Random Initial Data and Average Shock Time in the Fermi-Pasta-Ulam-Tsingou Chain

Matteo Gallone,[1,*] Ricardo Grande,[1,†] Antonio Ponno,[2,‡] Stefano Ruffo,[3,1,4,§] and Erwan Druais[5,‖]
[1]*Scuola Internazionale di Studi Superiori Avanzati, Via Bonomea 265, 34136 Trieste, Italy*
[2]*Department of Mathematics "T. Levi-Civita," University of Padova, Via Trieste 63, 35131 Padova, Italy*
[3]*ISC-CNR, via Madonna del Piano 10, 50019 Sesto Fiorentino, Firenze, Italy*
[4]*INFN, Sezione di Firenze, 50019 Sesto Fiorentino, Italy*
[5]*École Normale Supérieure de Lyon, 15 parvis René Descartes, 69342 Lyon, France*

(Received 11 September 2025; accepted 1 May 2026; published 27 May 2026)

We investigate the dynamics of the Fermi-Pasta-Ulam-Tsingou chain with long-wavelength random initial data. When the energy per particle is small, thermal equilibrium is not reached on a fast timescale, and the system enters prethermalization. The formation of the prethermal state is characterized by the development of a Burgers-type shock and the onset of a turbulentlike spectrum with a time dependent exponent $\zeta(t)$ in the inertial range. We perform a significant step forward by demonstrating that these features are robust under generic long-wavelength random initial conditions. By employing advanced probabilistic techniques inspired by the works of Dudley and Talagrand, we derive a sharp asymptotic expression for the average shock time in the thermodynamic limit. For large p, this time scales as $(p\sqrt{\log p})^{-1}$, where p is the number of excited modes, proving that it is an intensive quantity up to a logarithmic correction in the size of the system.

DOI: 10.1103/qxzs-1t1l

Introduction—Understanding how isolated physical systems approach thermal equilibrium is a central open problem in statistical mechanics. The microscopic mechanisms by which macroscopic systems redistribute energy among their degrees of freedom are complex and not yet fully understood [1]. This complexity was brought to light by the numerical simulations of Fermi, Pasta, Ulam, and Tsingou (FPUT) [2]. In their pioneering work, they integrated numerically a one-dimensional chain of particles and followed the time evolution of the Fourier energy spectrum (FES). Rather than reaching equipartition, which is a necessary condition for *thermal equilibrium*, the system exhibited unexpected recurrent dynamics. Since then, FPUT-like recurrences have been reported across a wide range of physical settings, including holography [3], graphene resonators [4], nonlinear phononic lattices [5], photonic systems [6], and trapped cold atoms [7–9].

Random initial data drawn from the Gibbs measure have played a crucial role in the study of the dynamics of coupled harmonic oscillators [10]. Initial conditions close to thermal equilibrium have been employed when studying anomalous conduction in one-dimensional systems of coupled nonlinear oscillators [11–14]. With a different perspective, the robustness of dynamical behavior with respect to random perturbations of the invariant measure has been examined [15–17]. By contrast, initial data far from statistical equilibrium or systems subject to strong random perturbations have more rarely been considered. These regimes, which capture how typical nonequilibrium configurations relax toward thermal equilibrium, are the focus of this Letter.

Concerning the time evolution of the FPUT system, numerical studies have shown that, when energy is above a certain threshold, the system undergoes a quick thermalization, and the FPUT recurrence is not present [18]. The process by which a system reaches thermal equilibrium is called *thermalization* [19]. While approaching thermal equilibrium, the system may become trapped for a long time in a quasistationary *prethermal* state. Indeed, this is what happens in the FPUT system when the initial energy is below the threshold mentioned above [20]. This initial stage, known as *prethermalization*, has attracted broad interest beyond the FPUT dynamics, including quantum systems [21–24], because of its deep experimental and technological relevance [19,25,26]. We emphasize that prethermalization corresponds to the relaxation to a quasistationary state different from thermal equilibrium, preceding the eventual onset of thermal equilibrium on longer timescales [19]. In the FPUT system, prethermalization is

*Contact author: mgallone@sissa.it
†Contact author: rgrandei@sissa.it
‡Contact author: ponno@math.unipd.it
§Contact author: ruffo@sissa.it
‖Contact author: erwan.druais@ens-lyon.fr

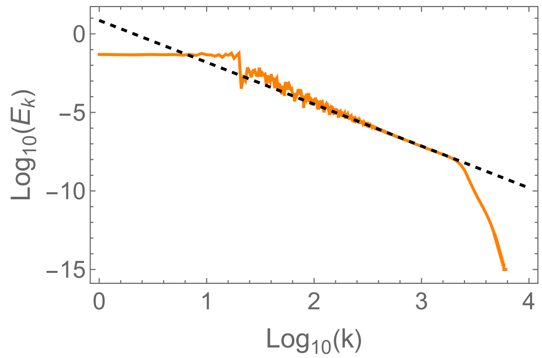

FIG. 1. Fourier energy spectrum (full orange line) at the shock time with $\epsilon = 10^{-3}$ for $\alpha = 1$, $\beta = 0.1$, and $N = 16\,384$. Twenty modes are initially excited. The black dashed line is the theoretical prediction $\log E_k = -8/3 \log k + 0.86$.

strongly dependent on the choice of the initial condition. Numerical evidence and heuristic arguments suggest that, for certain classes of out-of-equilibrium random initial data, prethermalization is governed solely by intensive parameters [27–29]. We will indeed show that, up to a logarithmic correction, the onset of prethermalization in the FPUT system is uniquely governed by the specific energy and the fraction of initially excited modes with random phases.

