

Hybrid Monte Carlo algorithm for sampling rare events in space-time histories of stochastic fieldsG. Margazoglou,^{1,2} L. Biferale,¹ R. Grauer,³ K. Jansen,⁴ D. Mesterházy,^{5,*} T. Rosenow,⁶ and R. Tripiccione⁷¹*Department of Physics, University of Rome Tor Vergata and INFN-Tor Vergata, 00133 Rome, Italy*²*Computation-based Science and Technology Research Center, Cyprus Institute, 2121 Nicosia, Cyprus*³*Institut für Theoretische Physik I, Ruhr-University Bochum, 44780 Bochum, Germany*⁴*NIC, DESY, 15738 Zeuthen, Germany*⁵*Institute for Theoretical Physics, University of Bern, 3012 Bern, Switzerland*⁶*Institut für Physik, Brandenburg University of Technology Cottbus-Senftenberg, 03046 Cottbus, Germany*⁷*Department of Physics, Università di Ferrara and INFN-Ferrara, 44122 Ferrara, Italy*

(Received 5 August 2018; revised manuscript received 27 January 2019; published 7 May 2019)

We introduce a variant of the Hybrid Monte Carlo (HMC) algorithm to address large-deviation statistics in stochastic hydrodynamics. Based on the path-integral approach to stochastic (partial) differential equations, our HMC algorithm samples space-time histories of the dynamical degrees of freedom under the influence of random noise. First, we validate and benchmark the HMC algorithm by reproducing multiscale properties of the one-dimensional Burgers equation driven by Gaussian and white-in-time noise. Second, we show how to implement an importance sampling protocol to significantly enhance, by orders of magnitudes, the probability to sample extreme and rare events, making it possible to estimate moments of field variables of extremely high order (up to 30 and more). By employing reweighting techniques, we map the biased configurations back to the original probability measure in order to probe their statistical importance. Finally, we show that by biasing the system towards very intense negative gradients, the HMC algorithm is able to explore the statistical fluctuations around instanton configurations. Our results will also be interesting and relevant in lattice gauge theory since they provide unique insights into reweighting techniques.

DOI: [10.1103/PhysRevE.99.053303](https://doi.org/10.1103/PhysRevE.99.053303)**I. INTRODUCTION**

Intermittency and anomalous scaling are two key features of turbulent flows important for fundamental questions of both out-of-equilibrium systems and applied flow configurations [1,2]. Although these phenomena have been subjects of research for decades, it is fair to say that we are still far from understanding their origin and controlling their statistical properties from first principles. Intermittency is connected to the strong non-Gaussian nature of turbulent energy dissipation, which is dominated by localized quasisingular structures. Anomalous scaling is connected to intermittency via the inertial-range turbulent energy cascade, which proceeds from large to small scales, breaking self-similarity, with power-law correlation functions that do not follow dimensional scaling. The two phenomena are correlated, with the small-scale energy dissipation being the result of the inertial-range energy transfer [2]. The problem is therefore how to characterize the statistical properties of intense but rare hydrodynamical fluctuations, an issue that is difficult to attack with brute force forward-in-time evolution of the underlying partial differential equations due to the unpredictability and sparsity of such events. This sobering state of affairs prompted repeated speculations whether techniques developed for quantum field theory (QFT) might eventually turn out to be useful to attack the existence of these (quasi)singular structures in a nonperturbative way, free from any modeling assumptions [3–10].

The way to proceed is to use the Janssen–de Dominicis [11,12] path-integral approach based on the seminal work by Martin *et al.* [5,13–17] to describe the space-time flow configuration when stirred by a random external forcing. This formalism is based on the introduction of an action that depends on the flow configuration and constructs the measure as a weighted sum of all possible flow realizations. This opens up the possibility to address Navier-Stokes equations using Markov chain Monte Carlo methods well known from lattice QFT and/or statistical mechanics by sampling full space-time histories. Although computationally challenging, this provides a unique perspective on the problem of turbulence in the sense that it allows us to consider systematic improvements of the importance sampling in regions of the phase space where standard (forward-in-time) numerical integration faces difficulties, e.g., due to insufficient statistics. In particular, it allows us to address questions regarding the probability of rare events associated with exceptionally large fluctuations, which are at the focus of turbulence research and often attacked by semianalytical tools based on instanton calculus and large-deviation theory.

Instantons were introduced in turbulence theory in [18], where the probability densities of positive velocity gradients and increments (smooth ramps) were calculated analytically. The calculation of the probability densities of negative velocity gradients and increments (shocks) were performed in [19], where the asymptotic behavior could be determined utilizing the Cole-Hopf transformation [20,21]. Using instantons in the calculation of rare irregular transitions between different attractors in fluid flows was presented in [22,23].

*Corresponding author: mesterh@itp.unibe.ch

The development of numerical methods for the investigation of rare events is a long-standing effort that has been pursued in many disciplines. Important examples are the adaptive multilevel splitting techniques (see [24–26] and references therein), transition path sampling [27,28], and the cloning algorithm [29]. Significant advances have also been established in the field of Molecular Dynamics [30,31], where interesting rare events, such as protein folding, occur on disparate timescales [32]. See Ref. [33] for a recent study of extreme heat waves in climate models. Comparison of these methods with our path-integral-based approach is envisaged for future studies.

The objective of this work is to implement, test, and employ a Hybrid Monte Carlo (HMC) algorithm [34,35] for hydrodynamic turbulence. The HMC algorithm was developed to tackle outstanding problems in the theory of strong interactions [36] and is advantageous for problems where the classical action involves nonlocal terms. We address the case of the one-dimensional random-noise-driven Burgers equation [37], which is widely considered the perfect test bed for new ideas in turbulence [38]. A previous attempt based on the path integral for hydrodynamical systems was explored in [39–41], based on a local successive overrelaxation algorithm [42,43].

From a methodological point of view, our first important result is the validation of the HMC against pseudospectral (PS) forward-time-integration techniques that are widely used in simulations of the random-noise-driven Burgers equation. We clearly stress that while the HMC is certainly not competitive with standard PS methods whenever the interest is confined to low-order flow moments, e.g., the total mean energy and total mean energy dissipation, it becomes unavoidable if the focus is on very large fluctuations, e.g., either high-order moments of velocity increments or extreme events for the space-time distribution of the energy dissipation. Indeed, the main quantitative result about the properties of the Burgers equation is the implementation of an importance sampling technique to steer the HMC algorithm to explore the phase-space region where rare and extreme fluctuations happen. We show later that, due to several technical improvements of the basic HMC algorithm, we are able to probe fluctuations 30 (and more) standard deviations away from the mean for the velocity gradient probability distribution function (PDF), something that would be simply impossible to achieve with standard time-advancing algorithms.

The outline of this article is as follows. In Sec. II we briefly discuss the phenomenology of the random-noise-driven Burgers equation. In Sec. III we introduce the path integral for stochastic dynamics. Section IV introduces the HMC algorithm and details the individual steps of our implementation. In Sec. V we show that the HMC algorithm successfully reproduces the results of a standard PS forward-time-integration method [hereafter also referred to as direct numerical simulation (DNS)] at the example of the stochastic Burgers equation. In Sec. VI we investigate different boundary conditions and constraints in space and time. In Sec. VI A we impose periodic boundary conditions in time, while in Sec. VI B we show how the HMC is capable of consistently enhancing the sampling of extreme and rare events by imposing field-force constraints to systematically support the occurrence of strong negative

velocity gradients. Here we will also discuss the significant performance improvements by the HMC compared to a standard DNS method in regard to the sampling of the tails of the probability distribution function of observables. In Sec. VII we emphasize the significance of instantons for the theory of turbulence and derive the instanton configuration for the Burgers equation. We also show numerical results associated with the methods developed in Sec. VI B to support the relevance of instantons in extreme events. We summarize in Sec. VIII.

II. STOCHASTIC BURGERS EQUATION: A SIMPLE MODEL FOR HYDRODYNAMIC TURBULENCE

In this work we are concerned with the one-dimensional random-noise-driven Burgers equation [37], which can be seen as a prototype system for compressible hydrodynamic turbulence and is given by

$$\partial_t v + v \partial_x v - \nu \partial_x^2 v = \eta. \quad (1)$$

Specifically, we consider the time evolution of the scalar velocity field v in a periodic spatial domain $x \in [-L/2, L/2]$ on a given time interval $t \in [t_0, t_f]$ of length $T = t_f - t_0$; ν is the kinematic viscosity and the random noise $\eta = \eta(x, t)$ is assumed to be centered and Gaussian distributed. Thus, the random noise can be fully characterized in terms of two-point correlations

$$\langle \eta(x, t) \eta(x', t') \rangle \equiv \int \mathcal{D}\eta \mathcal{P}_\eta \eta(x, t) \eta(x', t'), \quad (2)$$

where $\mathcal{P}_\eta \equiv \mathcal{P}[\eta]$ is the random-noise probability distribution functional and the integration $\int \mathcal{D}\eta \dots$ is taken over all field configurations $\eta = \eta(x, t)$.

Generally, in the following, $\mathcal{D}\phi$ will denote a functional measure associated with the field ϕ . Path integrals $\int \mathcal{D}\phi \dots$ will always be supplied with boundary conditions in field space. Furthermore, where appropriate, ensemble averages will be denoted by angular brackets. For example, for an observable $\mathcal{O}_\phi \equiv \mathcal{O}[\phi]$ we have $\langle \mathcal{O}_\phi \rangle = \int \mathcal{D}\phi \mathcal{P}_\phi \mathcal{O}_\phi$ and $\int \mathcal{D}\phi \mathcal{P}_\phi = 1$. Depending on context, we might drop the index indicating the field degrees of freedom to be averaged over.

In this paper we restrict our attention to the case where the random noise is self-similar in space and δ correlated in time, with the corresponding two-point Fourier correlation given by

$$\langle \eta(k, t) \eta(k', t') \rangle = \Gamma(k) \delta_{k+k', 0} \delta(t - t'), \quad (3)$$

where $k, k' \in \mathbb{Z}$ and $\Gamma(k) = \Gamma_0 |k|^\beta$ with a negative power-law exponent β that controls the scale-by-scale energy injection.¹

From Eq. (1) it is easy to derive the evolution equation for the energy spectrum $E(k, t) = |v(k, t)|^2$,

$$\partial_t E(k, t) = T(k, t) - 2\nu k^2 E(k, t) + 2 \operatorname{Re}\{v(k, t)^* \eta(k, t)\}, \quad (4)$$

where $T(k, t) = (k/L) \sum_{k'} \operatorname{Im}\{v(k, t)^* v(k', t) v(k - k', t)\}$ is the energy transfer [2]. Equation (4) can be further simplified

¹In practice, when $\beta < -1$, the correlator will be regularized by an infrared cutoff $k_{\text{IR}} \sim 1/L$ and $k_{\text{IR}} > 0$.

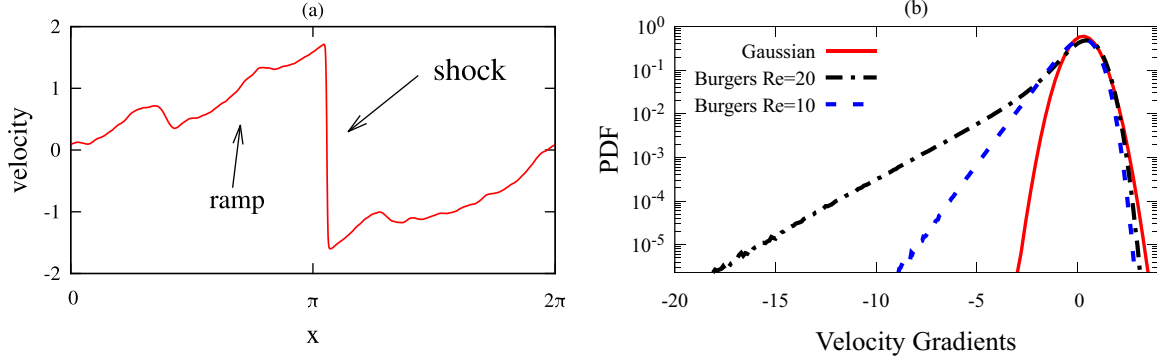


FIG. 1. (a) Shock wave as a particular realization of the Burgers equation. Its two main features are the ramps and jumps of finite width. (b) PDF of velocity gradients of the Burgers equation for two different Reynolds numbers compared with a Gaussian distribution. The finite jumps in the shock wave profile are responsible for the heavy left tail, while the right tail is due to the ramps.

if we average over noise realizations and assume stationarity:

$$\langle T(k) \rangle - 2\nu k^2 \langle E(k) \rangle + 2\Gamma(k) = 0. \quad (5)$$

The ensemble-average cumulative energy injection due to the stochastic forcing η is given by $\langle \varepsilon_{\text{in}}(k) \rangle = (2/L) \sum_{|k'| \leq k} \Gamma(k')$ and is dominated by the infrared regime only if $\beta < -1$. Thus, in order to mimic the standard large-scale injection we will always keep $\beta = -3$ in this paper (see [44–47] for a detailed investigation of the statistical properties at changing the forcing slope).

It is well known that the evolution to the Burgers equation is characterized by the formation of quasisingular shocks, i.e., localized events with a steep negative velocity gradient where all the dissipation is concentrated. In the small-viscosity limit the typical width of the shock becomes smaller, but the ensemble-average mean energy dissipation $\langle \varepsilon_{\text{diss}} \rangle = (2\nu/L) \sum_k k^2 \langle E(k) \rangle$ remains nonvanishing. Since $T(k)$ only transfers the energy between different modes but does not contribute to the total energy, the total energy injection matches the energy dissipation:

$$\lim_{k \rightarrow \infty} \langle \varepsilon_{\text{in}}(k) \rangle = \langle \varepsilon_{\text{diss}} \rangle. \quad (6)$$

Writing Eq. (1) in a dimensionless way reveals that the problem has only one control parameter, the Reynolds number Re . This is made manifest by introducing characteristic scales of length L_0 and velocity V_0 and a time scale $T_0 = L_0/V_0$. We change x , t , v , and η according to

$$\begin{aligned} x &\mapsto xL_0, & v &\mapsto vV_0, & t &\mapsto tT_0 = tL_0/V_0, \\ \eta &\mapsto \eta V_0/T_0 = \eta V_0^2/L_0 \end{aligned} \quad (7)$$

to obtain the dimensionless stochastic Burgers equation

$$\partial_t v + v \partial_x v - \frac{1}{\text{Re}} \partial_x^2 v = \eta, \quad (8)$$

with $\text{Re} \equiv L_0 V_0 / \nu$. Consequently, in the remainder of this paper we will speak about the large-Reynolds-number and small-viscosity limits interchangeably.

In Fig. 1 two of the most characteristic elements of Burgers turbulence are shown. Figure 1(a) depicts the shock formation as a solution of the Burgers equation, which is described by the finite-width jumps and approximately linear ramps. In Fig. 1(b) we show the probability distribution function of the

velocity gradients, defined as

$$P(w) = \langle \delta(\partial_x v(x, t) - w) \rangle. \quad (9)$$

The localized jumps at the shock are the source of intermittency in this model and contribute to the heavy left tail of the PDF, while the ramps are related to the right tail (we refer the reader to [38] for an in-depth review of Burgers turbulence). From the previous discussion the question we ultimately want to address becomes clear: Is it possible to develop algorithms which are able to focus specifically on the phenomenon of shock formation by exploring only the far left tail of the PDF shown in Fig. 1(b)? This will be the aim of the HMC approach we propose.

