Estimation of diffusion time with the Shannon entropy approach

Pablo M. Cincotta [*] and Claudia M. Giordano [†]
Grupo de Caos en Sistemas Hamiltonianos, Facultad de Ciencias Astronómicas y Geofísicas, Universidad Nacional de La Plata and Instituto de Astrofísica de La Plata (CONICET), B1900FWA La Plata, B1900FWA, Argentina

(Received 2 March 2023; accepted 12 May 2023; published xxxxxxxxxx)

The present work revisits and improves the Shannon entropy approach when applied to the estimation of an instability timescale for chaotic diffusion in multidimensional Hamiltonian systems. This formulation has already been proved efficient in deriving the diffusion timescale in 4D symplectic maps and planetary systems, when the diffusion proceeds along the chaotic layers of the resonance's web. Herein the technique is used to estimate the diffusion rate in the Arnold model, i.e., of the motion along the homoclinic tangle of the so-called guiding resonance for several values of the perturbation parameter such that the overlap of resonances is almost negligible. Thus differently from the previous studies, the focus is fixed on deriving a local timescale related to the speed of an Arnold diffusion-like process. The comparison of the current estimates with determinations of the diffusion time obtained by straightforward numerical integration of the equations of motion reveals a quite good agreement.

DOI: 10.1103/PhysRevE.00.004100

I. INTRODUCTION

The diffusion time, $T_{\rm D}$, is a relevant timescale in dynamical systems since it drives the evolution of the phase space configuration. For instance, in stable domains the actions do not experience any evolution at all since the motion is confined to invariant tori, the action space does not evolve with time, and thus $T_{\rm D} \to \infty$. On the other hand, in a connected chaotic region of the phase space, the actions could exhibit large variations and the finite value of $T_{\rm D}$ provides the timescale in which such changes take place.

In near-integrable Hamiltonian systems analytical estimates of $T_{\rm D}$ could be obtained only when the perturbation acting onto the integrable part is rather small, and thus their application is somewhat limited. On the other hand, numerical determinations of the diffusion time are in general derived under the assumption of a nearly normal diffusion process, that is, when the variance of the actions scales almost linearly with time, its rate being proportional to the diffusion coefficient D and therefore $T_{\rm D} \sim D^{-1}$. This approach is largely discussed and applied to investigate the global diffusion process in many different dynamical systems, such as in [1–10].

Alternatively, $T_{\rm D}$ can be computed from straight numerical simulations, such as the motion time after which the actions escape from a given domain of phase space as was done in [11–15]. In this direction, in [14], the diffusion time was estimated in the Arnold Hamiltonian [16,17] for the motion along the homoclinic tangle or stochastic layer of the so-called guiding resonance. The computed values of $T_{\rm D}$ were then compared first with the analytical estimates provided by Chirikov [17] for small values of the parameters, and later on a relationship between this timescale, and the Lyapunov time was investigated for a wide range of parameter values.

It has been shown [18,19] that the assumption of a nearly normal diffusion process in the Arnold model is not well sustained, at least for moderate motion times ($t \lesssim 5 \times 10^6$). Thus the classical approach of looking at the variance evolution to derive the diffusion coefficient does not provide good estimates of the diffusion time. The authors of [19] explored another way to derive the timescale for diffusion, the Shannon entropy approach, which afterwards was successfully applied to different dynamical systems, from multidimensional symplectic maps to multiplanetary dynamics (see [11–13,20]).

All these studies focus on the diffusion in multiplets of resonances or resonance crossings, so the derived diffusion time is macroscopical, when the chaotic motion proceeds over the resonance web. However, reports concerning the use of this technique to estimate the diffusion speed along a single resonance are still lacking.

In this effort we review and improve the theoretical formulation of the entropy approach. Later we implement it in the Arnold model to estimate a nearly local instability timescale for the diffusion along the stochastic layer of the so-called guiding resonance (similarly to an Arnold diffusion process) for different values of the parameters. Finally we compare the obtained results with those presented in [14] that, as mentioned, were obtained by direct numerical integration of the equations of motion.

This work is organized as follows: in Sec. II the entropy formulation is revisited and improved; in Sec. III the Arnold model is briefly discussed; in Sec. IV a single experiment is presented in order to illustrate the evolution of the entropy and its associated diffusion coefficient; while in Sec. V a global numerical experiment for different values of the perturbation parameter is presented. Finally, in Sec. VI we summarize the main conclusions of this research.

[*] pmc@fcaglp.unlp.edu.ar
[†] giordano@fcaglp.unlp.edu.ar

II. FORMULATION OF THE SHANNON ENTROPY APPROACH TO DIFFUSION

Here we review and extend the main derivations given in [12,13,19] regarding the Shannon entropy as an efficient technique to estimate the diffusion rate in action space of multidimensional dynamical systems. For a general background on Shannon entropy we refer to [21–23].

A. Shannon entropy

Let us consider a volume-preserving N-dimensional dynamical system of discrete or continuous time defined through action-angle variables $(I_1,\dots,I_N,\vartheta_1,\dots,\vartheta_N)$, $I_j\in\mathbb{R}$, $\vartheta_j\in\mathbb{S}$. In several problems, the dynamics could be analyzed considering different pairs of action variables (and their conjugate angles) such as was shown in [11–13,20]. We focus then on the dynamics restricted to a given pair, say, (I_1,I_2), on the section $\mathcal{S}:\{\vartheta_1=\vartheta_1^0,\vartheta_2=\vartheta_2^0\}$.

For a given initial condition $(I_1(0),I_2(0),\vartheta_1^0,\vartheta_2^0)$ and a total motion time T a finite trajectory of the system on $\mathcal{S}$, $\gamma=\{(I_{1,l},I_{2,l}),\,l=1,\dots,N_s\gg 1\}$ leads to a (discrete) distribution density $\rho_\gamma(I_1,I_2)$ defined in $\mathcal{G}\subset\mathcal{S}$. Since N_s is finite, then $\mathrm{diam}(\mathcal{G})=d$ is bounded.

Introduce a partition on $\mathcal{G}$, $\alpha=\{a_k,\,k=1,\dots,q\}$, $q\gg 1$, a collection of q bidimensional cells that cover $\mathcal{G}$. The elements a_k are assumed to be measurable and disjoint. The measure of a_k is

$$\mu_\gamma(a_k)=\int_{a_k}\rho_\gamma(I_1,I_2)\,dI_1\,dI_2=\frac{n_k}{N_s},\qquad(1)$$

where n_k is the number of action values (I_1,I_2) of γ in the cell a_k.

The entropy of γ for the partition α is defined as

$$S(\gamma,\alpha)=-\sum_{k=1}^{q_0}\mu_\gamma(a_k)\ln[\mu_\gamma(a_k)]$$
$$=\ln N_s-\frac{1}{N_s}\sum_{k=1}^{q_0}n_k\ln n_k,\qquad(2)$$

where $1\ll q_0\leqslant q$ denotes the nonempty elements of the partition.