Prior to tackling the case of random initial conditions, a transient turbulent regime was identified in a recent work for single-mode initial conditions [30,31]. This regime highlights a quantitative mechanism for the redistribution of energy from low to high modes of the Fourier spectrum. This regime is present for large-wavelength initial conditions and at sufficiently short timescales. The key idea was to show that, on such a timescale, the FPUT dynamics can be well approximated by the inviscid Burgers equation. Solutions of the Burgers equation exist up to a finite time t_s, the shock time. At t_s, the FPUT displays a turbulentlike FES with an inertial-range scaling $E_k \sim k^{-8/3}$. Thus, the shock time emerged as the fundamental timescale for the formation of the prethermal state in the FPUT chain [30].

In this Letter, we show analytically and confirm numerically that the Burgers turbulence scenario is robust for out-of-equilibrium long-wavelength random initial data (see Fig. 1). Developing advanced probabilistic techniques inspired by the works of Dudley and Talagrand [32,33], we derive an explicit analytic formula for the average shock time in the thermodynamic limit for a finite and large fraction of excited modes. The theory rigorously explains the observed differences in scaling behavior between *coherent* initial data, where the phases of the Fourier modes are all equal or equispaced, and *incoherent* initial data, where the phases are taken randomly on the unit circle [29]. In addition to the power-law scaling, we analytically predict a logarithmic correction, which we confirm numerically (Fig. 3).

Model, initial conditions, and main result—The FPUT model consists of N masses interacting nonlinearly on a one-dimensional lattice. The Hamiltonian is

$$H = \sum_{j=1}^{N} \left[\frac{p_j^2}{2} + V(q_{j+1} - q_j) \right], \tag{1}$$

where $V(z) = z^2/2 + \alpha z^3/3 + \beta z^4/4$, q_j is the displacement from equilibrium of the jth mass, and p_j is its momentum. Without loss of generality, we take the masses and the harmonic constant equal to one, and we consider periodic boundary conditions $p_{N+1} = p_1$, $q_{N+1} = q_1$.

When $\alpha = \beta = 0$, the system is a linear harmonic chain that is completely integrable. The dynamics of the FES becomes trivial, and no exchange of energy among normal modes is allowed. When $\alpha \neq 0$ or $\beta \neq 0$, the nonlinearity drives thermalization by coupling the modes and allowing for energy exchange. Although the latter is very common, the nonlinearity by itself is not sufficient to drive thermalization [34–36].

The energy of the mode k is defined as

$$E_k(t) = \frac{|\hat{p}_k(t)|^2}{2} + \frac{\omega_k^2 |\hat{q}_k(t)|^2}{2}, \tag{2}$$

where $\omega_k = |2\sin(\pi k/N)|$, and $\hat{q}_k$ and $\hat{p}_k$ are the Fourier coefficients, e.g., $\hat{q}_k = (1/\sqrt{N}) \sum_{j=1}^{N} q_j e^{i2\pi kj/N}$. We consider the initial data

$$q_j(0) = \sum_{k=1}^{p} \frac{A_k}{\sqrt{p}} \sin(\psi_k) \cos\left(\frac{2\pi kj}{N} + \phi_k\right), \tag{3}$$

$$p_j(0) = \sum_{k=1}^{p} \frac{\omega_k A_k}{\sqrt{p}} \cos(\psi_k) \sin\left(\frac{2\pi jk}{N} + \theta_k\right), \tag{4}$$

where A_k is the amplitude of the kth mode, and ψ_k, ϕ_k, θ_k are the phases. The ratio between kinetic and potential energy in the mode k is determined by ψ_k, while ϕ_k and θ_k are relative spatial offsets between the normal modes.

