underline{q})$ (simply defined by $I(\underline{p}, \underline{q}) = (-\underline{p}, \underline{q})$) such that $IS = S^{-1}I$;

2. in the *Hamiltonian* case ($\underline{E} = \underline{0}$ and $\alpha = 0$, Eq.s 39) it is *volume preserving.*

The first property ensures that *assuming the Chaotic Hypothesis* the Fluctuation Relation holds for the map S. The second property ensures that at equilibrium the discretization algorithm conserves the phase space volume.

C. Details of the simulation

In the simulation, we chose the external force of the form $E_i = E\, u_i$, where the unit vectors u_i were parallel to the x direction but with different orientation: half of them were oriented in the positive direction, and half in the negative direction, *i.e.* $u_i = (-1)^i \widehat{x}$, in order to keep the center of mass fixed. We considered two different systems, selecting interaction potentials widely used in numerical simulations (for the purpose of making easier possible future independent checks and rederivations of our results):

1. (*model I*) the first investigated system is made by $N = 8$ particles of equal mass m in $d = 2$. The interaction potential is a sum of pair interactions, $V(\underline{q}) = \sum_{i<j} v(|q_i - q_j|)$, and the pair interaction is represented by a WCA potential, *i.e.* a Lennard-Jones potential truncated at the minimum:
$$v(r) = \begin{cases} 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6\right] + \epsilon \ , & r \le \sqrt[6]{2}\sigma \ ; \\ 0 \ , & r > \sqrt[6]{2}\sigma \ . \end{cases}$$
The reduced density was $\rho = N\sigma^2/L^2 = 0.95$ (that determines L), the kinetic temperature was fixed to $T = 4\epsilon$ and the *time step* to $dt = 0.001 t_0$, where $t_0 = \sqrt{m\sigma^2/\epsilon}$. In the following, all the quantities will be reported in units of m, ϵ and σ (*LJ units*). This system was already studied in the literature, see *e.g.* [6, 25]. We investigated different values of the external force E ranging from $E = 0$ to $E = 25$.

2. (*model II*) the second system is a binary mixture of N=20 particles (16 of type A and 4 of type B), of equal mass m, in $d = 3$, interacting via the same WCA potential of model I; the pair potential is
$$v_{\alpha\beta}(r) = \begin{cases} 4\epsilon_{\alpha\beta} \left[\left(\frac{\sigma_{\alpha\beta}}{r}\right)^{12} - \left(\frac{\sigma_{\alpha\beta}}{r}\right)^6\right] + \epsilon_{\alpha\beta} \ , \\ \qquad\qquad\qquad\qquad r \le \sqrt[6]{2}\sigma_{\alpha\beta} \ ; \\ 0 \ , \qquad\qquad\qquad\quad r > \sqrt[6]{2}\sigma_{\alpha\beta} \ ; \end{cases}$$
α and β are indexes that specify the particle species ($\alpha, \beta \in [A, B]$). The parameters entering the potential are the following: $\sigma_{AB} = 0.8\sigma_{AA}$; $\sigma_{BB} = 0.88\sigma_{AA}$; $\epsilon_{AB} = 1.5\epsilon_{AA}$; $\epsilon_{BB} = 0.5\epsilon_{AA}$. Similar potentials have been studied, [26, 27], as models for liquids in the supercooled regime (*i.e.*, below the melting temperature). For this system the *LJ units* are m, ϵ_{AA}, and σ_{AA}; the unit of time is then $t_0 = \sqrt{m\sigma_{AA}^2/\epsilon_{AA}}$. The reduced density was $\rho = N\sigma_{AA}^3/L^3 = 1.2$ and the integration step was $dt = 0.001 t_0$. The unit vectors u_i are chosen such that half of the A particles and half of the B particles have positive force in the x direction, and the remaining particles have negative force in the x direction. For this system we investigated different values of external force $E \in [0, 10]$ and temperature $T \in [0.5, 3]$.

For each system and for each chosen value of T and E, we simulated a very long trajectory ($\sim 2 \cdot 10^9 dt$) starting from a random initial data; we recall that in both systems we chose $dt = 0.001 t_0$, t_0 being the natural unit time introduced in items (1) and (2) above. After a short transient ($\sim 10^3 dt$), still much bigger than the decay time τ_0 of self-correlations (that appears to be $\tau_0 = 10^2 dt$), the system reached stationarity, in the sense that the instantaneous values of observables (e.g. potential energy, Lyapunov exponents) agree with the corresponding asymptotic values within the statistical error of the asymptotic values themselves. After this transient we started recording values p_i, $i = 1, \dots, \mathcal{N}$, of the variable $p(x)$ (defined in Eq. 6), integrating the entropy production rate (Eq. 37) on adjacent segments of trajectory of length $\tau_0 = 100dt = 0.1t_0$. Note that the length of the time interval over which we averaged the entropy production rate was chosen as equal to the mixing time, consistently with the discussion in Remark (4) of sec. II D.

In conclusion, from each simulation run at fixed T and E we obtain $\mathcal{N} \sim 10^7$ values p_i of $p(x)$ which are the starting point of our data analysis. The value of σ_+ is estimated by averaging the entropy production rate over the whole trajectory.

From a shorter simulation run we measured also the Lyapunov exponents of the map S using the standard algorithm of Benettin *et al.* [25, 28].

D. Remarks

To conclude this section, we note that the WCA potential has a discontinuity in the second derivative. Thus, one should be concerned with the possibility that the error of our discretization is not $O(dt^4)$ over the q_i's on a single time step, as it should be for potentials $V \in C^4$. To check that this is not the case (or that at least this does not affect our results) we made two independent tests:

1. we simulated a system similar to model I but with a potential $V \in C^4$ and we obtained qualitatively the same results;

2. we simulated model I using an *adaptive step size* algorithm [24]; this kind of algorithms adapt the step size dt during the simulation in order to keep constant the difference between a single step of size dt and two steps of size $dt/2$. If the precision of our discretization changed at the singular points of the potential, the time step should change abruptly during the simulation, while we observed a practically constant time step during the simulation.

Hence, we have evidence of the fact that the (isolated) singularities of the potentials do not produce relevant effects on our observations; this is probably due to the fact that the set of singular points of the total potential energy $V(\underline{q})$ has zero measure w.r.t. the SRB measure.

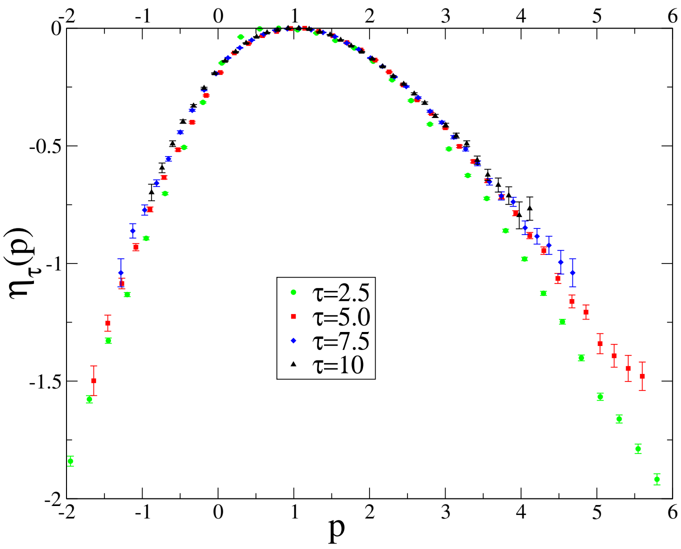

FIG. 1: Model I at $E = 5$: the function $\eta_\tau(p) = \zeta_\infty(p) + O((p-1)^2/\tau)$ for different values of τ.

