Predictability in the large: an extension of the concept of Lyapunov exponent

E. Aurell
Department of Mathematics, Stockholm University S–106 91 Stockholm, Sweden

G. Boffetta
Dipartimento di Fisica Generale, Università di Torino, Via Pietro Giuria 1, I-10125 Torino, Italy
Istituto Nazionale Fisica della Materia, Unità di Torino

A. Crisanti
Dipartimento di Fisica, Università di Roma "La Sapienza", P.le Aldo Moro 2, I-00185 Roma, Italy
Istituto Nazionale Fisica della Materia, Unità di Roma

G. Paladin
Dipartimento di Fisica, Università dell'Aquila, Via Vetoio, Coppito I-67100 L'Aquila, Italy
Istituto Nazionale Fisica della Materia, Unità dell'Aquila

A. Vulpiani
Dipartimento di Fisica, Università di Roma "La Sapienza", P.le Aldo Moro 2, I-00185 Roma, Italy
Istituto Nazionale Fisica della Materia, Unità di Roma

(TNT 96-ShPrePap-V.4, 27 June 1996)

We investigate the predictability problem in dynamical systems with many degrees of freedom and a wide spectrum of temporal scales. In particular, we study the case of $3D$ turbulence at high Reynolds numbers by introducing a finite-size Lyapunov exponent which measures the growth rate of finite-size perturbations. For sufficiently small perturbations this quantity coincides with the usual Lyapunov exponent. When the perturbation is still small compared to large-scale fluctuations, but large compared to fluctuations at the smallest dynamically active scales, the finite-size Lyapunov exponent is inversely proportional to the square of the perturbation size. Our results are supported by numerical experiments on shell models. We find that intermittency corrections do not change the scaling law of predictability. We also discuss the relation between finite-size Lyapunov exponent and information entropy.

PACS numbers 47.27.Gs, 05.45.+b, 47.27.Jv

I. INTRODUCTION

The ability to predict has been the single most important qualifier of what constitutes scientific knowledge, all since the successes of Babylonian and Greek astronomy. Indeed, the famous statement of Laplace that an intelligent being with complete knowledge of the present and of the laws of Nature will know the future for all time, assumes that the future is completely predicated by the past, and that perfect prediction would in principle be possible. In more mathematical terms one can say that in the physical sciences, whether in the classical or the quantum regime, one believes that Nature is ultimately described by differential equations, and if one knows them and how to solve them, one knows all there is to know about the world [1].

Laplacian determinism is always conditioned by the fact that in the real world initial conditions can never be known to arbitrary accuracy. More recent is the general appreciation of the fact that in the presence of deterministic chaos, predictability is even more severely limited, because small errors typically grow exponentially in time [2]. Most sufficiently complex systems in the world display chaos. Therefore, most sufficiently complex systems can only be predicted for a finite time. However, there may be some aspects of a system that are stable, while others vary. To take a familiar example, weather prediction is possible, typically for about ten days on temperate latitudes, but how the wind blows on the corner of the street is in practice unpredictable from one moment to the next [3].

The words predictability and prediction are rather empty by themselves: one has to ask for predictability of what feature against what perturbation, in particular, against a perturbation of what size. In complex and spatially extended systems one can typically talk about large-scale features and small-scale features. The predictability of a small-scale feature against a small-scale perturbation is typically shorter than the predictability of a large-scale feature against a large-scale perturbation. Certainly this is not always true, and we will consider a counter example, $2D$ turbulence, below. But it seems to be a very common situation. That successful large-scale prediction essentially can mean just assuming that things do not change, is not *per se* an argument against prediction in the large. Also in weather prediction the assumption that the weather tomorrow will be like today is a fairly good one.

In this work, a further development of results presented in the a recent brief report [4], we will introduce a quantity which measures predictability in the large, and apply it to hydrodynamic turbulence. Before we proceed to the definitions, let us first recall some facts about predictability in the small, i.e. the effects of dynamical chaos. A system is said to be chaotic if small – i.e. infinitesimal – perturbations grow exponentially in time. If the initial perturbation is of size δ, and the accepted error tolerance is Δ, still small, then a rough estimate gives that the predictability time is

$$T_{\mathrm{p}} \sim \frac{1}{\lambda_{\max}} \ln\left(\frac{\Delta}{\delta}\right), \tag{1}$$

where $\lambda_{\max}$ is the leading Lyapunov exponent [5].

Already within the framework of infinitesimal perturbations there are important modifications to (1) [6,7]. In fact, in typical chaotic systems, (1) is not quite true [8]. The exponent $\lambda_{\max}$ is a global quantity which measures the *average* exponential rate of separation of nearby trajectories, and fluctuations of the local exponential grow should be taken into account [7], but these effects are not what concerns us here. In this paper we shall address the problem of predictability in systems with many characteristic times, e.g. the case of fully developed turbulence where a hierarchy of different eddy-turnover times do exist, or when the threshold δ is not small. In these cases the predictability time T_{p} is determined by the details of the nonlinear mechanism responsible for the growth of the error [3,9]. In particular, T_{p} may have no relation with the maximum Lyapunov exponent governed by the linearized equations for the infinitesimal error. In general, in this case the predictability time strongly depends on the details of the system [3,8].

According to Oseledec theorem, the leading eigenvalue of the linearized equations of motion is $\exp(\lambda_{\max} t)$, except on a set of points of measure zero. The sub-leading eigenvalues have the form $e^{\lambda_i t}$, and, taken together, the leading and sub-leading Lyapunov exponents measure the growth rate of d-dimensional volumes spanned by d infinitesimal vectors, where d can range from one to the dimensionality of the space where the motion takes place.

In dynamical systems, in addition to Lyapunov exponents, an important dynamical characterization is given by the Kolmogorov-Sinai entropy, which measures the bandwidth necessary to observe a system over time, so that it could later be faithfully reproduced from the observations [5]. Arguing heuristically, new observations are necessary if an error grows in time, and the necessary rate of accumulation of information is the growth rate of the error, and therefore one expects Pesin theorem,

$$h_{\mathrm{KS}} = \sum_{\lambda_i > 0} \lambda_i, \tag{2}$$

to hold true for a large class of systems.

It was realized by Shannon that with finite error tolerance, the relevant quantity is the bandwidth necessary to observe a system such that it could later by reproduced within this error, not to arbitrarily high accuracy. This quantity was called by "source entropy with respect to a fidelity criterion" by Shannon [10], and "ϵ-entropy" by Kolmogorov [11], who found analytic formulae, valid for Gaussian variables and Gaussian stationary processes. Recently the concept of ϵ-entropy was taken up by Gaspard and Wang, who computed it from experimental data in thermal turbulence [12], and for a very large variety of model problems in stochastic processes and statistical physics [13]. We especially recommend their lucid and remarkably complete review [13].

However, the ϵ-entropy does not say all one wants to know about predictability in the large. Just as the Lyapunov exponent is often more relevant, and more easily computable, than the Kolmogorov-Sinai entropy, so the predictability time with respect to a finite perturbation should be determined by a quantity analogous to the Lyapunov exponent, and not by the ϵ-entropy.

The natural starting point in looking for such a quantity is the time it takes for a perturbation to grow from an initial size δ to a tolerance Δ. We call this the (δ, Δ) predictability time and denote it by $T(\delta, \Delta)$. Generally speaking, the predictability time will fluctuate. The natural definition of the finite-size Lyapunov exponent is therefore an average of some function of the predictability time, such that if both δ and Δ are in the infinitesimal range, we will recover the usual Lyapunov exponent, and an obvious choice is then

$$\lambda(\delta, \Delta) = \left\langle \frac{1}{T(\delta, \Delta)} \right\rangle \ln\left(\frac{\Delta}{\delta}\right). \tag{3}$$

In appendix A we discuss other possible definitions for the finite-size Lyapunov exponent and the relation with the ϵ-entropy.

In contrast to infinitesimal perturbations, for finite perturbations the threshold Δ is typically not to be taken much larger than the perturbation δ. What is interesting, and what makes finite-size Lyapunov exponents different from Lyapunov exponents for infinitesimal perturbations, is the dependence on δ.

This paper is organized as follows: in section II we recall the multifractal approach to turbulence, and Lorenz approach to the predictability problem within the Kolmogorov theory. We show that there are no multifractal corrections to the results of Lorenz, but that the scaling range for the finite-size Lyapunov exponents is shorter. In section III we describe numerical experiments on predictability in shell models for $3D$ and $2D$ fully developed turbulence. In section IV we present the results from the eddy damped quasi normal Markovian (EDQNM) approximation for the shell model, and compare them with the results of section III. In section V we summarize our results and present conclusions. In appendix A we discuss alternative ways of defining the finite-size Lyapunov exponent. Appendix B contains a derivation of the Kolmogorov results of ϵ-entropy for Gaussian processes and Gaussian random fields. In appendix C we apply the results of appendix B to space-time Gaussian fields with spectra as in Kolmogorov 1941 theory of $3D$ turbulence, and to fictitious 0-dimensional fields that describe shell models. We show that the finite-size Lyapunov exponent is the the same for any dimensionality, but the ϵ-entropy largely depends on the dimensionality-dependent density of degrees of freedom. The results of sections II and III were also presented in less generality in [4].

II. MULTIFRACTALS AND MULTISCALING

Our understanding of high Reynolds-number turbulence is still mainly based on the fundamental contribution of Kolmogorov in 1941 [14]. We will here just discuss on a phenomenological level the predictions of the Kolmogorov theory and of its multifractal generalizations [15–17].

Turbulence is a statistically stationary state of matter on macroscopic scales maintained by external forces. One considers only effects that are captured in a hydrodynamic level of description, that is the time evolution is supposed to be completely described by the macroscopic Navier-Stokes equations [18]

$$\partial_t \boldsymbol{v} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{v} = -\nabla P + \nu\nabla^2\boldsymbol{v}, \tag{4}$$

$$\nabla\cdot\boldsymbol{v} = 0. \tag{5}$$

From the typical length scale L, the typical fluctuations of velocity on that scale V, and the viscosity ν, we can form the Reynolds number,

$$\mathrm{Re} = \frac{LV}{\nu}, \tag{6}$$

which characterizes the flow.

The multifractal model [15–17,8] consists in assuming that at scales much less than L, however sufficiently large that the action of viscosity is weak, the velocity differences assume a scaling form

$$|\boldsymbol{u}(\boldsymbol{x}+\boldsymbol{l}) - \boldsymbol{u}(\boldsymbol{x})| = u_l \sim V\left(\frac{l}{L}\right)^h. \tag{7}$$

Different values of h are assumed to occur according to a probability distribution, which also takes a scaling form

$$\mathrm{Prob}\left\{\frac{u_l}{V} \in \left[\left(\frac{l}{L}\right)^h, \left(\frac{l}{L}\right)^{h+dh}\right]\right\} \sim \left(\frac{l}{L}\right)^{3-D(h)} dh. \tag{8}$$

The function $D(h)$ is the fractal dimension of the subset with scaling exponent h. The moments of the velocity differences on length scale l can be computed as

$$\langle|\boldsymbol{u}(\boldsymbol{x}+\boldsymbol{l}) - \boldsymbol{u}(\boldsymbol{x})|^q\rangle \sim V^q \int \left(\frac{l}{L}\right)^{qh} \left(\frac{l}{L}\right)^{3-D(h)} dh, \tag{9}$$

and, for small-l (i.e. in the inertial range), the integral in (9) can be evaluated by saddle point method:

$$\langle|\boldsymbol{u}(\boldsymbol{x}+\boldsymbol{l}) - \boldsymbol{u}(\boldsymbol{x})|^q\rangle \sim V^q \left(\frac{l}{L}\right)^{\zeta_q}, \tag{10}$$

$$\zeta_q = \min_h\left[qh + 3 - D(h)\right]. \tag{11}$$

The model is physically reasonable for a large set of possible choices of the function $D(h)$, but not entirely arbitrary. By normalization, the value of $D(h)$ must always be less or equal than three, and the maximum must be obtained for some h. That is

$$3 - D(h) \geq 0. \tag{12}$$

The function $D(h)$ can have support only at positive h, because a negative value of h implies that velocity fluctuations in a local inertial frame of size l increase without limit as l tends to zero. The Navier-Stokes equations are derived under the assumption that all velocities are much smaller than the velocity of sound, and this condition would then no longer hold [17]. Furthermore, an exact result of Kolmogorov assures that $\zeta_3 = 1$, so that

$$3h + 3 - D(h) \geq 1, \tag{13}$$

where equality is obtained for at least one value of h [17]. The inequality (13) is the analogous for turbulence of the inequality $f(\alpha) \leq \alpha$ for multifractal measures [8].