We consider initial conditions where $E_k(0) = \epsilon/p$ for $1 \leq k \leq p$ and $E_k(0) = 0$ otherwise. We then prove the existence of a timescale t_s at which the FES has a window that exhibits a power-law spectrum

$$E_k(t_s) \sim k^{-\frac{8}{3}}, \qquad k_1 \leq k \leq k_c. \tag{5}$$

The lower bound k_1 is empirically determined, while k_c is placed at the beginning of the exponential cutoff of the spectrum. The width of the window is intensive in N and t_s and scales with the number of excited modes p, depending on the relation between phases. When the phases ψ_k, θ_k, ϕ_k are coherent (e.g., all vanishing or all equispaced), then t_s is a deterministic variable, and we find that $t_s^{-1} \sim (\alpha\sqrt{\epsilon}/N) p^{\frac{3}{2}}$ (see Fig. 2). If the phases are chosen randomly, then t_s becomes a random variable, and we show that t_s is much longer on average, with

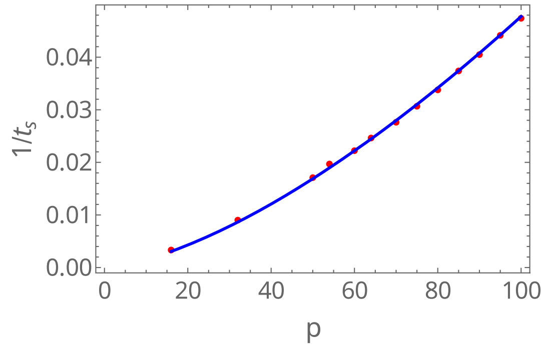

FIG. 2. Numerical estimate of the inverse shock time as a function of the number of excited modes p with all phases equal to zero (points), Eqs. (3) and (4). The solid blue line is the interpolation curve $4.774 \times 10^{-5} p^{\frac{3}{2}}$. $\alpha = 1$, $\beta = 0.1$, $\epsilon = 10^{-3}$, $N = 2048$.

$$\langle t_s^{-1} \rangle \sim \frac{\alpha\sqrt{\epsilon} p \sqrt{\log p}}{N}, \tag{6}$$

for large p (see Fig. 3). Here, $\langle\rangle$ denotes the average over the phases.

When a finite, small fraction ν of modes is initially excited, i.e., $p = \nu N$, then Eq. (6) yields $\langle t_s^{-1}\rangle \sim \alpha\sqrt{\epsilon}\nu\sqrt{\log N}$, which means that, apart from a logarithmic correction $\sqrt{\log N}$, $\langle t_s^{-1}\rangle$ is ruled by intensive parameters only.

Continuum approximation—In order to study the evolution of the initial condition, Eq. (3), for small specific energy and a large number of particles, we introduce the two fields $Q(x,\tau)$ and $P(x,\tau)$ of spatial period one, such that $q_j(t) = NQ(j/N, t/N)$, $p_j(t) = P(j/N, t/N)$. Following [30], we note that $\epsilon = (1/N)\sum_{k=1}^{p} E_k(0)$, up to higher order corrections in ϵ. We introduce the left and right fields $L = (Q_x + P)/(\sqrt{2\epsilon})$ and $R = (Q_x - P)/(\sqrt{2\epsilon})$, where partial derivatives are denoted by subscripts. Neglecting higher order terms in ϵ and $(1/N^2)$, the evolution equations are

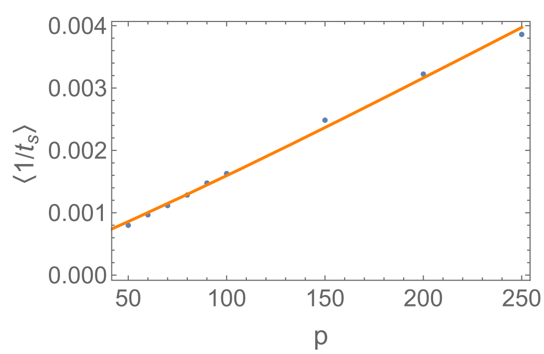

FIG. 3. Average of the inverse shock time $\langle t_s^{-1}\rangle$ vs the number of initially excited modes p (points), Eqs. (3) and (4). Each point is the average over 1000 realizations of the phases. The solid orange line is the theoretical interpolation curve $p\sqrt{\log p}$. $\alpha = 1$, $\beta = 0.1$, $\epsilon = 10^{-3}$, $N = 16\,384$.

$$L_\tau = \left(L + \frac{\alpha\sqrt{\epsilon}}{2\sqrt{2}}(L+R)^2\right)_x, \tag{7}$$

$$R_\tau = -\left(R + \frac{\alpha\sqrt{\epsilon}}{2\sqrt{2}}(L+R)^2\right)_x. \tag{8}$$

Since the equations for L and R are nonlinearly coupled, their analysis is simplified by means of a close-to-identity canonical transformation $\mathcal{C}_{\sqrt{\epsilon}}$: $(L,R) \mapsto (\lambda,\rho)$, which decouples the equations for the new fields λ and ρ at order $\sqrt{\epsilon}$. Neglecting terms of order ϵ, the transformed equations are

$$\lambda_\tau = \left(1 + \frac{\alpha\sqrt{\epsilon}}{\sqrt{2}}\lambda\right)\lambda_x, \qquad \rho_\tau = -\left(1 + \frac{\alpha\sqrt{\epsilon}}{\sqrt{2}}\rho\right)\rho_x \tag{9}$$

with initial conditions $\big(\lambda(x,0), \rho(x,0)\big) = \big(L(x,0), R(x,0)\big) + O(\sqrt{\epsilon})$. Equations (9) are a couple of inviscid Burgers equations, whose solution is well defined for times $\tau < \tau_s$, where τ_s is the scaled shock time of the system, i.e., $\tau_s = \min\{\tau_s^\rho, \tau_s^\lambda\}$ and $t_s = N\tau_s$.