For the given partition and any γ, the entropy is always bounded, $0\leqslant S(\gamma,\alpha)\leqslant\ln q_0$. The entropy takes its minimum when γ is confined to a single element of α, $\mu(a_j)=1$, $\mu(a_i)=0,\,\forall i\neq j$, i.e., motion on a torus, while its maximum is reached when the nonempty elements have the very same measure, $\mu(a_i)=1/q_0$, i.e., ergodic motion.

The estimation of the last sum in (2) is simple if we assume random motion. Let $\gamma^r=\{(I_{1,l},I_{2,l})=(I_{1,l}^r,I_{2,l}^r),\,l=1,\dots,N_s\gg 1\}\subset\mathcal{S}$, where $I_{i,l}^r$ are random values; then the n_k follow a Poisson distribution with mean value (and variance) $\lambda=N_s/q_0$.

If $\lambda\gg 1$, the distribution is strongly peaked around $n_k=\lambda$ so we can write $n_k=\lambda+\xi_k$, $|\xi_k|\ll\lambda$, and since $\sum_{k=1}^{q_0}\xi_k=0$ due to the normalization condition, then up to $O((\xi_k/\lambda)^2)$

$$\sum_{k=1}^{q_0}n_k\ln n_k=N_s\ln N_s-N_s\ln q_0+\frac{1}{2\lambda}\sum_{k=1}^{q_0}\xi_k^2,\qquad(3)$$

and the entropy reduces to

$$S(\gamma^r,\alpha)\approx\ln q_0-\frac{q_0}{2N_s^2}\sum_{k=1}^{q_0}\xi_k^2.\qquad(4)$$

By the central limit theorem, $\sum_{k=1}^{q_0}\xi_k^2=q_0\lambda=N_s$, so (4) reads

$$S^r(\alpha)\equiv S(\gamma^r,\alpha)\approx\ln q_0-\frac{1}{2\lambda},\qquad(5)$$

i.e., for random motion $|S^r-\ln q_0|=O(1/\lambda)$, and thus

$$S^r(\alpha)\approx S_0\equiv\ln q_0.\qquad(6)$$

For a given strong chaotic trajectory γ we assume that the above approximation partially holds in the sense that the n_k distribution still presents a sharp maximum around $n_k=\lambda$, so writing again $n_k=\lambda+\tilde{\xi}_k$ and introducing β such that

$$\sum_{k=1}^{q_0}\tilde{\xi}_k^2=\beta\sum_{k=1}^{q_0}\xi_k^2,\qquad\beta=\frac{\langle\tilde{\xi}_k^2\rangle}{\lambda},$$

the entropy of γ results in

$$S(\gamma,\alpha)\approx\ln q_0-\frac{\beta}{2\lambda}.\qquad(7)$$

If γ presents weak correlations, $\beta/\lambda\ll 1$ and $S(\gamma,\alpha)\approx S^r(\alpha)\approx S_0$.

Let us mention that if $\mathcal{G}$ is compact, then the n_k distribution approaches a $\delta(n_k-\lambda)$ function, so $|\tilde{\xi}_k|\approx 1/2$ and (see [24])

$$|S(\gamma,\alpha)-\ln q_0|\approx\frac{1}{8\lambda^2}.\qquad(8)$$

B. Entropy-like diffusion coefficient

As was shown, for a given motion time $t<T$, the entropy for chaotic motion can be approximated by $S(t)\approx\ln q_0(t)$, where $q_0(t)$ denotes the cells visited by γ after a time t. Actually this estimate for the entropy is true provided that $\lambda(t)=N_s(t)/q_0(t)\gg 1$ where $N_s(t)$ denotes the number of intersections with the given section at time t. Thus the approximation for the entropy applies for $t>t_c$, where t_c is some transient time.

The variation of S over a finite but small time interval $\Delta t\ll T$ reads

$$\frac{\Delta S}{\Delta t}\approx\frac{1}{q_0(t)}\frac{\Delta q_0}{\Delta t}\qquad(9)$$

and involves the rate $\Delta q_0/\Delta t$ in the interval $(t,t+\Delta t)$.

Changes in the number of occupied cells in this interval are due to the diffusion of I_1,I_2 in Δt, so we introduce the following assumption: the mean-square displacements of both I_1 and I_2 in $(t,t+\Delta t)$ provide a measure of $\Delta q_0(t)$. Denoting with $\langle\Delta I_1^2(t)\rangle$ and $\langle\Delta I_2^2(t)\rangle$ such displacements we set

$$\Delta q_0(t)\propto\langle\Delta I_1^2(t)\rangle+\langle\Delta I_2^2(t)\rangle.$$

Let Σ be the measure of $\mathcal{G}$ (area) where the partition of q cells is defined, then

$$\Delta q_0(t)\approx\frac{q}{\Sigma}\big[\langle\Delta I_1^2(t)\rangle+\langle\Delta I_2^2(t)\rangle\big].$$

In the interval Δt we assume that the distribution density of the action values f_i satisfies a 1D diffusion equation of the form $\partial_t f_i = D_t^{(i)}\partial_{xx} f_i$ where x denotes either I_1 or I_2 and $D_t^{(i)}$ the corresponding diffusion coefficient in the interval Δt. Thus, in the normal diffusion approximation the mean-square displacements in each direction satisfy

$$\langle \Delta I_1^2(t)\rangle \approx 2D_t^{(1)}\Delta t, \quad \langle \Delta I_1^2(t)\rangle \approx 2D_t^{(2)}\Delta t, \qquad (10)$$

where $D_t^{(i)} \equiv D^{(i)}(I_1(t), I_2(t))$ is a local diffusion coefficient when the trajectory γ is restricted to the domain $(I_1(t), I_2(t)) \times (I_1(t+\Delta t), I_2(t+\Delta t))$, then

$$\frac{\Delta q_0(t)}{\Delta t} \approx 4\frac{q}{\Sigma}D_t, \quad D_t = \frac{1}{2}\big(D_t^{(1)} + D_t^{(2)}\big), \qquad (11)$$

and therefore from (9) it turns out that the diffusion coefficient is related to the time variation of the entropy.

Following the above discussion, an entropy-like diffusion coefficient in the interval $(t, t+\Delta t)$ can be defined as

$$D_S(\gamma, t) := \frac{1}{4}\frac{\Sigma}{q}q_0(t)\frac{\Delta S}{\Delta t}(t). \qquad (12)$$

Let $L = [T/\Delta t]$ be the number of intervals where $D_S(\gamma, t)$ is computed, then a diffusion coefficient for γ can be defined as

$$D_S(\gamma) := \frac{1}{L}\sum_{k=1}^{L} D_S(\gamma, t_k) = \langle D_S(\gamma, t_k)\rangle, \qquad (13)$$

or alternatively, $D_S(\gamma) := D_s(\gamma, t_L)$.

Finally an instability or diffusion time is given by

$$T_{\rm inst} = K\frac{\Delta^2}{D_S}, \qquad (14)$$

where Δ^2 denotes a given mean-square displacement, the squared distance between the initial and boundary values of the actions, and K a numerical factor of the order of 1. Indeed, K should be included since in case of fully anisotropic diffusion, i.e., when the diffusion proceeds only along one direction, I_2, for instance, $D_t \approx D_t^{(2)}/2$ implying $K = 2$.