In terms of the multifractal model, the Kolmogorov theory is formulated by supposing that the function $D(h)$ has support only at a single point. From (12) and (13) it then follows that this point must be $h = 1/3$, and that $D(1/3) = 3$. It further follows the Kolmogorov law

$$\langle |\boldsymbol{u}(\boldsymbol{x}+\boldsymbol{l}) - \boldsymbol{u}(\boldsymbol{x})|^p \rangle \sim V^p \left(\frac{l}{L}\right)^{p/3}. \tag{14}$$

Energy dissipation per unit mass and time, ϵ, has dimension V^3/L. We could therefore also write the right-hand side in (14) in the more familiar form $(\epsilon l)^{p/3}$

From (14) it follows by balancing in (4), that viscous forces become comparable with inertial forces at the Kolmogorov scale η which marks the lower end of the inertial range:

$$\eta = L\,\mathrm{Re}^{-3/4}. \tag{15}$$

If there exists more than one value of h then each h selects a different damping scale $\eta(h)$. By using (7) and balancing, one gets [19]

$$\eta(h) = L\,\mathrm{Re}^{-1/(1+h)}. \tag{16}$$

Lorenz investigated the predictability problem within the framework of the Kolmogorov theory [20]. Let us assume that we have a disturbance on scale l, and that it grows at a characteristic rate given by the turn-over time at this scale:

$$\tau(l) = \frac{l}{\sqrt{\langle u_l^2 \rangle}} \sim \frac{L}{V}\left(\frac{l}{L}\right)^{2/3}. \tag{17}$$

We can turn around (17) and say that after a time t a disturbance will have grown large on all scales smaller than

$$l(t) \sim L\left(\frac{Vt}{L}\right)^{3/2}. \tag{18}$$

The size of the disturbance will then be $\langle u_{l(t)}^2 \rangle^{1/2}$, since all the smaller scales contribute relatively little. If we call the difference between two fields δ, we can rewrite (17) as the predictability time with respect to a perturbation of size δ:

$$\tau(\delta) \sim \frac{L}{V}\left(\frac{\delta}{V}\right)^2. \tag{19}$$

In other words, the predictability time of a perturbation of size δ grows as δ^2 in Lorenz scenario. The finite-size Lyapunov exponent thus decreases with the error threshold as δ^{-2}. Finally, we can insert δ in (18), and find how the error grows with time:

$$\delta(t) \sim V\sqrt{\frac{Vt}{L}}. \tag{20}$$

The upshot of these simple estimates is that finite error growth and predictability in high Reynolds number turbulence are characterized by *algebraic* laws, very different form the *exponential* laws characteristic of infinitesimal perturbations in chaotic dynamical systems.

We now turn to possible consequences of a spectrum of h's to the predictability problem, in direct analogy with the Lorenz theory. We start by assuming the inverse of (8) that is

$$\mathrm{Prob}\left\{\frac{\delta}{V}\in\left[\left(\frac{l}{L}\right)^{h},\left(\frac{l}{L}\right)^{h+dh}\right]\right\}\sim\left(\frac{\delta}{V}\right)^{\frac{3-D(h)}{h}}dh, \tag{21}$$

where we have identified u_l with δ, and used (7) to relate l and δ. The length l is now assumed a fluctuating quantity, and has the interpretation of a scale such that two fields are uncorrelated on all smaller scales, given that the distance of the two configurations is δ.

The finite-size Lyapunov exponent is, according to (3), proportional to the expectation value of the inverse predictability time:

$$\left\langle\frac{1}{T(\delta)}\right\rangle=\left\langle\frac{u_l}{l}\right\rangle\sim\frac{V}{L}\int\left(\frac{\delta}{V}\right)^{1-\frac{1}{h}}\left(\frac{\delta}{V}\right)^{\frac{3-D(h)}{h}}dh. \tag{22}$$

In the small error limit we thus expect the finite-size Lyapunov exponent to scale as a power of the error size:

$$\left\langle\frac{1}{T(\delta)}\right\rangle\sim\frac{V}{L}\left(\frac{\delta}{V}\right)^{\chi}, \tag{23}$$

$$\chi=\min_h[1+\frac{2-D(h)}{h}]. \tag{24}$$

The exponent χ is always equal to the Lorenz value of -2. This follows easily from (13), which can be rewritten in the form

$$1+\frac{2-D(h)}{h}\geq-2\qquad\text{for all } h, \tag{25}$$

where the equality holds for the exponent h_3 which dominates the third order structure function.

As far as we know this result is new. One could therefore conclude that the exponent χ for the scaling of the finite-size Lyapunov exponent with error threshold is a new invariant of the multifractal approach to turbulence, and that the law $\langle 1/T(\delta)\rangle\sim\delta^{-2}$ can be easily observed in numerical experiments. This is not quite simple, due to the influence of the fluctuating cut-off (16). The smallest fluctuation in a field scaling with exponent $\tilde{h}$ is

$$\delta(\tilde{h})=V\,\mathrm{Re}^{-\tilde{h}/(1+\tilde{h})}, \tag{26}$$

which inversely determines the smallest value of h contributing to a fluctuation of size δ. A modified version of (22) therefore reads

$$\left\langle\frac{1}{T(\delta)}\right\rangle=\left\langle\frac{u_l}{l}\right\rangle\sim\frac{V}{L}\int_{\tilde{h}}^{h_{\max}}\left(\frac{\delta}{V}\right)^{1-1/h}\left(\frac{\delta}{V}\right)^{(3-D(h))/h}dh. \tag{27}$$

The integral is dominated by h_3 as long as δ is much larger than $\delta^*\sim V\mathrm{Re}^{-h_3/(1+h_3)}$. For smaller δ values, the integral is dominated by the lower end-point in (27), which leads to an intermediate dissipative range, in the sense of Frisch and Vergassola [21]. As a consequence, we have

$$\left\langle\frac{1}{T(\delta)}\right\rangle\sim\begin{cases}\lambda_{\max} & \text{for } \delta<\delta(h_{\max}),\\ \delta^{\chi(\delta)} & \text{for } \delta(h_{\max})<\delta<\delta^*\\ \delta^{-2} & \text{for } \delta>\delta^*,\end{cases} \tag{28}$$

with

$$\chi(\delta)=1+\frac{2-D(\tilde{h})}{\tilde{h}}, \tag{29}$$

where $\tilde{h}$ and δ are related via (26). From (13) follows that h_3 is less or equal to $1/3$ in the multifractal approach, and therefore the bottom of the scaling range of the finite-size Lyapunov exponent is larger than the corresponding error size in the Kolmogorov theory. The scaling range for the finite-size Lyapunov exponent thus is generally shorter in a multifractal model.

III. ERROR GROWTH IN SHELL MODELS

Simplified dynamical models of fluid turbulence with relatively few degrees of freedom, collectively referred to as *shell models*, have been studied since the seventies. These models by construction typically include a Richardson cascade of energy from large to small scales. Some of the models are dynamically stable, with a fixed point which reproduces the Kolmogorov law for the energy spectrum, $E(k) \sim k^{-5/3}$. An historical overview, with references to much of the early work, can be found in a recent monograph [22].

More interestingly, other models are dynamically unstable, with chaotic motion taking place on a strange attractor where the Kolmogorov 5/3 law holds to good accuracy, but not exactly. One of the simplest example is the family of models introduced by Gledzer [23], and Yamada and Ohkitani [24], now commonly called the GOY models [25]. They have recently been the subject of several investigations [26,27,7,28,29]. An in-depth description of this work can also be found in [22].

The GOY models are defined as follows: Fourier space is divided into $n = 1, \ldots, N$ shells, labeled by the wave-vector modulus $k_n = k_0\, 2^n$, where k_0 is a constant. The velocity difference over a length scale $l_n \sim k_n^{-1}$ are represented, each by one complex variable u_n, which obey the following system of coupled ordinary differential equations:

$$\frac{d}{dt} u_n = -\nu k_n^2 + i g_n + f\, \delta_{n,4} \tag{30}$$

$$g_n = a\, k_n\, u^*_{n+1}\, u^*_{n+2} + b\, k_{n-1}\, u^*_{n-1}\, u^*_{n+1} + c\, k_{n-2}\, u^*_{n-2}\, u^*_{n-1} \tag{31}$$

where f is the strength of the external force, acting on large scales, and ν the viscosity. For any values of the three coefficients a, b and c, phase space volume in preserved in the force-free inviscid limit.

The restricted number of degrees of freedom is both the main advantage and disadvantage of shell models. It is an advantage, because it allows simulations at much lower viscosity and for much longer time than in the full Navier-Stokes equations. But it is also a severe departure. All spatial structure of the field is ignored.

One of the coefficients in (31), say a, can be scaled to one and the condition of energy conservation fixes one more, such that, in terms of one parameter ϵ, b is equal to $-\epsilon$ and c to $-(1-\epsilon)$. With ϵ greater than one the GOY equations conserve one more positive definite quantity besides energy, i.e. an analogous situation to $2D$ hydrodynamics. The dynamical behavior of the GOY models in this range is rather far removed from $2D$ turbulence as was shown in recent papers [30,31]. We include below a study of predictability in such shell models, but just as a simplified model to demonstrate one possible scaling behavior of the finite-size Lyapunov exponent.

With parameter ϵ in the range between and zero and one, the GOY models also preserve another invariant, but which is not positive definite. In the following we will look at the GOY model with the standard choice of ϵ equal to 1/2, i.e.

$$a = 1, \qquad b = -\frac{1}{2}, \qquad c = -\frac{1}{2}. \tag{32}$$

The second invariant then has physical dimension of helicity [28]. Presumably that is the reason why this particular model has turned out to be so close to numerical and experimental data on Navier-Stokes turbulence in $3D$ [32,33].

Before we turn to the numerical experiments, let us summarize some salient features of the system defined by (30)-(32). Energy is pumped into the system by the force, which acts only on shells with low values of n, and is removed at high shells by viscosity. From k_0, ν and the typical fluctuations of velocity on large scales, V, we can form a Reynolds number $\mathrm{Re} = V/k_0\nu$. We consider the situation where Re is large, such that there is a wide range in n where the external and viscous forces are both negligible compared to the inertial forces. In this inertial range, we have $\langle |u_n|^2\rangle \sim k_n^{-\zeta_2}$, where the exponent ζ_2 is close to 2/3 [27].

An estimate of the smallest excited scale is in analogy with the Kolmogorov scale $k_{n^*} \sim k_0 \mathrm{Re}^{3/4}$. The slowest dynamical scale is the time-scale of the the shells containing most energy, about $1/k_0 V$, and the fastest is $\tau_{n^*}^{-1} \sim k_{n^*} \langle |u_{n^*}|^2\rangle^{1/2}$, or, about $k_0 V \mathrm{Re}^{1/2}$. From the fastest time-scale follows by dimensional analysis in the Kolmogorov theory that the leading Lyapunov exponent should grow with Reynolds number as $\mathrm{Re}^{1/2}$, a prediction due to Ruelle [34]. In the multifractal picture there are corrections to this estimate and that the leading Lyapunov exponent of the GOY shell model scales as Re^{α} where

$$\alpha = \max_h \left[\frac{D(h) - 1 - 2h}{1 + h} \right]. \tag{33}$$

It is indeed numerically observed to scale as $\mathrm{Re}^{0.495}$, in good agreement with a computation of α staring from a function $D(h)$, obtained by a parametric fit of measured values of the scaling exponents ζ_q in experiments [16,7].

The mean square fluctuations at the Kolmogorov scale are $\langle |u_{n^*}|^2\rangle \sim \mathrm{Re}^{-1/2}$. If we compute the distance between two shell variable configurations as

$$|u-u'| = \sqrt{\sum_n |u_n - u'_n|^2}, \tag{34}$$

an error smaller than $O(\mathrm{Re}^{-1/4})$ is relatively small all over the inertial range. It can be taken to be infinitesimal, and its growth rate will be the fastest linear growth rate.