III. PATH INTEGRAL FOR STOCHASTIC DYNAMICS

The path integral for stochastic dynamics was first introduced in Refs. [48–50]. To make our exposition self-consistent, however, we briefly repeat the main steps of its derivation. While we employ the same notation as in Eq. (1), we emphasize that the following reasoning is in principle applicable to any stochastic (partial) differential equation (SPDE) driven by Gaussian random noise, δ correlated in time. We will denote these SPDEs by the shorthand notation

$$F(x, t, v, \partial_x^m v, \partial_t^n v) = \eta, \quad (10)$$

with $m, n \in \mathbb{N}_0$. Here $F \equiv F(x, t, v, \partial_x^m v, \partial_t^n v)$ should be interpreted as a (nonlinear) differential operator, which acts on the dynamical field $v = v(x, t)$ with m and n denoting the highest-order of spatial and temporal derivatives occurring in a finite linear combination of derivatives in F . We will only make some minimal assumptions regarding its form, namely, that it should yield a well-posed initial value problem. By well-posed we mean that for any given random noise realization η , there exists one and only one solution v to Eq. (10) in the domain $-L/2 \leq x \leq L/2$ and for finite times $0 \leq t \leq T$.

To derive the path integral associated with Eq. (10) we define the partition sum \mathcal{Z} by integrating \mathcal{P}_η over all noise realizations. Since η is Gaussian and white in time, we have

$$\mathcal{P}_\eta \propto \exp \left(-\frac{1}{2} \int dt \int dx \eta(x, t) \int dx' \Gamma^{-1}(x - x') \eta(x', t) \right), \quad (11)$$

where Γ^{-1} is the inverse of the correlation function of the noise defined in Eq. (3). Accordingly, we define the partition sum as

$$\mathcal{Z} = \int \mathcal{D}\eta \exp\left(-\frac{1}{2} \int dt (\eta, \Gamma^{-1} * \eta)\right), \quad (12)$$

where the binary operator $*$ denotes the convolution, i.e., $(f * g)(x) = \int dx' f(x')g(x-x')$, and by (\cdot, \cdot) we designate the integral over the (bounded) spatial domain $[-L/2, L/2]$, i.e., $(f, g) \equiv \int dx f(x)g(x)$, with $\|f\|^2 \equiv (f, f)$. Changing the integration in Eq. (12) from η to v modifies the functional measure as

$$\mathcal{D}\eta = \mathcal{D}v |\det(\delta F/\delta v)|, \quad (13)$$

where $\mathcal{J} = |\det(\delta F/\delta v)|$ is the Jacobian associated with the map $v \mapsto \eta$. The latter is assumed to be nonsingular and therefore $\mathcal{J} > 0$.

Putting everything together, we may write the partition sum in the form of a path integral over v ,

$$\mathcal{Z} = \int \mathcal{D}v \mathcal{J} \exp\left(-\frac{1}{2} \int dt (F, \Gamma^{-1} * F)\right) \equiv \int \mathcal{D}v e^{-\mathcal{S}}, \quad (14)$$

with the action

$$\mathcal{S} = \frac{1}{2} \int dt (F, \Gamma^{-1} * F) - \ln \mathcal{J}, \quad (15)$$

associated with the stochastic partial differential equation (PDE) (10). Note that the action bears resemblance to the well-known Onsager-Machlup functional [51]. More generally, one arrives at the following action:

$$\mathcal{S} = -\ln\{\mathcal{Z}\mathcal{P}_\eta[\eta = F]\} - \ln \mathcal{J}. \quad (16)$$

The probability distribution functional \mathcal{P}_v for the dynamical field v is given by $\mathcal{P}_v = \mathcal{Z}^{-1} e^{-\mathcal{S}}$ and satisfies the normalization condition $\int \mathcal{D}v \mathcal{P}_v = 1$. Specifically, for the Burgers equation $F = \partial_t v + v\partial_x v - \nu\partial_x^2 v$ and $\mathcal{J} = \text{const}$ (which holds for causal forward-time propagation; see, e.g., [52] and Sec. V of this paper), the action (15) takes the form

$$\begin{aligned} \mathcal{S} &= \frac{1}{2} \int dt \int dx (\partial_t v + v\partial_x v - \nu\partial_x^2 v) \\ &\quad \times \int dx' \Gamma^{-1}(x-x') (\partial_t v + v\partial_x v - \nu\partial_x^2 v), \end{aligned} \quad (17)$$

where we dropped the constant contribution from the Jacobian.

IV. HYBRID MONTE CARLO ALGORITHM

The Hybrid Monte Carlo algorithm, originally introduced in [34], has become a standard computational tool to tackle demanding numerical simulations of quantum field theories in the path-integral formulation (see [53,54] for reviews). It belongs to the broad class of Markov chain Monte Carlo methods and uses artificial Hamiltonian dynamics, frequently termed Molecular Dynamics (MD), to advance the dynamical degrees of freedom in Monte Carlo time to generate unbiased field samples. A main feature of the HMC is that dynamical fields and their conjugate momenta can be evolved in parallel

in a given time step of the evolution, if, e.g., a leapfrog-type integrator is used. This makes the HMC most suitable for problems where the classical action of the theory features nonlocal terms. They may arise from the stochastic equation itself (as the pressure term in the Navier-Stokes equation) or, in addition, from the convolution with the inverse force correlator (as in the present case).

In this work we apply the HMC algorithm for a stochastically driven PDE to the example of the Burgers equation. In order to be self-contained, we will first briefly review its basic elements. Then we proceed to discuss important improvements to the HMC algorithm, which allow for a significant enhancement of performance to sample various statistical estimators in a stable and consistent way.

In the HMC algorithm a set of momenta is introduced which are conjugate to the, in our case, velocity fields. Adding these momenta to the partition sum of Eq. (14) leads to an (artificial) Hamiltonian, which governs the dynamics in a fictitious Molecular Dynamics time s via Hamilton's equations of motion. In practice, the numerical solution of Hamilton's equations starting from some initial MD time, say, $s = 0$, to a final time $s = \tau$ has to be performed in discrete steps, which leads to the fact that the energy of the artificial Hamiltonian system is not conserved. This can be repaired by adding a global reject or accept step which makes the algorithm exact and guarantees the convergence to the desired probability distribution. See [55] for a general review of the HMC algorithm.

To be more concrete, the HMC algorithm starts by generating a set of Gaussian-distributed momenta $\pi = \pi(x, t)$ such that the partition sum is modified as

$$\mathcal{Z} \propto \int \mathcal{D}\pi \exp\left(-\frac{1}{2} \int dt \|\pi(t)\|^2\right) \int \mathcal{D}v e^{-\mathcal{S}}. \quad (18)$$

Identifying $\mathcal{K} = \frac{1}{2} \int dt \|\pi(t)\|^2$ as the kinetic term and \mathcal{S} as the potential, we may interpret $\mathcal{H} = \mathcal{K} + \mathcal{S}$ as the Hamiltonian of the system, with the probability distribution functional $\mathcal{P}_{v,\pi} \propto e^{-\mathcal{H}}$. Since $\int \mathcal{D}\pi \mathcal{P}_{(v,\pi)} = \mathcal{P}_v$, the ensemble average of any velocity-dependent observable \mathcal{O}_v remains unaltered. The so constructed Hamiltonian system can now be evolved using Hamilton's equations of motion. In this evolution, the role of time is played by s . In order to make the dependence on s explicit, we will introduce $v_s(x, t)$ and $\pi_s(x, t)$, where the subscript indicates the MD time. Hamilton's equations for the Hamiltonian \mathcal{H} , with the action as in Eq. (17), are then given by

$$\frac{dv_s}{ds} = \frac{\delta \mathcal{H}}{\delta \pi_s(x, t)} = \pi_s(x, t), \quad (19a)$$

$$\frac{d\pi_s}{ds} = -\frac{\delta \mathcal{H}}{\delta v_s(x, t)} = -\frac{\delta \mathcal{S}}{\delta v_s(x, t)}. \quad (19b)$$

In the case of the one-dimensional Burgers equation, the Molecular Dynamics forces $\Phi_\pi \equiv -\frac{\delta \mathcal{S}}{\delta v_s(x, t)}$ acting on the conjugate momenta are given by

$$\begin{aligned} \Phi_\pi(x, t) &= (\partial_t + v\partial_x + \nu\partial_x^2) \\ &\quad \times \int dx' \Gamma^{-1}(x-x') (\partial_t v + v\partial_x v - \nu\partial_x^2 v). \end{aligned} \quad (20)$$

A. The HMC implementation

The equations of motion (19) are solved for (v_s, π_s) , $0 \leq s \leq \tau$, starting at MD time $s = 0$ and integrating up to $s = \tau$; τ defines the trajectory length. We apply a symmetric symplectic integrator (leapfrog scheme) with step size $\Delta\tau = \tau/N_{\Delta\tau}$, with $N_{\Delta\tau}$ a parameter that is the discrete number of steps of the Hamiltonian evolution for a single HMC iteration, so that the trajectory of length τ is completed.

Due to the finite integration step-size error, the Hamiltonian reached at $s = \tau$ will be different from the initial Hamiltonian. To correct for this deficiency, we apply a global Metropolis accept or reject step of the proposed new momentum and velocity-field configuration: The new field configuration is accepted with probability $p = \min(1, e^{-\Delta\mathcal{H}})$, where $\Delta\mathcal{H} = \mathcal{H}[v_\tau, \pi_\tau] - \mathcal{H}[v_0, \pi_0]$, i.e., the difference of the Hamiltonian at the beginning and the end of the trajectory. If the proposal is rejected, we resample the conjugate momenta and restart from the old set $v = v_0$. After each completed HMC iteration, we resample the momenta, regardless the outcome of the Metropolis step. This is necessary to satisfy ergodicity. Other important requirements for the HMC algorithm to be exact are the preservation of the phase-space volume and the reversibility in the fictitious time s . These are inherent properties of the Hamiltonian dynamics.

Reversibility, which is a necessary condition of detailed balance, is in practice measured by first performing a Hamiltonian evolution $(v_0, \pi_0) \mapsto (v_\tau, \pi_\tau)$ and then negating the momenta $\pi_\tau \mapsto -\pi_\tau$. Now, starting from $(v_\tau, -\pi_\tau)$ and performing another Hamiltonian evolution, we return to (v'_0, π'_0) . In any numerical implementation we expect violations of reversibility which are quantified by $\max(|v'_0 - v_0|)/v_{\text{rms}}$. Large reversibility violations will spoil the invariance of the desired distribution (here $e^{-\mathcal{S}}$) under the HMC updates. Therefore, reversibility violations need to be monitored in the actual simulation. Indeed, we checked that reversibility violations are negligible in our simulations, i.e., of order 10^{-12} – 10^{-14} .

Let us finally briefly summarize, and illustrate in Fig. 2, the three basic steps of the HMC algorithm for a single HMC iteration.

1. *Momentum heat bath.* Sample π_0 according to the Gaussian distribution

$$\mathcal{P}_\pi \propto \exp\left(-\frac{1}{2} \int dt \|\pi_0(t)\|^2\right). \quad (21)$$

2. *Hamiltonian evolution.* Use a symplectic integrator to numerically solve the system of equations (19) starting from (v_0, π_0) and propose (v_τ, π_τ) .

3. *Metropolis step.* Accept the proposed field configuration (v_τ, π_τ) with probability

$$p = \min(1, e^{-\Delta\mathcal{H}}), \quad (22)$$

where $\Delta\mathcal{H} = \mathcal{H}[v_\tau, \pi_\tau] - \mathcal{H}[v_0, \pi_0]$.

The generated ensemble is a Markov chain of configurations $v_s(x, t)$ that results from numerous HMC iterations, by repeating steps 1–3. In our simulations the parameters τ and $N_{\Delta\tau}$ are tuned such that acceptance rate originating from step 3 is close to 90% and more. This ensures that the autocorrelation time does not become too large and also avoids too many rejected velocity configurations.

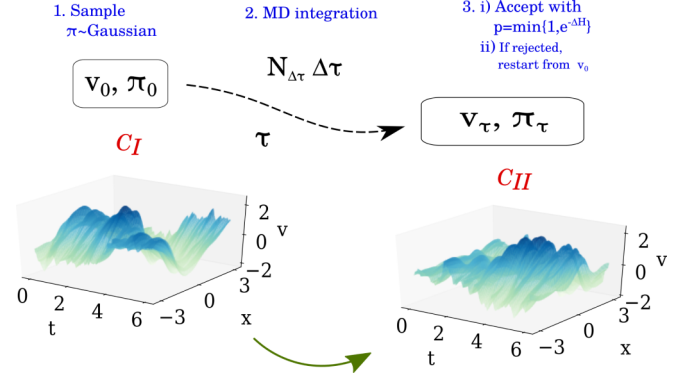


FIG. 2. Schematic description of a single HMC iteration of trajectory length τ . The dashes in the top arrow represent the $N_{\Delta\tau}$ number of intermediate steps of size $\Delta\tau$, so $\tau = N_{\Delta\tau} \Delta\tau$ holds. After the numerical integration of Eqs. (19), a configuration v_τ is proposed and is accepted with probability $p = \min(1, e^{-\Delta\mathcal{H}})$. If accepted, we use v_τ as the initial field for the next iteration; otherwise we restart from v_0 . In both cases, we discard π_τ and resample the momenta according to the Gaussian distribution for the next HMC iteration. (Note that the conjugate momentum fields are not plotted.)

B. Fourier acceleration

We observe that the application of the standard HMC, based on Eq. (18), leads to very large autocorrelation times. The problem with the large autocorrelation time is essentially due to the multiscale nature of the stochastic forcing, which in turn means that different Fourier modes are forced with different intensity. In order to deal with this problem, we made use of a well-known approach from the area of lattice field theory, i.e., the method of Fourier acceleration [56–59]. The latter assigns different effective trajectory lengths to the evolution of the Fourier modes. Indeed, this technique proved highly effective in our approach and it improved the performance of the HMC algorithm by considerably decreasing autocorrelation effects.

In practice, we apply the Fourier acceleration by introducing the space-time-dependent kernel $\Omega(x, t)$ to multiply the momenta $\pi_s(x, t)$. This gives rise to the following effective Hamiltonian:

$$\mathcal{H}^{\text{eff}} = \frac{1}{2} \int dt (\pi_s, \Omega * \pi_s) + \mathcal{S}. \quad (23)$$

It is important to note that we consider $\Omega(x, t)$ as being independent of the MD time s . The introduction of the kernel $\Omega(x, t)$ and the redefinition of the Hamiltonian do not affect the physical results, as the redefined kinematic term is still independent of the velocity field and can be factored out of the path integral (18).

In our HMC implementation, we propose to define Ω as

$$\Omega(x, t) \propto \frac{1}{\langle |\Phi_\pi(x, t)| \rangle^2}, \quad (24)$$

with $\Phi_\pi(x, t) = -\frac{\delta\mathcal{S}}{\delta v_{s=\tau}(x, t)}$ the MD forces recorded at $s = \tau$ and defined in Eq. (20). For a detailed derivation see Appendix A. The ensemble average on the right-hand side of Eq. (24) denotes the MC average and implies that the Fourier acceleration scheme needs to be adjusted dynamically during

the initial convergence phase of the Markov chain. Initially, we use the ansatz

$$\Omega(k, t) \propto \Gamma^{-1}(k)\delta(t - t'). \quad (25)$$

Then, after a fixed number of HMC steps, a tuning stage follows in which we measure the forces Φ_π . This tuning stage can be composed of several cycles, while at the end of each cycle we set Ω to the new forces. As soon as Ω has converged, the tuning stage is completed and we can start the measurement of physical observables. Typically, we require five to ten reset cycles of about 10^3 – 10^4 iterations to achieve an optimal choice of Ω . Finally, since the stochastic forcing is δ correlated in time, it turns out in practice that the final choice for Ω is approximately time independent, i.e., $\Omega(k, t) \approx \Omega(k)\delta(t - t')$. The technical details regarding the implementation of the Fourier acceleration will be discussed elsewhere [60] and are beyond the scope of the present paper.