C. Dependence on the partition

This formulation depends on the partition on $\mathcal{G}$. If $d = \mathrm{diam}(\mathcal{G})$ is not known beforehand, take a given domain $\mathcal{G}_0 \subset \mathcal{G}$ where $d_0 = \mathrm{diam}(\mathcal{G}_0) \leqslant d$ is known. Introduce a partition of q cells in $\mathcal{G}_0$ and redefine Σ as

$$\Sigma = (I_{2\max} - I_{2\min})(I_{1\max} - I_{1\min}),$$

where $I_{i\max}, I_{i\min} \in \mathcal{G} \setminus \mathcal{G}_0$ denotes the maximum and minimum values attained by the actions. This is the right procedure when dealing with symplectic maps where the actions are, in general, defined on a torus as was discussed in [12,13]. Notice, however, that this renormalization of Σ for each trajectory leads to different sizes of the elements of the partition in action units.

The selection of q depends on the total number of intersecting points with the section $\mathcal{S}$, N_s, and on its density distribution that is determined by the dynamics as discussed in [12].

For the entropy computation the restriction $\lambda = N_s/q \gg 1$ allows small values of q, however, we are interested in the time variation of S where we assume that the nonempty elements of the partition $q_0(t)$ grow with time. Thus for the computation of D_S, q should be large enough such that the time variation of the entropy is always positive in the case of unstable chaotic motion, so $q_0(t) \ll q$.

A suitable selection of q could be $q \gtrsim N_s$ with the restriction $q_0(t) \ll N_s \lesssim q$. Then the final value of the (normalized) entropy satisfies

$$\hat{S} \approx \frac{\ln q_0}{\ln q} \ll \frac{\ln N_s}{\ln q} \lesssim 1, \quad \rightarrow \quad N_s \lesssim q \ll N_s^{1/\hat{S}}. \qquad (15)$$

In the case of a nearly uniform distribution $\rho(I_1, I_2) \approx \rho_0$, D_S is invariant under a change of the partition. Indeed, if α is defined through $q = m \times m$ elements and $\tilde{\alpha}$ through $\tilde{q} = pm \times rm$ elements with p, r positive rational numbers and since the definition of D_S in (12) involves q_0/q, this ratio is the same for both partitions. On the other hand, the entropy depends on the partition being their relation

$$\tilde{S} = \frac{\eta + S}{1+\eta}, \quad \eta = \frac{\ln(pr)}{\ln q},$$

that for small η reduces to $\tilde{S} \approx S + \eta$. If $pr > 1$, $\tilde{S} > S$ while $\tilde{S} < S$ whenever $pr < 1$, but $\tilde{S} \approx S$ for a wide range of values of pr and large enough q.

In [12] several numerical experiments are shown regarding the dependence of both S and D_S on the parameters of the method considering a 4D symplectic map and a multidimensional Hamiltonian system modeling a planar nonrestricted three-body problem.

III. THE ARNOLD MODEL

Let us consider the Arnold model [16,17], introduced *ad hoc* to report the Arnold diffusion. Here we briefly summarize the discussion given in [14], where it is defined through the Hamiltonian

$$\begin{aligned} H(I_1, I_2, \vartheta_1, \vartheta_2, t; \varepsilon, \mu) &= \tfrac{1}{2}\big(I_1^2 + I_2^2\big) \\ &\quad + \varepsilon(\cos\vartheta_1 - 1)[1 + \mu B(\vartheta_2, t)], \\ B(\vartheta_2, t) &= \sin\vartheta_2 + \cos t, \quad I_1, I_2 \in \mathbb{R}, \\ \vartheta_1, \vartheta_2, t &\in \mathbb{S}; \quad \varepsilon\mu \ll \varepsilon \ll 1. \end{aligned} \qquad (16)$$

For $\varepsilon \neq 0$, $\mu = 0$, the Hamiltonian (16) reduces to

$$\begin{aligned} H_0(I_1, I_2, \vartheta_1; \varepsilon) &= H_1(I_1, \vartheta_1; \varepsilon) + H_2(I_2) \\ &= \tfrac{1}{2}I_1^2 + \varepsilon(\cos\vartheta_1 - 1) + \tfrac{1}{2}I_2^2, \end{aligned} \qquad (17)$$

and the system has two global integrals,

$$H_1(I_1, \vartheta_1; \varepsilon) = \tfrac{1}{2}I_1^2 + \varepsilon(\cos\vartheta_1 - 1), \quad I_2 = \omega_2. \qquad (18)$$

Here H_1 is a pendulum model for the resonance $\omega_1 = 0$ with small oscillation frequency $\omega_0^2 = \varepsilon$. Following Chirikov, we refer to this resonance as the guiding resonance.

From (18) the energy level $H_1 \equiv h_1 = -2\varepsilon$ corresponds to the exact resonance or stable equilibrium point at $(I_1, \vartheta_1) = (0, \pi)$, while $h_1 = 0$ leads to the unstable equilibrium point at $(I_1, \vartheta_1) = (0, 0) \equiv (0, 2\pi)$, and of course the same energy level corresponds to the separatrix.

The guiding resonance $\omega_1 = 0$ has an amplitude ε, half-width $(\Delta I_1)^r = 2\sqrt{\varepsilon}$, so changes of I_1 are bounded by $|\Delta I_1| \leqslant 2\sqrt{\varepsilon}$ while I_2 remains constant.

For $\varepsilon \neq 0$, $\mu \neq 0$ the full system (16) can be rewritten as

$$H(I_1, I_2, \vartheta_1, \vartheta_2, t; \varepsilon, \mu) = H_0(I_1, I_2, \vartheta_1; \varepsilon) + \mu V(\vartheta_1, \vartheta_2, t; \varepsilon),$$
$$\mu V = \varepsilon\mu(\sin\vartheta_2 + \cos t)(\cos\vartheta_1 - 1), \tag{19}$$

where H_0 is given by (17) and $\vartheta_2(t) = \omega_2 t + \vartheta_2^0$. Therefore the full Hamiltonian is a pendulum model for the guiding resonance $\omega_1 = 0$ and a free rotator of constant frequency ω_2, coupled by the perturbation $\mu V(\vartheta_1, \vartheta_2, t; \varepsilon)$ that leads to further resonances.

Since V depends on ϑ_1, ϑ_2, and t, its main effect is to modify the unperturbed separatrix of the guiding resonance giving rise to the stochastic layer of finite width, i.e., motion across the layer (in I_1). On the other hand the dependence of V on ϑ_2 causes changes not only in I_1 but also in I_2, and then motion along the stochastic layer would proceed. Because the dynamics inside the layer is highly chaotic, the variation of I_2 is also chaotic, giving rise then to a diffusion in I_2. Therefore I_2 could change without any bound, and an instability could set up when considering large enough motion times. These are the main arguments provided by Chirikov [17] to qualitatively explain the Arnold diffusion, which also included a more rigorous formulation in terms of the so-called transition chain.