Statistics of the shock time—Using Formula (7) and Section V-A in [31], the large k asymptotics of the FES at τ_s for initial data of the form Eqs. (3) and (4) is $E_k(\tau_s) \sim k^{-\frac{8}{3}}$.

The scaled time τ_s^η is given by

$$(\tau_s^\eta)^{-1} = \frac{\alpha\sqrt{\epsilon}}{\sqrt{2}} \max_{x\in[0,1]} \eta_x(x),$$
$$\eta(x) = \lambda(x,0), \rho(x,0). \tag{10}$$

Using this formula, the computation of the shock time reduces to a maximization problem for each realization of the initial data. For coherent initial phases in Eqs. (3) and (4) (e.g., all equal to zero), the maximization reduces to finding the maximum of $1/\sqrt{p}\sum_{k=1}^{p} 2\pi k\cos(2\pi k x)$, which occurs at $x = 0$ and scales asymptotically as $p^{\frac{3}{2}}$ for large p. For random initial phases, τ_s becomes a random variable, and computing its average reduces to studying the extremal statistics of the stochastic process, $\eta_x(x)$. We show that, on average, the scaled shock time is significantly longer than in the coherent case, with $\langle\tau_s^{-1}\rangle \sim \alpha\sqrt{\epsilon}p\sqrt{\log p}$. Further details on the argument are given in Supplemental Material [37], while a streamlined version is presented below.

Since the maximum in Eq. (10) is taken over infinitely many points, we first restrict to a finite set, for which the statistics can be bounded as

$$\left\langle \max_{m=1,\ldots,M} \eta_x(x_m)\right\rangle = \int_0^{+\infty} \mathrm{Prob}\Big(\max_{m=1,\ldots,M}\eta_x(x_m) > \lambda\Big)d\lambda$$
$$\leq M\int_0^{+\infty}\mathrm{Prob}\Big(\eta_x(x_1) > \lambda\Big)d\lambda. \tag{11}$$

Because the maximum we want to evaluate is taken over an infinite set, these simple bounds diverge as $M \to \infty$.

However, η_x is a smooth function, so its values at nearby points are not statistically independent. These correlations can, thus, be exploited to improve the bound. In particular, we are able to derive the following sub-Gaussian bound,

$$\mathrm{Prob}\Big(|\eta_x(x)-\eta_x(y)|>\lambda\Big)\lesssim \exp\left(-\frac{\lambda^2 p}{4d_p(x,y)^2}\right), \qquad (12)$$

where $d_p(x,y)^2 = p\langle|\eta_x(x)-\eta_x(y)|^2\rangle$ is the desired measure of correlation. For our process, we derive

$$d_p(x,y)^2 = \frac{1}{2}\sum_{k=1}^{p} k^2\Big(1-\cos\big(2\pi k(x-y)\big)\Big). \qquad (13)$$

In order to carefully combine the bounds of Eq. (11) and exploit correlations, we introduce a family of point lattices $\mathcal{N}_m$ such that, for any $m'<m$, $\mathcal{N}_{m'}$ forms a sublattice of $\mathcal{N}_m$. The optimal way to create this sublattice is to select points that are equally spaced with respect to the distance d_p. The mesh size r^m must decrease exponentially fast, i.e., r^m for $r\in(0,\frac{1}{8})$ is sufficiently small.

The next step consists of a chaining argument: for any $x\in[0,1)$, we construct a sequence of points $x_m\in\mathcal{N}_m$ converging monotonically to x and such that

$$\eta_x(x) = \eta_x(0) + \sum_{m=1}^{\infty}[\eta_x(x_m)-\eta_x(x_{m-1})]. \qquad (14)$$

Then, taking the max on both sides, we find

$$\max_{x\in[0,1]}|\eta_x(x)-\eta_x(0)| \le \sum_{m=1}^{\infty}\max_{x\in[0,1]}|\eta_x(x_m)-\eta_x(x_{m-1})|. \qquad (15)$$