V. BENCHMARKING THE HMC AGAINST A FIRST-ORDER EULER-MARUYAMA EXPLICIT SOLVER

Fixed and open boundary conditions in time

As this is an effective approach for the sampling of stochastic PDEs, we took considerable care to benchmark the HMC with standard numerical methods that are employed in computational fluid dynamics. In the following we will demonstrate that our simulations match results obtained via a first-order Euler-Maruyama explicit solver (DNS) for a wide range of viscosities [61,62].

In both implementations, the spatiotemporal domain is discretized uniformly and the Burgers equation is expressed in Fourier space, with the nonlinear term being written in a flux-conservative form, i.e., $v\partial_x v = \frac{1}{2}\partial_x(v^2)$. We apply the pseudospectral method, i.e., first v^2 is measured in real space and afterward transformed to Fourier space so that the partial derivative can be conveniently treated as $\partial_x \mapsto ik$. Therefore, the nonlinear term is calculated as $\frac{ik}{2}\mathcal{F}(v^2)$, where by \mathcal{F} we denote the (forward) Fourier transform. To further ensure stability we apply two further steps in the numerics. First, we transform $v(k, t) \rightarrow \exp(-\nu k^2 \Delta t)v(k, t)$, which corresponds to an exact integration of the viscous term. It relaxes the restriction on the time step Δt by the diffusive term and significantly improves the convergence for large wave numbers. Second, we effectively remove the aliasing error by setting $v(2\pi|k|/L \geq N_x/3, t) = 0$ (see [63]).

Here we present three different runs, with the parameters summarized in Tables I and II. Both the DNS and the HMC share the same setup, i.e., the same forcing correlation function, the same discretization, and the same periodic boundary conditions in space. As for the HMC, we choose fixed and open boundary conditions in time, corresponding to a standard initial-value problem. Note that this choice yields a Jacobian \mathcal{J} that is field independent [52], which therefore can be neglected for the purposes of importance sampling.

In Fig. 3(a) we compare the HMC and DNS temporal evolution of the mean kinetic energy $\bar{\varepsilon}_{\text{kin}}(t) = \|v(t)\|^2/L$ for configurations corresponding to three different viscosities. As one can see, the overall intensities of fluctuations are very similar. More quantitatively, in Fig. 4(a) we show the temporal

TABLE I. Parameters and observables of the numerical simulations for fixed and open boundary conditions. Here we employ the following parameters: the number of grid points in time $N_t = 1056$, the number of grid points in space $N_x = 128$, $T = 6$, and $L = 2\pi$. The Reynolds number is defined as $\text{Re} = \frac{v_{\text{rms}}L}{\nu}$, with root-mean-square velocity $v_{\text{rms}} = \langle \sqrt{\|v\|^2}/L \rangle$. Here $\ell_d = (\nu^3/\langle \bar{\varepsilon}_{\text{diss}} \rangle)^{1/4}$ defines the Kolmogorov dissipation length scale and $\ell = \frac{\bar{\varepsilon}_{\text{kin}}^{3/2}}{\langle \bar{\varepsilon}_{\text{diss}} \rangle}$ is the integral length scale with $\bar{\varepsilon}_{\text{kin}} = v_{\text{rms}}^2$. In addition, $\langle \bar{\varepsilon}_{\text{diss}} \rangle$ denotes the ensemble-average mean energy dissipation, i.e., $\langle \bar{\varepsilon}_{\text{diss}} \rangle = 2\nu\langle \|\partial_x v\|^2 \rangle/L$, and $T_\ell = \ell/v_{\text{rms}}$ is the large-eddy turnover time.

ν	Re	v_{rms}	ℓ_d	ℓ	$\langle \bar{\varepsilon}_{\text{diss}} \rangle$	T_ℓ
0.08	90	1.14	0.15	1.5	1	1.31
0.1	70	1.12	0.18	1.4	1	1.24
0.2	30	1.03	0.3	1.1	1	1.03

evolution of the ensemble average of the mean kinetic energy, i.e., $\langle \bar{\varepsilon}_{\text{kin}}(t) \rangle$, starting from $v(x, t_0) = 0$. Around time $t_s \approx 3$ the system reaches stationarity, meaning that the dissipative and injection forces are balanced and the system is driven to a nonequilibrium steady state; beyond t_s , the $\langle \bar{\varepsilon}_{\text{kin}}(t) \rangle$ is constant in time. Figure 4(b) shows the temporal evolution of the ensemble-average mean energy dissipation $\langle \bar{\varepsilon}_{\text{diss}}(t) \rangle$, where $\bar{\varepsilon}_{\text{diss}}(t) = 2\nu\|\partial_x v(t)\|^2/L$, while in Fig. 5(a) we consider the ensemble average of the energy spectrum $\langle \bar{E}(k) \rangle$, which is averaged in time t , i.e., $\bar{E}(k) = \frac{1}{T'} \int_{t_s}^{t_f} dt E(k, t)$, $T' = t_f - t_s$.

In Fig. 5(b) we show the probability distribution function of the velocity gradients. In practice, the PDF is approximated by determining the counts of a fixed number of bins ($w_{\text{min}}, w_{\text{max}}$) with equal width δw . The velocity gradients measured on the generated ensemble are counted only if $t > t_s$. The resulting histogram is normalized by dividing with the total number of counts, i.e.,

$$\sum_i \int_{w_i - \delta w/2}^{w_i + \delta w/2} dw P(w) = 1. \quad (26)$$

From Figs. 3–5 we conclude that the HMC produces the same results as the DNS. Furthermore, we identify the same discretization effects in both implementations, which can be

TABLE II. Technical parameters and observables of the HMC simulations for the three different sets of runs, using fixed and open boundary conditions. Here τ is the trajectory length and $N_{\Delta\tau}$ is the number of steps of the symplectic integrator for a single HMC iteration. The fourth column gives the ensemble size per viscosity or, in other words, the number of HMC iterations performed. Finally, τ_{int} is the integrated autocorrelation time evaluated here for the kinetic energy and measured in units of τ [64]. The ratio of ensemble size to τ_{int} estimates the effective statistically uncorrelated ensemble size.

ν	τ	$N_{\Delta\tau}$	Ensemble size	τ_{int}
0.08	1024	20480	2×10^5	70
0.1	1024	20480	2×10^5	40
0.2	1024	20480	10^5	10

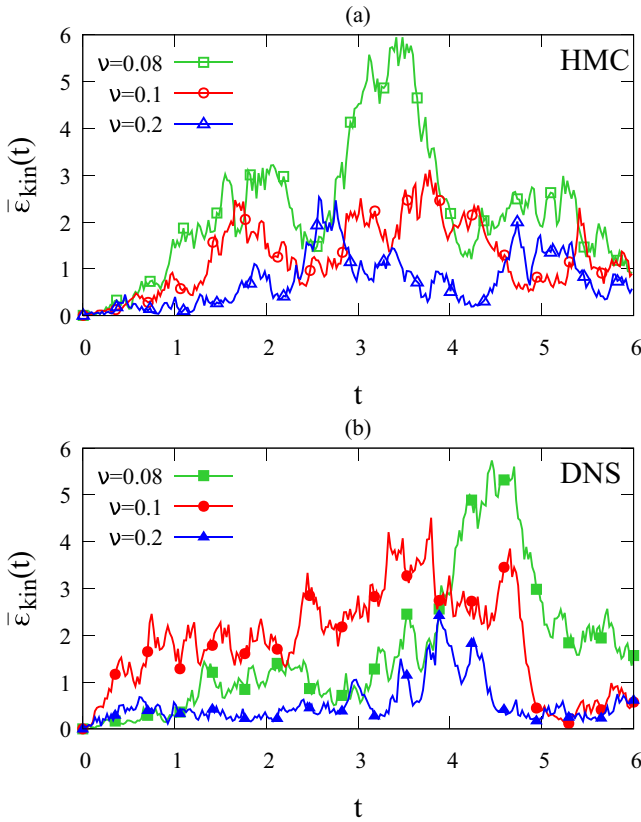


FIG. 3. Kinetic energy as a function of time for three samples corresponding to three different viscosities for the (a) HMC and (b) DNS.

removed by taking the continuum limit. This has been thoroughly checked, but we skip this discussion here.

There are two interesting remarks regarding the behavior of the HMC and in connection with Table II. First, we notice that for fixed resolution and trajectory length τ , the integrated autocorrelation time τ_{int} increases with decreasing viscosity. Second, we manage to perform highly efficient simulations simply by increasing the trajectory length τ while keeping $\Delta\tau$ fixed. Contrary to common practice in lattice QCD, where τ

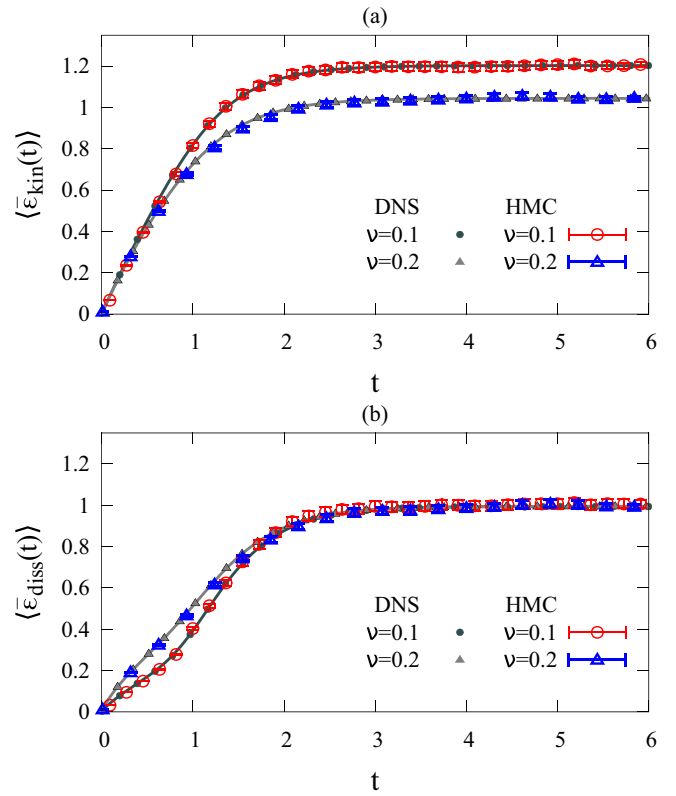


FIG. 4. (a) Temporal evolution of the ensemble-average mean kinetic energy $\langle \bar{\epsilon}_{\text{kin}}(t) \rangle$. (b) Temporal evolution of the ensemble-average mean energy dissipation $\langle \bar{\epsilon}_{\text{diss}}(t) \rangle$. Note that $\langle \bar{\epsilon}_{\text{diss}}(t > t_s) \rangle = 1$ because the energy injection is fixed at $\lim_{k \rightarrow \infty} \langle \bar{\epsilon}_{\text{in}}(k) \rangle = 1$. Closed gray symbols correspond to the DNS results (lines denote their interpolation) and open colored symbols correspond to the HMC results.

is kept of $O(1)$, to avoid energy and reversibility violations [65], in our case it proved a safe and beneficial choice to set τ of order $\tau \approx 10^2$ or $\tau \approx 10^3$ without introducing significant effects of reversibility violations or loss of acceptance rate. This allowed us to significantly decrease autocorrelation

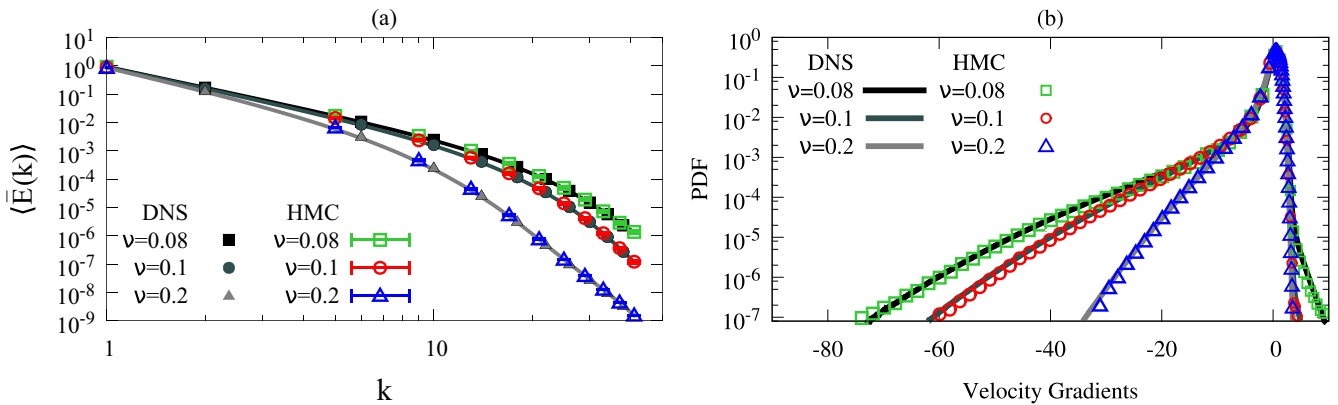


FIG. 5. (a) Ensemble-average energy spectra $\langle \bar{E}(k) \rangle$. Closed gray symbols correspond to the DNS results (lines denote their interpolation) and open colored symbols correspond to the HMC results. (b) Probability distribution functions of velocity gradients. The gray dashed line corresponds to the DNS results, while open colored symbols correspond to the HMC results.

times and avoided the disposal of many generated configurations between measurements. In principle we can increase the trajectory length to higher values that will allow us to generate statistically independent configurations at each HMC iteration, as it is done for the runs in Sec. VI, but we did not check this systematically for the present section.

Finally, a key element of a Monte Carlo–based approach is a rigorous error analysis, for which there are well-established methodologies [66–68]. Since Markov chain Monte Carlo simulations are known to be prone to autocorrelation effects, we had to go through a thorough investigation of the integrated autocorrelation times τ_{int} for each observable. Therefore, as a postproduction step, we used the data analysis package provided in [64] as a tool to estimate the errors of the observables, which takes into account the corresponding autocorrelation effects. This recipe for the error calculation will be followed throughout this article.

VI. CONSTRAINED SPACE-TIME EVOLUTION USING HMC

Now that we have benchmarked the HMC against a standard DNS algorithm, we will present the features and advantages that this path-integral-based approach can bring to the numerical studies of turbulent models, as well as stochastic PDEs in general. First, since the HMC considers the full temporal evolution of the field, this provides an additional flexibility towards the choice of boundary conditions in time. Therefore, in Sec. VIA we show, for instance, that one can apply periodic boundary conditions in time, i.e., $v(x, t) \equiv v(x, t + T)$. Then in Sec. VIB we turn towards the motivation for this article. That is to introduce field constraints, which will affect the Monte Carlo sampling in a controlled way, in order to favor the generation of specific configurations that will comply with the imposed constraint. More specifically, as a first application, we apply a protocol to systematically generate configurations where a large negative velocity gradient is produced at a prescribed space-time point. This also provides with some insight into the underlying dynamics of how the system evolved in time t to reach this extreme condition.