As mentioned, in the Hamiltonian (19), $\omega_1 = 0$ is just one of the six first-order resonances. Using simple trigonometric relations in the expression of μV, the set of primary resonances is

$$\mathcal{O}(\varepsilon) : \{\omega_1 = 0\}; \quad \mathcal{O}(\mu\varepsilon) : \{\omega_2 = 0, \quad \omega_1 \pm \omega_2 = 0 \quad \omega_1 \pm 1 = 0\}, \tag{20}$$

where $\mathcal{O}$ denotes the amplitude of the resonance. Notice that all the resonances involved in μV have the same half-width, $(\Delta I)^{\rm r} = \sqrt{2\mu\varepsilon} \ll 2\sqrt{\varepsilon}$, much smaller than the half-width of the guiding resonance whenever $\mu \ll \varepsilon$.

The full set of resonances is then a linear combination of the three involved frequencies

$$m_1\omega_1 + m_2\omega_2 + m_3 = 0, \quad \forall m_1, m_2, m_3 \in \mathbb{Z} \setminus \{\mathbf{0}\}, \tag{21}$$

where ω_1 is the pendulum frequency and ω_2 is, at first order, constant.

Figure 1 displays the final value of the MEGNO (Mean Exponential Growth factor of Nearby Orbits) contour plot for $\varepsilon = 0.25$, $\mu = 0.010$ on the section defined by $\vartheta_1 = \pi$, $\vartheta_2 = 0$. White and light gray denote stable motion (periodic or quasiperiodic), and dark colors indicate highly chaotic dynamics. Let us mention that the MEGNO is a fast dynamical indicator that provides in an efficient way the maximum Lyapunov characteristic number of an orbit (see, for instance, [25–27] for a general description or [14] for a brief explanation). The contour plot includes a grid of 1000×1000 initial values of (I_1, I_2) in the range $|I_1| \leqslant 1.5$, $|I_2| < 2.15$.

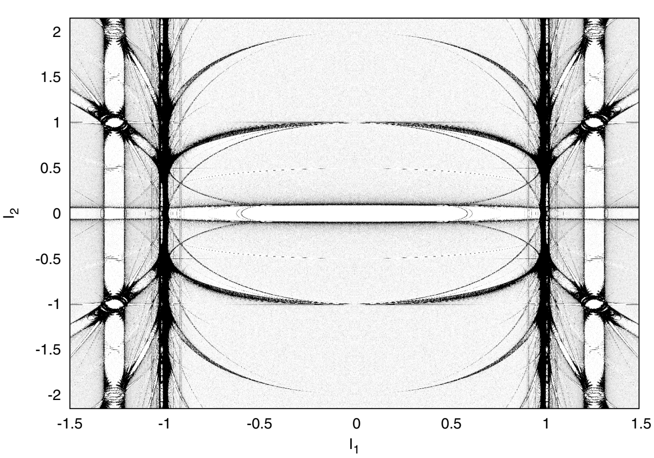

FIG. 1. MEGNO contour plot revealing the actual resonance web of the Arnold Hamiltonian (19) for $\varepsilon = 0.25$, $\mu = 0.010$ on the section $\vartheta_1 = \pi$, $\vartheta_2 = 0$.

The figure illustrates the guiding resonance, $\omega_1 = 0$, whose center appear at $I_1 = 0$, its stochastic layer centered at $|I_1| = 2\sqrt{\varepsilon} = 1$, all the primary resonances given in (20) of half-width $2\sqrt{\varepsilon\mu} = 0.1$ as well as many other high-order resonances of the form (21).

In all the numerical experiments presented in this work the integrations were carried out with a Runge-Kutta 7/8th-order integrator, the so-called DOPRI8 routine [28,29], where the local tolerance was set to 10^{-13}.

IV. ILLUSTRATIVE EXAMPLE

This section includes numerical experiments concerning the entropy approach in order to show the temporal evolution of both $S(t)$ and $D_S(t)$ as well as other relevant parameters for a given orbit in the Arnold model. To this end we consider $\varepsilon = 0.25$, $\mu = 0.010$, the same values of the parameters adopted to produce Fig. 1, and take an ensemble of n_p random initial values of the actions of size 10^{-7} centered at the chaotic layer of the guiding resonance, $I_1(0) = 2\sqrt{\varepsilon} = 1$, $I_2(0) = \omega_2$, while the angles are all fixed to $\vartheta_1(0) = \pi$, $\vartheta_2(0) = 0$.

Figure 2 shows the diffusion along the chaotic layer of the guiding resonance in the Arnold model for a small ensemble of $n_p = 400$ centered at $I_1(0) = 1$, $\omega_2 = 0.01\sqrt{3}$ represented as a green point, on the section (or slice) $|\vartheta_1 - \pi| + |\vartheta_2| \leqslant 0.02$ after a motion time 4×10^6. Since only the intersections with this section are considered, an ensemble of initial conditions is required in order to get large enough values of intersecting points N_s.

The figure reveals that for $|I_1| < 1$ the density distribution is nearly continuous, while for $|I_1| > 1$, i.e., at large times, the distribution reveals its discrete character. For the adopted values of the parameter model and motion time, the diffusion spreads along the homoclinic tangle of the guiding resonance up to $|I_2| \approx 1.5$, and just a few intersecting points appear on the chaotic layers of the resonances $\omega_1 = \pm\omega_2$ and some other high-order resonances. If instead a larger value of μ is considered, the diffusion is not confined to the layer of the resonance $\omega_1 = 0$ but spreads over the nearby ones.

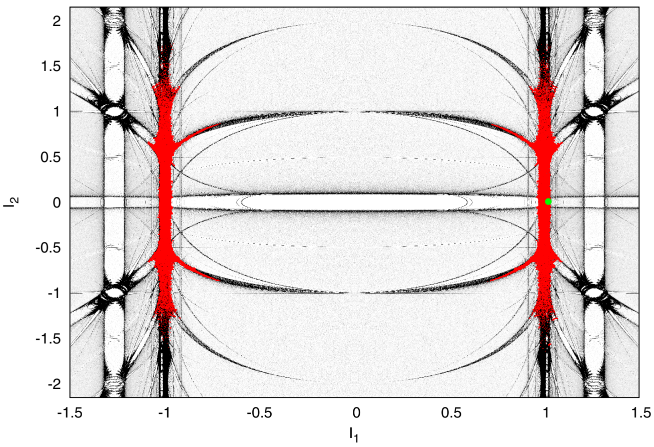

FIG. 2. An initial ensemble indicated as a green point is followed onto the MEGNO contour plot for $\epsilon = 0.25$, $\mu = 0.010$; white and light gray denote stable motion, and black indicates strong chaotic dynamics. The concomitant trajectories for the initial ensemble that intersect the section $|\vartheta_1 - \pi| + |\vartheta_2| \leqslant 0.02$ are depicted in red.

We set as $\mathcal{G}$ the region where the diffusion takes place, $(I_1, I_2) \in (-1.5, 1.5) \times (-2.4, 2.4)$, which defines Σ, the measure (area) of $\mathcal{G}$. If the diffusion spreads beyond this domain, the orbital points are discarded.