If, however, the error is larger than $O(\mathrm{Re}^{-1/4})$, it could be larger than the typical size of the fluctuations at the Kolmogorov scale. Such an error would have to be concentrated on larger scales, since otherwise we have that for some n u_n and/or u'_n is much larger than the typical size. In other words, a physical perturbation larger than $O(\mathrm{Re}^{-1/4})$ cannot be obtained by a random perturbation of that size uniformly distributed over all the shells.

In the Appendix A we discuss some possible definitions for the finite Lyapunov-exponents. In what follows we adopt the following procedure: after a long integration time to let the system relax towards the statistically stationary state, we introduce a very small error. This is done by generating a new shell variable configuration u'_n differing from u_n by a small fraction of $\langle |u_n|^2\rangle^{1/2}$. Another possible approach would be, that if we want an initial error of size ϵ, we determine a shell n_ϵ such that $\langle |u_{n_\epsilon}|^2\rangle^{1/2}$ is about ϵ, and we concentrate the perturbation on shells above n_ϵ.

We then iterate u_n and u'_n (perturbed system) for again a long time, such that the error has grown to a threshold, which is still small compared to $V\mathrm{Re}^{-1/4}$. We thus have two realizations of configurations in the statistically stationary state, which only differ by a small error, which we call δ_0.

Further we define a series of thresholds $\delta_n = r^n\delta_0$, and we measure the times it takes for the error to grow from δ_0 to δ_1, and so on. For brevity we will call these times error doubling times, even if r can be different from two. The threshold rate r should not be taken too large, because then the error has to grow through several different scales before reaching the next threshold. On the other hand, the rate r can not be too close to one, so a sensible threshold rate is on the order of two. The most convenient choice of r clearly depends on the how the fluctuations in the shell variable depend on n, and in what way we measure the error. For our model (31) with error measured by (34), the range of error sizes in the inertial range is not large, scaling as $\mathrm{Re}^{-1/4}$. For 35 shells, which is the largest systems we simulate, we can take Re equal to 10^{10}, which gives an error range of about 300. With error threshold rate equal to two, that would give about 8 data points, with some points lost on both ends due to boundary effects. For practical reasons we therefore take r equal to $\sqrt{2}$.

When we have performed N error doubling experiments, we can form an estimate of the expectation value of some quantity A:

$$\langle A\rangle_e = \frac{1}{N}\sum_{i=1}^{N} A_i. \tag{35}$$

This is not the same as taking a time-average, since different error doubling experiments may take different times. Indeed, we have

$$\langle A\rangle_t = \frac{1}{T}\int_0^T A(t)\,dt = \frac{\sum_i A_i\,\tau_i}{\sum_i \tau_i} = \frac{\langle A\,\tau\rangle_e}{\langle\tau\rangle_e}. \tag{36}$$

A particular case of the above relation concerns the mean error doubling times themselves. Let $T_r(\delta_n)$ be the time it takes for an error to grow from threshold δ_n to δ_{n+1}. Then

$$\lambda(\delta_n) = \left\langle \frac{1}{T_r(\delta_n)}\right\rangle_t \ln r = \frac{1}{\langle T_r(\delta_n)\rangle_e}\ln r, \tag{37}$$

where we have used the definition of (3).

The finite-size Lyapunov exponents, $\lambda(\delta_n)$, can be compared with shell turn-over times as follows: we first select a shell n_δ such that $\langle |u_{n_\delta}|^2\rangle^{1/2}$ is about δ, and then estimate τ_δ^{-1} as $k_{n_\delta}\delta$, which scales as δ^{-2}. This argument for typical error growth times is the same as Lorenz argument for $3D$ turbulence [20], discussed above in section II.

In Fig.1 we compare error doubling times and shell turn-over times, as a function of size of the perturbation and of the typical fluctuations in the corresponding shell. Below the Kolmogorov scale, the turn-over times increase: we are here in the dissipation range, where the shell amplitudes decrease quickly. On the other hand, the doubling times tend to a constant as the error threshold is small. We are here in the infinitesimal range, and the constant is approximatively the inverse of the Lyapunov exponent. At the Kolmogorov scale, there is rather large discrepancy between the Lyapunov exponent and the turn-over time. This observation, that the Lyapunov exponent obeys a scaling law with a sizeable numerical pre-factor, has been made before [27], but without a plausible explanation. We here find nice agreement with our prediction that the inertial range for the finite-size Lyapunov exponent is shorter than the spectral inertial range, because the first is limited from below by the scaling exponent h_3, as in equation (26), while the second is limited from below by the scaling exponent h_2 [21].

In Fig. 2 we compare the error doubling times for different Reynolds numbers. For small thresholds the doubling times scale as the Lyapunov exponent, i.e. as $\mathrm{Re}^{-1/2}$. We also observe that the bend away from the infinitesimal growth rate occurs at smaller error scales for larger Reynolds numbers. This suggests that a simple scaling ansatz can be sought in the following form: times and errors are scaled with the turn-over time and the typical scale of fluctuations at the Kolmogorov scale, that is by $\mathrm{Re}^{-1/2}$ and $\mathrm{Re}^{-1/4}$, respectively. In Fig. 3 we show such re-scaled data. The data collapse is reasonable.

To improve the data collapse, taking into account multifractal corrections, we made a scaling based on multiscaling [21], i.e. of the form

$$\frac{\ln\langle 1/T_r(\delta v)\rangle}{\ln(\mathrm{Re}/\mathrm{R}_o)} = f\big(\ln(\delta v/V_o)/\ln(\mathrm{Re}/\mathrm{R}_o)\big) \tag{38}$$

where R_o, V_o are parameter to be fixed, and $f(x)$ is the scaling function. According the argument at the end of section II, we have $f(x) \sim x^{-2}$ for large x, while $f(x)$ is constant for small values of x. In the intermediate regime $f(x)$ has a nontrivial form which depends on the shape of $D(h)$, as follows from eqs. (28)-(29). The result is shown in Fig. 4. The data collapse is clearly improved.

We conclude this section discussing the case of $2D$ turbulence. Two dimensional Euler equation has the peculiar property of an infinite number of invariants. Two of them are retained in a finite Fourier discretization, the energy and the average square vorticity, or enstrophy. As we previously discussed, the second conserved quantity in the GOY shell model depends on the choice of the parameter ϵ. With the choice $\epsilon = 5/4$, leading to

$$a = 1, \qquad b = -\frac{5}{4}, \qquad c = \frac{1}{4}, \tag{39}$$

equations (30) conserve in the unforced and inviscid limit, in addition to the energy, the enstrophy here defined as

$$Z = \frac{1}{2}\sum_n k_n^2\,|u_n|^2. \tag{40}$$

Despite the fact that the two-dimensional shell model has superficially the same physical justification of its three-dimensional corresponding model, it has been demonstrated that it has little to do with turbulence [30,31]. Moreover, all the numerical simulations of two-dimensional Navier-Stokes equation at sufficiently high Reynolds number have demonstrated the dynamical relevance of coherent structures which emerge spontaneously from the turbulent flow. The predictability problem, which is more relevant for geophysical flows in this case than in $3D$ turbulence, is also ruled by coherent vortex motion in the physical space, rather than modes dynamics in Fourier space [35].

With this limitations, the study of the predictability problem, as addressed in the present paper, in two-dimensional shell model is nevertheless interesting because of the different scaling behavior with respect to the $3D$ situation. Dimensional analysis [36] shows that in the enstrophy cascade one expects constant – i.e. independent on the scale – turn-over times. Hence an argument similar to that of section II shows that

$$\left\langle \frac{1}{T_r(\delta v)} \right\rangle \sim \mathrm{const} = \lambda_{\mathrm{max}}/\ln r, \tag{41}$$

where λ_{max} is the largest Lyapunov exponent. The predictability time for two-dimensional shell model is thus determined by a single value $1/\lambda_{\mathrm{max}} \sim \epsilon_Z^{-1/3}$, where ϵ_Z is the enstrophy flux toward small scales, up to a perturbation of the order of the large scale velocity field, where saturation effects are dominant. Figure 5 shows the finite-size Lyapunov exponent $\lambda(\delta v)$ for a simulation with $N = 24$ shells. The forcing term is now $f = 5 \times 10^{-4}(1+i)$ and $\nu = 10^{-8}$. Because of the inverse energy cascade we introduce in the equation (30) an artificial large-scale dissipation ν'/k_n ($\nu' = 10^{-5}$) in the first shells $n \le 3$ [26].

The velocity field shows a pseudo-cascade power law, see [30], $\langle |u_n|^2\rangle^{1/2} \sim \epsilon_Z^{1/3} k_n^{-1}$ in a wide range $1 \le k_n \le 10^3$. The mean enstrophy flux in this range is $\epsilon_Z = 5 \times 10^{-6}$ which is in agreement with the dimensional evaluation of the eddy turnover time $\tau_n \sim \epsilon_Z^{1/3}$.

The data plotted in figure 5 are obtained by using the method described in appendix A for computing the size dependent Lyapunov exponent. The same results, not reported, can be obtained from the doubling time algorithm.

We stress once more that, in the light of the results discussed in [30], two-dimensional shell model is not a good model for two-dimensional turbulence. As a consequence a discussion of the predictability problem for two-dimensional turbulence requires the direct study of Navier-Stokes equations [35]. Preliminary results show that this scenario remains, nevertheless, valid in the direct cascade [37].

IV. CLOSURE APPROXIMATION

In this Section we describe the results obtained from the eddy damped quasi-normal Markovian approximation (EDQNM) for the shell model. The basic idea of closure approximations is quite simple: write down the Reynolds hierarchy for moments of the shell variables and truncate the chain to the lowest sensible order. The important point is that in the closure approximation intermittent effects are washed out, so we can directly test if the relevant mechanism is due to the existence of many characteristic times. We do not report the derivation of the EDQNM equations for the shell model. The interested reader can find it in Ref. [38].

We consider two independent realizations of the shell model field, u_n and v_n, with the same energy spectrum $E_n = \langle u_n\, u_n^* \rangle = \langle v_n\, v_n^* \rangle$, and both evolving according the shell model equations (30)-(32). The distance between the two fields can be obtained from, cfr. (34), the energy difference at shell n:

$$\Delta_n = \frac{1}{2}\, \langle (u_n - v_n)\,(u_n^* - v_n^*) \rangle = (E_n - \Re\, W_n), \tag{42}$$

where $W_n = \langle u_n\, v_n^* \rangle$, and $\Re$ denotes the real part. . From the definition it follows

$$\delta v(t) = \left[\sum_n \Delta_n(t) \right]^{1/2} . \tag{43}$$

The evolution equations of E_n and W_n in the EDQNM approximation read:

$$\begin{aligned} \left(\tfrac{d}{dt} + 2\nu k_n^2\right)\, E_n = 2\ & \Big[k_n^2\, \theta(n,t)\, (E_{n+1}\, E_{n+2} - \tfrac{1}{2}\, E_n\, E_{n+2} - \tfrac{1}{2}\, E_n\, E_{n+1}) \\ & - \tfrac{1}{2}\, k_{n-1}^2\, \theta(n-1,t)\, (E_n\, E_{n+1} - \tfrac{1}{2}\, E_{n-1}\, E_{n+1} - \tfrac{1}{2}\, E_{n-1}\, E_n) \\ & - \tfrac{1}{2}\, k_{n-2}^2\, \theta(n-2,t)\, (E_{n-1}\, E_n - \tfrac{1}{2}\, E_{n-2}\, E_n - \tfrac{1}{2}\, E_{n-2}\, E_{n-1}) \Big] \\ & + 2\, \epsilon\, \delta_{n,4}, \end{aligned} \tag{44}$$

and

$$\begin{aligned} \left(\tfrac{d}{dt} + 2\nu k_n^2\right)\, W_n = 2\ & \Big[k_n^2\, \theta(n,t)\, (W_{n+1}^*\, W_{n+2}^* - \tfrac{1}{2}\, W_n\, E_{n+2} - \tfrac{1}{2}\, W_n\, E_{n+1}) \\ & - \tfrac{1}{2}\, k_{n-1}^2\, \theta(n-1,t)\, (W_n\, E_{n+1} - \tfrac{1}{2}\, W_{n-1}^*\, W_{n+1}^* - \tfrac{1}{2}\, E_{n-1}\, W_n) \\ & - \tfrac{1}{2}\, k_{n-2}^2\, \theta(n-2,t)\, (E_{n-1}\, W_n - \tfrac{1}{2}\, E_{n-2}\, W_n - \tfrac{1}{2}\, W_{n-2}^*\, W_{n-1}^*) \Big] \\ & + 2\, \epsilon\, \delta_{n,4}, \end{aligned} \tag{45}$$

where

$$\theta(n,t) = \frac{1 - e^{-[\nu(k_n^2 + k_{n+1}^2 + k_{n+2}^2) + \mu_n + \mu_{n+1} + \mu_{n+2}]\, t}}{\nu(k_n^2 + k_{n+1}^2 + k_{n+2}^2) + \mu_n + \mu_{n+1} + \mu_{n+2}}, \tag{46}$$

and

$$\mu_n \equiv \mu(k_n, E_n) = \alpha\, k_n\, E_n^{1/2}. \tag{47}$$

We have one free parameter, the dimensionless constant α. It should be adjusted such that the spectrum is as similar as possible to the spectrum obtained in simulations of the full equation. The energy spectrum of the shell model in the EDQNM approximation must therefore obey $E_n \simeq C(\alpha) \epsilon^{2/3} k_n^{-2/3}$ in the inertial range. The undetermined function $C(\alpha)$ is the Kolmogorov constant.