TABLE III. Parameters and observables of the numerical simulations for periodic boundary conditions of the HMC. Here the fixed parameters for both implementations are $N_t = 544$, $N_x = 64$, $T = 6$, and $L = 2\pi$. Also for the HMC $\tau = 128$, and $N_{\Delta\tau} = 2560$. See also Table I for definitions.

ν	Re	v_{rms}	ℓ_d	ℓ	$\langle \bar{\varepsilon}_{\text{diss}} \rangle$	T_ℓ
0.3	20	0.93	0.40	0.81	1	0.87
0.6	7	0.64	0.68	0.26	0.99	0.41
1.4	1	0.31	1.29	0.03	0.97	0.09

A. Time-periodic boundary conditions

As a first application, we discuss the use of periodic boundary conditions in time. Under this scenario, we observe that after the system has equilibrated (to the desired target distribution), the ensemble consists of configurations that have reached stationarity at any time $t \in [t_0, t_f]$. This can be better understood by looking at Fig. 6, where the ensemble average of the mean kinetic energy [Fig. 6(a)] and of the mean energy dissipation [Fig. 6(b)] are constant in time in the example of the HMC (colored symbols). We also show the results of the DNS (gray lines and symbols) using zero initial conditions as a further comparison. The parameters used for the three different runs are summarized in Table III.

The use of periodic boundary conditions in time leads to a field-dependent Jacobian \mathcal{J} [52] and therefore we must expect it to affect the importance sampling. Nevertheless, in this work, we have consistently neglected the evaluation of the Jacobian (which, in the lattice field theory literature, is often referred to as the quenched limit). To get a better impression of the systematic error associated with this approximation, we have chosen to compare our results with periodic boundary conditions to the case of fixed and open boundary conditions. As can be seen from Fig. 6, our results overlap with the stationary regime attained by using fixed and open boundary conditions with reasonable accuracy. In fact, it is possible to show that the error by neglecting the Jacobian vanishes in the limit $T \rightarrow \infty$. Thus, for those cases considered here, the systematic error is likely negligible. We defer the evaluation of the Jacobian determinant to future work.

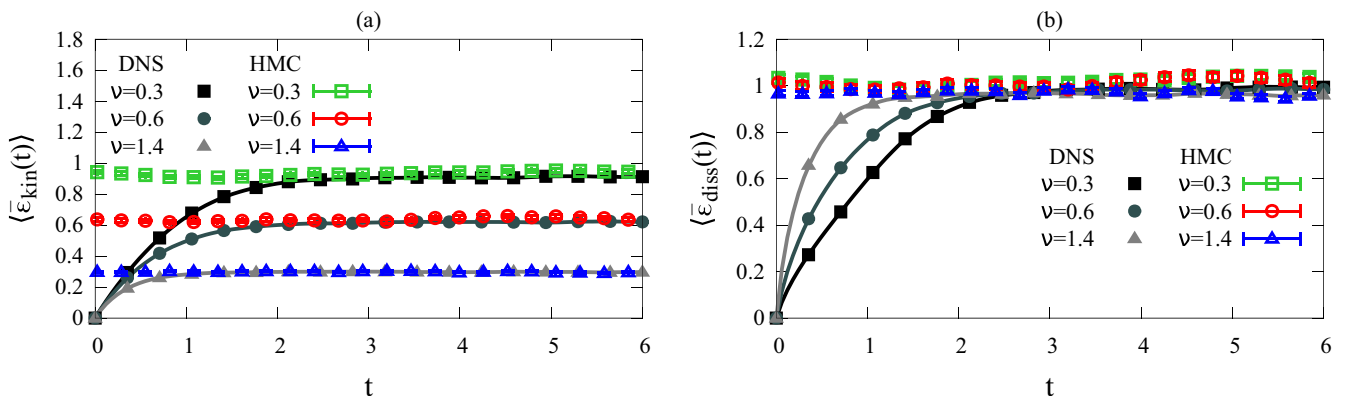


FIG. 6. Results for the HMC using periodic boundary conditions in time. (a) Temporal evolution of the ensemble-average mean kinetic energy $\langle \bar{\varepsilon}_{\text{kin}}(t) \rangle$. (b) Temporal evolution of the ensemble-average mean energy dissipation $\langle \bar{\varepsilon}_{\text{diss}}(t) \rangle$. Closed gray symbols denote the DNS results (lines correspond to their interpolation) and open colored symbols denote the HMC results.

B. Enhanced sampling of extreme and rare events

We will now describe the important steps towards constraining the sampling of the HMC to generate a large negative velocity gradient at a specified space-time point. Also, we will explain how to directly compare the observables obtained from the constrained ensemble with the ones related to an unconstrained ensemble, by using reweighting techniques, and therefore estimate their relative importance with respect to the typical statistics of the system. We note two important points. First, we will use the same boundary conditions as in Sec. V, i.e., periodic in space and fixed and open in time. Second, the statistics of the DNS will be referred to as the ones related to the unconstrained system. We could use the corresponding ones from the HMC with unconstrained sampling, but another purpose of ours is to demonstrate the benefits of employing this method for the purpose of systematically sampling extreme and rare events, compared to a standard DNS implementation, where such instances are a matter of chance.

1. Reweighting

Reweighting is a standard technique introduced in [69] that has proved very helpful in the study of phase transitions and critical phenomena. In short, it allows one to exploit the information of a generated ensemble of a single Monte Carlo simulation performed at a certain parameter (e.g., at fixed inverse temperature β) and obtain results for a range of nearby parameters (e.g., β_i). Reweighting can also provide a way to modify the sampling in a Monte Carlo simulation, which is how we use it here by constraining the sampling of the HMC to enhance the generation of strong negative gradients. What is common in both cases is that we include a reweighting factor in the ensemble averages to obtain the desired ensemble (see [70] for a review on the topic).

In our application, reweighting is employed at the post-production stage as a means to relate the observable $\langle \mathcal{O} \rangle'$, measured using the ensemble which is generated by sampling with respect to the action S' , to $\langle \mathcal{O} \rangle$, measured on the ensemble sampled with the action S . We briefly revisit here the standard steps of the derivation. Note that any probability density functional $\mathcal{P}'_v = e^{-S'}/\mathcal{Z}'$ can be related to another \mathcal{P}_v for the same field configuration $v = v(x, t)$ via

$$\mathcal{P}_v = \frac{1}{\mathcal{Z}} e^{-S} = \frac{\mathcal{Z}'}{\mathcal{Z}} e^{-(S-S')} \mathcal{P}'_v. \quad (27)$$

The expectation value of an observable $\langle \mathcal{O} \rangle$ using S is given by

$$\begin{aligned} \langle \mathcal{O} \rangle &= \int \mathcal{D}v \mathcal{P}_v \mathcal{O} \\ &= \frac{\mathcal{Z}'}{\mathcal{Z}} \int \mathcal{D}v \mathcal{P}'_v e^{-(S-S')} \mathcal{O} \\ &= \frac{\mathcal{Z}'}{\mathcal{Z}} \langle e^{-(S-S')} \mathcal{O} \rangle', \end{aligned} \quad (28)$$

where the notation $\langle \dots \rangle'$ implies that the expectation value is evaluated with the action S' . From the identity $\langle 1 \rangle = 1$ we may derive the relation $\mathcal{Z}/\mathcal{Z}' = \langle e^{\Delta S} \rangle'$, with $\Delta S = S' - S$,

whereby

$$\langle \mathcal{O} \rangle = \frac{\langle e^{\Delta S} \mathcal{O} \rangle'}{\langle e^{\Delta S} \rangle'}. \quad (29)$$

As the next step we determine the error of the estimator $\langle \mathcal{O} \rangle$. Note that in Eq. (29) both the numerator and denominator have fluctuations. Furthermore, as they are calculated from the same ensemble we expect that both errors are correlated. As explained in [71], to estimate the error $\delta \langle \mathcal{O} \rangle$ of $\langle \mathcal{O} \rangle$ we employ the propagation of error of two dependent variables including the covariance and the cross covariance of the nominator and the denominator. Simplifying Eq. (29) to $\langle \mathcal{O} \rangle = A/B$, the final expression is

$$\delta \langle \mathcal{O} \rangle = \langle \mathcal{O} \rangle \sqrt{\left(\frac{\delta A}{A}\right)^2 + \left(\frac{\delta B}{B}\right)^2 - 2\left(\frac{\delta(AB)}{AB}\right)^2}, \quad (30)$$

where $\delta(AB) = \langle AB \rangle - \langle A \rangle \langle B \rangle + 2 \sum_{i,j>i} (A_i - \langle A \rangle)(B_j - \langle B \rangle)$.

2. Implementation of sampling constraints

The idea is to define a different action S' to sample via the HMC, which consists of the original S [Eq. (15)] in addition to a constraint functional ΔS :

$$S' = S + \Delta S. \quad (31)$$

The choice of ΔS cannot be arbitrary. If there is a tiny overlap of the distributions e^{-S} and $e^{-S'}$, the reweighting procedure will most likely not work. Therefore, it is not clear from the beginning for which parameter values a successful reweighting can be performed. We remark that in cases where reweighting fails, one could attempt to insert intermediate reweighting steps as explained in [72]. We also need to stress that any constraint functional ΔS will contribute to the MD forces Φ_π through the functional derivative $\delta S'/\delta v = \delta S/\delta v + \delta \Delta S/\delta v$ and this contribution needs to be evaluated exactly.

Nevertheless, the observables of the HMC are not directly comparable to the DNS. In the following, we will explain how to directly compare the statistics of the HMC using the action S' , with the typical unconstrained statistics using the action S , by utilizing reweighting techniques.

In order to demonstrate the application of Eq. (29), we first discuss the example of the ensemble-average mean kinetic energy before and after reweighting. This is also a sufficient step to further ensure the consistency with the unconstrained statistics, meaning that after reweighting the observable measured by the constrained ensemble should collapse, within error bars, with the corresponding unconstrained one. Following Eq. (29), the reweighted ensemble-average mean kinetic energy will be

$$\langle \bar{\epsilon}_{\text{kin}}(t) \rangle = \frac{\langle e^{\Delta S} \bar{\epsilon}_{\text{kin}}(t) \rangle'}{\langle e^{\Delta S} \rangle'}. \quad (32)$$

As a first attempt we tried a series of local constraint functionals, with a suitable shape, that enhance the probability to produce a large negative velocity gradient at a certain point in the middle of the spatial domain at the last time slice, i.e., ($x = 0, t = t_f$). The parameters that we used for the HMC are

TABLE IV. Parameters for HMC simulations with constrained sampling. The integral length-scale Reynolds number is defined as $\text{Re}_\ell = \frac{v_{\text{rms}}\ell}{\nu}$ and the large-scale Reynolds as $\text{Re} = \frac{v_{\text{rms}}L}{\nu}$. The results for the HMC have been reweighted and the temporal interval for averaging corresponds to the stationary regime. The number in parentheses gives the error of the last digit of the mean. Here the fixed parameters for both DNS and HMC are $N_t = 144$, $N_x = 64$, $T = 6$, $L = 2\pi$, $\nu = 0.5$, and $\ell_d = 0.59$. In addition, κ is defined in Eq. (37). Specifically for the HMC, $\tau = 128$ and $N_{\Delta\tau} = 2560$.

only HMC							
c_1	w_1	Re_ℓ	Re	v_{rms}	$\langle \bar{\epsilon}_{\text{diss}} \rangle$	ℓ	κ
1.2	1	1.02(3)	10(2)	0.8(1)	0.96(3)	0.61(2)	1.12
1.6	1	1.0(2)	10(4)	0.8(3)	0.9(1)	0.6(1)	1.21
1.9	1	0.7(4)	9(5)	0.7(4)	0.7(2)	0.5(2)	1.93
c_2	w_2	Re_ℓ	Re	v_{rms}	$\langle \bar{\epsilon}_{\text{diss}} \rangle$	ℓ	κ
80	12	0.83(5)	10(2)	0.8(2)	0.79(4)	0.55(3)	418
80	18	0.8(2)	9(4)	0.8(4)	0.8(2)	0.6(1)	2.6×10^5
160	24	1.4(6)	12(6)	1.0(5)	1.2(3)	0.7(3)	5×10^9
160	30	1.7(6)	14(7)	1.1(5)	1.6(4)	0.8(2)	2.6×10^{11}
c_3	w_3	Re_ℓ	Re	v_{rms}	$\langle \bar{\epsilon}_{\text{diss}} \rangle$	ℓ	κ
80	12	1.5(4)	12(6)	1.0(5)	1.3(3)	0.8(2)	3.9×10^5
80	18	1.0(3)	11(5)	0.8(4)	1.0(2)	0.6(1)	5.4×10^8
80	24	1.2(4)	11(5)	0.9(4)	1.1(3)	0.7(2)	5.2×10^{11}
120	30	1.2(3)	11(5)	0.9(4)	1.0(2)	0.7(2)	4.5×10^{16}
only DNS							
		Re_ℓ	Re	v_{rms}	$\langle \bar{\epsilon}_{\text{diss}} \rangle$	ℓ	
		1.01	10.5	0.83	0.95	0.61	

summarized in Table IV. A general way to define the local functional $\Delta\mathcal{S}$ is

$$\Delta\mathcal{S}_i = c_i \int dt \int dx g_i(\partial_x v/w_i) \delta(x) \delta(t - t_f), \quad (33)$$

where c_i is a prefactor to characterize the strength of the functional and w_i is an imposed velocity gradient value around which we want our simulation to sample at $(x = 0, t = t_f)$. With the index i we label the different choices of g_i for which we have tested

$$g_1(z) = z, \quad (34a)$$

$$g_2(z) = (z + 1)^2, \quad (34b)$$

$$g_3(z) = (z^2 - 1)^2. \quad (34c)$$

Note that the chosen constraints $\Delta\mathcal{S}_i$ result in singular derivatives in the Hamiltonian (23). In our case the regularization of the δ function happens through the space-time grid and its finite lattice spacing. In the approximation of the integral, which is a finite sum, the δ function's approximation appears as a properly normalized Kronecker δ . Our discretization of the δ function in space, $\delta = \delta(x_i - x_j)$, is $\delta_{i,j}/\Delta x$. Accordingly, in time we discretize $\delta = \delta(t_m - t_n)$ by $\delta_{m,n}/\Delta t$.

The HMC will sample around the region where $e^{-\mathcal{S}'}$ is maximal, i.e., where \mathcal{S}' is minimal, and the constraint functionals $\Delta\mathcal{S}_i$ contribute towards this procedure. In particular, the constraints imposed by $\Delta\mathcal{S}_2$ and $\Delta\mathcal{S}_3$ are of a localization nature in the sense that the generated configurations comply with the constraint by sampling in a narrow region around

the imposed gradient w_i , where $\Delta\mathcal{S}_2$ and $\Delta\mathcal{S}_3$ are minimal. In the same spirit, as $\Delta\mathcal{S}_1$ is a linear function of $\partial_x v$, then for any negative $\partial_x v$ it will have a negative contribution to the action, which will favor the sampling towards this direction. It therefore allows us to sample across a wider range of negative velocity gradients. Nevertheless, we can redefine $\Delta\mathcal{S}_1$, as in this case w_1 can be absorbed by c_1 . Thus, we set $w_1 = 1$ and present only values of c_1 . Accordingly, if we wish to sample positive gradients at $(x = 0, t = t_f)$ in the same manner, it is sufficient to consider negative values of c_1 . That is also possible for $\Delta\mathcal{S}_2$, where according to Eq. (34b), if we consider $w_2 < 0$, then the algorithm will preferably sample positive gradients at $(x = 0, t = t_f)$. In contrast, for $\Delta\mathcal{S}_3$, due to its symmetry around $\partial_x v = 0$, one could expect the sampling of both positive and negative gradients around w_3 . However, the physics of the Burgers equation favors the generation of strong negative gradients instead of their positive counterparts, i.e., notice the asymmetric PDF of velocity gradients [Fig. 1(b)]. Hence, by choosing $\Delta\mathcal{S}_3$, negative gradients will be preferred at $(x = 0, t = t_f)$.