In $\mathcal{G}$ a partition of $q = 500 \times 500$ elements is introduced, and then the time evolution of the entropy is computed as well as $D_S(\gamma, t)$ given in (12). For the determination of the entropy variation, the interval $\Delta t = 4 \times 10^4$ is adopted, so $L = [T/\Delta t] = 100$. In this numerical experiment two different values of n_p are taken, $n_p = 400,\ 800$ in order to see any dependence of this approach on N_s.

Figure 3 (left) displays the time evolution of the mean value, $\lambda(t) = N_s(t)/q_0(t)$, for both values of n_p. After a transient time of about 5×10^5, $\lambda(t)$ increases almost linearly with time rather slowly. The changes in $\lambda(t)$ are small, about $\Delta\lambda \approx 5, 10$ for $n_p = 400,\ 800$ respectively over a time span larger than 3×10^6. Notice that $\lambda(t)$ does not attain quite large values, $2 < \lambda(t) < 8$ for $n_p = 400$, while in the case of $n_p = 800$, $4 < \lambda(t) < 15$. This result suggests that in (13) the average should be computed over $1 < k_0 < k \leqslant L$.

The initial bump in λ is due to a change in the diffusion rate. At early times the speed of the diffusion is higher than in the rest of the time span as Fig. 3 (middle) reveals, where the evolution of I_2, starting at $I_2(0) = 0.01\sqrt{3}$ is drawn. At $t \approx 5 \times 10^5$, $|I_2|$ increases to 1 while it takes values below 1.8 for $5 \times 10^5 < t < 4 \times 10^6$. The distribution of the I_2 variables is presented in the right panel where we observe a nearly normal distribution, and some departures are observed at both tails revealing a stickiness effect.

This is an expected behavior since, as discussed in [14,17], the diffusion along the layer could be described by a time-dependent whisker-like map whenever the parameters are small. Recall that the whisker map models the motion across the stochastic layer of a nonlinear resonance (in the direction of I_1) of width $w_s \sim \mu$ and where its central region of size $\sim w_s/4$ around the unperturbed separatrix looks ergodic, while the external one exhibits stability domains due to resonances of the map as Fig. 3 (middle) shows for the larger values of $|I_2|$. These stability islands are responsible of the stickiness observed in Fig. 3 (right).

In this numerical example the initial ensemble is taken around the unperturbed separatrix and $I_2(0) \approx 0$, so the diffusion at small times is fast, close to free, but at larger times the motion proceeds close to the borders of the layer where the resonances lead to phase correlations that reduce or prevent the free diffusion, and in this direction Chirikov introduced a reduction factor, of the order of the relative size of the central region of the chaotic layer ($R \approx 1/4$), in order to take into account somehow this fact (see [14,17]).

Figure 4 (left) shows the results of $\hat{S}(t), \hat{S}_0(t) = \ln q_0(t)/\ln q$ corresponding to the given initial ensemble and both values of n_p. Notice that in any case the entropy shows a logarithmic trend and $|\hat{S}(t) - \hat{S}_0(t)| \approx 0.035$ for $t > t_c \approx 5 \times 10^5$, consistent with the estimate (7).

For instance, in the case of $n_p = 400$, the final value of the entropy is $\hat{S}_0 \approx 0.75$ and $N_s(T) \approx 6.5 \times 10^4$, so from (15) we observe that this choice of q satisfies such a condition.

On the other hand, Fig. 4 (right) displays the evolution of $D_s(t)$ also for $n_p = 400\,800$ revealing a weak dependence of the entropy diffusion coefficient with $N_s(t)$, and both approach an asymptotic value close to 1×10^{-7}, and for $t \gtrsim 10^6$ the change in $D_s(t)$ is, at most, about a factor 4 over a time span 3×10^6. At short times D_S takes larger values in the region where the diffusion is fast, as expected. The irregularities in D_S are due to the computation of $\Delta S/\Delta t$ sampled at $\Delta t = 4 \times 10^4$ whose amplitude decreases with time. In this particular example, since the diffusion is almost 1D, the value of $D_s(t)$ displayed in the figure is taken as half of the one given in (12).

In order to get an independent rough estimate of the diffusion rate, the ensemble variance over the $n_p = 400$ values of

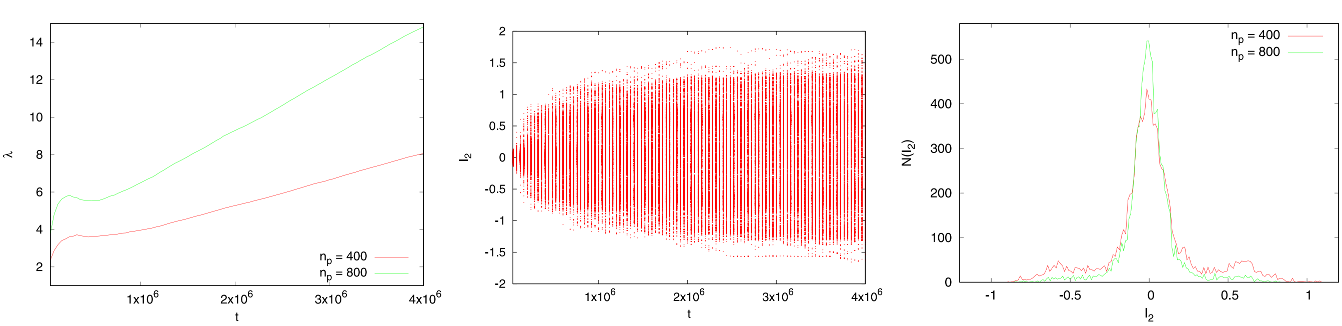

FIG. 3. Left: Evolution of the mean value $\lambda(t) = N_s(t)/q_0(t)$ for both values of n_p. Middle: Evolution of I_2 for $n_p = 800$. Right: Distribution of the I_2 values binned in 150 intervals.