On the other hand it has become clear in several independent investigations that intermittency corrections exist in shell models. The energy spectrum is therefore in reality more closely described by $E_n \sim F(\epsilon) k_n^{-\zeta_2}$, where the exponent ζ_2 has been estimated to be 0.70 [27]. The function F that gives the prefactor to the power law in the inertial range should not depend on viscosity, but depends on the forcing through ϵ, the mean dissipation of energy per unit time, or, equivalently, the mean energy input into the system from the force. In a really large inertial range the two power-laws are not good approximations to one other. The best that can be done is to demand that the spectra agree as closely as possible at the upper end of the inertial range. A reasonable agreement is obtained for $\alpha = 0.06$, leading to $C(\alpha) = 1.5$ which is the value observed both in simulations of the shell model and in experiments [39].

The procedure described in the previous Section to compute the scale dependent Lyapunov exponent for the shell model can be adopted here for the closure equations. In practice after a long iteration time, to have a well stabilized energy spectrum E_n, we take a small initial distance $\delta v(0)$ and perform the doubling experiment similar to those of the previous section iterating equations (45).

In Fig. 6 we show $\langle 1/T_r(\delta v)\rangle/\mathrm{Re}^{1/2}$ as a function of the rescaled distance $\delta v/\mathrm{Re}^{-14}$, for different $\mathrm{Re} = \nu^{-1}$. The other parameters are $N = 32$ shells, $k_0 = 0.05$, integration step 10^{-6} and $r = 2^{1/2}$. From Fig. 6 we see that the closure approximation leads to the same scenario observed for the shell model, confirming that this is due to the existence of many characteristic scales. We note that the slope of the curve for $\delta v/\mathrm{Re}^{-1/4} > 10$ is the Lorenz value -2 since in the EDQNM approximation there are not intermittent corrections.

We note that in this case, since there is not intermittency, the effective inertial range roughly coincides with the inertial range. In Fig. 7 we compare $\langle 1/T_r(\delta v)\rangle$ and the inverse of the turnover time $\tau^{-1}(n) = k_n E_n^{1/2}$ as a function of the distance δv. In the figure we used $r = 2$.

V. CONCLUSION

We have introduced a generalization $\lambda(\delta)$ of the leading Lyapunov exponent to finite perturbations of size δ. Unlike the Lyapunov exponent, $\lambda(\delta)$ contains direct information on the predictability time for an extended chaotic system and it is particularly useful in presence of several characteristic times. Moreover, it is computationally no more expensive than the standard Lyapunov exponent.

In the limit of infinitesimal perturbation δ, the finite-size Lyapunov exponent gives the leading Lyapunov exponent. The way in which this limit is reached depends on the details of the particular system and gives informations about the characteristic time scales. In the case of $3D$ fully developed turbulence we have found an universal scaling law $\lambda(\delta) \sim \delta^{-2}$ where the value -2 of the exponent is an invariant of the multifractal approach.

The scaling law is confirmed by extensive numerical simulations on shell model for turbulence at very high Reynolds numbers, but we warn that it would be very difficult to observe such a scaling in experimental data because of the reduction of the scaling range in presence of intermittency.

We have also shown and discussed the relationship between our approach and the ϵ-entropy results in literature. In conclusion, we think that it would be useful to compute the finite-scale Lyapunov exponent in other complex dynamical systems, as a new tool for investigating the presence of characteristic times and the predictability properties.

VI. ACKNOWLEDGMENTS

This work was supported by the Swedish Natural Science Research Council through grant S-FO-1778-302 (E.A.), by INFN (*Iniziativa Specifica Meccanica Statistica FI11*) and by CNR (Progetto Coordinato *Predicibilitá in turbolenza geofisica*). E.A. thanks Dipartimento di Fisica, Università di Roma “La Sapienza” for hospitality. G. B. thanks the “Istituto di Cosmogeofisica del CNR”, Torino, for hospitality.

APPENDIX A: FINITE SIZE LYAPUNOV EXPONENTS AND ϵ-ENTROPY

Here we discuss some alternative ways of computing the finite-size Lyapunov exponent beyond the definition (3).

The first method is a modification of the standard technique [40,41]. We integrate two trajectories $\boldsymbol{u}(t)$ and $\boldsymbol{u}'(t)$ with initial Euclidean distance $\delta u(0) = \delta$ until their separation becomes larger, at a given time T_r than $r\delta$ where r is a given constant coefficient. The perturbed trajectory $\boldsymbol{u}'(T_r)$ is then rescaled at the original distance δ, keeping the direction $\boldsymbol{u}' - \boldsymbol{u}$ constant. The possible problem with this definition is that it assumes that the statistically stationary state of the system is homogeneous with respect to perturbations of finite size. One may plausibly argue that the structure of the attractor in phase space on which the motion takes place may be fractal, and not at all equally dense at all distances from a given point. The procedure outlined would then not necessarily sample in a faithful way the motion on the attractor. In practice, as it is showed by the numerical experiments, we do not find any difference with the numerical method presented in the main text, so the effect must be quite small. For the usual Lyapunov exponent the problem does not exist since we only use small finite perturbations as an approximation to infinitesimal perturbations.

The finite-size Lyapunov exponent at scale δ is obtained by averaging the divergence rate

$$\lambda(\delta) = \left\langle \frac{1}{T_r} \ln \left(\frac{\delta u(T_r)}{\delta u(0)} \right) \right\rangle = \frac{1}{\langle T_r \rangle_e} \ln r, \tag{A1}$$

along the unperturbed trajectory $\boldsymbol{u}(t)$. The average $\langle \cdots \rangle_e$ is over many error-doubling experiments.

In the case of infinitesimal error δ, and not too large factor r, this definition leads to the maximal Lyapunov exponent $\lambda_{\max}$.

For maps, the above methods have to be slightly modified since the distance between two states is not continuous in time. From eq. (37) readily follows

$$\lambda(\delta) = \frac{1}{\langle T_r \rangle_e} \left\langle \ln \left(\frac{\delta(T_r)}{\delta} \right) \right\rangle_e , \tag{A2}$$

where T_r is the minimum time such that the distance between two realizations is larger or equal to $r\delta$ and $\delta(T_r)$ is the distance at that time.

Definition (A1) is valid for any value of r, which however is generally assumed not too large to avoid the interference of different scales. In particular, one could think of removing the threshold condition used for defining T_r and simply compute the average error growth rate at every time step. In other words, at every time step δt in the integration, the perturbed trajectory $\boldsymbol{u}'(t)$ is rescaled to the original distance δ, keeping the direction $\boldsymbol{u} - \boldsymbol{u}'$ constant. The finite-size Lyapunov exponent is given by the average of the one-step exponential divergence:

$$\lambda(\delta) = \frac{1}{\delta t} \left\langle \ln \left(\frac{\delta u(t+\delta t)}{\delta u(t)} \right) \right\rangle_t , \tag{A3}$$

which is equivalent to the above definition (A1). This method is indeed the one used for computing $\lambda(\delta)$ in the case of $2D$ shell model. The one-step method (A3) has the advantage that it can be easily generalized to compute the sub-leading finite-size Lyapunov exponent following the standard orthonormalization method [40]. One introduces k perturbed trajectories $\boldsymbol{u}^{(1)}, \ldots, \boldsymbol{u}^{(k)}$ each at distance δ from $\boldsymbol{u}$ and such that $\boldsymbol{u}^{(k)} - \boldsymbol{u}$ are orthogonal each to the others. At every time step, any difference $\boldsymbol{u}^{(k)} - \boldsymbol{u}$ is rescaled at the original value and orthogonalized, while the corresponding finite size Lyapunov exponent is accumulated according to (A3). Here we have again the problem of the implicitly assumed homogeneity of the attractor, but also a problem of isotropy when we re-orthogonalize the perturbations. We note that this could be a more serious problem, which will not be discussed here any further.

For systems with only one positive Lyapunov exponent the size dependent Lyapunov exponent $\lambda(\delta)$, definition (A2), or (3), coincides with the ϵ-entropy widely described below in Appendix B and by Gaspard and Wang [12,13]. We consider now few examples.

1. Map with noise

Consider a chaotic deterministic map perturbed by a Gaussian term:

$$x(n+1) = f(x(n)) + \sigma y(n), \tag{A4}$$

where $y(n)$ is a stochastic variable with Gaussian distribution of zero mean and unit variance. The $y(n)$ for different times n are independent. When $\sigma = 0$ the map $f(x)$ is chaotic with positive Lyapunov exponent λ. A simple computation shows that

$$\delta x(n) \sim \delta x(0)\, e^{\lambda n} + \sigma\, \delta y(n-1). \tag{A5}$$

For $|\delta x(0)| \ll \sigma$ the noise term is negligible, so that $\lambda(|\delta x(0)|) \simeq \lambda$. In the opposite limit, $|\delta x(0)| \gg \sigma$, the first term in (A5) can be neglected so that $\delta x(n) \sim \sigma\, \delta y(n-1)$. Thus in one iteration of the map the distance between the two trajectories grows to $O(\sigma)$, larger than the tolerance. From (A2) we have $\lambda(|\delta x(0)|) \sim \ln(\sigma/|\delta x(0)|)$. These are the same result obtained for the ϵ-entropy, see Sect. 3.5 of Ref. [13].

2. Ornstein-Uhlenbeck process and Yaglom noises

Consider the Gaussian process described by the Langevin equation

$$\frac{dx}{dt} = -ax + c\eta, \tag{A6}$$

where $a > 0$ and η is a white noise with zero mean and correlation $\langle \eta(t)\,\eta(t')\rangle_\eta = \delta(t-t')$. The formal solution of (A6) reads

$$x(t) = e^{-at}x(0) + c\int_0^t e^{-a(t-t')}\,\eta(t')\,dt', \tag{A7}$$

so that the distance between two different process behaves in time as

$$\delta x(t) = e^{-at}\delta x(0) + c\int_0^t e^{-a(t-t')}\,\delta\eta(t')\,dt'. \tag{A8}$$

This implies that $\delta x(t) \sim \sqrt{ct}$, and thus the predictability time $T(\delta x(0),\Delta)$ behaves as $T(\delta x(0),\Delta) \sim (\Delta/c)^2$ so that $\lambda(\delta x(0),\Delta) \sim (c/\Delta)^2 \sim (c/\delta x(0))^2$, as found for the ϵ-entropy, see below or Sect. 3.6.2 of [13].

Similar computations can be done for the Yaglom noise, see Sect. 3.6.3 of [13].

3. Model of deterministic diffusion

The one-dimensional map

$$x(n+1) = x(n) + p\sin(2\pi x(n)), \tag{A9}$$

presents deterministic diffusion. For example for $p = 0.8$ the diffusion coefficient is $D \simeq 0.18$. The size dependent Lyapunov exponent for this map can be computed numerically using the definition (A2). From Fig. 8 we see that $\lambda(\delta) \simeq \lambda$ for small δ, while $\lambda(\delta) \sim \delta^{-2}$ for large values of δ. The ϵ-entropy shows the same behavior, see Fig. 25.b of Ref. [13]. Note that in Ref. [13] the ϵ-entropy is measured in unit of digit/iteration, so there is a multiplicative factor $\ln 10 \simeq 2.3$ with $\lambda(\delta)$ of Fig. 8.