IV. DATA ANALYSIS

In this section we will discuss in detail the procedure we followed to analyze the numerical data. As an example, we will discuss the data obtained from the simulation of model I at $E = 5$. As discussed in the previous section, from the simulation run we obtain a set $\mathcal{P}_0 = \{p_i\}_{i=1\dots\mathcal{N}}$ of values of the variable $p(x)$ that correspond to $\tau = \tau_0$ and are measured on adjacent segments of trajectory. We recall again that $\tau_0 = 0.1 = 100dt$ is of the order of the *mixing time*, *i.e.* the time scale over which the correlation functions (*e.g.* of density fluctuations) decay to zero.

Probability distribution function - From the dataset $\mathcal{P}_0$ we construct the histograms $\pi_\tau(p)$ for different values of $\tau = n\tau_0$ as follows: the values of $p(x)$ for $\tau = n\tau_0$ are obtained by averaging n subsequent entries of the dataset $\mathcal{P}_0$; we obtain a new dataset $\mathcal{P}_n = \{p_j^{(n)}\}_{j=1\dots\mathcal{N}/n}$ such that $p_j^{(n)} = n^{-1} \sum_{i=nj+1}^{n(j+1)} p_i$. Finally, from the dataset $\mathcal{P}_n$ the histogram of $\pi_\tau(p)$ is constructed for $\tau = n\tau_0$; the errors are estimated as the square roots of the number of counts in each bin. The function $\zeta_\tau(p)$ is then defined as $\zeta_\tau(p) = \tau^{-1} \log \pi_\tau(p)$.

Shifting of the maximum - By fitting the function $\zeta_\tau(p)$ in $p \in [-1, 3]$ with a sixth-order polynomial we determine the position of the maximum $\widetilde{p}_\tau$ within an error that, since δp is the length of a bin, we estimate to be $\delta p/2$. Then, we construct the function $\eta_\tau(p) = \zeta_\tau(p - 1 + \widetilde{p}_\tau)$ (see Eq. 32) which is expected to approximate the limiting function $\zeta_\infty(p)$ with error $O((p-1)^2/\tau)$. The functions $\eta_\tau(p)$ are reported in Fig. 1 for different values of τ. We observe a very good convergence for $\tau \gtrsim 5.0 = 50\tau_0$.

By a fourth-order fit of the so-obtained limiting function $\zeta_\infty(p)$ around $p = 1$ we extract the coefficients

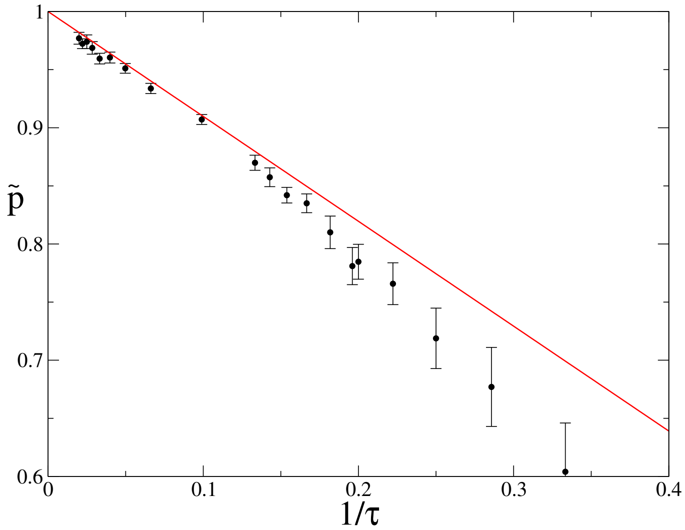

FIG. 2: Model I at $E = 5$: the maximum $\widetilde{p}_\tau$ of $\zeta_\tau(p)$ as a function of $1/\tau$. The full line is the prediction of Eq. 30, $\widetilde{p} = 1 - \zeta_\infty^{(3)}/[2\tau(\zeta_\infty^{(2)})^2]$.

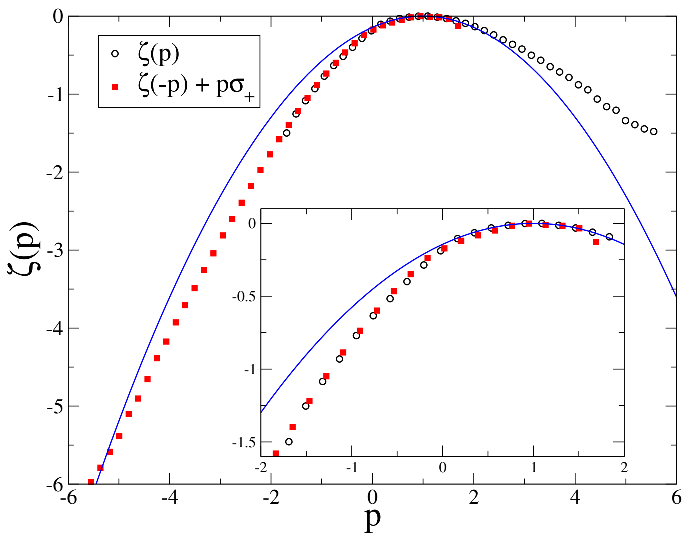

FIG. 3: Model I at $E = 5$: the estimate of the function $\zeta_\infty(p)$ (open circles). In the same plot $\zeta_\infty(-p) + p\sigma_+$ (filled squares) is reported. In the inset, the interval $p \in [-2, 2]$ where the data overlap is magnified. The full line is the Gaussian approximation, $\frac{1}{2}\zeta_\infty^{(2)}(p-1)^2$. The plot shows that the Gaussian is not a good approximation in the interval $[-2, 2]$. The validity of the Fluctuation Relation in the same interval is shown by the overlap of the open circles and filled squares.

$\zeta_\infty^{(2)} = -0.287$ and $\zeta_\infty^{(3)} = 0.149$ in order to test the correctness of Eq. 30. In Fig. 2 we report $\widetilde{p}_\tau$. The full line is the prediction of Eq. 30, that is indeed verified for $\tau \gtrsim 10$. This result confirms the analysis of section II.

Graphical verification of the fluctuation relation - From the previous analysis we can conclude that the function $\eta_\tau(p)$ for $\tau = 5.0$ provides a good estimate of the function $\zeta_\infty(p)$ for $p \in [-2, 4]$ (see Fig. 1); thus, we can use this function to test the fluctuation relation, Eq. 4, in this range of p. In Fig. 3 we report the estimated functions $\zeta_\infty(p)$ and $\zeta_\infty(-p) + p\sigma_+$. An excellent agreement between the two functions is observed in the interval $p \in [-2, 2]$ where our data allows the computation of both $\zeta_\infty(p)$ and $\zeta_\infty(-p)$. Note that in this range of p the function $\zeta_\infty(p)$ is not Gaussian, see the inset of Fig. 3.

Quantitative verification of the fluctuation relation - The translation of the function $\zeta_\tau(p)$ is crucial to obtain a correct estimate of the limit $\zeta_\infty(p)$ and to verify the fluctuation relation. In this section we will try to quantify this observation; as the discussion will be very technical, the reader who is satisfied with Fig. 3 should skip to next section.

The histogram $\pi_{n\tau_0}(p)$ derived from the dataset $\mathcal{P}_n$ is constructed assigning the number of counts π_α in the α-th bin to the middle of the binning interval, that we call p_α (the latter will be an *increasing* function of α). The statistical error $\delta\pi_\alpha$ on the number of counts is $\sqrt{\pi_\alpha}$. Our histograms are constructed in such a way that if p_α is the center of a bin, also $-p_\alpha$ is the center of a bin; we call $\overline{\alpha}$ the bin such that $p_{\overline{\alpha}} = -p_\alpha$. There exists a value p_m such that for $p_\alpha < p_m$ the number of counts in the bin α is smaller than m (we choose $m = 4$). Let us indicate by p_{α_m} the smallest value of $p_\alpha > p_m$. Hence, the histogram is characterized by:

1. a *bin size* δp;

2. the bin α_m corresponding to the minimum value of p_α such that the number of counts in the bin is at least m;

3. the total number M of bins such that $\alpha \in [\alpha_m, \overline{\alpha}_m]$; for these values of p_α, both $\pi_\tau(p)$ and $\pi_\tau(-p)$ can be computed and they can be used to verify the fluctuation relation.