For each fixed m, the sum involves a maximum over a finite set of points with cardinality $|\mathcal{N}_m||\mathcal{N}_{m-1}|$, which allows the application of Eq. (11). We then define $\sigma_m = (8r)^m\sqrt{\log(2^m|\mathcal{N}_m||\mathcal{N}_{m-1}|)}$ and $\sigma=\sum_{m=1}^{\infty}\sigma_m$. We can then show that $\max_{x\in[0,1]}\eta_x(x)$ satisfies the following sub-Gaussian bound:

$$\begin{aligned}
&\mathrm{Prob}\left(\max_{x\in[0,1]}|\eta_x(x)-\eta_x(0)|>\frac{\lambda\sigma}{\sqrt{p}}\right)\\
&\le \mathrm{Prob}\left(\sum_{m=1}^{\infty}|\eta_x(x_m)-\eta_x(x_{m-1})|>\frac{\lambda\sigma}{\sqrt{p}}\right)\\
&\lesssim \sum_{m=1}^{\infty}|\mathcal{N}_m||\mathcal{N}_{m-1}|\exp\left(-\frac{\lambda^2\sigma_m^2}{8d_p(x_m,x_{m-1})^2}\right)\\
&\lesssim 2^{-\lambda^2}.
\end{aligned} \qquad (16)$$

Using Fubini's theorem [43],

$$\begin{aligned}
&\langle \max_{x\in[0,1]}|\eta_x(x)-\eta_x(0)|\rangle\\
&= \int_0^{\infty}\mathrm{Prob}\left(\max_{x\in[0,1]}|\eta_x(x)-\eta_x(0)|>\lambda\right)d\lambda\\
&= \frac{\sigma}{\sqrt{p}}\int_0^{\infty}\mathrm{Prob}\left(\max_{x\in[0,1]}|\eta_x(x)-\eta_x(0)|>\frac{\lambda\sigma}{\sqrt{p}}\right)d\lambda\\
&\le \frac{\sigma}{\sqrt{p}}.
\end{aligned} \qquad (17)$$

It thus remains to estimate σ, which can be bounded as follows:

$$\sigma \lesssim \int_0^{+\infty}\sqrt{\log N_p(\varepsilon)}\,d\varepsilon, \qquad (18)$$

where $N_p(\varepsilon)$ is the packing number associated with the distance d_p; that is, the maximum number of points in [0, 1] whose d_p distance is at least ε. The main contribution to the integral comes from the small ε region. A similar bound was established by Dudley for Gaussian processes [32]. In order to compute the packing number, we estimate Eq. (13) as follows:

$$d_p(x,y)^2 \le 2\pi^2|x-y|^{(\log p)^{-1}}\sum_{k=1}^{p}k^{2+(\log p)^{-1}}. \qquad (19)$$

This quantitative comparison between the canonical distance d_p and the Euclidean distance allows us to compare their respective packing numbers, the latter being explicitly computable. This leads to an upper bound on $N_p(\varepsilon)$ which, when combined with Eq. (18), gives $\sigma\lesssim p^{\frac{3}{2}}\sqrt{\log p}$. Together, the estimates Eqs. (10) and (17) imply an upper bound for the inverse of the shock time.

Concerning the lower bound, following the pioneering work of Talagrand [33,44], we show that

$$\langle \max_{x\in[0,1]}\eta_x(x)\rangle \gtrsim \frac{1}{\sqrt{p}}\int_0^{+\infty}\sqrt{\log N_p(\varepsilon)}\,d\varepsilon. \qquad (20)$$

For $p\gg 1$ and $\varepsilon \ll p^{\frac{3}{2}}$, we explicitly construct $p^{\frac{5}{2}}\varepsilon^{-1}$ points in [0, 1] whose distance is at least ε when measured with respect to d_p, thereby realizing the packing number $N_p(\varepsilon)$.

The construction of this large set of ε-separated points relies on two key ideas. First, a local approximation $d_p(x,0)\sim p^{\frac{3}{2}}[px+o(px)]$ allows us to construct a cluster of $p^{\frac{3}{2}}\varepsilon^{-1}$ points in the interval $[0,(1/p)]$ that are ε separated. Summing the series Eq. (13), one finds that $d_p(x,y)>p^{\frac{3}{2}}$ whenever $|x-y|\gtrsim p^{-1}$. This enables the construction of additional clusters of $p^{\frac{3}{2}}\varepsilon^{-1}$ points, provided they are p^{-1} separated in the Euclidean distance. Altogether, this yields a set of $p^{\frac{5}{2}}\varepsilon^{-1}$ points that are ε separated with respect to the distance d_p. This yields the lower bound $N_p(\varepsilon)\gtrsim p^{\frac{5}{2}}\varepsilon^{-1}$, which, together with Eqs. (10)

and (20), provides a sharp lower bound for the inverse of the shock time.

Conclusions—In this Letter, we have shown that the Burgers turbulence scenario in the FPUT model is robust when p long-wavelength modes are initially excited with random phases. By developing advanced probabilistic techniques based on [32,33], we obtained an analytic expression for the large p asymptotics of the average shock time. This computation is crucial for understanding the formation of the prethermal state for typical initial conditions.