APPENDIX B: THE KOLMOGOROV RESULT ON ϵ-ENTROPIES OF GAUSSIAN PROCESSES

The purpose of this appendix is to derive systematically on a physical level of rigor the results of Kolmogorov on the ϵ-entropy of Gaussian variables and Gaussian processes and random fields. These results were stated without proof by Kolmogorov [11], and in the review by Tikhomirov [42]. The published proofs we are aware of are either not easily accessible, or carried out in such generality that the simple underlying idea may be lost to the less mathematically inclined reader. Hence the interest of including a simple derivation here. Let us just for completeness refer to comparatively recent paper [43].

The general setting is that of two random variables ξ and η, taking values in spaces X and Y. The values in two realizations of ξ and η are denoted x and y, respectively, where x lies in X and y lies in Y. When X and Y are continuous the probability distributions of ξ and η will be denoted $P_\xi(dx)$ and $P_\eta(dy)$. Except for the trivial case, the variables ξ and η are not independent, but characterized by the joint probability distribution $P_{\xi,\eta}(dx,dy)$.

The single-variable probabilities are determined from the joint distribution as

$$P_\xi(dx) = \int_y P_{\xi,\eta}(dx,dy) \qquad P_\eta(dy) = \int_x P_{\xi,\eta}(dx,dy) \tag{B1}$$

The conditional probabilities are

$$P_{\xi|\eta}(dx) = \frac{P_{\xi,\eta}(dx,dy)}{P_\eta(dy)} \qquad P_{\eta|\xi}(dy) = \frac{P_{\xi,\eta}(dx,dy)}{P_\xi(dx)} \tag{B2}$$

1. Preliminaries from information theory

We suppose in this subsection that the spaces X and Y are discrete, and consist of finitely many points, $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$, and associated probabilities $p_1, \ldots, p_n$ and $q_1, \ldots, q_m$.

The *entropy* of the random variable ξ is

$$H(\xi) = -\sum_{i=1}^{n} p_i \ln p_i \tag{B3}$$

If we want to code a message of N letters, which are given by N consecutive independent realizations of ξ, so taking values in $x_1, \ldots x_n$, this can be done, in the limit when N is large, by using $H(\xi)/\ln 2$ bits per letter [10]. The pair ξ and η specify both a source signal and the output after sending the signal through a channel of transmission. The joint probability distribution $P_{\xi,\eta}$ can be decomposed either to the pair ξ (an input) and $\eta|\xi$ (an output, given the input), or to the pair η and $\xi|\eta$. In the discrete case these second conditional probabilities are denoted $p_{i|j}$, where i ranges from one to n, and j ranges from one to m. The *equivocation* is

$$\langle H(\xi|\eta)\rangle_y = \sum_j p_j \left(-\sum_i p_{i|j} \ln p_{i|j} \right). \tag{B4}$$

Shannon [10] considered a *gedanken* experiment where we send an error-correcting message parallel to the transmission of ξ to η, and asked for the number of bits in the error-correcting message needed to transmit N letters in the original message virtually without error. In the limit of large N, the answer is $\langle H(\xi|\eta)\rangle_y / \ln 2$ bits per letter.

Suppose further that we have some source of information which we can recode into letters from ξ, then transmit to η, and then observe. The rate of transmission of information is not the source entropy, since we must correct for the equivocation, but the *mutual information*:

$$I(\xi,\eta) = H(\xi) - \langle H(\xi|\eta)\rangle_y. \tag{B5}$$

If we introduce the discrete joint probabilities $p_{i,j}$ and rearrange terms, we see that (B5) can be rewritten as in a more symmetrical way, namely:

$$I(\xi,\eta) = \sum_{i,j} p_{i,j} \ln \frac{p_{i,j}}{p_i p_j} \tag{B6}$$

A channel of transmission of information can be considered as a collection of pairs ξ and η, where the ξ's are the possible inputs, and the η's the corresponding outputs. The *capacity* of the channel is the maximum rate of information transfer:

$$C = \max_{\xi,\eta} I(\xi,\eta) \tag{B7}$$

where the maximization is performed over the collection of pairs ξ and η that describe the channel. The fundamental Shannon theorem says that a source signal can be transmitted over a channel, if the source entropy is less than the channel capacity.

Suppose finally that we have an input ξ and we wish that the output η contains only some partial information about ξ, saying that some fidelity criterion is fulfilled. We consider all channels consisting of the fixed source ξ and different outputs η. The least rate of information transfer needed to specify ξ in this way is *source entropy with respect to the fidelity criterion*:

$$R = \min_{\eta} I(\xi,\eta) \qquad P_\xi \text{ fixed} \tag{B8}$$

where the minimization is performed over all η which satisfy the criterion with ξ.

In practice many fidelity criteria can be written

$$G(\xi,\eta) \leq \epsilon, \tag{B9}$$

where G is some suitable function and ϵ is a measure of the fidelity. In this case the source entropy is naturally considered as a function of ϵ, and we have the Kolmogorov ϵ-entropy

$$H(\xi,\epsilon) = \min_{\eta} I(\xi,\eta), \quad \text{where} \quad G(\xi,\eta) \leq \epsilon \quad \text{and } P_\xi \text{ fixed}. \tag{B10}$$

2. Continuous random variables

Suppose we have a random variable with a continuous distribution, that is $P_\xi(dx) = p_\xi(x)dx$ with continuous density $p_\xi(x)$. Then (B3) is infinite.

As stressed by Kolmogorov, the mutual information, (B5), the capacity, (B7), and, in particular, the rate of information production with respect to a fidelity criterion, (B8,B10), are all well-defined also in the continuous case. In other words, although a real number observed with infinite accuracy contains an infinite amount of information, a real number observed with finite accuracy contains only a finite amount of information.

We can see this in a simple heuristic way as follows: if we introduce a discretization of the space X into boxes with diameter δ, we have a new random variable that we can call ξ_δ with entropy

$$H(\xi_\delta) = d\ln\left(\frac{1}{\delta}\right) - \int p_\xi(x)\ln p_\xi(x)dx + \mathcal{O}(\delta), \tag{B11}$$

where d is the dimension of X.

Similarly, from the conditional variable $\xi|\eta$ we have a new random variable $\xi_\delta|\eta$ with equivocation

$$\langle H(\xi_\delta|\eta)\rangle_y = d\ln\left(\frac{1}{\delta}\right) - \int p_\eta(y)\left(\int p_{\xi|\eta}(x,y)\ln p_{\xi|\eta}(x,y)dx\right)dy + \mathcal{O}(\delta). \tag{B12}$$

Rearranging terms as in (B6), the mutual information between ξ_δ and η is thus

$$I(\xi_\delta, \eta) = \int p_{\xi,\eta}(x,y)\,\ln\left(\frac{p_{\xi,\eta}(x,y)}{p_\xi(x)p_\eta(y)}\right)\,dxdy + \mathcal{O}(\delta). \tag{B13}$$

As the discretization tends to zero the mutual information tends to a finite value, provided that the joint probability $P_{\xi,\eta}(dx,dy)$ is not singular with respect to the product $P_\xi(dx)P_\eta(dy)$, a result due to Yaglom and Gel'fand [11].

We can give a meaning to the most random continuous distribution with respect some given constraints by maximizing the entropy of a discretized continuous variable, as in (B11), subject to these constraints, and then letting the discretization tend to zero. The most random distribution with given first and second moments is thus, not surprisingly, a Gaussian distribution with the same first and second moments. The entropy of a discretization of a d-dimensional real Gaussian random variable ξ with correlation matrix C is

$$H(\xi_\delta) = d\ln\left(\frac{1}{\delta}\right) + \ln\sqrt{(2\pi e)^d \det C} + \mathcal{O}(\delta). \tag{B14}$$

All these results on the Gaussian distributions are due to Shannon [10].

3. The ϵ-entropy result

The spaces X and Y are now identical. We consider first a one-dimensional Gaussian random variable, and then a finite-dimensional variable and show that the Kolmogorov formula holds in these cases. In the next subsection, we then observe that the Fourier components of Gaussian processes and Gaussian random fields are independent Gaussian random variables, and so, taking the the formal infinite-dimensional limit, we find the desired result.

The fidelity criterion will be the mean square deviation:

$$\langle(\xi - \eta)^2\rangle \le \epsilon^2 \tag{B15}$$

Let us first estimate the answer by dimensional arguments. If the finite-dimensional normal variable ξ is decomposed in its principal components, a fluctuation in direction i will be of typical size σ_i. To estimate the fluctuation in direction i with an accuracy ϵ we need about $\ln(\sigma_i/\epsilon)$ bits. The answer should therefore be of the form

$$H(\xi, \epsilon) = A\sum_{\sigma_i > \epsilon}\ln\left(\frac{\sigma_i}{\epsilon}\right) + B. \tag{B16}$$

The result of the analysis will be that A is one and B is zero.

Let ξ be a Gaussian one-dimensional random variable with variance σ^2. The random variable $\xi|\eta$ is defined on the support of P_η. We suppose it to have mean $m(y)$ and variance $\alpha(y)$. We have the obvious inequality

$$\int \alpha(y)\, P_\eta(dy) \le \langle \xi^2 \rangle, \tag{B17}$$

and the equality

$$\int \alpha(y)\, P_\eta(dy) = \langle (\xi - \eta)^2 \rangle. \tag{B18}$$

On the other hand, the equivocation, (B12), with fixed outcome y of η and given mean $m(y)$ and variance $\alpha(y)$ is maximized if the variable is Gaussian with variance $\alpha(y)$. Therefore, the mutual information between ξ and η is bounded from below by

$$I(\xi, \eta) \ \ge \int \frac{1}{2} \ln \left(\frac{\sigma^2}{\alpha(y)} \right) \ P_\eta(dy). \tag{B19}$$

We can therefore pose a constrained minimization problem over the auxiliary positive function $\alpha(y)$. Strictly speaking, $H(\xi, \epsilon)$ is only bounded from below by the minimization, but, as we will see, the solution can be realized in terms of variables ξ and η with desired properties.

$$H(\xi, \epsilon) = \min_{\alpha(y)} \int \frac{1}{2} \ln \left(\frac{\sigma^2}{\alpha(y)} \right) \ P_\eta(dy), \tag{B20}$$

$$\epsilon^2 \ge \int \alpha(y)\, P_\eta(dy), \tag{B21}$$

$$\sigma^2 \ge \int \alpha(y)\, P_\eta(dy). \tag{B22}$$

Either one or the other of the two constraints will be satisfied at the minimum in (B20). A variation with Lagrange multipliers gives that $\alpha(y)$ must be constant, either equal to ϵ^2 or to σ^2. We therefore have

$$H(\xi, \epsilon) = \max \left[\ln \left(\frac{\sigma}{\epsilon} \right), 0 \right]. \tag{B23}$$

The derivation also shows clearly that the solution can be realized as follows: if σ is less than ϵ, then $\xi|\eta$ is Gaussian with mean zero and variance σ^2, and η a delta function with support at the origin. If, on the other hand, σ is greater than ϵ, then $\xi|\eta$ is Gaussian with mean $m(y)$ equal to y and variance ϵ^2, and η is Gaussian with mean zero and variance $\sigma^2 - \epsilon^2$. By the semigroup property of Gaussian kernels follows in both cases that ξ is Gaussian with mean zero and variance σ^2.