The function $\zeta_\tau(p)$, derived from the histogram, is specified by a set of values $(p_\alpha, \zeta_\alpha, \delta\zeta_\alpha)$ for each bin α, where $\zeta_\alpha = \tau^{-1}\log\pi_\alpha$ and the error $\delta\zeta_\alpha$ has been defined by

$$\delta\zeta_\alpha = \frac{1}{\tau}\frac{\delta\pi_\alpha}{\pi_\alpha} = \frac{1}{\tau\sqrt{\pi_\alpha}} \,. \tag{44}$$

A quantitative verification of Eq. 4 is possible defining the following χ^2 function:

$$\chi^2 \equiv \frac{1}{M}\sum_{\alpha=\alpha_m}^{\overline{\alpha}_m} \frac{(\zeta_\alpha - \zeta_{\overline{\alpha}} - p_\alpha\sigma_+)^2}{(\delta\zeta_\alpha)^2 + (\delta\zeta_{\overline{\alpha}})^2} \tag{45}$$

The value of χ is the average difference between $\zeta_\tau(p)$ and $\zeta_\tau(-p) + p\sigma_+$ in units of the statistical error. Translating p of a quantity $a\delta p/2$, $a \in \mathbb{Z}$, corresponds to shifting the histogram, *i.e.* to consider a new histogram $(p_\alpha + a\delta p/2, \zeta_\alpha, \delta\zeta_\alpha)$. This preserves the property that if p_α is the center of a bin, also $-p_\alpha$ is the center of a bin; we call $\overline{\alpha}(a)$ the new value of α such that $p_{\overline{\alpha}(a)} + a\delta p/2 =$

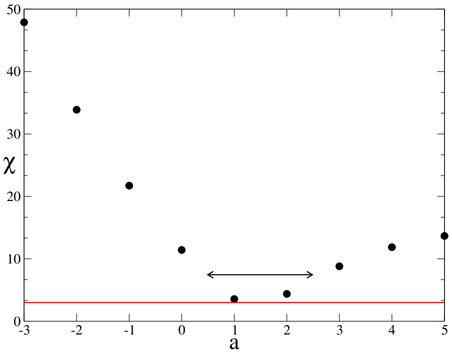

FIG. 4: Model I at $E = 5$: the function $\chi(a)$. The full line corresponds to $\chi = 3$. The arrow indicates the interval $\delta_0 \pm \delta p/2$ (note that its length is 2 in units of a) into which the minimum of χ can be located within the accuracy of the histogram.

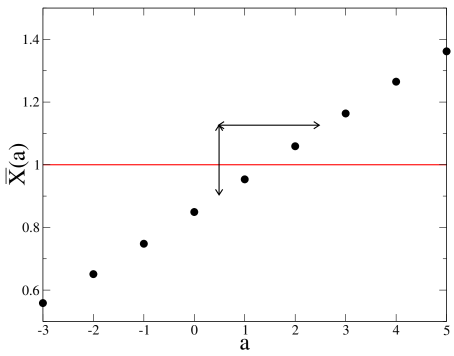

FIG. 5: Model I at $E = 5$: the function $\overline{X}(a)$. The horizontal arrow marks the interval where the minimum of χ is located, see Fig. 4. The vertical arrow indicates the error δX on the value $X = 1$ which is estimated as $\delta X = 2(\overline{X}(2) - \overline{X}(1))$. The optimal slope of the fluctuation relation without the translation would have been $\overline{X}(a = 0) \sim 0.85$.

$-(p_\alpha + a\delta p/2)$. Also, the number M_a of bins such that $\alpha(a) \in [\alpha_m, \overline{\alpha}_m(a)]$ depends on a. We define

$$\chi^2(a) \equiv \frac{1}{M_a} \sum_{\alpha=\alpha_m}^{\overline{\alpha}_m(a)} \frac{\left(\zeta_\alpha - \zeta_{\overline{\alpha}(a)} - (p_\alpha + a\delta p/2)\sigma_+\right)^2}{(\delta\zeta_\alpha)^2 + (\delta\zeta_{\overline{\alpha}(a)})^2} \tag{46}$$

We shall use the criterion that the fluctuation relation is satisfied if $\chi \leq 3$, which means that $\zeta_\infty(p)$ and $\zeta_\infty(-p) + p\sigma_+$ differ, *on average*, by less than 3 times the statistical error $\sqrt{\left(\delta\zeta(p)\right)^2 + \left(\delta\zeta(-p)\right)^2}$. The function $\chi(a)$ for the case of model I at $E = 5$ is reported in Fig. 4.

The minimum of χ is assumed between $a^* = 1$ and $a^* + 1 = 2$ and an upper limit for the value of χ at the minimum is $\chi(1) = 3.5$. We estimate the translation that minimizes χ as $\delta_0 = (a^* + 0.5)\delta p/2 = 1.5 \cdot 0.093 = 0.140$, and to this estimate we attribute an error $\pm\delta p/2$, where $\delta p = 0.186$ is the size of a bin. On the other hand, we have seen above that, in order to shift the maximum of $\zeta_\tau(p)$ in $p = 1$, one has to translate p by a quantity $\delta \equiv 1 - \widetilde{p} = 0.215$. The consistency of our analysis requires that δ and δ_0 coincide within their errors, i.e. that the intervals $\delta \pm \delta p/2$ and $\delta_0 \pm \delta p/2$ overlap, or in other words $|\delta - \delta_0| < \delta p$. In the present case $0.075 = |\delta - \delta_0| < \delta p = 0.186$, then δ and δ_0 coincide within the errors. This means that the translation of p brings the maximum of $\zeta_\tau(p)$ in $p = 1$ and, *at the same time*, minimizes the difference between $\eta_\tau(p)$ and $\eta_\tau(-p) + p\sigma_+$, where η_τ is our finite time estimate of $\zeta_\infty(p)$. The value $\chi(a^*)$ quantifies this difference and is a first estimate of the precision of our analysis.

Another estimate of the precision of our analysis can be obtained as follows. We define a parameter X as the slope of $\zeta_\infty(p) - \zeta_\infty(-p)$ as a function of $p\sigma_+$:

$$\zeta_\infty(p) = \zeta_\infty(-p) + Xp\sigma_+ \tag{47}$$

The fluctuation theorem predicts $X = 1$, but other values of X are possible under different hypothesis, see [7, 10, 29, 30]. We define a function $\chi^2(a, X)$ as

$$\chi^2(a, X) \equiv \frac{1}{M_a} \sum_{\alpha=\alpha_m}^{\overline{\alpha}_m(a)} \frac{\left(\zeta_\alpha - \zeta_{\overline{\alpha}(a)} - X(p_\alpha + a\delta p/2)\sigma_+\right)^2}{(\delta\zeta_\alpha)^2 + (\delta\zeta_{\overline{\alpha}(a)})^2} \tag{48}$$

and for each value of a we calculate the optimal value of X, $\overline{X}(a)$, by minimizing $\chi^2(a, X)$. The function $\overline{X}(a)$ is reported in Fig. 5. As the shift of the maximum δ is between $a = 1$ and $a = 2$, we see that the slope X is compatible with one. Moreover, as the natural error on p is the size of a bin δp, we assign to the value $X = 1$ a statistical error $\delta X = 2(\overline{X}(2) - \overline{X}(1)) = 0.22$. Note again that without the translation of p the optimal slope would be $X \sim 0.85$, incompatible with Eq. 4.

Discussion - From the present analysis, we can conclude that:

1. the translation shifting the maximum of $\zeta_\tau(p)$ to $p = 1$ at the same time minimizes the difference between $\eta_\tau(p)$ and $\eta_\tau(-p) + p\sigma_+$, where η_τ is our finite time estimate of ζ_∞; this proves the consistency of our theory of finite time corrections;

2. without the translation of p (that corresponds to $a = 0$), the function $\zeta_\tau(p)$ for $\tau \sim 5.0$ *do not satisfy the fluctuation relation*, as $\chi(a = 0) = 11$ and $\overline{X}(a = 0) = 0.85$;

3. the function $\eta_\tau(p) = \zeta_\tau(p - \delta)$ satisfies the fluctuation relation with $\chi \sim 3$ and an error of about 20% on the slope X: both quantities measure the accuracy of our analysis.