As for the numerical stability, we note that the grid resolution should always be sufficient to “fit” the strong shock. Therefore, we cannot increase c_i and w_i unconditionally for a fixed resolution. In practice, for a particular discretization, there is a threshold beyond which the HMC is not reliable anymore.

To identify the impact of constraining the sampling of the HMC on the generated configurations, we show three independent samples in Fig. 7. A large negative velocity gradient at $(x = 0, t = t_f)$ is achieved in all cases. The general idea here is that we provide the HMC with a certain constraint, local or global, by which the HMC will consider all the possible realizations in the configuration space to fulfill the corresponding condition on the velocity field. In the case of extreme and rare events, for instance, the HMC provides a systematic way to sample the fluctuations around a particular extreme event (e.g., the occurrence of a strong velocity gradient).

Focusing now on the constraint functionals of Eqs. (34), Fig. 8(a) shows the ensemble average of the velocity field for the final time slice $\langle v(x, t = t_f) \rangle'$ at changing c_1 . It further indicates the functionality of c_1 and the effect it has on the sampled configurations, i.e., the larger c_1 is, the more negative the sampled gradient will be. This can also be justified from Fig. 8(b), which, for different c_1 , depicts the PDF of the velocity gradients measured only at the point that we constrain, i.e., at $(x = 0, t = t_f)$. It is defined as

$$P'(w) = \langle \delta(\partial_x v(0, t_f) - w) \rangle', \quad (35)$$

where w is the value of the bin which is incremented according to the value of the velocity gradient $\partial_x v(0, t_f)$ and is generated using the action \mathcal{S}' . In this plot we see that by increasing c_1 , the peak of the histogram moves to the left towards larger negative velocity gradients.

As for the prefactors c_2 and c_3 , they have a slightly different behavior with respect to c_1 . In fact, as we increase c_2 and c_3 , the HMC will sample more systematically around the prescribed velocity gradient w_i . In Fig. 9 we show $P'(w)$ at varying c_i and w_i , with $i = 2, 3$. In Fig. 9(a) the functional $\Delta\mathcal{S}_2$ has been used, and in Fig. 9(b) the functional $\Delta\mathcal{S}_3$.

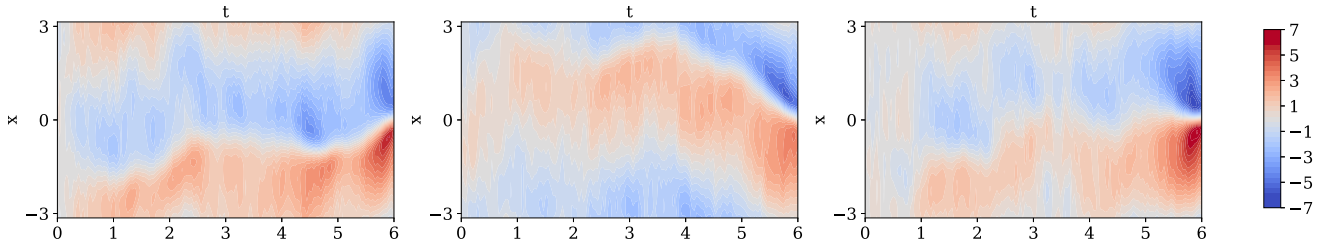


FIG. 7. Sample configurations from the HMC simulation using the constraint ΔS_1 with $c_1 = 1.9$, enforcing a negative velocity gradient maximization at $(x = 0, t = t_f)$. The color spectrum corresponds to the intensity of the velocity field.

For the same parameters, the quartic functional ΔS_3 has a slightly better performance towards sampling the prescribed velocity gradient w than the quadratic functional ΔS_2 . Notice that in Fig. 8(b) and both plots of Fig. 9 we also include the PDF of the velocity gradients of the DNS (black line) to give a qualitative description of how the constrained sampling compares with the original statistics.

To be more specific, we first refer to Fig. 10(a), where we show the nonweighted ensemble-average mean kinetic energy, defined as $\langle \bar{\epsilon}_{\text{kin}}(t) \rangle'$ at changing c_1 , using the functional ΔS_1 , and we compare it with the ensemble-average kinetic energy of the DNS (black line, unconstrained statistics). The larger the value of c_1 is, the more pronounced the kinetic

energy will be closer to the final time $t = t_f$, where the constraint is applied. Figure 10(b) depicts the corresponding reweighted data, i.e., $\langle \bar{\epsilon}_{\text{kin}}(t) \rangle$, by using (32), where both the DNS and the reweighted HMC collapse within error bars.

We remark two points. First, through Fig. 10(a), we can also get an estimate of how important the constraint is as a function of time. For instance, on average, at time $t \approx 3$ the effects of ΔS_1 seem to have decayed. Second, for the particular observable, by increasing here c_1 we get increased error bars after reweighting. For instance, in the case of $\langle \bar{\epsilon}_{\text{kin}}(t) \rangle$ for $c_1 = 1.2$ we notice small error bars and very good agreement with the DNS, while for $c_1 = 1.9$ the $\langle \bar{\epsilon}_{\text{kin}}(t) \rangle$ has much more pronounced error bars. This is related to a previous comment on the applicability of reweighting, for which we stated that

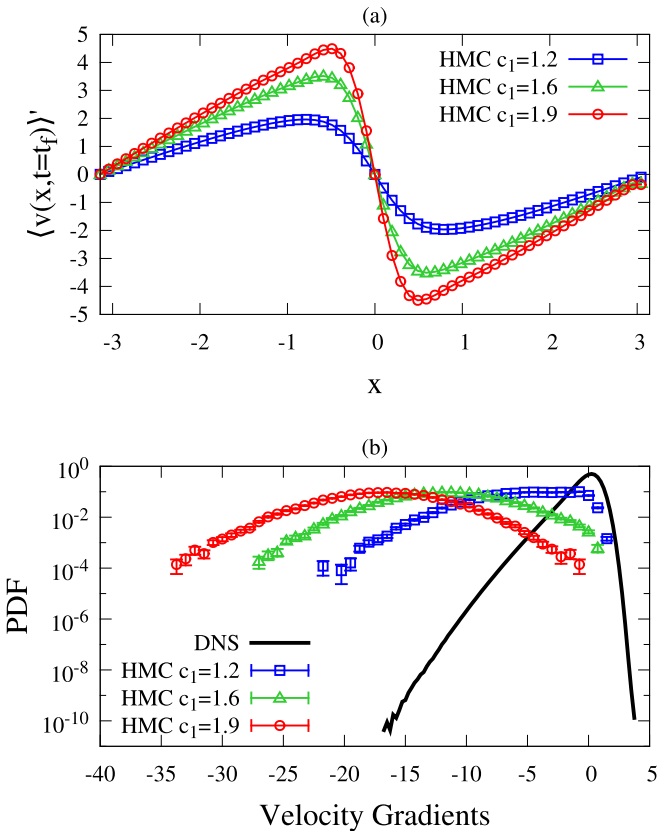


FIG. 8. (a) Ensemble average of the velocity field $v(t = t_f)$ using the HMC with the action $S' = S + \Delta S_1$ for different values of c_1 . (b) PDF of velocity gradients (DNS versus the HMC). For the HMC we measure $P'(w)$ only at the space-time point where we constrain the ensemble, i.e., at $(x = 0, t = t_f)$.

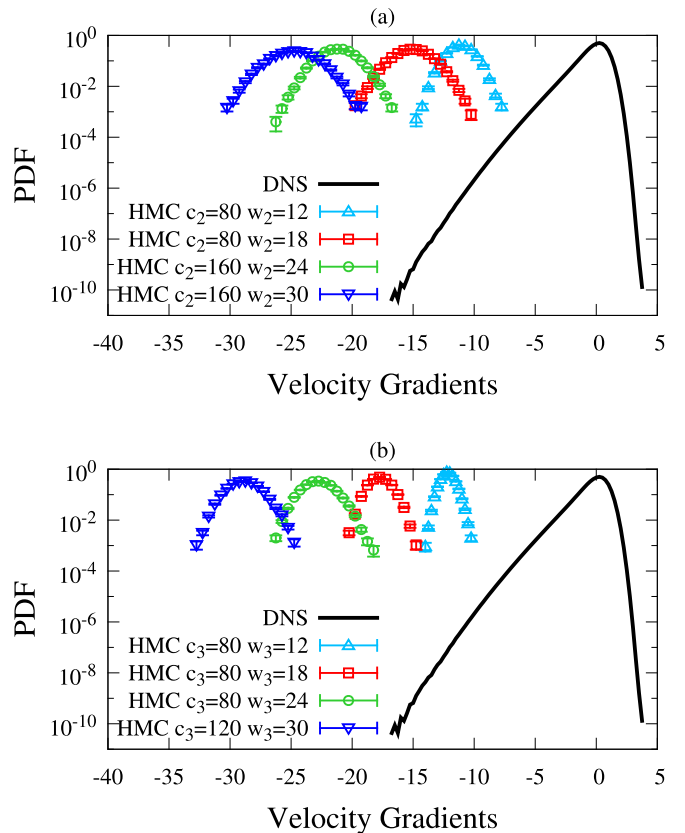


FIG. 9. PDF of velocity gradients (DNS versus the HMC), i.e., $P'(w)$, generated using the action S' and measured only at the point that we constrain, i.e., at $(x = 0, t = t_f)$, for different values of c_i and w_i , $i = 2, 3$, using (a) ΔS_2 and (b) ΔS_3 .

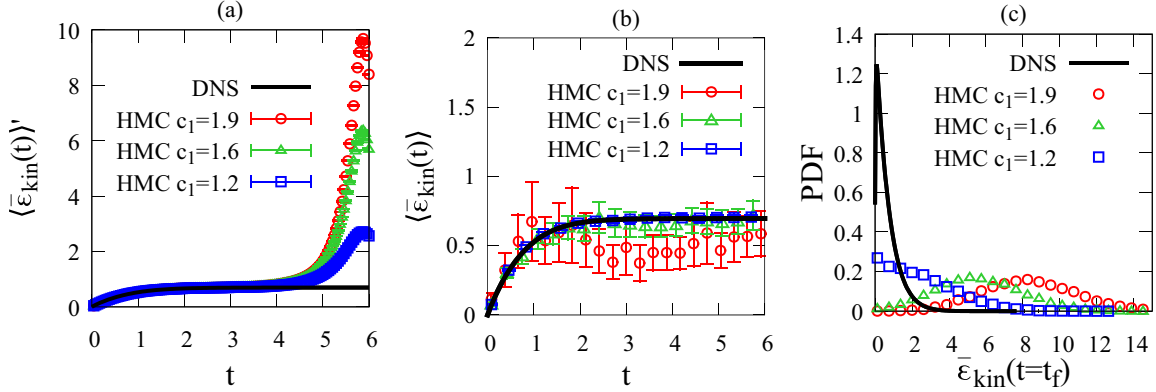


FIG. 10. Ensemble-average mean kinetic energy of the HMC vs DNS using ΔS_1 for different c_1 (a) before reweighting and (b) after reweighting. Notice that the error bars for the $c_1 = 1.9$ case after reweighting are pronounced as for this choice the fluctuations introduced by the reweighting factor become significantly large. (c) PDF of the nonreweighted mean kinetic energy $\bar{\epsilon}_{\text{kin}}(t)$ measured only for the last time slice $t = t_f$, in the case of the HMC, and using ΔS_1 for different c_1 .

the distributions e^{-S} and $e^{-S'}$ should have a sufficient overlap. In this example, for $c_1 = 1.2$, the distribution of $\bar{\epsilon}_{\text{kin}}(t)$, for $t = t_f$, of the constrained ensemble and the distribution of $\bar{\epsilon}_{\text{kin}}(t)$, for $t > t_s$, of the unconstrained system do overlap considerably, as can be seen in Fig. 10(c) (blue and black lines accordingly), which leads to the resulting collapse of the data [same colors in Fig. 10(b)]. The difference with $c_1 = 1.9$ [red line in Fig. 10(a)] is that the corresponding overlap with the DNS is marginal. Also $c_1 = 1.9$ favors more the sampling of extreme velocity gradients $\partial_x v$, which, together with a (finite) characteristic dissipation scale ℓ_d , implies large values of $v_d \sim (\partial_x v)\ell_d$ [see Fig. 8(a)]. The averaged kinetic energy is a global observable, which is mostly related to the bulk of the statistics of v and consequently not sensitive to very strong and rare fluctuations. Therefore, if we want to improve the behavior of $\langle \bar{\epsilon}_{\text{kin}}(t) \rangle$ for $c_1 = 1.9$, we should simply increase the statistics of the particular constrained ensemble to capture, by chance, events with smaller v that are more representative of the unconstrained ensemble. This translates to the fact that for the constrained ensemble, a rare event can be an event which, for the unconstrained ensemble, is a typical one.

To sum up, reweighting of the ensemble-average kinetic energy is a sufficient but not a necessary condition to determine whether the particular constrained ensemble is representative of the original system. In fact, here it was a simple demonstration of the reweighting technique (29) in our application. As we will see in the following, we can achieve a very-well-behaved reweighting for the PDF of the velocity gradients for any c_i and w_i , considering that the latter are appropriately chosen, as stated earlier, so that the HMC is numerically stable.

C. Velocity gradient statistics

To assess the performance of generating extreme and rare events, we compare the HMC, when using sampling constraints, with the DNS by studying the statistics related to the velocity gradients, such as their PDF. We note that, in the following, the observables that we consider are measured only at the single point that we constrain, i.e., at $(x = 0, t = t_f)$. This is related to the introduction of the local constraint

ΔS , which breaks the space-time symmetry of the system. In principle, after applying Eq. (29) we restore the symmetries of the system, in the limit of infinite statistics, but in practice this is not the case. However, for histogram reweighting, by considering only the site on which the local constraint acted, we restore homogeneity and we will show that it is sufficient to obtain a systematic comparison with the unconstrained statistics, regardless of the mutual overlap of the nonreweighted histogram and the unconstrained histogram (e.g., of the DNS). For instance, in Fig. 9(b) the case of the HMC with $c_3 = 120$ and $w_3 = 30$ has no overlap with the DNS (even though there would be if we increased the statistics to infinite), yet in the following we will demonstrate that this particular PDF, together with other similar cases, will be successfully reweighted to the unconstrained statistics. Nevertheless, if we consider other sites, we encounter problems similar to the ones discussed in the preceding section, e.g., for the kinetic energy, where by increasing c_i and w_i we notice increasing error bars.