The generalization to higher-dimensional case is as follows. The variable ξ is Gaussian with second moments C_{ij}, with principal components in a diagonal basis $(\sigma_1^2, \ldots, \sigma_d^2)$. At a point $y_1, \ldots, y_d$ in the support of P_η we consider the variable $\xi|\eta$ with first moments $m_i(y_1, \ldots, y_d)$ and second moments $\alpha_{ij}(y_1, \ldots, y_d)$. We have, in analogy with the one-dimensional case,

$$\int \alpha_{ii}(y_1, \ldots, y_d)\, P_\eta(dy) \le \langle \xi_i^2 \rangle = \sigma_i^2, \tag{B24}$$

and

$$\int \sum_i \alpha_{ii}(y_1, \ldots, y_d)\, P_\eta(dy) = \langle (\xi - \eta)^2 \rangle \le \epsilon^2. \tag{B25}$$

The mutual information between ξ and η is bounded by

$$I(\xi, \eta) \ \ge \int \frac{1}{2} \left(\ln \det C_{ij} - \ln \det \alpha_{ij} \right) \ P_\eta(dy) \tag{B26}$$

Some of the inequalities (B24) may hold as equalities in the solution, some not. It does however follow from the diagonal structure of (B24) and (B25) and the relation

$$\delta \ln \det \alpha = \mathrm{Tr}\ \frac{\delta \alpha}{\alpha}, \tag{B27}$$

that when (B26) is minimized, $\alpha_{ij}(y)$ must be constant in y and diagonal in i and j. Let the diagonal elements of α be D_i^2. We then have the following simpler discrete constrained minimization problem:

$$H(\xi,\epsilon) = \min \sum_i \ln\left(\frac{\sigma_i}{D_i}\right), \tag{B28}$$

$$\epsilon^2 \geq \sum_i D_i^2, \tag{B29}$$

$$\sigma_i^2 \geq D_i^2. \tag{B30}$$

The minimum can be found by starting with all D_i's very small and increasing them proportionally to the gradient, that is, at the same rate. When one of the D_i's hits the constraint (B30) we keep it constant from thereon and increase the others. The constraint (B29) will eventually be fulfilled when the D_i's that still change are equal to θ, and there we stop. The solution is thus implicitly given in terms of θ, which can be interpreted as a cut-off threshold for modes where normal fluctuations are small:

$$H(\xi,\epsilon) = \sum_i \max\left[\ln\left(\frac{\sigma_i}{\theta}\right), 0\right], \qquad \epsilon^2 = \sum_i \min\left[\sigma_i^2, \theta^2\right] \tag{B31}$$

The solution can be realized by taking η a random variable with independent components, diagonal in the same basis as ξ, and letting the i'th component of ξ only depend on the i'th component of η. The variables η_i and $\xi_i|\eta_i$ are then constructed as in the one-dimensional case, with the only difference that the parameter θ substitutes for the error ϵ.

4. Gaussian random fields

Let us now consider a scalar Gaussian random field in D-dimensional space. A particular example ($D = 1$) is Gaussian processes. We begin by the approximation that the field is periodic with period L in all directions. Fourier components of Gaussian random fields are independently distributed Gaussian random variables. Therefore (B31) can be rewritten, using the volume element Δk, equal to $(2\pi/L)^D$:

$$H(\xi,\epsilon)\left(\frac{2\pi}{L}\right)^D = \sum_k \max\left[\ln\left(\frac{\sigma_k}{\theta}\right), 0\right] \Delta k, \tag{B32}$$

$$\epsilon^2\left(\frac{2\pi}{L}\right)^D = \sum_k \min\left[\sigma_k^2, \theta^2\right] \Delta k. \tag{B33}$$

In the limit when L tends to infinity the right hand sides turn into integrals, and the left hand sides to entropy and mean-square distance per unit volume. Therefore we have

$$h^{\text{volume}}(\xi,\epsilon) = \left(\frac{1}{2\pi}\right)^D \int \max\left[\frac{1}{2}\ln\left(\frac{E(\boldsymbol{k})}{\theta^2}\right), 0\right] d^D k, \tag{B34}$$

$$\epsilon^2 = \left(\frac{1}{2\pi}\right)^D \int \min\left[E(\boldsymbol{k}), \theta^2\right] d^D k. \tag{B35}$$

which, for D equal to one, is the Kolmogorov result for the ϵ-entropy per unit time of a Gaussian random process.

We note that θ is still a cut-off to eliminate modes such that the energy of the fluctuations in these modes is less than θ^2. In fields with a power-law spectrum, such as turbulence, θ can simply be substituted for a wave-number cut-off K.

5. Other distributions and other fidelity criteria

The mean square fidelity criterion is convenient for analytical computations, but it is not the only one possible. Shannon lists also the maximum distance, another choice would be the mean absolute distance.

More interestingly, there may be cases where the most appropriate fidelity criteria is not simply proportional to the distance in the mean or in maximum value. If the signal ξ is a Lévy-process it will have a non-negligible probability of changing by a large amount over a short interval of time [44]. In many applications one would then like an approximating signal η to capture well the large jumps, but one would be less interested in a very precise approximation when ξ changes comparatively little [45]. A reasonable fidelity criterion would then be that there is only a small probability that the distance between ξ and η is large.

As far as we know no investigations have been performed of the entropy of such sources with respect to such fidelity criteria. Of course, for a Lévy process a mean square error function is not possible, since the second moment is infinite. It would be possible to use mean absolute error, as was done implicitly by Gaspard and Wang in [13], but the relevance of such computation would have to be motivated in each case by the application at hand.

APPENDIX C: SPACETIME PROCESSES

In this appendix we want to discuss and compare the finite-size Lyapunov exponent and the ϵ-entropy for a turbulent flow and for the shell model. We assume that the flow and the shell model are both Gaussian as in appendix B, and we assume that the spectrum follows the Kolmogorov theory, as described in section II. Non Gaussian effects and intermittency correction to the spectrum are not taken into account.

At a given error threshold δ, we have seen that the finite-size Lyapunov exponent scales as $\lambda(\delta)\sim\delta^{-2}$, and this holds as well for the shell model as for $3D$ turbulence. Let us now consider the ϵ-entropy, and first give a simple dimensional estimate. It will then be seen that the expressions (B34) (B35) just reproduce this result. The dimensional estimate starts from the time scales in the Kolmogorov theory, $\tau(k)\sim k^{-\frac{2}{3}}$. The distance between two fields that only differ in wave numbers greater than k is $\delta^2\sim k^{-\frac{2}{3}}$, that is, it does not depend on the dimensionality of space, but only of the form of the spectrum $E(k)\sim k^{-\frac{5}{3}}$. On the other hand, the number of degrees of freedom at wave numbers less than or equal to k is proportional to k^D. The system must be observed at a rate $\tau(k)^{-1}$ to capture all the motion in wave numbers less than k. The amount of information per unit time and space needed to describe the system up to an error tolerance δ is thus $k^D\tau(k)^{-1}$. Translating this into a functional dependence on δ we have

$$h^{\text{space,time}}(\delta)\sim\delta^{-2-3D} \tag{C1}$$

which, for $D=3$, leads to $h^{\text{space,time}}(\delta)\sim\delta^{-11}$. As far as we can see there is an inconsistency between this result, also derived in [13], and the result stated in [12], which we believe is a misprint resulting from using an argument that really applies to the finite-size Lyapunov exponent and not the ϵ-entropy. The very different scaling laws $h\sim\delta^{-11}$ and $\lambda(\delta)\sim\delta^{-2}$ is also a further motivation why introducing the quantity finite-size Lyapunov exponent.

This straight-forward dimensional analysis can easily be generalized to a generic stationary Gaussian processes with spectral density (see section 6.2.6 of [13]),

$$\Phi(\mathbf{k},\omega)\sim k^{-y}F\left(\frac{\omega}{k^z}\right), \tag{C2}$$

The function F is supposed to vanish for argument much larger and much smaller than unity, and its integral is on the order of unity. The energy spectrum is therefore

$$E(k)\sim k^{D-1}\int\Phi(\mathbf{k},\omega)d\omega\sim k^{z-y+D-1}. \tag{C3}$$

We will here check that the previous dimensional estimates can also be obtained from the Kolmogorov formulae (B34) and (B35), that we write in this case as follows:

$$h^{\text{space,time}}(\delta)=\left(\frac{1}{2\pi}\right)^{D+1}\int\max\left[\frac{1}{2}\ln\left(\frac{\Phi(\mathbf{k},\omega)}{\theta^2}\right),0\right]d^Dk\,d\omega, \tag{C4}$$

$$\delta^2=\left(\frac{1}{2\pi}\right)^{D+1}\int\min\left[\Phi(\mathbf{k},\omega),\theta^2\right]d^Dk\,d\omega. \tag{C5}$$

The first observation is that the integral over ω in (C5) gives a contribution of order k^{z-y} if k^{-y} is less than θ^2, but otherwise a contribution of order $k^z\theta^2$. The error is thus determined by a cut-off wave number K such that

$$\delta^2\sim K^{z-y+D}\qquad K^{-y}\sim\theta^2 \tag{C6}$$

The second observation is that the integral over ω in (C4) gives a vanishing contribution is k^{-y} is less than θ^2, that is, for wave numbers larger than the cut-off K. For smaller wave numbers the integration over ω gives a contribution of the order of k^z, up to a logarithmic correction that we don't take into account. The ϵ-entropy with cut-off wave number K is thus

$$h^{\text{space,time}}(\delta) \sim K^{z+D} \qquad K^{-y} \sim \theta^2 \tag{C7}$$

Combining (C5) and (C4) we find

$$h(\delta) \sim \delta^{2\,(D+z)/(z+D-y)} \tag{C8}$$

which is equation (6.14) of [13]. By inserting the values $D = 3$, $y = 13/3$ and $z = 2/3$ of the Kolmogorov theory we reproduce the result (C1).

---

[1] V.I. Arnold, Preface to *Geometrical Methods in the Theory of Ordinary Differential Equations*, Grundlehren der mathematischen Wissenschaften **250**, Springer-Verlag (1983).
[2] E. Lorenz, J. Atmos. Sci. **20**, 130 (1963).
[3] A. Monin, *Weather prediction as a problem in physics.* (MIT Press, 1973).
[4] E. Aurell, G. Boffetta, A. Crisanti, G. Paladin and A. Vulpiani, "Growth of non-infinitesimal perturbations in turbulence", chao-dyn/9604007.
[5] J.-P. Eckmann and D. Ruelle, Rev. Mod. Physics **53**, 617 (1985).
[6] A. Pikovsky, Chaos **3**, 225 (1993).
[7] A. Crisanti, M.H. Jensen, G. Paladin and A. Vulpiani, Phys. Rev. Lett. **70**, 166 (1993); A. Crisanti, M.H. Jensen, G. Paladin and A. Vulpiani, J. Phys A **26**, 6943 (1993).
[8] G. Paladin and A. Vulpiani, Phys. Rep. **156**, 147 (1987).
[9] G. Paladin and A. Vulpiani, J. Physics A **27**, 4911 (1994).
[10] C. Shannon and W. Weaver, *The mathematical theory of communication*, The University of Illinois Press, Urbana, (1949).
[11] A. N. Kolmogorov, IRE Trans. Inf. Theory **1**, 102 (1956).
[12] P. Gaspard and X.-J. Wang, Phys. Rev. **46 A**, R3000 (1992).
[13] P. Gaspard and X.-J. Wang, Physics Reports **235**, 291-343 (1993).
[14] A. N. Kolmogorov, Dokl. Akad. Nauk SSSR **30**, 9-13 (1941).
[15] G. Parisi and U. Frisch, in *Turbulence and predictability in geophysical flows and climatic dynamics*, ed. N. Ghil, R. Benzi and G. Parisi (North-Holland, 1985).
[16] R. Benzi, G. Paladin, G. Parisi and A. Vulpiani, J. Phys. **A17**, 3521 (1984).
[17] U. Frisch, *Turbulence after 50 years: the legacy of A.N.Kolmogorov*, Cambridge University Press (1995).
[18] L. Landau and E. Lifshitz, *Course On Theoretical Physics, Volume 6: Fluid Mechanics*, Pergamon Press (1959).
[19] G. Paladin and A. Vulpiani, Physical Review **A35**, 1971 (1987)
[20] E. Lorenz, Tellus **21**, 3 (1967); D.K. Lilly, in *Dynamic Meteorology*, ed. P. Morel, Riedel, Boston (1973); C.E. Leith and R.H. Kraichnan, J. Atmos. Sci. **29**, 1041 (1972).
[21] U. Frisch and M. Vergassola, Europhys. Lett. **14**, 439 (1991); M.H. Jensen, G. Paladin and A. Vulpiani; Phys. Rev. Lett. **67**, 208 (1991).
[22] T. Bohr, M. Jensen, G. Paladin and A. Vulpiani, *Dynamical Systems Approach to Turbulence*, Cambridge Nonlinear Science Series, Cambridge University Press (to appear).
[23] E.B. Gledzer, Sov. Phys. Dokl. **18**, 216 (1973).
[24] M. Yamada and K. Ohkitani, J. Phys. Soc. of Japan **56**, 4210 (1987); M. Yamada and K. Ohkitani, Phys. Rev. Lett. **60**, 983 (1988).
[25] D. Pisarenko, L. Biferale, D. Courvoisier, U. Frisch, M. Vergassola, Phys. Fluid A **5**, 2533 (1993).
[26] M. Yamada and K. Ohkitani, Progr. Theo. Phys. **79**, 1265 (1988).
[27] M.H. Jensen, G. Paladin and A. Vulpiani, Phys.Rev. A **43**, 798 (1991); M.H. Jensen, G. Paladin and A. Vulpiani, Phys.Rev. A **45**, 7214 (1992).
[28] L. Kadanoff, D. Lohse, J. Wang and R. Benzi, Phys. Fluids **7**, 617 (1995).
[29] L. Biferale, A. Lambert, R. Lima and G. Paladin, Physica D **80**, 105 (1995).
[30] E. Aurell, G. Boffetta, A. Crisanti, P. Frick, G. Paladin and A. Vulpiani, Phys. Rev. E **50**, 4705 (1994).
[31] P. D. Ditlevsen and I. A. Morgensen, Phys. Rev E **53**, 4785 (1996).
[32] L. Biferale and R. Kerr, Phys. Rev. E **52**, 6113 (1995).
[33] R. Benzi, L. Biferale, R. Kerr and E. Trovatore, Phys. Rev. E **53**, 3541 (1996).
[34] D. Ruelle, Phys. Lett. A **72**, 81 (1979).
[35] G. Boffetta, A. Celani, A. Crisanti and A. Vulpiani, Phys. Fluids A (submitted, 1996).