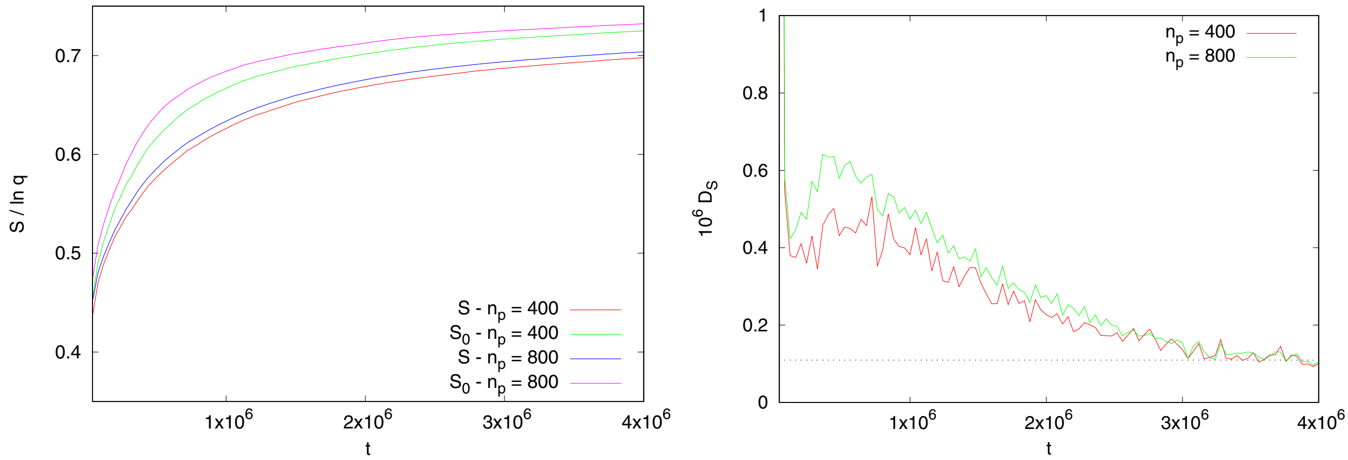

FIG. 4. Evolution of the normalized entropies $\hat{S}(t)$, $\hat{S}_0(t)$ and $D_S(t)$, for $\epsilon = 0.25$, $\mu = 0.010$ and two different values of n_p.

I_2 is computed at every t as

$$\mathrm{var}_e(t) = \frac{1}{n_p}\sum_{k=1}^{n_p}[I_2(t) - I_2(0)]^2. \quad (22)$$

Also the variance over the section $|\vartheta_1 - \pi| + |\vartheta_2| \leqslant 0.02$ is calculated, defined as

$$\mathrm{var}_s(t_l) = \frac{1}{n_l}\sum_{k=1}^{n_l}\left[I_2^{(k)}(\tau_l) - I_2^{(k)}(0)\right]^2, \quad (23)$$

where $t_l = l\,\Delta t$ and n_l denotes the number of intersecting points at times $\tau_l \in (t_{l-1}, t_l]$.

Figure 5 (left) shows the time evolution of the two variances revealing a similar trend; however, it is not linear along the whole time span. A least-square fit in the range $[0, 4 \times 10^6]$ of a power law $\mathrm{var}_e(t) = Dt^b$ leads to $b \approx 0.68$, $D \approx 2 \times 10^{-5}$ showing an anomalous diffusion process with a large value of D, at least for the motion time considered. Certainly correlations due to the presence of stability domains in the external region of the stochastic layer lead to a subnormal diffusion. In contrast, at short motion times both var_e, var_s expose the nearly free diffusion already discussed, while at $t \approx 10^6$ the change in the diffusion regime is observed. An effective diffusion coefficient, in the sense of (14), to estimate an instability timescale over the full time span would lead to $D^{1/b} \approx 1.2 \times 10^{-7}$.

Therefore we proceed in a different way to estimate D. In Fig. 5 (right) four linear fits to var_e of the form $\mathrm{var}_e(t) = 2Dt + a$ are performed leading to $D \approx 1.4 \times 10^{-7}$ in the range $[0, 8 \times 10^5]$, $D \approx 8 \times 10^{-8}$ in $(8 \times 10^5, 2 \times 10^6]$, $D \approx 5.4 \times 10^{-8}$ in $(2 \times 10^6, 3 \times 10^6]$, while in the interval $(3 \times 10^6, 4 \times 10^4]$, $D \approx 4.5 \times 10^{-8}$. These values of D agree, in order of magnitude, with those of D_S shown in Fig. 4 for $n_p = 400$.

Thus the anomalous diffusion observed in the whole time span could be well approximated by a nearly normal diffusion process at different time intervals. In terms of the dynamics of the systems this is clear; as has been already discussed, during some time interval (δt_1) the motion takes place in a region of the phase space where correlations are negligible, and then it proceeds almost freely. But in a subsequent time interval (δt_2)

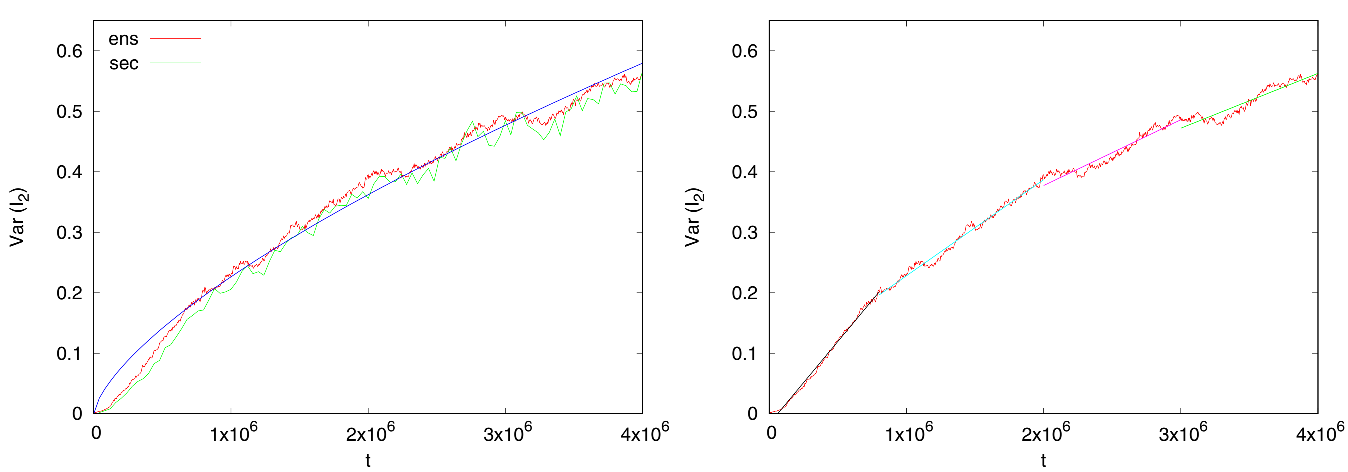

FIG. 5. Evolution of var_s and var_e for $\epsilon = 0.25$, $\mu = 0.010$, $n_p = 400$. Left: Fit of the form $\mathrm{var}_e(t) = Dt^b$ is included in blue with $b \approx 0.68$ and $D \approx 10^{-5}$. Right: Four different linear fits to var_e are performed in the ranges $[0, 8 \times 10^5]$ in black, $(8 \times 10^5, 2 \times 10^6]$ in sky blue, $(2 \times 10^6, 3 \times 10^6]$ in magenta, and $[3 \times 10^6, 4 \times 10^6]$ in green.