Thus, the check of the fluctuation relation relies crucially on the translation of the function $\zeta_\tau(p)$ that has been discussed in section II. By considering larger values of τ one could avoid this problem (as $\delta \sim \tau^{-1}$); however, as one can see from Fig. 1, for $\tau > 5.0$ the negative tails of $\zeta_\tau(p)$

are not accessible to our computational resources. The computation of the finite time corrections is mandatory if one aims to test the fluctuation relation at high values of the external driving force.

Summary of the data analysis - To conclude, we summarize the procedure we follow to analyze the data of a given simulation run:

1. we determine a value of τ such that $\zeta_\tau(p)$ appear to be close to the asymptotic limit $\zeta_\infty(p)$;

2. we determine the maximum $\widetilde{p}$ of $\zeta_\tau(p)$ by a sixth-order polynomial fit around $p = 1$, in an interval as big as possible compatibly with the request that the χ^2 from the fit is less than ~ 10;

3. we shift the histogram of an integer multiple a of the half bin size $\delta p/2$ and compute the function $\chi(a)$ according to Eq. 46. We determine the value a^* such that the minimum of $\chi(a)$ is assumed in the interval $[a^*, a^*+1]$: the consistency of our analysis requires that $\delta = 1 - \widetilde{p}$ and $\delta_0 = (a^* + 0.5)\delta p/2$ coincide within their errors (i.e. $|\delta - \delta_0| < \delta p$);

4. The value $\chi^* = \min[\chi(a^*), \chi(a^*+1)]$ is an upper limit for the value of χ at the minimum. The number of bins $\min\{M_{a^*}, M_{a^*+1}\}$ involved in this estimate will be called M^*;

5. we compute the error $\delta X = 2(\overline{X}(a^*+1) - \overline{X}(a^*))$.

The relevant quantities τ, δ, δ_0, $|\delta - \delta_0|$, δp, M^*, χ^* and δX for model I are reported in table I for different values of the external force E.

V. NUMERICAL SIMULATION OF MODEL I

We will now discuss systematically the numerical data obtained from the simulation of model I (defined in section III) at different values of the driving force E. In Fig. 6 we report the *mobility* $\mu(E) = T\langle J\rangle_E/(NE)$, *i.e.* the l.h.s. of Eq. 11 times T/N, as a function of E. The current $J(\underline{p}, \underline{q})$ has been defined in Eq. 38. From the Green-Kubo relation, Eq. 8, we have [9]

$$\lim_{E\to 0} \mu(E) = \frac{D}{T} \ , \tag{49}$$

E	τ	σ_+	δ	δ_0	$\lvert\delta-\delta_0\rvert$	δp	M^*	χ^*	δX
2.5	5.0	0.194	0.272	0.183	0.089	0.244	43	2.2	0.24
5.0	5.0	0.810	0.215	0.139	0.076	0.187	20	3.5	0.22
7.5	4.0	1.945	0.197	0.116	0.081	0.116	18	2.8	0.18
10.0	2.5	4.044	0.262	0.151	0.111	0.122	17	4.4	0.20
12.5	2.5	7.090	0.257	0.137	0.120	0.111	8	3.5	0.28

TABLE I: Model I: results of the data analysis for some selected values of E. All the quantities are defined in section IV. For $E > 12.5$ the negative tails of the distribution are not accessible to our numerical simulation.

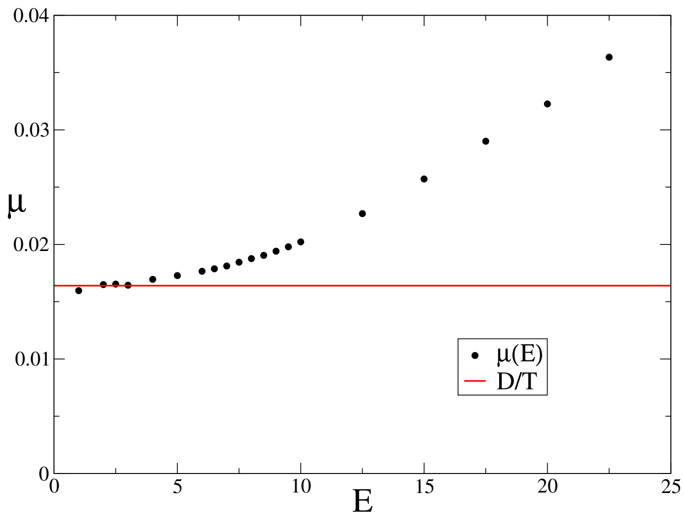

FIG. 6: Model I: mobility μ as a function of the driving force E. The full line is the equilibrium diffusion coefficient D divided by the temperature. Deviations from the linear response are observed around $E = 5$. The error bars are of the order of the dimension of the symbols. Studying $\mu(E)$ for values of E bigger than those shown in the figure, one can verify that the mobility increases up to a value μ_{max}, reached in correspondence of $E \sim 45$. For values of E bigger than $E \sim 45$, the mobility begins to decrease essentially following the limiting curve $TJ_T/(NE)$, where $J_T = \sqrt{T(d-1/N)}N/T$ is the maximum allowed value of the current (saturation value).

where D is the equilibrium diffusion coefficient,

$$D = \lim_{t\to\infty} \frac{1}{2Nd} \sum_i \langle |q_i(t) - q_i(0)|^2 \rangle_{E=0} \tag{50}$$

Deviations from the linear response are observed and $\mu(E) \sim D/T + O(E^2)$ above $E = 5$.

In table I we report the main parameters that result from the data analysis (as discussed in the previous section) for some selected values of E. The value $|\delta - \delta_0|$ is always less than δp, consistently with our discussion above, except for $E = 12.5$ where, however, the relative difference between the two quantities is small ($\sim 9\%$). It can be noted that δ is systematically bigger than δ_0. This could be due to the fact that the error terms $O((p-1)^2/\tau)$ or $o(1/\tau)$ that we are discarding likely produce a systematic shift in δ or in δ_0; or that the velocity of convergence of $\zeta_\tau(p)$ is not the same on the negative or on the positive side (because numerically is much more difficult to observe big negative fluctuations of σ than the positive ones – and the Fluctuation Relation provides a quantitative estimate of the relative probabilities). At the moment, because of the level of precision of our simulations, we are not able to investigate this problem in more detail, see also Remark (3) in section II D. On increasing the value of E, we are forced to decrease the value of τ we use for the analysis as, for longer τ, the negative tail of the distribution $\zeta_\tau(p)$ becomes unobservable. This can be seen as the number M^* of bins used for the compu-

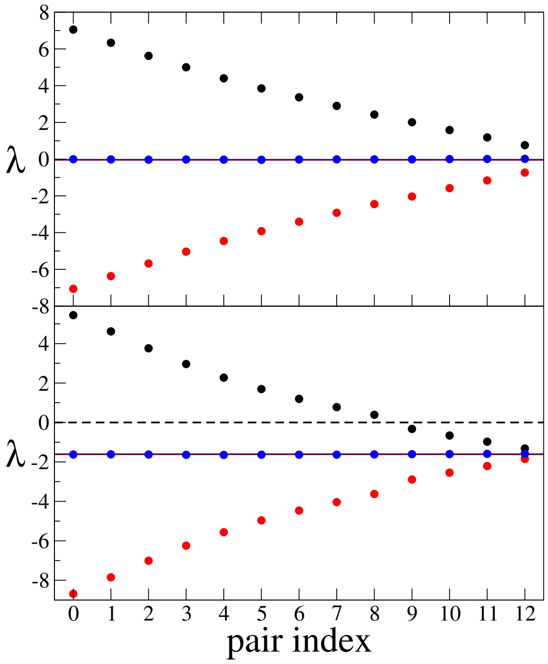

FIG. 7: Model I: Lyapunov exponents for $E = 5$ (top) and for $E = 25$ (bottom). For each panel, the upper and lower dots are the two paired exponents $\lambda_j^{(+)}$ and $\lambda_j^{(-)}$, and the middle dot is their average $(\lambda_j^{(+)} + \lambda_j^{(-)})/2$. The full line is $\sigma_+/2Nd$, the dashed line is at $\lambda = 0$.

tation of χ decrease on increasing E; above $E = 12.5$ it is impossible to find a value of τ such that $\zeta_\tau(p)$ is close to the asymptotic limit and the negative tail is observable. Thus, the fluctuation relation cannot be tested above $E = 12.5$ with our computational power. However, we are able to check the fluctuation relation in the region $E > 5$ where deviations from the linear response are observed. Moreover, the estimated distributions $\zeta_\infty(p)$ are very similar to the one reported in Fig. 3: in particular, they are not Gaussian in the investigated interval of p (also for $E < 5$, in the linear response regime).