To reweight the PDF of the velocity gradients $P'(w) = \langle \delta(\partial_x v(0, t_f) - w) \rangle'$, we use Eq. (29) to get

$$P(w) = \frac{\langle \delta(\partial_x v(0, t_f) - w) e^{\Delta S} \rangle'}{\langle e^{\Delta S} \rangle'}, \quad (36)$$

where, in practice, for each measurement i of the ensemble, we increment the bin w by $e^{\Delta S_i}$. In Fig. 11 we apply (36) on the ensemble that generated the nonreweighted PDF $P'(w)$ (open blue squares) to produce the reweighted histogram $P(w)$ (open red circles) and compare it with the corresponding PDF of the DNS (black line). In Fig. 11(a) we identify a slight discrepancy between the $P(w)$ and the DNS (seen more clearly in the inset), while the trend is similar. This is related to the fact that the HMC is constrained to systematically sample large negative velocity gradients (far left tail) and therefore the support on the right tail is limited. As a result, by strictly applying (36), and since it normalizes the area under $P(w)$ to 1, the comparison between the HMC and the DNS is not straightforward, as $P(w)$ is actually an excerpt of the original PDF of velocity gradients, which is assumed to be the curve of the DNS here. For the same reason, $P(w)$ cannot

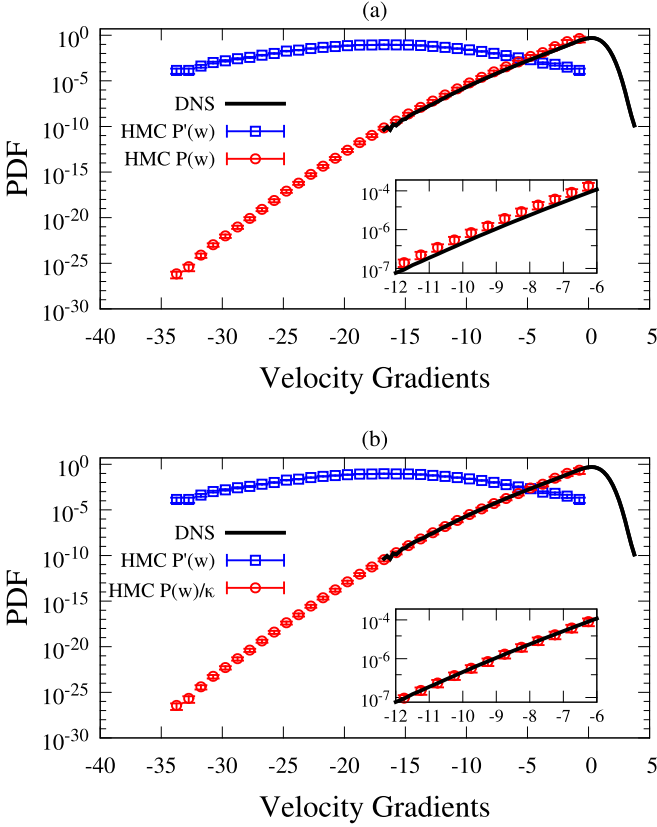


FIG. 11. PDF of the velocity gradients for the HMC (blue and red symbols) versus DNS (black line) (a) without rescaling and (b) rescaling the reweighted HMC histogram by dividing it by $\kappa = 1.93$. For the HMC we use the constraint $\Delta\mathcal{S}_1$ with $c_1 = 1.9$, and $P'(w)$ is measured considering only the site ($x = 0, t = t_f$) on which the constraint is enforced. The inset shows the chosen interval $[-12, -6]$.

be considered as a PDF. What is missing is the rescaling of $P(w)$ with an appropriate factor κ so that both the DNS and the HMC calculate the same probability $p(a, b)$ to sample in a particular interval (a, b) of velocity gradients. By definition, $p(a, b) = \sum_a^b P(w)\delta w$, with δw the bin width, so κ is defined as the ratio of the two probabilities measured by the HMC and the DNS,

$$\kappa = \frac{p(a, b)_{\text{HMC}}}{p(a, b)_{\text{DNS}}}, \quad (37)$$

where we have tested that by increasing the statistics of the HMC, $\kappa \rightarrow 1$. We also assume that the DNS has enough support in both tails to be claimed as a PDF and therefore to be considered as a reliable benchmark for the rescaling of $P(w)$. In Fig. 11(b) we show the rescaled $P(w)/\kappa$, with $\kappa = 1.93$. Also here κ is measured in the interval $[-12, -6]$ for the rescaling. In this way we achieve a collapse of the HMC and the DNS data. What is striking, in this example, is the unique ability of the HMC to systematically sample intense gradients that are up to $\sim 30\sigma$ and more, with $\sigma = 0.99$, far from the mean. For a similar discussion on the resultant statistical efficiency of the chosen constraint functionals we refer the reader to [73,74].

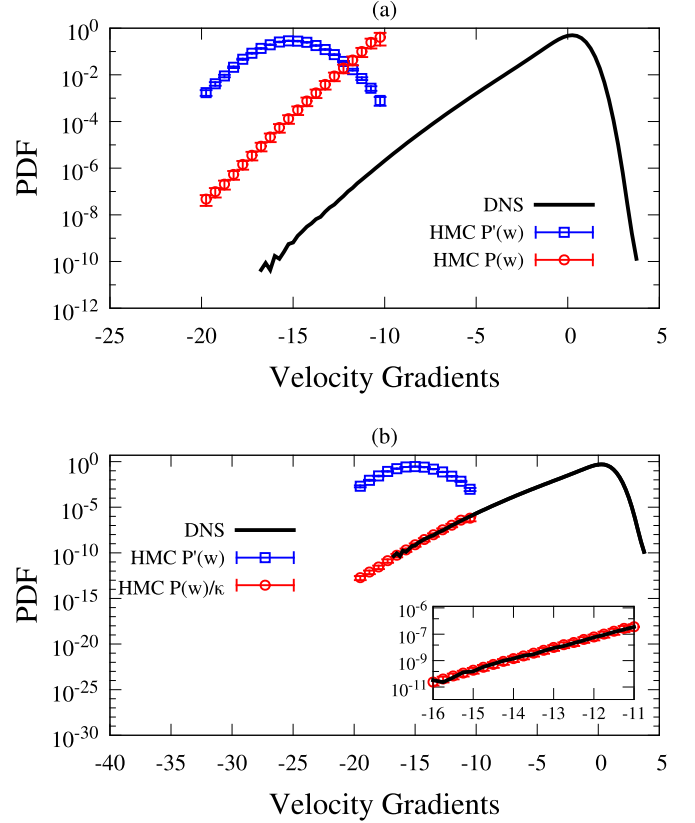


FIG. 12. PDF of velocity gradients for the HMC (blue and red symbols) against DNS (black line) (a) without rescaling and (b) rescaling the reweighted HMC histogram by dividing it by $\kappa = 2.63 \times 10^5$. We consider only the extracted histogram from the lattice point on which the constraint $\Delta\mathcal{S}_2$ acted, i.e., ($x = 0, t = t_f$), in the case of the HMC, with $c_2 = 80$ and $w_2 = 18$. The inset shows the chosen interval $[-16, -11]$ for the rescaling.

Another example where the need to further treat the reweighted velocity gradients histogram $P(w)$, by rescaling it with an appropriate factor κ , becomes more evident is when we consider one of $\Delta\mathcal{S}_2$ or $\Delta\mathcal{S}_3$. In Fig. 12 we show $P'(w)$ (open blue squares) and $P(w)$ (open red circles) using the functional $\Delta\mathcal{S}_2$, with $c_2 = 80$ and $w_2 = 18$ in the case of the HMC, against the DNS (black line). In Fig. 12(a), $P(w)$ (red symbols) is derived by applying (36) to the PDF of the HMC (blue symbols). As before, the area below $P(w)$ is equal to 1. However, by considering $P(w)$, the probability $p(a, b)_{\text{HMC}}$ to sample within an interval (a, b) of velocity gradients does not correspond to the one of the DNS, $p(a, b)_{\text{DNS}}$, so we employ again (37) to get $\kappa = 2.63 \times 10^5$. In Fig. 12(b) we plot $P(w)/\kappa$ (red symbols) instead, in order to achieve the collapse with the PDF of the DNS.

Now that we have clarified how to derive $P(w)$ and explained the need for a further rescaling with a constant, we can do the same procedure for all the different runs using the three different constraint functionals of Eqs. (34).² This is done in Fig. 13, where we compare $P(w)$ for different combinations

²Note that for the rest of article, when referring to $P(w)$, it is implied that $P(w)$ is rescaled with an appropriate κ .

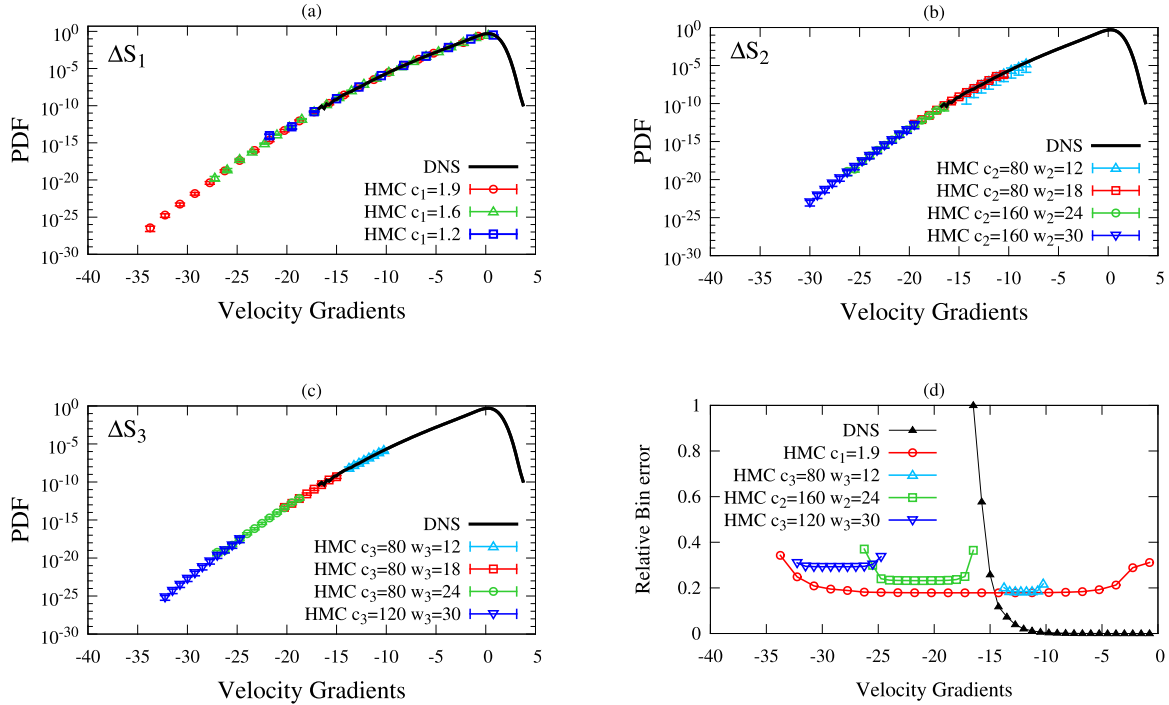


FIG. 13. PDF of velocity gradients $P(w)$ for the HMC and DNS. We consider here only the extracted histogram from the lattice point on which the constraint ΔS acted, i.e., $(x = 0, t = t_f)$, in the case of the HMC. We show the effect of reweighting for different parameters of the constraints: (a) ΔS_1 , (b) ΔS_2 , (c) ΔS_3 , and (d) relative bin error $\delta P(w)/P(w)$, with the error being evaluated using Eq. (30). Regarding rescaling, for those $P(w)$ of which the overlap with the DNS was marginal or nonexistent, the rescaled $P(w)$ for $c_1 = 1.9$ was used.

of ΔS_i , c_i , and w_i with the velocity gradients PDF of the DNS (black line).

In Fig. 13(a) we show $P(w)$ for ΔS_1 and different c_1 , in Fig. 13(b) the results correspond to $P(w)$ for ΔS_2 and different c_2 and w_2 , and Fig. 13(c) depicts the $P(w)$ for ΔS_3 and different c_3 and w_3 . An important remark is that for those cases of the reweighted histogram $P(w)$, where the overlap with the DNS is marginal or absent, we used the rescaled $P(w)$ for ΔS_1 and $c_1 = 1.9$ as a guide to rescale them. For instance, this was necessary for $w_i = 24, 30$. Furthermore, the different κ that were used for each case are shown in Table IV. Finally, Fig. 13(d) shows the relative bin error $\delta P(w)/P(w)$ as a measure of the statistical efficiency of each different constraint ΔS_i . For the HMC, we used Eq. (30) to measure $\delta P(w)$, while for the DNS it is simply equal to $\delta P(w) = 1/\sqrt{\text{counts}}$. Interestingly, the HMC has a constant ratio for extreme values of the velocity gradients, while the DNS quickly diverges as soon as the statistics are limited. Note that a typical ensemble size of the HMC runs is of the order of 10^4 , while that of the DNS is of the order of 10^9 .

To further quantify the performance of the HMC for the purpose of systematically sampling very intense velocity gradients we provide Fig. 14. Figures 14(a)–14(c) shows $P(w)w^q$, i.e., the reweighted and rescaled histogram of the velocity gradients multiplied by a moment w^q . The idea is that the higher the power q is, the more we focus towards larger negative gradients. If the statistics of $P(w)$ are sufficient in the corresponding “focused” region, then $P(w)w^q$ has a clear peak and shape. Figures 14(d)–14(f) depicts the computational cost that the ensemble running average of a moment of a velocity gradient $\langle (\partial_x v)^q \rangle$ requires in order to stabilize

at a certain value and stop fluctuating. Here, for the HMC, we used the functional ΔS_1 , for $c_1 = 1.9$, and we consider only the velocity gradient at the point $(x = 0, t = t_f)$. The data here are the same as the red and black data sets of Fig. 11 for the HMC and the DNS, respectively. Also, the observable is reweighted according to Eq. (29) so that the comparison is equivalent. Finally, for visualization purposes, we normalize to one the observables by dividing them by the final value of the stabilized line (depending on q , this might be the line of either the HMC or DNS).

The plots in Fig. 14 are complementary, as a specific power q is chosen for each column. The plots in Figs. 14(a) and 14(d) are for a small $q = 6$. In this region the DNS performs better as here the data of the HMC are only measured on the site on which the constraint acts and therefore the appearance of relatively small negative gradients is suppressed. For $q = 16$ [Figs. 14(b) and 14(e)] we see that both the HMC and the DNS are equivalent in terms of the computational cost and quality of the statistics. Finally, for $q = 30$ [Figs. 14(c) and 14(f)] the HMC significantly outperforms the DNS, as it immediately converges to the expectation value, while for the DNS we would have to remarkably increase the computational cost to achieve comparable statistics. Note that the data for both the HMC and the DNS in Figs. 13 and 14 required the same computational cost to be produced, using the same processors. Overall, Fig. 14 summarizes the ability of the HMC to consistently sample intense negative gradients that belong in the large-deviation regime and furthermore gives a qualitative measure of the computational performance gained over a standard DNS method.