Our approach explains the previously observed differences in the scaling behavior of the prethermal state for incoherent versus coherent phases [29]. Up to a logarithmic correction, when an extensive fraction of modes is initially excited, two distinct behaviors emerge: for random phases, the average shock time, Eq. (6), is governed solely by intensive parameters, whereas for coherent phases, it vanishes as $N \to \infty$. This implies that the scenario in which the shock time vanishes corresponds to a measure-zero set of synchronized initial phases and is, thus, not relevant for the average from a statistical mechanics point of view. The derivation of the logarithmic corrections relies on sophisticated probabilistic techniques, detailed in Supplemental Material [37].

To place our results in context, we recall that the evolution of integrable systems with random initial data can mimic certain aspects of hydrodynamic turbulence, such as power-law spectra and the formation of coherent structures [45–47]. Within the approximation considered here, the dynamics under long-wavelength random initial data fits naturally within the framework of integrable turbulence. In fact, our techniques also allow one to derive the full statistics of the shock time for the Burgers equation when random initial data are considered. The statistical properties of shocks in the Burgers equation with various classes of random initial data have been the subject of extensive study [48–51]. Rigorous results on the spatial statistics of shocks in the zero-dissipation limit have also been obtained [50,52]. To our knowledge, however, the statistics of the first shock time—corresponding to the onset of the turbulence regime in FPUT—has not been previously studied. Addressing this question requires the technical apparatus developed in this Letter.

The discrepancy in the spatial profile between Burgers dynamics and the lattice FPUT model arises shortly before the shock time. At this time, the dispersion of the lattice model is responsible for the generation of a train of short-wavelength oscillations. As time passes, the short-wavelength oscillations cover all the space profile of the wave and, as a result, if the energy per particle is large enough, thermalization is observed. Such a phenomenology resembles the one of tygers in the Galerkin-truncated Burgers equation [53–55], with a crucial difference: the tygers appear far away from the spatial point of the shock, while in our case, the short-wavelength oscillations are originated at the shock point. However, the route to thermalization after the shock time is qualitatively similar, and possible analogies deserve further investigations.

Our approach extends well beyond the FPUT problem. Adapting the arguments of [54], it yields the formation time of tygers in the Galerkin-truncated Burgers equation when a large number of Fourier modes is initially excited (see [37]). This probabilistic approach naturally generalizes to the study of extremal statistics in nonlinear systems, both at thermal equilibrium and far from equilibrium, as in the case considered extensively in this Letter.

Thermalization in the FPUT chain remains an open and debated problem. Several approaches have been proposed to describe its long-time dynamics. For certain classes of initial data, heuristic arguments based on wave turbulence theory predict specific thermalization timescales [56], whereas other initial conditions generate periodic solutions, known as q-breathers [57,58], which never thermalize. The picture becomes even subtler in near-integrable regimes, where the network of interactions among actions and phases has been characterized [59].

The probabilistic framework presented here offers a novel perspective on this long-standing problem by leveraging recent advancements in the statistical characterization of random functions. Indeed, these methodologies are applicable to a broad spectrum of nonlinear problems, provided a functional mapping can be established between the initial probability distribution and its time-evolved counterpart.

Acknowledgments—This work was funded by the European Union—Next Generation EU. M. G., R. G., and A. P. acknowledge support by INdAM (GNFM and GNAMPA). R. G. acknowledges financial support from the GNAMPA project "Deterministic and probabilistic evolution of out-of-equilibrium Hamiltonian systems." S. R. acknowledges support from the MUR PRIN2022 Project No. BECQuMB Grant No. 20222BHC9Z. E. D. acknowledges ENS-Lyon for financially supporting his visit to SISSA and SISSA for the hospitality.

Views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

Data availability—The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

[1] R. Balescu, *Statistical Dynamics: Matter Out of Equilibrium* (Imperial College Press, London, 1997).

[2] E. Fermi, J. Pasta, and S. Ulam, in *Collected Papers of Enrico Fermi*, Vol. 2, edited by E. Segré (University of Chicago Press, Chicago, 1965), pp. 978–988.

[3] V. Balasubramanian, A. Buchel, S. R. Green, L. Lehner, and S. L. Liebling, Holographic thermalization, stability of anti–de Sitter space, and the Fermi-Pasta-Ulam paradox, Phys. Rev. Lett. **113**, 071601 (2014).

[4] D. Midtvedt, A. Croy, A. Isacsson, Z. Qi, and H. S. Park, Fermi-Pasta-Ulam physics with nanomechanical graphene resonators: Intrinsic relaxation and thermalization from flexural mode coupling, Phys. Rev. Lett. **112**, 145503 (2014).

[5] L. Cao, D. Qi, R. Peng, M. Wang, and P. Schmelcher, Phononic frequency combs through nonlinear resonances, Phys. Rev. Lett. **112**, 075505 (2014).

[6] D. Pierangeli, M. Flammini, L. Zhang, G. Marcucci, A. Agranat, P. Grinevich, P. Santini, C. Conti, and E. DelRe, Observation of Fermi-Pasta-Ulam-Tsingou recurrence and its exact dynamics, Phys. Rev. X **8**, 041017 (2018).