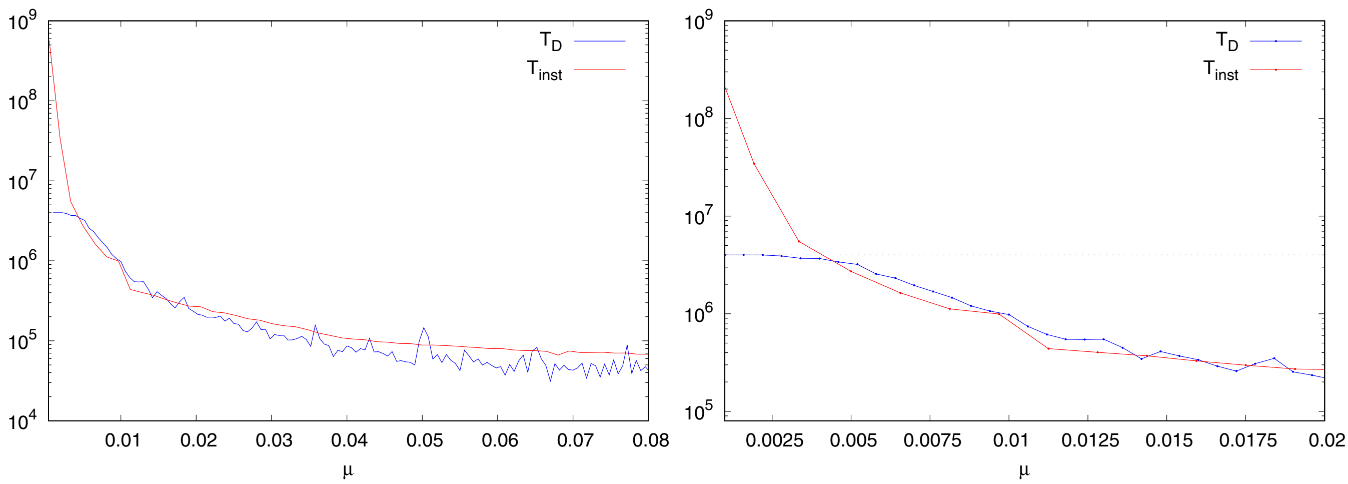

FIG. 6. Left: Diffusion and instability time in blue and red, respectively, in logarithmic scale for $\varepsilon = 0.25$, $0.0005 \leqslant \mu \leqslant 0.08$. Right: Zoom for $0.0005 \leqslant \mu \leqslant 0.02$.

the diffusion rate would be governed by the dynamical objects present in this new domain that could, for instance, diminish the diffusion rate.

Therefore, if we are interested in the mean diffusion rate over the whole time interval, the average of $D_S(t)$ as given in (13) should be considered, while if the long-range diffusion is the relevant feature, its final value (or its average over the last time intervals) should be adopted in the computation of T_{inst}.

V. COMPUTATION OF THE DIFFUSION TIMESCALE

To determine the instability time we perform similar numerical experiments as before with $n_p = 400, I_1(0) \approx 2\sqrt{\varepsilon}$, $I_2(0) = \omega_2 = 0.01\sqrt{3}$, $\mathcal{G}$ defined by $(I_1, I_2) \in (-1.5, 1.5) \times (-2.4, 2.4)$, $q = 500 \times 500$, $T = 4 \times 10^6$, and $\Delta t = 4 \times 10^4$. Finally, T_{inst} given by (14) is taken as the average over the L values but discarding the first five ones in order to reduce any noise in the computation of D_S introduced by the relatively small value of λ as Fig. 4 (right) shows.

In [14] a diffusion time along the stochastic layer of the guiding resonance is defined as the required motion time for a small ensemble around $I_1(0), I_2(0)$ on the section $\vartheta_1 = \pi, \vartheta_2 = 0$ to reach $I_2(T_{\mathrm{D}}) = I_2(0) \pm \delta$, where $\delta \sim O(1)$. Therefore in (14) the mean-square displacement should be taken as $\Delta \approx \delta$.

We adopt $\delta = 0.5$ following [14], where the motivation of setting this particular value is discussed in detail. Briefly we are interested in the diffusion along the chaotic layer of the guiding resonance, and, such as Fig. 1 illustrates, for $|I_2| > 0.5$ a resonance crossing occurs (between the guiding and $\omega_2 = \pm\omega_1$ resonances). Thus the diffusion could proceed over a different resonance set, and then the computed value of the instability time would be largely affected by the dynamics in the resonance junction.

In the above mentioned work it was shown that the relevant parameter in the Arnold model is μ, so we take only one value of ε, $\varepsilon = 0.25$, and μ will be taken in such a way that the resonance overlap is almost negligible, in the range $0.0005 \leqslant \mu \leqslant 0.080$ with step 0.0015, below the theoretical expected one for a first-order resonance overlap at this value of ε. Indeed, in [14] it was shown that the theoretical critical value of $\mu_c(\varepsilon = 0.25)$ for an overlap of the guiding resonance with the resonance $\omega_1 = \pm 1$ is about 0.1, but numerically it turns out to be smaller.

Regarding the parameter K, as mentioned, it is introduced in order to take into account the diffusion spread in the action space. Accordingly to the above discussion it would be expected that for small μ the motion takes place essentially in I_2, so we set $K = 2$ for $\mu \leqslant 0.010$ while $K = 1$ for $\mu > 0.010$ since the diffusion proceed in both directions; see, for instance, Fig. 4 (right) in [14] for $\varepsilon = 0.25$ and $\mu = 0.025$.

Figure 6 shows the results of T_{inst} for the given values of ε and μ, and, for comparison, the results given in [14] for the diffusion time, T_{D}, obtained by direct numerical integration of the equations of motion are included. For the smaller values of μ, $T_{\mathrm{D}} \approx 4 \times 10^6$, similar to the total motion time, and thus the diffusion is slow enough such that the motion along the layer does not reach the prescribed bound ($|I_2| \approx 0.5$). For $0.0007 \lesssim \mu \lesssim 0.010$ the computed diffusion time (defined as the average over the ensemble) could be overestimated because some of the n_p trajectories could not exceed $|I_2| \approx 0.5$ leading then to an increase of the ensemble average.

On the other hand, for $\mu < 0.0035$, T_{inst} reaches much larger values than the considered motion time revealing that, for this range of μ values, the speed of the diffusion is quite slow as expected, while for $0.0035 < \mu < 0.010 T_{\mathrm{inst}}$ takes values larger than 10^6, close to those of T_{D}.

Meanwhile for $\mu \gtrsim 0.01$, T_{D} decreases up to $\mu \approx 0.06$, and for the largest values of this parameter it reaches a nearly constant value about $T_{\mathrm{D}} \approx 4.5 \times 10^4$. Notice that several fluctuations appear, for instance, around $\mu = 0.05$, $\Delta T_{\mathrm{D}} \approx 2 \times 10^4$ while T_{inst} presents a rather smooth behavior over the full range of μ, but in any case for $\mu \geqslant 0.005$, the agreement between both estimations of the diffusion time is quite remarkable.

VI. DISCUSSION

The Shannon entropy approach has been shown to be an efficient technique to display the local dynamics of a multidimensional system as well as to provide an accurate estimate of the diffusion speed. Both its theoretical formulation and its computation are quite simple. Numerically, it requires a counting box scheme while integrating the equations of motion of the system for a given ensemble of initial conditions. In fact, the computational effort is similar to the one to estimate the diffusion coefficient through the evolution of the action's variances. Nevertheless, the entropy approach provides much better results, since as Fig. 5 (left) reveals, the diffusion is not normal, at least for moderate motion times, so the estimation of the instability time through the variance evolution leads to poor results.