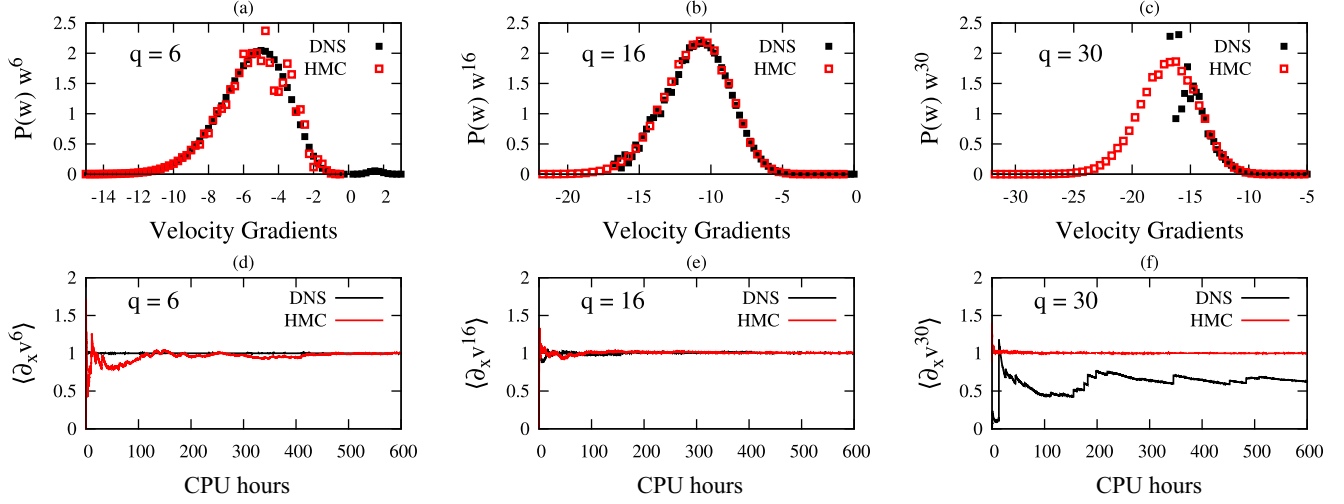


FIG. 14. (a)–(c) Velocity gradient PDF multiplied by a moment w^q . The $P(w)$ of the HMC is reweighted, rescaled with κ , and we consider only the lattice point at which the constraint acts. Here we used $\Delta\mathcal{S}_1$ for $c_1 = 1.9$. (d)–(f) Computational time to the stabilized running average of the velocity gradient moment $\langle (\partial_x v)^q \rangle$, divided with respect to the final stabilized value. Regarding DNS, any site belonging to the stationary regime is considered.

VII. RELEVANCE OF INSTANTONS IN EXTREME EVENTS

The application of instantons in turbulent flows was first proposed in [18], where the instanton contribution to the right tail of the velocity increment PDF was calculated for Burgers turbulence, while in a succeeding work [19], the left tail of the increment PDF was studied using the instanton approach. These works paved the way to other hydrodynamical models, such as the advection of a passive scalar by a turbulent velocity field [75,76], shell models [77,78], geophysical flows [22,23,79], and atmospheric and oceanic flows [80,81] (see also [82] and references therein).

A. Derivation of the instanton configuration

In order to calculate ensemble averages of observables $\langle \mathcal{O}_v \rangle$ as, e.g., the probability distribution of the gradient $P(\partial_x v = w) = \langle \delta(\partial_x v(x=0, t=t_f) - w) \rangle$, we utilize the path-integral formulation introduced in Sec. III:

$$\begin{aligned} P(w) &\propto \int \mathcal{D}v \delta(\partial_x v(x=0, t=t_f) - w) e^{-\mathcal{S}} \\ &= \int \mathcal{D}v \int_{-i\infty}^{i\infty} d\lambda e^{-\mathcal{S}'(\lambda)}. \end{aligned} \quad (38)$$

Here $\mathcal{S}' = \mathcal{S}'(\lambda)$ contains both the Onsager-Machlup action \mathcal{S} [cf. Eq. (15)] and the contribution of the observable $\delta(\partial_x v(x=0, t=t_f) - w)$:

$$\begin{aligned} \mathcal{S}' &= \mathcal{S} + \lambda(\partial_x v(0, t_f) - w) \\ &= \int_{t_0}^{t_f} dt \left\{ \frac{1}{2}(F, \Gamma^{-1} * F) \right. \\ &\quad \left. + \lambda(\partial_x v(x, t) - w, \delta(x))\delta(t - t_f) \right\} - \ln \mathcal{J}. \end{aligned} \quad (39)$$

Instanton configurations are “classical” solutions that extremize the action and therefore dominate the path integral of the stochastic Burgers equation (18). They can be computed by Laplace’s method or alternatively, as in many

applications, instantons are found by numerically minimizing the action directly (see, e.g., [23]). Here, where the observable is evaluated only at the final time $t = t_f$, it is advantageous to switch to another equivalent formulation by applying a Hubbard-Stratonovich transformation [83,84], which leads to the alternative representation of the partition sum

$$\mathcal{Z} \propto \int \mathcal{D}v \mathcal{D}\mu \exp \left(\int dt [i(\mu, F) - \frac{1}{2}(\mu, \Gamma * \mu)] + \ln \mathcal{J} \right), \quad (40)$$

which prompts us to define

$$\mathcal{S}_{\text{MSRJD}} = - \int dt \{ i(\mu, F) - \frac{1}{2}(\mu, \Gamma * \mu) \} - \ln \mathcal{J}, \quad (41)$$

also known as Martin-Siggia-Rose–Janssen–de Dominicis action (MSRJD) [11,12]. At the expense of an additional auxiliary field μ , we have “linearized” the action with respect to the noise $\eta (=F)$. Furthermore, the force correlator Γ now appears directly and not through its inverse Γ^{-1} . This allows for the implementation of more general types of forcing as the power-law forcing considered in this paper. Now the corresponding expression for the PDF of velocity gradients reads

$$P(w) \propto \int \mathcal{D}v \mathcal{D}\mu \int_{-i\infty}^{i\infty} d\lambda e^{-\mathcal{S}'_{\text{MSRJD}}}, \quad (42)$$

with

$$\mathcal{S}'_{\text{MSRJD}} = \mathcal{S}_{\text{MSRJD}} + \lambda(\partial_x v(0, t_f) - w). \quad (43)$$

Before we proceed, we note that attempting to compute path integrals of the form of Eq. (40) is not straightforward and might be impossible for most cases. For instance, perturbative approaches might be helpful, depending on the problem. In the context of fluid dynamics, a diagrammatic approach (influenced by quantum field theory) was proposed by Wyld [6]. Using perturbation theory to expand the exponential in

Eq. (40) in powers of the nonlinear term [see also Eq. (1)] proves insufficient in the turbulent limit $\nu \rightarrow 0$, since the path integral is dominated by the nonlinear term forming strong shocks. Therefore, perturbative approaches must be abandoned, as a large parameter is required [18].

Nevertheless, the introduced Lagrange multiplier λ in Eq. (39) can be used as a large parameter. This allows the use of the saddle-point approximation, by which the variation of the integrand in Eq. (42) is equal to zero. In the case of Burgers turbulence, we obtain the instanton equations (minimizer of the action $\mathcal{S}'_{\text{MSRJD}}$)

$$\partial_t v + v \partial_x v - \nu \partial_x^2 v = -i \Gamma * \mu, \quad (44a)$$

$$\partial_t \mu + v \partial_x \mu + \nu \partial_x^2 \mu = i \lambda \delta'(x) \delta(t - t_f), \quad (44b)$$

where the term on the right-hand side in Eq. (44b) implements the boundary condition for μ at t_f according to which $\mu(x, t_f) = i \lambda \delta'(x)$. Recall that in the case of the Burgers equation $\mathcal{J} = \text{const}$ and therefore the Jacobian does not contribute to the saddle-point equations. In [85] an algorithm was proposed to numerically solve the above equations. In short, the sign in front of the viscous terms defines the temporal direction of the numerical integration, with v being integrated forward in time and μ backward. Using $\mu(x, t_f) = i \lambda \delta'(x)$ as an initial condition for some large value of t_f and starting by setting $v(x, t) = 0$, Eq. (44b) is first integrated backward until t_0 . Then the obtained $\mu(x, t)$ is used to integrate Eq. (44a) forward in time, with the whole procedure being iterated until convergence to the prescribed constraint $\partial_x v(0, t_f) = w$ is achieved. For more details see also [82,86,87], where the aforementioned methodology is revisited.

B. Numerical results

Instantons, strong field-force fluctuations and extremal points of the action $\mathcal{S}'_{\text{MSRJD}}$, may be considered as particular examples of extreme and rare events. Constraining the HMC to sample at large negative gradients, we observe that the generated configurations clearly resemble the classical instanton configurations determined via the saddle-point approximation. This will be checked directly via the averaged velocity-field profile and through the probability distribution function of velocity gradients.

Figure 15(a) compares the HMC ensemble average of the velocity field at the last time slice ($t = t_f$) with the velocity field obtained by performing the numerical integration of the instanton equations (44). The profile of the classical instanton at time $t = t_f$ is reproduced to a remarkable degree, implying that the ensemble average is equivalent to removing the fluctuations around the instanton. This confirms that instantons can be found in Burgers turbulence, as already shown in [86] using a postproduction filtering protocol to consider only events with strong gradients generated using DNS. Furthermore, the inset depicts the difference of the two velocity fields, which are on the order of statistical error. Similarly, in Fig. 16 we compare the whole averaged spatiotemporal domain of the HMC with the instanton velocity field in space and time. For Figs. 15 and 16 a resolution of $N_t = 576$ points in time was used for the HMC, the DNS, and the instanton, while the rest of the parameters are the same as in Table IV.

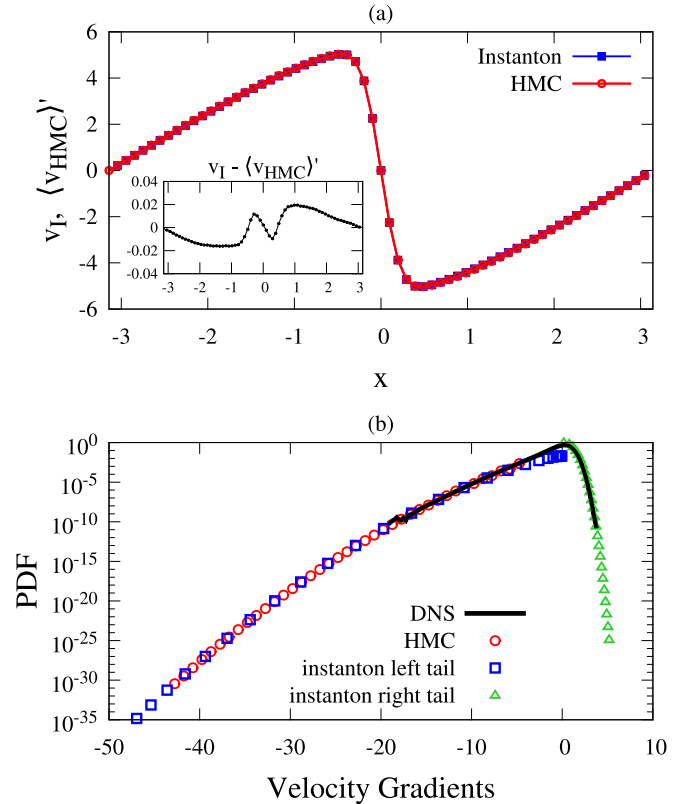


FIG. 15. (a) Ensemble average of velocity configurations generated by the HMC using ΔS_1 with $c_1 = 1.9$ compared to the classical instanton velocity-field profile generated for $\lambda = -1.148$ and $w = -24.23$. (b) PDF of the velocity gradients for the classical instanton (for a range of values of λ and w), HMC simulation, and DNS.

Figure 15(b) compares the PDF of the velocity gradients of the DNS, the HMC with constrained sampling, and the instanton. In the case of the instanton we plot $e^{-S_{\text{inst}}}$, with $S_{\text{inst}} = -\frac{1}{2} \int_{t_0}^{t_f} dt (\mu, \Gamma * \mu)$. We notice that the PDF predicted by the instanton follows the same trend as the HMC and the DNS and the agreement is extraordinary. However, in order to correctly interpret this result one should note the following. On the one hand, the PDF prediction for positive gradients is valid independently of the Reynolds number Re and is actually valid for all positive values besides small corrections near $\partial_x v = 0$. This is a result already obtained by Feigel'man [88] in the context of charge density waves and also confirmed by the instanton formalism [18]. On the other hand, the PDF of negative gradients depends on the Reynolds number and for a given Reynolds number Re_ℓ the instanton prediction is only valid for $|\partial_x v| > |\partial_x v^*(\text{Re}_\ell)|$. A precise estimate for $\partial_x v^*$ is given in [87] [see Eq. (17) in the same reference]. For the Reynolds number $\text{Re}_\ell = 1$ used in our simulation and depicted in Fig. 15, this means that the instanton prediction is valid only for $\partial_x v < -10$.

VIII. CONCLUSION

In this work we established how to apply Monte Carlo importance sampling for stochastic dynamics based on the Janssen-de Dominicis path integral, in order to address the

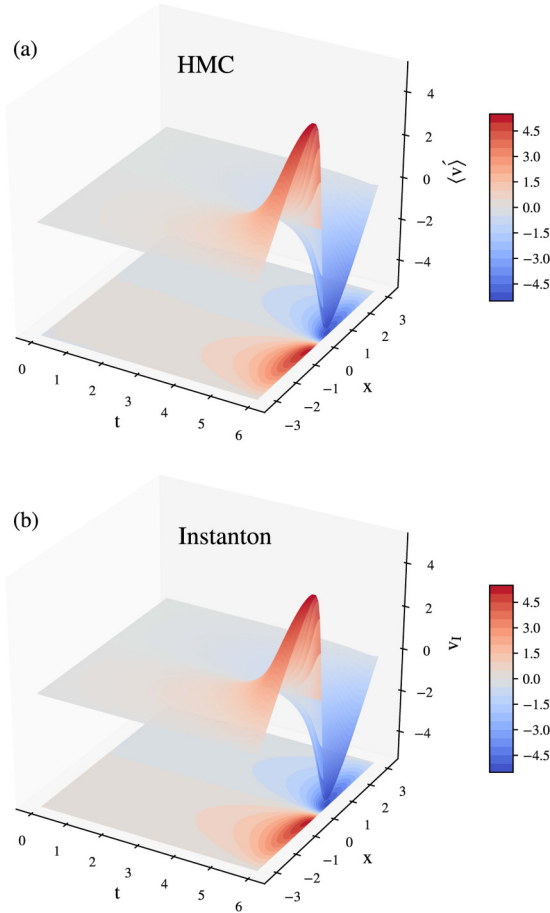


FIG. 16. (a) Ensemble average of velocity configurations generated by the HMC using $\Delta\mathcal{S}_1$ with $c_1 = 1.9$. (b) Instanton velocity field for $\lambda = -1.148$ and $w = -24.23$. By averaging over the HMC velocity-field ensemble the spatiotemporal shape of the classical instanton is restored.

statistics of large fluctuations in driven nonequilibrium systems. This approach allowed us to access the phase space of all possible field realizations of a stochastic system. Using reweighting techniques, we were able to systematically enhance the occurrence of extreme and rare events, by sampling in specific phase-space regions related to such events.

We have chosen to illustrate the HMC algorithm as an example of the random-noise-driven one-dimensional Burgers equation, which often is used as a model for benchmarking numerical methods in computational fluid dynamics. However, the HMC approach is generally applicable to any stochastic PDE and generally free from any modeling assumptions. Also, while the random forcing was chosen to be Gaussian, self-similar, and white in time, this is by no means necessary and other types of noises can be addressed within this approach. We thoroughly benchmarked our HMC implementation with a standard forward-time-integration pseudospectral method (see Figs. 4 and 5). By constraining the sampling of the HMC to generate a strong negative velocity gradient at a specific site we increased the statistics of the left tail of the PDF of velocity gradients significantly, producing gradients as intense as 30 (and more) times the rms value (see Fig. 13).