[7] T. Kinoshita, T. Wenger, and D. S. Weiss, A quantum Newton's cradle, Nature (London) **440**, 900 (2006).

[8] P. Villain and M. Lewenstein, Fermi-Pasta-Ulam problem revisited with a Bose-Einstein condensate, Phys. Rev. A **62**, 043601 (2000).

[9] I. Danshita, R. Hipolito, V. Oganesyan, and A. Polkovnikov, Quantum damping of Fermi-Pasta-Ulam revivals in ultracold Bose gases, Prog. Theor. Exp. Phys. **2014**, 43I03 (2014).

[10] G. W. Ford, M. Kac, and P. Mazur, Statistical mechanics of assemblies of coupled oscillators, J. Math. Phys. (N.Y.) **6**, 504 (1965).

[11] S. Lepri, R. Livi, and A. Politi, Heat conduction in chains of nonlinear oscillators, Phys. Rev. Lett. **78**, 1896 (1997).

[12] T. Prosen and D. K. Campbell, Momentum conservation implies anomalous energy transport in 1D classical lattices, Phys. Rev. Lett. **84**, 2857 (2000).

[13] S. Lepri, R. Livi, and A. Politi, Thermal conduction in classical low-dimensional lattices, Phys. Rep. **377**, 1 (2003).

[14] C. B. Mendl and H. Spohn, Dynamic correlators of Fermi-Pasta-Ulam chains and nonlinear fluctuating hydrodynamics, Phys. Rev. Lett. **111**, 230601 (2013).

[15] L.-S. Young, Stochastic stability of hyperbolic attractors, Ergodic Theory Dyn. Syst. **6**, 311 (1986).

[16] Y. Kifer, *Random Perturbations of Dynamical Systems* (Birkhäuser, Boston, 1988).

[17] J. Ferreira Alves and V. Araújo, in *Geometric Methods in Dynamics (I): Volume in Honor of Jacob Palis*, Astérisque No. 286, edited by W. de Melo, M. Viana, and J.-C. Yoccoz (Société mathématique de France, 2003), pp. 25–62.

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

[19] T. Mori, T. N. Ikeda, E. Kaminishi, and M. Ueda, Thermalization and prethermalization in isolated quantum systems: A theoretical overview, J. Phys. B **51**, 112001 (2018).

[20] G. Gallavotti, *The Fermi-Pasta-Ulam Problem* (Springer, Berlin, Heidelberg, 2008).

[21] J. Berges, S. Borsányi, and C. Wetterich, Prethermalization, Phys. Rev. Lett. **93**, 142002 (2004).

[22] D. J. Luitz, R. Moessner, S. Sondhi, and V. Khemani, Prethermalization without temperature, Phys. Rev. X **10**, 021046 (2020).

[23] A. Kyprianidis, F. Machado, W. Morong, P. Becker, K. S. Collins, D. V. Else, L. Feng, P. W. Hess, C. Nayak, G. Pagano, N. Y. Yao, and C. Monroe, Observation of a prethermal discrete time crystal, Science **372**, 1192 (2021).

[24] A. Stasiuk and P. Cappellaro, Observation of a prethermal $U(1)$ discrete time crystal, Phys. Rev. X **13**, 041016 (2023).

[25] T. Oka and S. Kitamura, Floquet engineering of quantum materials, Annu. Rev. Condens. Matter Phys. **10**, 387 (2019).

[26] G. He, B. Ye, R. Gong, Z. Liu, K. W. Murch, N. Y. Yao, and C. Zu, Quasi-Floquet prethermalization in a disordered dipolar spin ensemble in diamond, Phys. Rev. Lett. **131**, 130401 (2023).

[27] R. Livi, M. Pettini, S. Ruffo, M. Sparpaglione, and A. Vulpiani, Equipartition threshold in nonlinear large Hamiltonian systems: The Fermi-Pasta-Ulam model, Phys. Rev. A **31**, 1039 (1985).

[28] H. Kantz, R. Livi, and S. Ruffo, Equipartition thresholds in chains of anharmonic oscillators, J. Stat. Phys. **76**, 627 (1994).

[29] G. Benettin, R. Livi, and A. Ponno, The Fermi-Pasta-Ulam problem: Scaling laws vs. initial conditions, J. Stat. Phys. **135**, 873 (2008).

[30] M. Gallone, M. Marian, A. Ponno, and S. Ruffo, Burgers turbulence in the Fermi-Pasta-Ulam-Tsingou chain, Phys. Rev. Lett. **129**, 114101 (2022).

[31] M. Gallone, A. Ponno, and S. Ruffo, Energy cascade and burgers turbulence in the Fermi-Pasta-Ulam-Tsingou chain, Phys. Rev. E **110**, 054212 (2024).

[32] R. Dudley, The sizes of compact subsets of Hilbert space and continuity of Gaussian processes, J. Funct. Anal. **1**, 290 (1967).