Finally, in Fig. 7 we report the measured Lyapunov exponents of the model for $E = 5$ and $E = 25$. For this system, the Lyapunov exponents are known to be paired [25, 31, 32] like in Hamiltonian systems and the average of each pair is a constant equal to $\sigma_+/2Nd$. For $E = 5$, each pair is composed of a negative and a positive exponent. This means that the attractive set is dense in phase space [10, 30] and the chaotic hypothesis is expected to apply to the system yielding a slope $X = 1$ in the fluctuation relation, as confirmed by our numerical data. The same happens up to $E \sim 20$. Above $E = 20$, there is a number D of pairs composed by two negative exponents (for $E = 25$ we get $D = 4$, see Fig. 7). In this situation, the slope X in the fluctuation relation is expected to be given by $X = 1 - D/Nd$ [29, 30]. Thus, for $E = 25$ one expects $X \sim 0.75$. Unfortunately, as discussed above, above $E = 12.5$ we did not observe negative fluctuations of the entropy production, and this prediction could not be tested in our simulation.

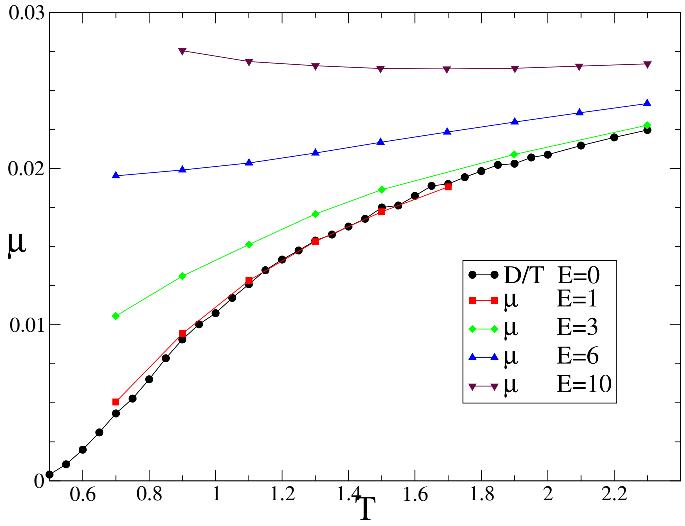

FIG. 8: Mobility as a function of the temperature T and of the driving force E for model II. The circles correspond to the equilibrium diffusion coefficient divided by the temperature. Deviations from the linear response are observed for $E \geq 3$; they become larger on lowering the temperature, as $D \to 0$.

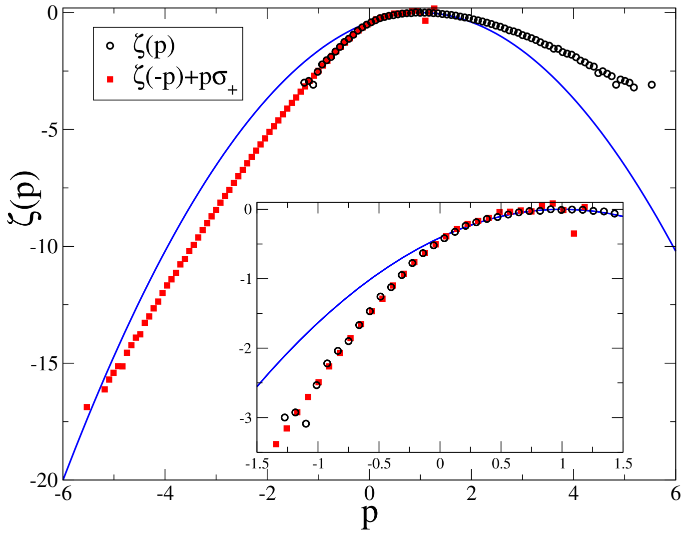

FIG. 9: The estimate of the function $\zeta_\infty(p)$ (open circles) for model II with $T = 1.1$ and $E = 3$. In the same plot $\zeta_\infty(-p) + p\sigma_+$ (filled squares) is reported. In the inset, the interval $p \in [-1.5, 1.5]$ where the data overlap is magnified. The full line is the Gaussian approximation, $\zeta_\infty(p) = \frac{1}{2}\zeta_\infty^{(2)}(p-1)^2$. The data have been obtained from the histogram of $\pi_\tau(p)$ with $\tau = 2.5$ (see table II).

VI. NUMERICAL SIMULATION OF MODEL II

Model II differs from model I in the dimension $d = 3$, in the larger number of particles $N = 20$, and because it is a binary mixture of two types of particles. Binary mixtures are frequently used as models for numerical simulations of supercooled liquids as they avoid crystallization also at very low temperature on the "physical" time scales (i.e. on the time scales of numerical experiments); for these systems, at low temperature deviations from the linear response are observed also for very low values of the external driving force.

In Fig. 8 we report the equilibrium diffusion coefficient D (divided by the temperature T) and the mobility (for different values of E) as functions of the temperature. Even though the number of particles is very small, on lowering the temperature the systems approaches the supercooled state and D becomes very small around $T \sim 0.5$. Slightly above this temperature, i.e. around $T = 1$, strong deviations from the linear response are observed for $E \geq 3$, where the entropy production σ_+ is still close to 0. Some values of σ_+ are reported in table II; to compare these values with those obtained for model I one should note that σ_+ is an *extensive* quantity. Thus, the entropy production *per degree of freedom*, $\sigma_+/2Nd$, is much smaller in model II than in model I.

In table II the results of the data analysis outlined in section IV are reported. For $E \leq 6$ we obtain a very good agreement of the data with the predictions of the fluctuation relation and with the theory of finite time corrections discussed in section II. For $E = 10$ it is very difficult to observe negative fluctuations of p with our computational power; see *e.g.* the result of the analysis for $E = 10$ and $T = 1.9$, where only $M^* = 7$ bins where available and we were forced to use $\tau = 0.2$, of the order of the mixing time τ_0. In Fig. 9 we report the estimated function $\zeta_\infty(p)$ obtained for $T = 1.1$ and $E = 3$ from the data with $\tau = 2.5$. Strong deviations from the Gaussian behavior are observed in the accessible range of p (see the inset of Fig. 9). A similar behavior of $\zeta_\infty(p)$ is observed in correspondence of all the values of E and T we investigated (those listed in Table II): in particular in all these cases highly non Gaussian behaviors are observed in the accessible range of p.

The Lyapunov spectrum for this system is very similar to the one reported in the upper panel of Fig 7. Pairs of two negative exponents were observed only for $E = 10$ at $T \leq 1.3$, where, as in the case of model I, σ_+ is too large to allow for a verification of the modified fluctuation relation expected in this case, see the discussion at the end of section V.