The entropy-like diffusion coefficient is almost independent of the transport process, since the assumption behind its definition is a normal diffusion behavior in short time intervals with different values of the local diffusion coefficient, as is illustrated in Fig. 5 (right). Finally D_S could be taken as its corresponding value in the last interval or that obtained as the average over the full range.

The numerical results here presented reveal a good agreement between the diffusion or instability time obtained in the Arnold model in comparison with the one computed by direct integration of the equations of motion. Moreover, for small values of μ the straight simulations reveal that $T_{\rm D}$ saturates to the total motion time while $T_{\rm inst}$ takes much larger values providing the expected diffusion time.

This report introduces insights concerning the use of the entropy-like approach to determine the speed of the diffusion along the chaotic layer of a single resonance, an instability process close to Arnold diffusion for small enough μ.

In previous successful applications of this tool such as in 4D symplectic maps or Hamiltonian systems that model multiplanetary dynamics, the involved parameters were kept fixed, only the location of the initial ensemble was changed so the dynamical model remains unchanged. Here instead the instability time is estimated for a wide range of values of μ for the same location of the initial ensemble. Then the dynamical model drastically changes from the smaller to the larger values of the perturbation parameter.

Summing up, the results here presented together with those already mentioned allow us to conclude that this approach would be very useful to derive the timescale of instabilities in very different dynamical systems.

ACKNOWLEDGMENTS

The authors are grateful to two anonymous reviewers for their valuable comments. P.M.C. and C.M.G. are supported by the Consejo Nacional de Investigaciones Científicas y Técnicas de la República Argentina (CONICET), Project No. 11220170100569, the Universidad Nacional de La Plata, Project No. 11G173, and Instituto de Astrofísica de La Plata.

The authors declare that they have no conflict of interest.

[1] P. M. Cincotta, C. Efthymiopoulos, C. M. Giordano, and M. F. Mestre, Chirikov and Nekhoroshev diffusion estimates: Bridging the two sides of the river, Physica D **266**, 49 (2014).

[2] C. Efthymiopoulos and M. Harsoula, The speed of Arnold diffusion, Physica D **251**, 19 (2013).

[3] C. Froeschlé, M. Guzzo, and E. Lega, Local and global diffusion along resonant lines in discrete quasi-integrable dynamical systems, Celest. Mech. Dyn. Astron. **92**, 243 (2005).

[4] C. Froeschlé, E. Lega, and M. Guzzo, Analysis of the chaotic behaviour of orbits diffusing along the Arnold web, Celest. Mech. Dyn. Astron. **95**, 141 (2006).

[5] N. Guillery and J. D. Meiss, Diffusion and drift in volume-preserving maps, Regul. Chaot. Dyn. **22**, 700 (2017).

[6] M. Guzzo, E. Lega, and C. Froechlé, First numerical evidence of global Arnold diffusion in quasi-integrable systems, Discrete Contin. Dyn. Syst. B **5**, 687 (2005).

[7] E. Lega and C. Froeschlé, Guzzo, Diffusion in Hamiltonian quasi-integrable systems, Lect. Notes Phys. **729**, 29 (2008).

[8] E. Lega, M. Guzzo, and C. Froeschlé, Detection of Arnold diffusion in Hamiltonian systems, Physica D **182**, 179 (2003).

[9] J. D. Meiss, N. Miguel, C. Simó, and A. Vieiro, Accelerator modes and anomalous diffusion in 3D volume-preserving maps, Nonlinearity **31**, 5615 (2018).

[10] N. Miguel, C. Simó, and A. Vieiro, On the effect of islands in the diffusive properties of the standard map, for large parameter values, Found. Comput. Math. **15**, 89 (2014).

[11] R. Alves Silva, C. Beuagé, S. Ferraz-Mello, P. M. Cincotta, and C. M. Giordano, Instability times in the HD 181433 exoplanetary system, Astron. Astrophys. **652**, A112 (2021).

[12] P. M. Cincotta, C. M. Giordano, R. Alves Silva, and C. Beuagé, The Shannon entropy: An efficient indicator of dynamical stability, Physica D **417**, 132816 (2021).

[13] P. M. Cincotta, C. M. Giordano, R. Alves Silva, and C. Beuagé, Shannon entropy diffusion estimates: Sensitivity on the parameters of the method, Celest. Mech. Dyn. Astron. **133**, 7 (2021).

[14] P. M. Cincotta, C. M. Giordano, and I. I. Shevchenko, Diffusion and Lyapunov timescales in the Arnold model, Phys. Rev. E **106**, 044205 (2022).

[15] K. Tsiganis, H. Varvoglis, and R. Dvorak, Chaotic diffusion and effective stability of Jupiter Trojans, Celestial Mech. Dyn. Astron. **92**, 71 (2005).

[16] V. I. Arnold, On the nonstability of dynamical systems with many degrees of freedom, Soviet Math. Dok. **5**, 581 (1964).

[17] B. V. Chirikov, A universal instability of many-dimensional oscillator systems, Phys. Rep. **52**, 263 (1979).

[18] P. M. Cincotta, C. M. Giordano, J. G. Martí, and C. Beaugé, On the chaotic diffusion in multidimensional Hamiltonian systems, Celest. Mech. Dyn. Astron. **130**, 23 (2018).

[19] C. M. Giordano and P. M. Cincotta, The Shannon entropy as a measure of diffusion in multidimensional dynamical systems, CMDA **130**, 35 (2018).

[20] E. Kővári, B. Érdi, and Z. Sándor, Application of the Shannon entropy in the planar (non-restricted) four-body problem: The long-term stability of the Kepler-60 exoplanetary system, Mon. Not. R. Astron. Soc. **509**, 884 (2022).

[21] V. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics*, 2nd ed. (Addison-Wesley, New York, 1989).

[22] A. Lesne, Shannon entropy: A rigorous notion at the crossroads between probability, information theory, dynamical systems and statistical physics, Math. Struct. Comput. Sci. **24**, e240311 (2014).
[23] C. Shannon and W. Weaver, *The Mathematical Theory of Communication* (University of Illinois Press, Urbana, 1949).
[24] P. M. Cincotta and C. M. Giordano, Phase correlations in chaotic dynamics: A Shannon entropy measure, Celest. Mech. Dyn. Astron. **130**, 74 (2018).
[25] P. M. Cincotta and C. M. Giordano, Theory and applications of the Mean Exponential Growth Factor of Nearby Orbits (MEGNO) method, Lect. Notes Phys. **915**, 93 (2016).
[26] P. M. Cincotta and C. Simó, Simple tools to study global dynamics in non-axisymmetric galactic potentials—I, Astron. Astrophys. Suppl. Ser. **147**, 205 (2000).
[27] P. M. Cincotta, C. M. Giordano, and C. Simó, Phase space structure of multi-dimensional systems by means of the mean exponential growth factor of nearby orbits, Physica D **182**, 151 (2003).
[28] E. Hairer, S. Norsett, and G. Wanner, *Solving Ordinary Differential Equations I: Nonstiff Problems* (Springer, Berlin, 1987).
[29] P. Prince and J. Dormand, High order embedded Runge-Kutta formulae, J. Comp. Appl. Math. **7**, 67 (1981).