[33] M. Talagrand, *Upper and Lower Bounds for Stochastic Processes: Modern Methods and Classical Problems* (Springer, Berlin, Heidelberg, 2014).

[34] M. Toda, Wave propagation in anharmonic lattices, J. Phys. Soc. Jpn. **23**, 501 (1967).

[35] M. Hénon, Integrals of the Toda lattice, Phys. Rev. B **9**, 1921 (1974).

[36] T. Goldfriend and J. Kurchan, Equilibration of quasi-integrable systems, Phys. Rev. E **99**, 022146 (2019).

[37] See Supplemental Material at http://link.aps.org/supplemental/10.1103/qxzs-1t11 for random initial data and average shock time in the Fermi-Pasta-Ulam-Tsingou chain. It includes details of the derivation of the main results and Refs. [38–42].

[38] M. Gallone and A. Ponno, Hamiltonian field theory close to the wave equation: From Fermi-Pasta-Ulam to water waves, in *Qualitative Properties of Dispersive PDEs* (Springer Nature, Singapore, 2022), pp. 205–244.

[39] M. Gallone, The role of integrability in Fermi-Pasta-Ulam-Tsingou-like models, arXiv:2509.17753.

[40] M. Avellaneda, Statistical properties of shocks in Burgers turbulence, II: Tail probabilities for velocities, shock-strengths and rarefaction intervals, Commun. Math. Phys. **169**, 45 (1995).

[41] C. Rampf, U. Frisch, and O. Hahn, Eye of the Tyger: Early-time resonances and singularities in the inviscid Burgers equation, Phys. Rev. Fluids **7**, 104610 (2022).
[42] D. Venkataraman and S. Sankar Ray, The onset of thermalization in finite-dimensional equations of hydrodynamics: insights from the Burgers equation, Proc. R. Soc. A **473**, 20160585 (2017).
[43] E. H. Lieb and M. Loss, *Analysis*, Graduate Studies in Mathematics (American Mathematical Society, Providence, RI, 2001).
[44] M. Talagrand, Majorizing measures: the generic chaining, Ann. Probab. **24**, 1049 (1996), https://www.jstor.org/stable/2244967.
[45] V. E. Zakharov, Turbulence in integrable systems, Stud. Appl. Math. **122**, 219 (2009).
[46] J. M. Soto-Crespo, N. Devine, and N. Akhmediev, Integrable turbulence and rogue waves: Breathers or solitons?, Phys. Rev. Lett. **116**, 103901 (2016).
[47] D. S. Agafontsev and V. E. Zakharov, Growing of integrable turbulence, Low Temp. Phys. **46**, 786 (2020).
[48] S. Kida, Asymptotic properties of Burgers turbulence, J. Fluid Mech. **93**, 337 (1979).
[49] J.-D. Fournier and U. Frisch, L'équation de Burgers déterministe et statistique, J. Méc. Théor. Appl **2**, 699 (1983).
[50] Y. G. Sinai, Statistics of shocks in solutions of inviscid Burgers equation, Commun. Math. Phys. **148**, 601 (1992).
[51] Z.-S. She, E. Aurell, and U. Frisch, The inviscid Burgers equation with initial data of Brownian type, Commun. Math. Phys. **148**, 623 (1992).
[52] M. Avellaneda and W. E, Statistical properties of shocks in Burgers turbulence, Commun. Math. Phys. **172**, 13 (1995).
[53] A. J. Majda and I. Timofeyev, Remarkable statistical behavior for truncated Burgers–Hopf dynamics, Proc. Natl. Acad. Sci. U.S.A. **97**, 12413 (2000).
[54] S. S. Ray, U. Frisch, S. Nazarenko, and T. Matsumoto, Resonance phenomenon for the Galerkin-truncated Burgers and Euler equations, Phys. Rev. E **84**, 016301 (2011).
[55] S. S. Ray, Thermalized solutions, statistical mechanics and turbulence: An overview of some recent results, Pramana **84**, 395 (2015).
[56] M. Onorato, L. Vozella, D. Proment, and Y. V. Lvov, Route to thermalization in the α-Fermi–Pasta–Ulam system, Proc. Natl. Acad. Sci. U.S.A. **112**, 4208 (2015).
[57] S. Flach, M. V. Ivanchenko, and O. I. Kanakov, q-breathers and the Fermi-Pasta-Ulam problem, Phys. Rev. Lett. **95**, 064102 (2005).
[58] H. Christodoulidi, C. Efthymiopoulos, and T. Bountis, Energy localization on q-tori, long-term stability, and the interpretation of Fermi-Pasta-Ulam recurrences, Phys. Rev. E **81**, 016210 (2010).
[59] C. Danieli, T. Mithun, Y. Kati, D. K. Campbell, and S. Flach, Dynamical glass in weakly nonintegrable Klein-Gordon chains, Phys. Rev. E **100**, 032217 (2019).