T	E	τ	σ_+	δ	δ_0	$\lvert\delta - \delta_0\rvert$	δp	M^*	χ^*	δX
0.9	1	3.0	0.209	0.453	0.334	0.119	0.223	68	1.9	0.19
0.9	3	3.0	2.615	0.286	0.264	0.024	0.132	15	1.0	0.23
1.1	1	4.0	0.233	0.231	0.126	0.105	0.126	79	1.7	0.24
1.1	3	2.5	2.493	0.217	0.238	0.021	0.087	30	1.0	0.12
1.1	6	1.5	13.32	0.113	0.230	0.117	0.092	7	1.1	0.21
1.5	1	3.0	0.230	0.179	0.140	0.039	0.140	86	0.9	0.13
1.5	3	2.5	2.227	0.145	0.123	0.022	0.082	33	4.7	0.18
1.5	6	0.5	52.14	0.074	0.130	0.056	0.052	11	0.6	0.10
1.7	1	3.0	0.221	0.127	0.141	0.014	0.283	49	1.0	0.26
1.9	3	2.5	1.981	0.106	0.122	0.016	0.122	26	0.8	0.12
1.9	6	0.4	43.52	0.078	0.126	0.048	0.085	14	1.7	0.11
1.9	10	0.2	139.0	0.079	0.135	0.056	0.039	7	0.8	0.10
2.1	6	0.4	40.48	0.074	0.110	0.036	0.110	11	1.0	0.15

TABLE II: Model II: results of the data analysis for some selected values of T and E. All the quantities are defined in section IV.

VII. CONCLUSIONS

We tested the fluctuation relation, in our opinion quite successfully, in a numerical simulation of two models of interacting particles subjected to an external nonconservative force and to a reversible mechanical thermostat. Our data satisfy the fluctuation relation with a $\chi \leq 3$ and an accuracy of the order of 20% also for very large values of the driving force, where strong deviations from the linear response are observed, and where the large deviation function is strongly non-Gaussian. The comparison of our numerical data with the predictions of the fluctuation relation is done by taking into account the (lowest order) finite time corrections to the distribution function for the fluctuations of the phase space contraction rate. This is crucial: if we did not take into account such corrections the fluctuation relation would be violated within the precision of our experiment.

In order to compute the finite time corrections, we proposed an algorithm which allows to reconstruct the asymptotic distribution function from measurable quantities at finite time, within a given precision. Our theory of the corrections relies on the symbolic representation of the chaotic dynamics, therefore it is applicable if one accepts the Chaotic Hypothesis.

Our interpretation of the numerical results is that the *chaotic hypothesis* can be applied to these systems, also very far from equilibrium, and in particular the fluctuation relation is satisfied even in regions where its predictions measurably differ from those of linear response theory.

A big open problem we are left with is trying to understand how the fluctuation relation is modified for values of the driving force so high that the attractive set is no longer dense in phase space. It is expected, [29], that in such a case $\zeta_\infty(p) - \zeta_\infty(-p)$ is still linear, but the slope is $X\sigma_+$, with X given by the ratio of the dimension of the attractive set and of that of the whole phase space. An estimate of such quantity can be given via the number of negative pairs of exponents in the Lyapunov spectrum [29, 30]. Unfortunately negative pairs begin to appear in the Lyapunov spectrum only for values of the external force so high that no negative fluctuations are observable anymore. We hope that future work will address this point.

APPENDIX A: A LIMIT THEOREM

In this section we prove Eq. 23–24. We reproduce in detail the proof in the case p is the average of independently distributed discrete variables σ_i^ε, assuming values in $\varepsilon\mathbb{Z}$, for some small mesh parameter ε; then we discuss how this can be applied and adapted to the situation considered in section IIC and subsequent sections.

Let us introduce some definitions. Let σ_i, $i \in \mathbb{N}$, be independent continuous random variables with identical distributions $\pi(d\sigma_i)$ with positive variance $\delta\sigma^2 > 0$, supported on the finite interval $[s_-, s_+]$. Let us assume that $\pi(d\sigma_i)$ gives positive probability to any finite interval contained in $[s_-, s_+]$. Let $\pi_\lambda(d\sigma)$ be the weighted distribution $\pi_\lambda(d\sigma) = e^{-\lambda\sigma}\pi(d\sigma)/\int e^{-\lambda\sigma}\pi(d\sigma)$ and let us define $z_\infty(\lambda) = -\log\int e^{-\lambda\sigma}\pi(d\sigma)$ and $\sigma_+ = z'_\infty(0)$. Note that the assumption that $\pi(d\sigma_i)$ gives positive probability to an interval of σ in $[s_-, s_+]$ implies that for any finite λ also $\pi_\lambda(d\sigma)$ has positive variance $-z''_\infty(\lambda) > 0$.

Also, given $\varepsilon > 0$ (with the property that $s_+ - s_- = N_\varepsilon\varepsilon$ for some integer N_ε), let us consider the discretization of σ_i on scale ε, call it σ_i^ε: σ_i^ε will be a discrete variable assuming the values $s_k^\varepsilon \overset{def}{=} s_- + (k - \frac{1}{2})\varepsilon$, $k = 1, \ldots, N_\varepsilon$, with probabilities $\pi^\varepsilon(s_k^\varepsilon) = \text{Prob}(\sigma_i^\varepsilon = s_k^\varepsilon) = \int_{s_k^\varepsilon \pm \frac{\varepsilon}{2}} \pi(d\sigma)$. The assumption that $\pi(d\sigma_i)$ gives positive probability to any finite interval contained in $[s_-, s_+]$ implies that $\pi^\varepsilon(s_k^\varepsilon) > 0$ for any ε and k. Let also $z_\varepsilon(\lambda) = -\log\sum_{k=1}^{N_\varepsilon} e^{-\lambda s_k^\varepsilon}\pi^\varepsilon(s_k^\varepsilon)$ and $\pi_\lambda^\varepsilon(s_k^\varepsilon) = \pi^\varepsilon(s_k^\varepsilon)e^{-\lambda s_k^\varepsilon + z_\varepsilon(\lambda)}$. Note that, since $\pi^\varepsilon(s_k^\varepsilon) > 0$ for any k, for any finite λ one has $-z''_\varepsilon(\lambda) > 0$.

If $p_\tau^\varepsilon = \frac{1}{\tau\sigma_+}\sum_{i=1}^\tau \sigma_i^\varepsilon$ and $\Pi_\tau(\varepsilon; I)$ is the probability that p_τ^ε belongs to the finite interval I, the following theorem holds.

Theorem: *Given a finite interval $I \subset (s_-, s_+)$, let σ_i^ε, π^ε and $\Pi_\tau(\varepsilon; I)$ be defined as above. Then, for a sufficiently small $\varepsilon > 0$, there exists an analytic "rate function" $\widetilde{\zeta}_\tau(p)$ such that*

$$\lim_{\tau\to\infty} \frac{\Pi_\tau(\varepsilon; I)}{\int_I dp e^{\tau\widetilde{\zeta}_\tau(p)}} = 1 \; . \tag{A1}$$

$\widetilde{\zeta}_\tau(p)$ is defined by:

$$\widetilde{\zeta}_\tau(p) + \frac{1}{\tau}\log\Big[\frac{\sinh[\varepsilon\lambda_p^\varepsilon/(2\sigma_+)]}{\varepsilon\lambda_p^\varepsilon/(2\sigma_+)}\Big] = \zeta_\tau^\varepsilon(p)$$
$$\zeta_\tau^\varepsilon(p) = -z_\varepsilon(\lambda_p^\varepsilon) + \lambda_p^\varepsilon p\sigma_+ - \frac{1}{2\tau}\log[\frac{2\pi}{\tau}\Big(-\frac{z''_\varepsilon(\lambda_p^\varepsilon)}{\sigma_+^2}\Big)] \tag{A2}$$

and λ_p^ε is the inverse of $p(\lambda) = z'_\varepsilon(\lambda)/\sigma_+$. The function $\zeta_\tau^\varepsilon(p)$ has the following property: if $\Delta \subset I$ is an interval of size $\frac{\varepsilon}{\tau\sigma_+}$ around a point p_Δ, then:

$$\lim_{\tau\to\infty} \frac{\Pi_\tau(\varepsilon; \Delta)}{|\Delta| e^{\tau\zeta_\tau^\varepsilon(p_\Delta)}} = 1 \tag{A3}$$