[36] R.H. Kraichnan and D. Montgomery, Rep. Prog. Phys. **43**, 547 (1980).
[37] G. Boffetta, A. Celani and A. Crisanti, preliminary results.
[38] E. Aurell, G. Boffetta, A. Crisanti, G. Paladin and A. Vulpiani, Phys. Rev. E **53**, 2337 (1996).
[39] S. A. Orszag, in *Fluid Dynamics*, ed. R. Balian and J. L. Peuble (Gordon and Breach, 1977).
[40] G. Benettin, L. Galgani, A. Giorgilli and J.M. Strelcyn, Meccanica **15**, 9 (1980).
[41] A. Wolf, J.B. Swift, H.L. Swinney and J.A. Vastano, Physica D **16**, 285 (1985).
[42] V.M. Tikhomirov *Some Aspects of Approximation Theory* (in Russian) (Izd. MGU, Moscow, 1976)
[43] M.S. Pinsker and L.B. Sofman, Problemy Peredachi Informatsii **19**, 52-67 (1983) [in Russian]; Problems of Information Transfer, 214-226 (1984) [English translation].
[44] P. Lévy, *Théorie de l'addition des variables aléatoires*, Gauthier-Villars, Paris, (1937:1954).
[45] J.-P. Bouchaud and A. Georges, Phys. Rep. **195**, 127 (1990).
[46] U. Frisch, Z.-S. She and O. Thual, J. Fluid Mech. **168**, 221 (1986).

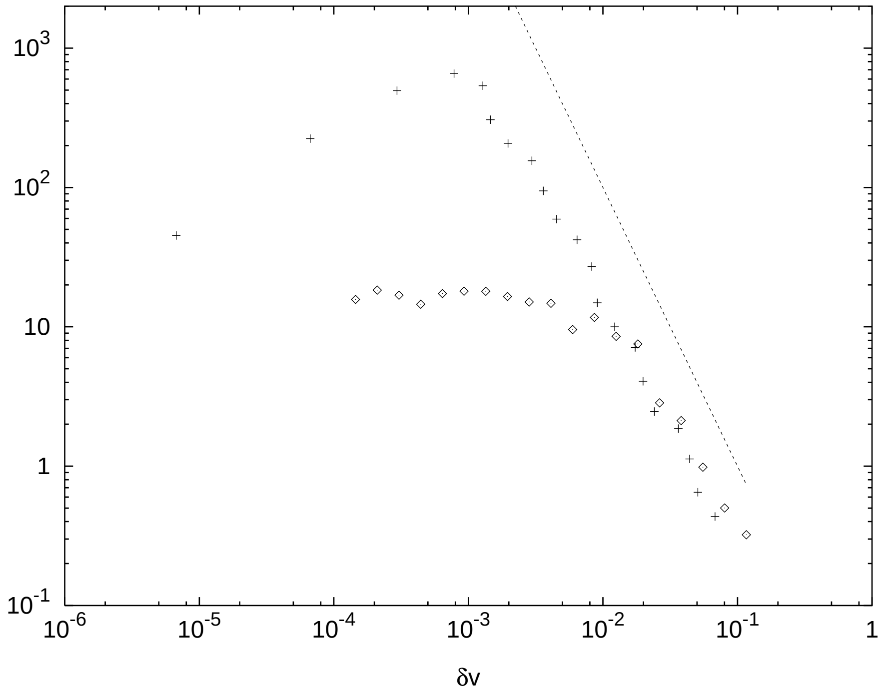

FIG. 1. Error doubling times (diamond) compared with shell turn-over times (plus). Number N of simulated shells is 27, and Reynolds number $\mathrm{Re} = \nu^{-1} = 10^9$, $k_0 = 0.05$ and $f = (1+i) \times 0.005$. The equations were integrated with a slaved-frog scheme [25,46], with constant time-step $2 \cdot 10^{-6}$. The initial perturbation was randomly uniform over all shells in the inertial range, with amplitude less than 10^{-6}. The perturbed and unperturbed configurations were integrated until the error reach the first threshold δ_0 at 10^{-4}. The error growth rate parameter r is $2^{1/2}$. The number of error doubling experiments was 400. The dashed line has slope -2.

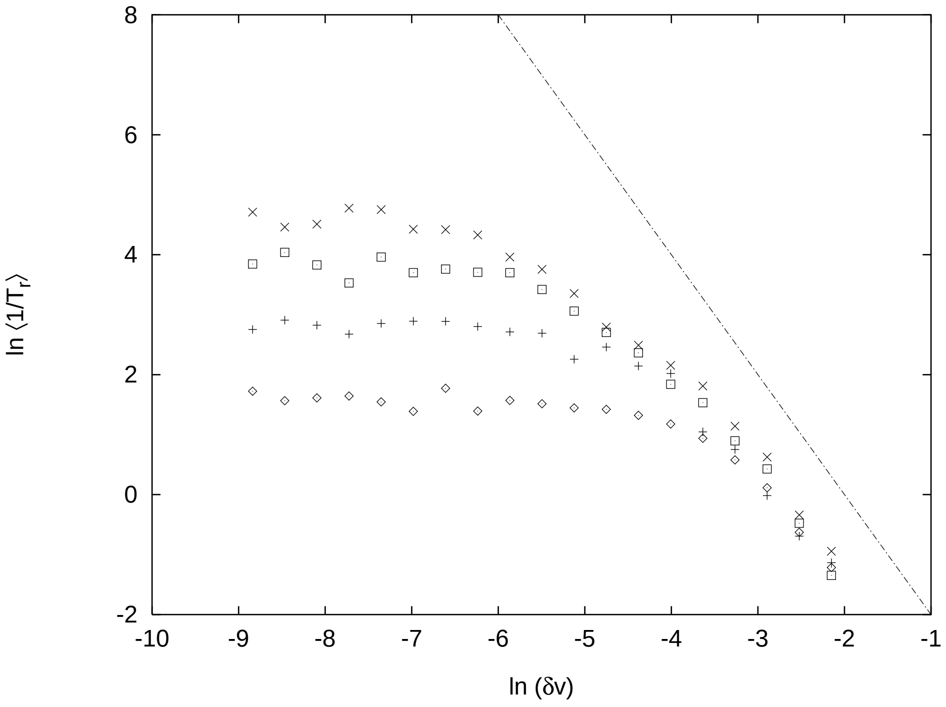

FIG. 2. $\ln\langle 1/T_r(\delta v)\rangle$ versus $\ln(\delta v)$ for different Reynolds numbers $\mathrm{Re} = \nu^{-1}$. Parameters as in Fig.1, except that the times-step has been adjusted to the changing viscosity. The different symbols refer to: $N = 24$ and $\nu = 10^{-8}$ (diamond); $N = 27$ and $\nu = 10^{-9}$ (plus); $N = 32$ and $\nu = 10^{-10}$ (square); $N = 35$ and $\nu = 10^{-11}$ (cross). The straight line has slope -2.

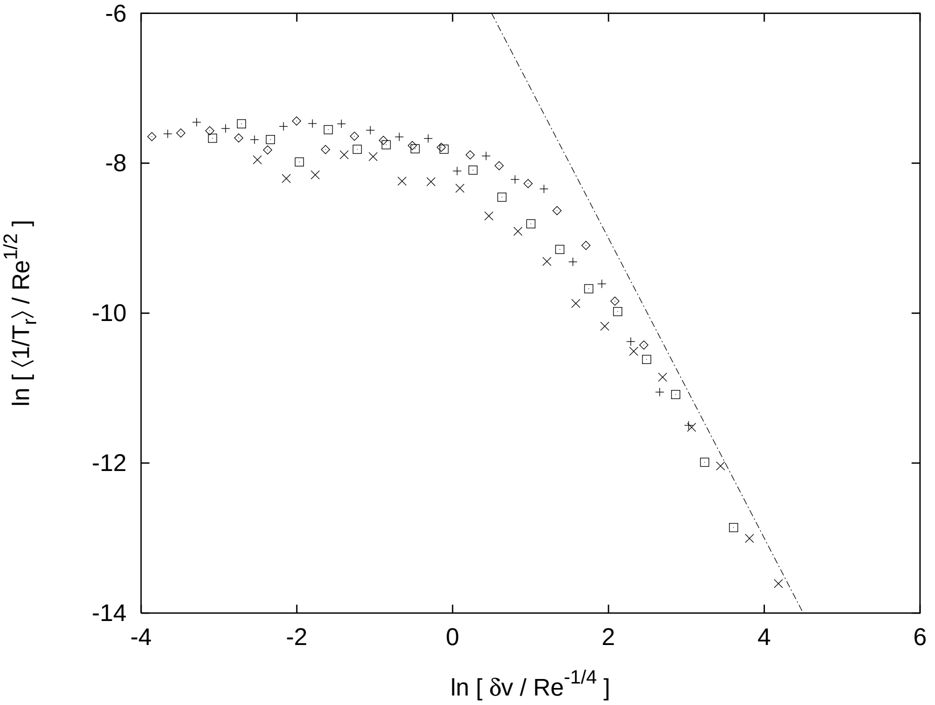

FIG. 3. $\ln\left[\langle 1/T_r(\delta v)\rangle/\mathrm{Re}^{1/2}\right]$ versus $\ln\left[\delta v/\mathrm{Re}^{-1/4}\right]$ at different Reynolds numbers $\mathrm{Re} = \nu^{-1}$. Parameters as in Fig.1, except that the times-step has been adjusted to the changing viscosity. The different symbols refer to: $N = 24$ and $\nu = 10^{-8}$ (diamond); $N = 27$ and $\nu = 10^{-9}$ (plus); $N = 32$ and $\nu = 10^{-10}$ (square); $N = 35$ and $\nu = 10^{-11}$ (cross). The straight line has slope -2.