Proof Let us introduce the auxiliary variable $q = \frac{1}{\tau\sigma_+}\sum_{i=1}^\tau \eta_i$, where η_i are i.i.d. discrete random variables, with distribution $\pi_\lambda^\varepsilon(s_k^\varepsilon)$. Let us call $\Pi_\tau^\lambda(\varepsilon; q_0)$ the probability that q assumes the value $q_0 \in I$, with $q_0\sigma_+ = s_k^\varepsilon/\tau$ for some $k \in \mathbb{N}$, and note that $\Pi_\tau^0(\varepsilon; q_0)$ is identical to the probability that $p_\tau = q_0$. By definition $\Pi_\tau^\lambda(q_0)$ and $\Pi_\tau^0(q_0)$ are related by:

$$\Pi_\tau^\lambda(\varepsilon; q_0) = \frac{e^{-\lambda q_0\sigma_+\tau}\Pi_\tau^0(\varepsilon; q_0)}{\big[\sum_k e^{-\lambda s_k^\varepsilon}\pi^\varepsilon(s_k^\varepsilon)\big]^\tau} \tag{A4}$$

Now, a local form of central limit theorem (Gnedenko's theorem, see pag. 211 of [33]) tells us that, if q is localized near its mean value, that is if $|q\sigma_+ - z'_\varepsilon(\lambda)| \le \frac{M\varepsilon}{\tau}$ for some finite M, then $\Pi_\tau^\lambda(\varepsilon; q_0)$ is asymptotically equivalent to the Gaussian with mean $z'_\varepsilon(\lambda)$ and variance $-z''_\varepsilon(\lambda)$, in the sense that

$$\Pi_\tau^\lambda(q_0) = \frac{\varepsilon}{\sqrt{2\pi\tau(-z''_\varepsilon(\lambda))}} e^{-\frac{(q_0\sigma_+ - z'_\varepsilon(\lambda))^2}{2(-z''_\varepsilon(\lambda))}\tau}(1 + o(1)) \; , \tag{A5}$$

for any q_0 s.t. $|q\sigma_+ - z'_\varepsilon(\lambda)| \le \frac{M\varepsilon}{\tau}$ [35].

So, given $\lambda_{q_0}^\varepsilon$ s.t. $z'_\varepsilon(\lambda_{q_0}^\varepsilon) = q_0\sigma_+$ (such $\lambda_{q_0}^\varepsilon$ exists, is unique and is an analytic function of q_0, by the remark that $-z''_\varepsilon(\lambda) > 0$ for any finite λ and $z_\varepsilon(\lambda)$ is an analytic function of λ), using Eq. A5 we see that Eq. A4 can be restated as:

$$\Pi_\tau^0(\varepsilon; q_0) = \frac{\varepsilon}{\sqrt{2\pi\tau(-z''_\varepsilon(\lambda_{q_0}^\varepsilon))}} e^{\lambda_{q_0}^\varepsilon q_0\sigma_+\tau - z_\varepsilon(\lambda_{q_0}^\varepsilon)}(1 + o(1)) \tag{A6}$$

Now, by the definition of $\zeta_\tau^\varepsilon(p)$ in Eq. A2, we see that the r.h.s. of the last equation is equal to $\frac{\varepsilon}{\tau\sigma_+}e^{\tau\zeta_\tau^\varepsilon(q_0)}(1 + o(1))$. Finally, the statement of the Theorem follows by the remark that

$$\frac{\varepsilon}{\tau\sigma_+}e^{\tau\zeta_\tau^\varepsilon(p_0)} = \int_{p_0 - \frac{\varepsilon}{2\tau\sigma_+}}^{p_0 + \frac{\varepsilon}{2\tau\sigma_+}} dp e^{\tau\tilde{\zeta}_\tau(p)}\Big(1 + o(1)\Big) \; . \tag{A7}$$

In fact the integral in the r.h.s. of the last equation is given by

$$e^{\tau\tilde{\zeta}_\tau(p_0)}\int_{p_0 - \frac{\varepsilon}{2\tau\sigma_+}}^{p_0 + \frac{\varepsilon}{2\tau\sigma_+}} dp e^{\tau\tilde{\zeta}'_\tau(p_0)(p - p_0)}\Big(1 + O(\frac{\widetilde{\zeta}''_\tau(p_0)\varepsilon^2}{\tau})\Big) =$$
$$= e^{\tau\tilde{\zeta}_\tau(p_0)}\frac{2\sinh[\tilde{\zeta}'_\tau(p_0)\varepsilon/(2\sigma_+)]}{\tau\tilde{\zeta}'_\tau(p_0)}\Big(1 + O(\frac{\widetilde{\zeta}''_\tau(p_0)\varepsilon^2}{\tau})\Big) \tag{A8}$$

and in the last expression one has to note that $\widetilde{\zeta}'_\tau(p_0) = [\zeta_\tau^\varepsilon]'(p_0) + O(\frac{1}{\tau}) = \lambda_{p_0}^\varepsilon + O(\frac{1}{\tau})$. ■

A first Remark to be done about the Theorem above is that, in order to define a "universal" rate function in terms of quantities depending only on $z_\infty(\lambda)$ (instead of quantities depending on the "non universal" function $z_\varepsilon(\lambda)$, which explicitly depends on the discretization step ε), it would desirable to perform (in a sense to be precised) the continuum limit $\varepsilon \to 0$. To this regard, we can note that the only point where in the proof above we really used the fact that ε is a constant (i.e. is independent of τ) was in using Gnedenko's Theorem, see [33]. However, by a critical analysis of the proof of Gnedenko's Theorem, one can realize that it is even possible to let $\varepsilon = \varepsilon_\tau$ go to 0 with τ; the velocity with which ε_τ is allowed to go to 0 depends on the details of the distribution $\pi(d\sigma)$. So we can even study the probability distribution of p_τ on a scale $\sim \varepsilon_\tau/\tau$: if we introduce bins Δ_τ of size $O(\varepsilon_\tau/\tau)$ and we define $\Pi_\tau(\Delta_\tau)$ to be the probability that $p_\tau = \frac{1}{\tau\sigma_+}\sum_i \sigma_i$ belongs to the bin Δ_τ centered in p_0, we can repeat the proof above to conclude that

$$\lim_{\tau\to\infty} \frac{\Pi_\tau(\Delta_\tau)}{|\Delta_\tau| e^{\tau\zeta_\tau(p_0)}} = 1 \tag{A9}$$

where ζ_τ satisfies the equation:

$$\zeta_\tau(p) = -z_\infty(\lambda_p) + \lambda_p p\sigma_+ - \frac{1}{2\tau}\log[\frac{2\pi}{\tau}\Big(- \frac{z''_\infty(\lambda_p)}{\sigma_+^2}\Big)] \tag{A10}$$

and λ_p is the inverse of $p(\lambda) = z'_\infty(\lambda)/\sigma_+$.

Another point to be discussed is that in the Theorem above we assumed the σ_i to be independent. This is not the case for the variables $\sigma(S^i\cdot)$ of sec. II. However, if, as discussed in Remark (3) of sec. II D, we choose the time unit to be of the order of the mixing time, the variables $\sigma(S^i\cdot)$ have (by construction) a decorrelation time equal to 1, and the analysis of previous theorem can be repeated step by step in order to construct the probability distribution of $p = \frac{1}{\tau\sigma_+}\sum_i \sigma(S^i\cdot)$. The only differences are that: (1) $\tau z_\infty(\lambda)$ should be replaced by $\tau z_\tau(\lambda) = -\log\int e^{-\lambda p\sigma_+\tau}\Pi_\tau(dp)$ throughout the discussion; (2) instead of Gnedenko's theorem one has to apply a generalization of Gnedenko's to short ranged Gibbs processes, to be proven via standard cluster expansion techniques (see for instance [34] for a proof of a generalization of Gnedenko's theorem to a short ranged Gibbs process in the context of non critical fluctuations of the phase separation line in the 2D Ising model).

The conclusion is that, if the bins Δ in sec. II C are chosen of size ε_τ/τ, the probability of the bin Δ centered in p_Δ is asymptotically given by $\pi(p\in\Delta) \simeq e^{\tau\zeta_\tau(p_\Delta)}$ (in the sense of Eq. 23) and $\zeta_\tau(p_\Delta)$ can be interpolated by an analytic function of p that in fact satisfies Eq. 24.

ACKNOWLEDGMENTS

We would like to thank Prof. Joel L. Lebowitz for his interest on this work and for a valuable discussion. Some of this work was completed while two of us (A.G. and G.G.) were visiting Rutgers University under invitation of Prof. Lebowitz. A.G. was partially supported by the NSF grant 4-23421 DMR 01-279-26.

We thank F. Bonetto for a useful discussion and his suggestions about the implementation of the numerical algorithm.

F.Z. wish to thank G. Ruocco and L. Angelani for many useful discussions.

The computations have been performed on the FDT cluster of the INFM-SOFT center in Rome. We thank S. Erriu for technical assistance in operating the cluster. The sources of the program we used for the simulations can be downloaded from *http://glass.phys.uniroma1.it/zamponi*.

[1] G. Gallavotti, *Statistical mechanics. A short treatise* (Springer Verlag, Berlin, 1999).

[2] G. Gallavotti and E.G.D. Cohen, *Dynamical ensembles in nonequilibrium statistical mechanics*, Phys. Rev. Lett. **74**, 2694–2697 (1995).

[3] G. Gallavotti, E. G. D. Cohen, *Dynamical ensembles in stationary states*, J. Stat. Phys. **80**, 931–970 (1995).

[4] G. Gallavotti, *Reversible Anosov maps and large deviations*, Mathematical Physics Electronic Journal, MPEJ, **1**, 1–12 (1995).

[5] D. Ruelle, *Smooth dynamics and new theoretical ideas in non-equilibrium statistical mechanics*, J. Stat. Phys. **95**, 393–468 (1999).

[6] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, *Probability of second law violations in shearing steady states*, Phys. Rev. Lett. **71**, 2401–2404 (1993).

[7] G. Gallavotti, *Entropy production in nonequilibrium stationary states: a point of view*, Chaos, **14**, 680–690, (2004).

[8] D. Ruelle, *Conversations on nonequilibriun physics with an extraterrestrial*, Physics Today, May, 48–53 (2004).

[9] D.J. Evans and G.P. Morris, *Statistical Mechanics of Nonequilibrium Liquids* (Academic Press, London, 1990).

[10] F. Bonetto, G. Gallavotti, and P. L. Garrido, *Chaotic principle: an experimental test*, Physica D **105**, 226–252 (1997).

[11] F. Bonetto, N. I. Chernov, J. L. Lebowitz, *(Global and local) fluctuations of phase-space contraction in deterministic stationary non-equilibrium*, Chaos **8**, 823–833 (1998).

[12] L. Biferale, D. Pierotti, and A. Vulpiani, *Time-reversible dynamical systems for turbulence*, J. Phys. A: Math. Gen. **31**, 21–32 (1998).

[13] G. Gallavotti, F. Perroni, *An experimental test of the local fluctuation theorem in chains of weakly interacting Anosov systems, preprint* chao-dyn/9909007.

[14] G. Gallavotti, L. Rondoni, and E. Segre, *Lyapunov spectra and nonequilibrium ensembles equivalence in 2D fluid mechanics*, Physica D **187**, 338–357 (2004).

[15] F. Zamponi, G. Ruocco, L. Angelani, *Fluctuations of en-*

tropy production in the isokinetic ensemble, J. Stat. Phys. **115**, 1655–1668 (2004).

[16] G. Gallavotti, D. Ruelle, *SRB states and nonequilibrium statistical mechanics close to equilibrium*, Comm. Math. Phys. **190**, 279–285 (1997).

[17] G. Gallavotti, *Extension of Onsager's reciprocity to large fields and the chaotic hypothesis*, Phys. Rev. Lett. **77**, 4334–4337 (1996); G. Gallavotti, *Chaotic hypothesis: Onsager reciprocity and fluctuation-dissipation theorem*, J. Stat. Phys. **84**, 899 (1996).

[18] Y. G. Sinai, *Markov partitions and C-diffeomorphisms*, Functional Analysis and Applications, **2**, n.1, 64–89 (1968); *Construction of Markov partitions*, Functional analysis and Applications, **2**, n.2, 70–80, (1968).

[19] Y. G. Sinai, *Lectures in ergodic theory*, Lecture notes in Mathematics, Princeton U. Press, Princeton, 1977.

[20] G. Gallavotti, *Fluid mechanics. Foundations*, Springer-Verlag, Berlin, 2002.

[21] G. Gallavotti, F. Bonetto, G. Gentile, *Aspects of the ergodic, qualitative and statistical properties of motion*, Springer Verlag, Berlin, 2004.

[22] S. de Groot, P. Mazur: *Non equilibrium thermodynamics*, Dover, 1984, (reprinted).

[23] D. Ruelle, *A remark on the equivalence of isokinetic and isoenergetic thermostats in the thermodynamic limit*, J. Stat. Phys. **100**, 757–763 (2000).

[24] M. P. Allen and D. J. Tildesley, *Computer simulation of liquids* (Oxford Science Publications, 1987).

[25] S. S. Sarman, D. J. Evans, P. T. Cumming, *Recent developments in non-Newtonian molecular dynamics*, Phys. Rep. **305**, 1–92 (1998).

[26] W. Kob and C. H. Andersen, *Scaling behavior in the β-relaxation regime of a supercooled Lennard-Jones mixture*, Phys. Rev. Lett. **73**, 1376–1379 (1994).

[27] C. De Michele, F. Sciortino, A. Coniglio, *Scaling in soft spheres: fragility invariance on the repulsive potential softness*, J. Phys.: Condens. Matter **16**, L489–L494 (2004).

[28] G. Benettin, L. Galgani, J.-M. Strelcyn, *Kolmogorov entropy and numerical experiments*, Phys. Rev. A **14**, 2338–2345 (1976).

[29] F. Bonetto, G. Gallavotti, *Reversibility, coarse graining and the chaoticity principle*, Communications in Mathematical Physics, **189**, 263–276, (1997).

[30] G. Gallavotti, *New methods in nonequilibrium gases and fluids*, Open Systems and Information Dynamics **6**, 101–136 (1999).

[31] D. J. Searles, D. J. Evans, D. J. Isbister, *The conjugate-pairing rule for non-Hamiltonian system*, Chaos **8**, 337–349 (1998).

[32] C. P. Dettmann and G. P. Morriss, *Proof of Lyapunov exponent pairing for systems at constant kinetic energy*, Phys. Rev. E **53**, R5545–R5548 (1996).

[33] M. Fisz, *Probability theory and mathematical statistics*, Wiley, New York, 1963.

[34] G. Gallavotti, *Phase separation line in the two–dimensional Ising model*, Comm. Math. Phys. **27**, 103–136 (1972).

[35] Note that Gnedenko's Theorem is *different* from the usual central limit theorem, stating instead that for $|q\sigma_+ - z'_\varepsilon(\lambda)| \le \frac{C}{\sqrt{\tau}}$ (C big) the sums of $\Pi^\lambda_\tau(\varepsilon;q)$ over intervals of amplitude $\frac{1}{\sqrt{\tau}}$ contained in $|q\sigma_+ - z'_\varepsilon(\lambda)| \le \frac{C}{\sqrt{\tau}}$ are asymptotically equal to the integrals of the Gaussian over the same intervals. That is, usual central limit theorem gives informations on the distribution in a bigger interval around the maximum, but on a rougher scale.