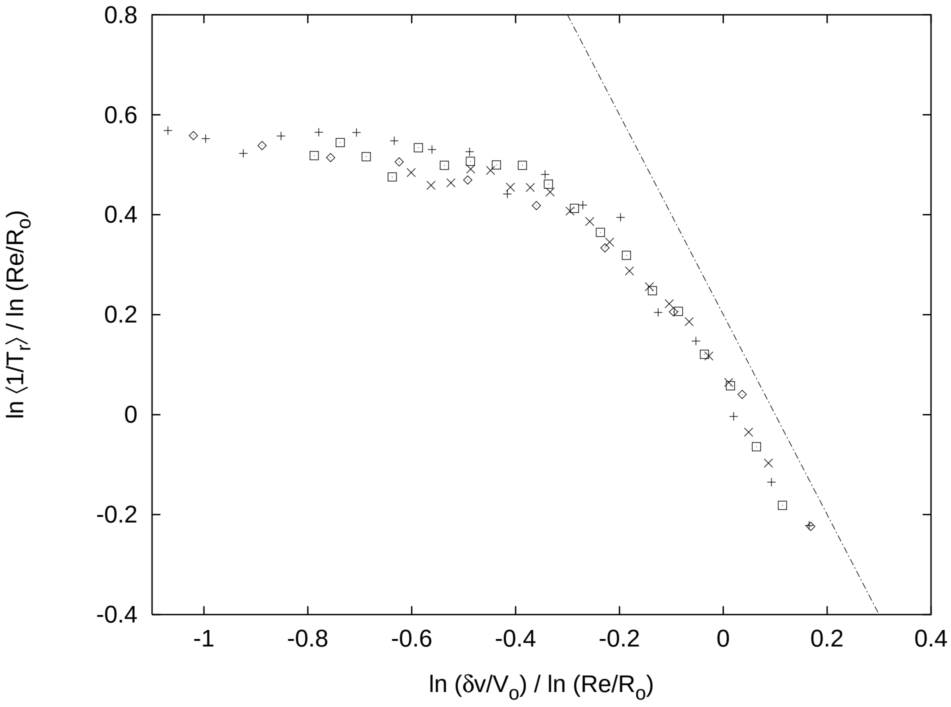

FIG. 4. Multiscaling data collapse [see eq. (38)]. Parameters as in Fig.1, except that the times-step has been adjusted to the changing viscosity. The different symbols refer to: $N = 24$ and $\nu = 10^{-8}$ (diamond); $N = 27$ and $\nu = 10^{-9}$ (plus); $N = 32$ and $\nu = 10^{-10}$ (square); $N = 35$ and $\nu = 10^{-11}$ (cross). The straight line has slope -2. The fitting parameters are $R_o = 6 \times 10^6$, $V_o = 5 \times 10^{-2}$, and $\mathrm{Re} = \nu^{-1}$.

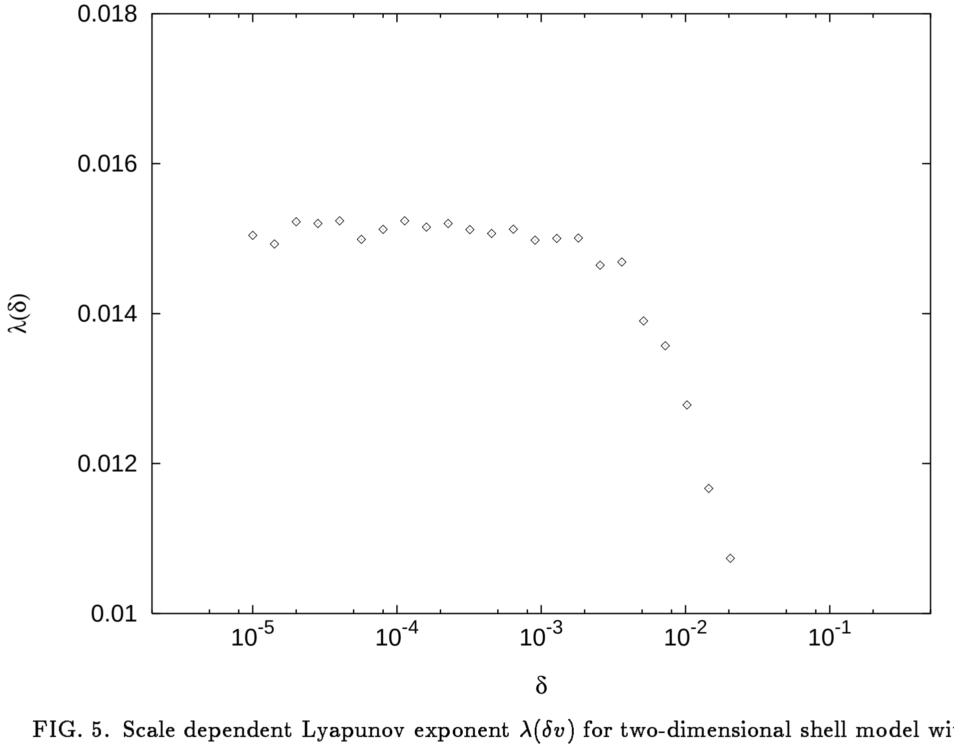

FIG. 5. Scale dependent Lyapunov exponent $\lambda(\delta v)$ for two-dimensional shell model with $N = 24$ shells, $k_0 = 0.05$, $\nu = 10^{-15}$ and $f = (1 + i) \times 0.005$. Note the linear scales on the ordinate axis.

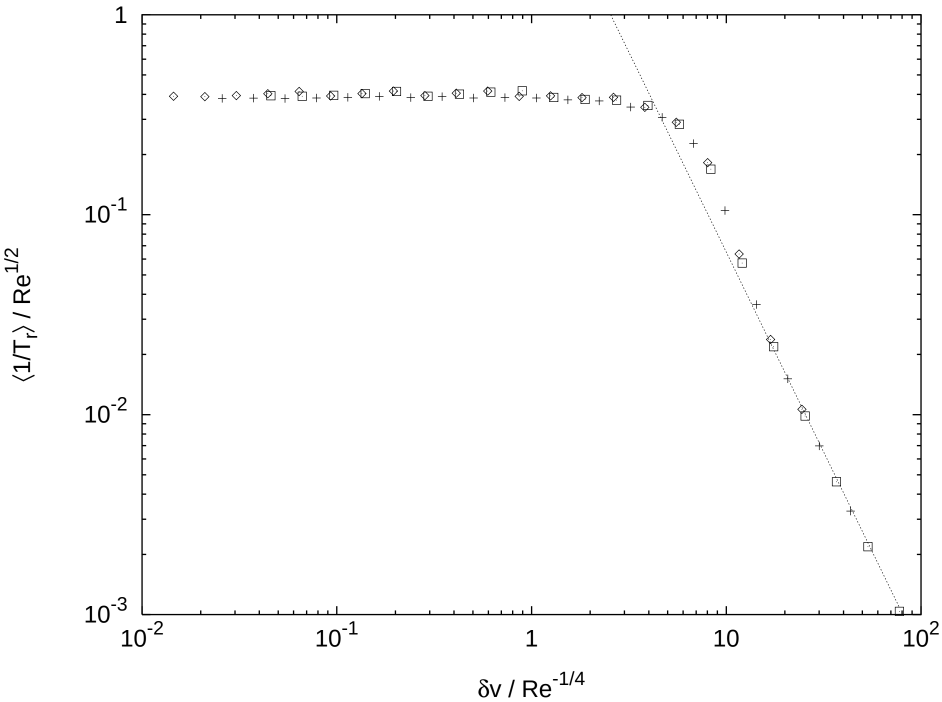

FIG. 6. $\langle 1/T_r(\delta v)\rangle/\mathrm{Re}^{1/2}$ as a function of $\delta v/\mathrm{Re}^{-1/4}$ for different values of the Reynolds number $\mathrm{Re} = \nu^{-1}$ for the EDQNM approximation. The different symbols refer to: $\nu = 10^{-8}$ (diamond), $\nu = 10^{-9}$ (plus) and $\nu = 10^{-10}$ (square). The line gives the Lorenz results -2. The other parameters are $N = 32$, $k_0 = 0.05$, $\epsilon = 1$ and $\alpha = 0.06$.

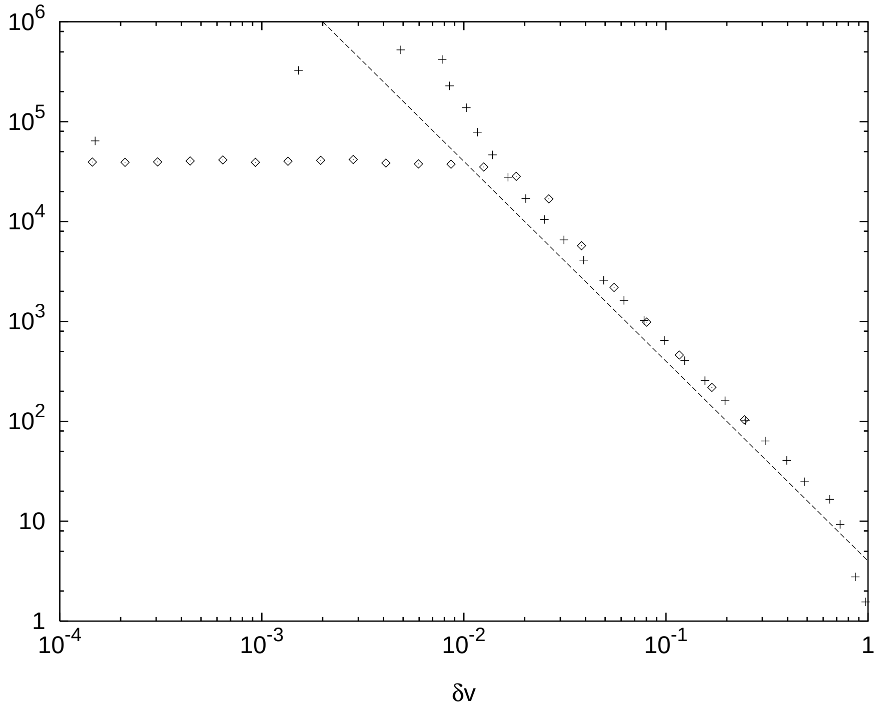

FIG. 7. $\langle 1/T_r(\delta v)\rangle$ (diamond) as a function of δv for the EDQNM approximation of GOY model with $N = 32$, $k_0 = 0.05$, $\epsilon = 1$ and $\nu = 10^{-10}$. The plus are the inverse of the eddy turn-over times $\tau^{-1}(\delta v) = k_n\, E_n^{1/2}$ versus $\delta v = (2NE_n)^{1/2}$. The straight line has slope -2.

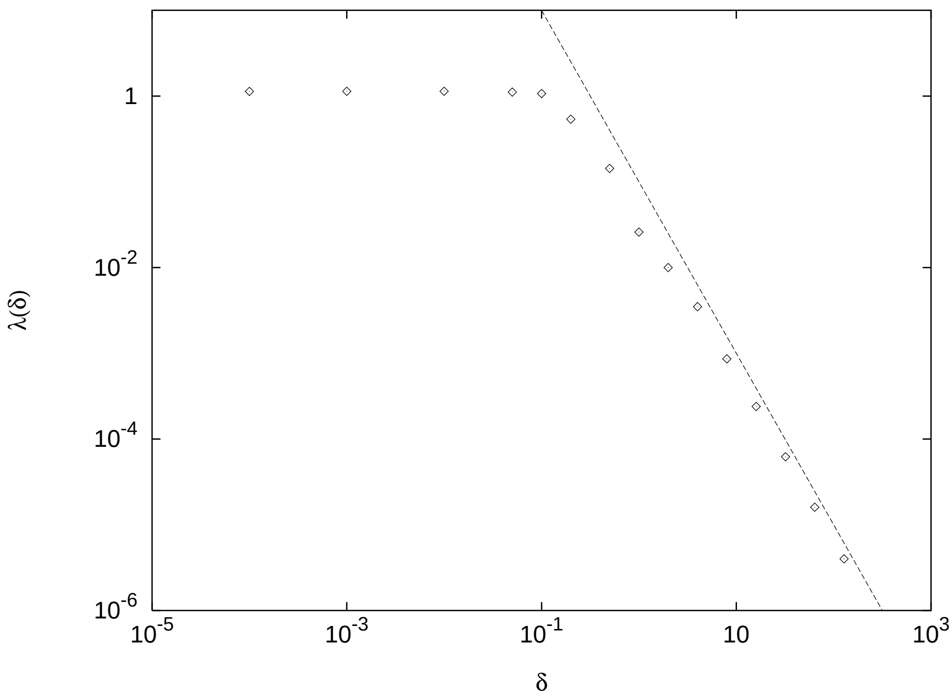

FIG. 8. Scale dependent Lyapunov exponent $\lambda(\delta)$ for the map (A9). The line has slope -2.