Although we restricted ourselves to the case of localized (in space and time) constraints, the technique can be easily extended to more general cases. Also, our constrained HMC sampling allowed us to decrease by order of magnitudes the time to the solution needed to collect sufficient statistics for high-order moments (up to order 30) if compared with DNS (see Fig. 14). We expect that the types of local constraints considered in this work might have an impact on similar studies in lattice gauge theories, where they may lead to new observables.

We demonstrated that instanton configurations can be found in Burgers turbulence. We have recovered the full shape of the classical instanton by averaging the generated ensemble of the constrained configurations, with the agreement of the HMC and the instanton being remarkable [see Figs. 15(a) and 16]. We further compared the PDF of the velocity gradients for a very large range of strong negative gradients and showed that, beyond a specific Reynolds-number-dependent threshold of applicability of the instanton method, both the HMC and the instanton produce the same left tail, which further ensures the relevance of instantons in Burgers turbulence [see Fig. 15(b)]. Thus, we established a one-to-one correspondence between the biased realizations of the HMC and the fluctuations around instantons. The present study focused on low-Reynolds-number turbulence. However, the present method is not restricted to this case and actually opens the possibility to explore the role of fluctuations around instantons with unmatched precision. We are confident that the suggested approach can find suitable applications in the diverse field of stochastic PDEs and related studies on extreme and rare events.

ACKNOWLEDGMENTS

This research has benefited from high-performance computing resources provided by the Jülich Supercomputing Center (Germany) and the CINECA Supercomputing Center (Italy), under the 13th PRACE call. D.M. acknowledges support from the Swiss National Science Foundation, the HPC-Europa2 Transnational Access Program, which funded research visits at the University of Rome Tor Vergata and the University of Ferrara, as well as the German Academic Exchange Service (DAAD), which provided financial resources to present part of this work at the Institute of Pure and Applied Mathematics, Rio de Janeiro (Brazil). G.M. acknowledges funding from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Grant Agreement No. 642069 (European Joint Doctorate Programme “HPC-LEAP”). L.B. acknowledges funding from the European Research Council under the European Union's Seventh Framework Programme, ERC Grant Agreement No. 339032.

APPENDIX A: IMPLEMENTATION OF THE LEAPFROG INTEGRATOR AND FOURIER ACCELERATION

The leapfrog integrator is a symplectic integrator that numerically integrates the Hamiltonian system of equations (19). It advances the fields $(v_0, \pi_0) \mapsto (v_\tau, \pi_\tau)$, up to corrections $O(\Delta\tau^2)$, along a trajectory on a hypersurface $\mathcal{H}^{\text{eff}}(v_s, \pi_s) =$

const, where the effective Hamiltonian is defined by $\mathcal{H}^{\text{eff}} = \frac{1}{2} \int dt (\pi_s, \Omega * \pi_s) + \mathcal{S}[v_s]$. We also note that in our application we treat the fields in Fourier space, i.e., $\mathcal{F}(v_s(x, t)) \mapsto v_s(k, t)$, where the forward Fourier transform is denoted by \mathcal{F} , as in Fourier space many calculations of our interest simplify. For instance, the Fourier transform of the convolution between two fields f and g becomes a simple multiplication among the corresponding individually Fourier transformed fields, i.e., $\mathcal{F}(f * g) \sim \mathcal{F}(f)\mathcal{F}(g)$. Therefore, the effective Hamiltonian can be simplified to

$$\mathcal{H}^{\text{eff}} = \frac{1}{2} \int dt \int dk \Omega(k, t) \pi_s^2(k, t) + \mathcal{S}[v_s]. \quad (\text{A1})$$

The numerical scheme of the leapfrog integrator starts with a half step $\Delta\tau/2$ evolution of the velocity, followed by a full step $\Delta\tau$ of the momenta and finally another half step $\Delta\tau/2$ of the velocity

$$v_{\Delta\tau/2}(k, t) = v_0(k, t) + \Omega(k, t) \pi_0(k, t) \frac{\Delta\tau}{2}, \quad (\text{A2})$$

$$\pi_{\Delta\tau}(k, t) = \pi_0(k, t) - \partial\mathcal{S}/\partial v(k, t)|_{\Delta\tau/2} \Delta\tau, \quad (\text{A3})$$

$$v_{\Delta\tau}(k, t) = v_{\Delta\tau/2}(k, t) + \Omega(k, t) \pi_{\Delta\tau}(k, t) \frac{\Delta\tau}{2}, \quad (\text{A4})$$

where the derivative $\frac{\partial\mathcal{S}}{\partial v}$ is evaluated at intermediate step $\Delta\tau/2$. The steps (A2)–(A4) are then repeated $N_{\Delta\tau}$ times until τ is reached.

The Fourier acceleration effectively prescribes different trajectory lengths to different Fourier modes of the velocity in an effort to balance the scaling of the forces. This is implemented with an appropriate choice of $\Omega(k, t)$, where k labels a particular wave number. Upon rescaling

$$\pi_s(k, t) \rightarrow \xi_s(k, t) = [\Delta t \Omega(k, t)]^{1/2} \pi_s(k, t), \quad (\text{A5})$$

$$\Delta\tau \rightarrow \Delta\tilde{\tau}(k, t) = \left(\frac{\Omega(k, t)}{\Delta t} \right)^{1/2} \Delta\tau, \quad (\text{A6})$$

we note that the new step size $\Delta\tilde{\tau}$ carries a k dependence. Then the equations of motion (A2)–(A4) transform as

$$v_{\Delta\tilde{\tau}/2}(k, t) = v_0(k, t) + \xi_0(k, t) \Delta\tilde{\tau}/2, \quad (\text{A7})$$

$$\xi_{\Delta\tilde{\tau}/2}(k, t) = \xi_0(k, t) - \Delta t \partial\mathcal{S}/\partial v(k, t)|_{\Delta\tilde{\tau}} \Delta\tilde{\tau}, \quad (\text{A8})$$

$$v_{\Delta\tilde{\tau}}(k, t) = v_{\Delta\tilde{\tau}/2}(k, t) + \xi_{\Delta\tilde{\tau}}(k, t) \Delta\tilde{\tau}/2. \quad (\text{A9})$$

We require that the fields should satisfy $\xi \sim O(1)$, with $\Delta\xi \sim O(\Delta\tau)$, i.e.,

$$\Delta t \frac{\partial\mathcal{S}}{\partial v}(k, t) \Delta\tilde{\tau}(k, t) = [\Delta t \Omega(k, t)]^{1/2} \frac{\partial\mathcal{S}}{\partial v}(k, t) \Delta\tau \sim O(\Delta\tau). \quad (\text{A10})$$

Then Eq. (A10) gives a relation for the kernel $\Omega(k, t)$,

$$\Omega^{-1}(k, t) = \Delta t \left\langle \left| \frac{\partial\mathcal{S}}{\partial v}(k, t) \right| \right\rangle^2. \quad (\text{A11})$$

APPENDIX B: CLASSICAL ACTION FOR FINITE APPROXIMATIONS OF THE BURGERS EQUATION

A numerical treatment of the path integral (14) relies on a proper regularization of the functional integration measure and weight. For this purpose we employ finite approximations of the stochastic dynamics of Eq. (1) using a uniform grid in space and time. To make contact with standard approaches employed in the explicit-time integration of the Burgers equation, we adopt a discretization in Fourier space, where the velocity field $v(k, t)$ is defined on a finite set of wave numbers $k = -N_x/2, -N_x/2 + 1, \dots, N_x/2 - 1$ and a discrete set of points in time $t = t_0 + n\Delta t$, with $n = 0, 1, \dots, N_t$, with $N_t \equiv M(t_f - t_0) = MT \in \mathbb{N}$ and $\Delta t \equiv T/N_t = 1/M$. Thus, we measure length in units of $L/2\pi$, time in units of $M\Delta t$, and velocity in units of $L/2\pi M\Delta t$. In the following we will simply set $L = 2\pi$ and $M\Delta t = 1$.

We also employ the initial condition $v(k, t_0) = 0$ for all wave numbers k and restrict the time evolution to a finite time interval $t_0 < t \leq t_f$ of length $T = t_f - t_0$. Correspondingly, the functional measure is given by

$$\int \mathcal{D}v \equiv \prod_{k=-N_x/2}^{N_x/2-1} \prod_{n=1}^{N_t} \int dv(k, t_0 + n\Delta t). \quad (\text{B1})$$

1. Finite approximation of the equation of motion

Passing from continuous space to a discrete finite number of Fourier modes, the original stochastic partial differential equation (1) becomes a high-dimensional set of coupled ordinary stochastic differential equations

$$\frac{d}{dt} v(k, t) = f^{(v)}(k, t) + \eta(k, t), \quad (\text{B2a})$$

$$f^{(v)}(k, t) \equiv \frac{-ik}{2(2\pi)} \sum_{l, m=-N_x/2}^{N_x/2-1} \{v(l, t)v(m, t)\delta_{k, l+m} - vk^2 v(k, t)\}. \quad (\text{B2b})$$

For later convenience, we have separated the equation of motion for each wave number into two parts that describe the deterministic and stochastic components of dv/dt , respectively. Note that $v(-k, t) \equiv v^*(k, t)$ and $\eta(-k, t) \equiv \eta^*(k, t)$, as well as $f^{(v)}(-k, t) \equiv [f^{(v)}(k, t)]^*$ for $k = 1, 2, \dots, N_x/2 - 1$, while $v(k = 0, t)$ and $v(k = N_x/2 - 1, t)$ are both real valued [and similarly $\eta(k = 0, t)$ and $\eta(k = N_x/2 - 1, t)$, etc].

2. Dealiasing

The nonlinear term leads to aliasing errors in the numerical integration of Eq. (B2a), which can be avoided by applying the 2/3 rule [89]. Thus, to correct for these artifacts we introduce the projection operator \mathbf{P} whose action is most conveniently defined in Fourier space

$$\mathbf{P}(f(k)) = \begin{cases} f(k), & |k| \leq N_x/3 \\ 0, & |k| > N_x/3 \end{cases} \quad (\text{B3})$$

for any function $f(k)$. The dealiased interaction term is then defined as

$$\hat{f}^{(v)}(k, t) \equiv \mathbf{P}(f^{(v=0)}(k, t)) - vk^2v(k, t) \quad (\text{B4})$$

and the corresponding equation of motion reads

$$\frac{d}{dt}v(k, t) = \hat{f}^{(v)}(k, t) + \eta(k, t). \quad (\text{B5})$$

The representation (B5) of the dynamics relies on standard approaches in the numerical treatment of partial differential equations, i.e., via spectral Galerkin or pseudospectral methods. We do not intend to advocate that these are in any way optimal in terms of their convergence properties (the interested reader is referred to [90]). We simply choose Eq. (B5) as our starting point to benchmark the performance of the HMC algorithm.

3. Discrete stochastic dynamics: Euler-Maruyama method

To arrive at a discrete-time representation of the dynamics (B2a) we employ the stochastic Taylor expansion in time for $v(k)$,

$$v(k, t + \Delta t) = v(k, t) + \hat{f}^{(v)}(k, t)\Delta t + \bar{\eta}(k)\sqrt{\Delta t} + O(\Delta t), \quad (\text{B6})$$

where we have used $\int dt \eta(k, t) = \bar{\eta}(k)\sqrt{\Delta t} + O(\Delta t)$. This corresponds to the weak first-order Euler scheme [62]. However, we may improve on the rate of convergence of the deterministic part $\hat{f}^{(v)}(k, t)\Delta t$ by considering the variable transformation

$$v'(k, t) = \mathcal{G}^{(v)}(k, t_0 - t)v(k, t), \quad (\text{B7})$$

where $\mathcal{G}^{(v)}(k, t) = \exp(-vk^2t)$. Taking the time derivative on both sides of Eq. (B7), we obtain

$$\frac{d}{dt}v'(k, t) = \mathcal{G}^{(v)}(k, t_0 - t)[\hat{f}^{(v=0)}(k, t) + \bar{\eta}(k)], \quad (\text{B8})$$

to which we may apply the stochastic Taylor expansion. Doing so we arrive at the result

$$v(k, t + \Delta t) = \mathcal{G}^{(v)}(k, \Delta t)[v(k, t) + \hat{f}^{(v=0)}(k, t)\Delta t + \bar{\eta}(k)\sqrt{\Delta t}] + O(\Delta t). \quad (\text{B9})$$

The exact integration of the viscous term [cf. Eq. (B7)] significantly improves the convergence for large wave numbers, provided the step size Δt is sufficiently small.

4. Finite approximation of stochastic noise

Here we consider finite approximations of the stochastic noise, which is assumed to be centered, Gaussian, and white in time. On a finite set of Fourier modes $k = -N_x/2, -N_x/2 + 1, \dots, N_x/2 - 1$, the second moment takes the form

$$\int \mathcal{D}\eta \mathcal{P}_\eta \eta(k, t)\eta(k', t') = \Gamma(k)\delta(k + k')\delta(t - t') \quad (\text{B10})$$

for $t, t' > t_0$. In Eq. (11) we showed that $\mathcal{P}_\eta = \exp[-\frac{1}{2} \int dt (\eta, \Gamma^{-1} * \eta)]$, by assuming that $\Gamma(k) \neq 0$.

The finite approximation of this expression is given by

$$-\ln \mathcal{P}_\eta = \sum_{n=1}^{N_t} \Delta t \left(\Gamma^{-1}(0)[\eta(0, t_n)]^2 + \Gamma^{-1}(-N_x/2)[\eta(-N_x/2, t_n)]^2 + \sum_{k=1}^{N_x/2-1} \eta(k, t_n)\Gamma^{-1}(k)\eta(-k, t_n) \right), \quad (\text{B11})$$

where $t_n = t_0 + n\Delta t$, and for symmetry reasons we consider half of the modes, i.e., $\sum_{k=-N_x/2}^{N_x/2-1} = 2 \sum_{k=0}^{N_x/2-1}$. Note that $\eta(0, t)$ and $\eta(-N_x/2, t)$ are both real valued, while $\eta(k, t)$, $|k| = 1, 2, \dots, N_x/2 - 1$, are generally complex. The expression in Eq. (B11) can be further simplified through applying a UV cutoff by considering the 2/3 dealiasing rule of Eq. (B3). Moreover, we consider $\eta(0, t) = 0$, i.e., the zero mode $k = 0$ is not forced. Finally, we note that we use a large-scale power-law forcing, i.e., $\Gamma(k) = \Gamma_0|k|^\beta$, with $\beta = -3$ throughout all the simulations in this article. Altogether we have

$$-\ln \mathcal{P}_\eta = \frac{1}{\Gamma_0} \sum_{n=1}^{N_t} \Delta t \sum_{k=1}^{N_x/3} \eta(k, t_n) k^{-\beta} \eta(-k, t_n). \quad (\text{B12})$$

This expression will be useful below, when we construct the finite-time-discretized approximations of the classical action \mathcal{S} , based on the regularized continuous-time stochastic dynamics.

5. Classical action

From Eq. (B9) we extract the deterministic part of the time-discrete representation of the equation of motion, i.e.,

$$F(k, t) = \frac{v(k, t)}{\Delta t} - \frac{\mathcal{G}^{(v)}(k, \Delta t)}{\Delta t} [v(k, t - \Delta t) + \Delta t \hat{f}^{(v=0)}(k, t - \Delta t)], \quad (\text{B13})$$

which enters the (reparametrized) classical action following Eqs. (B12) and (16),

$$\mathcal{S} = -\ln \mathcal{P}_v + \text{const} \quad (\text{B14a})$$

$$= \frac{1}{\Gamma_0} \sum_{n=1}^{N_t} \Delta t \sum_{k=1}^{N_x/3} F(k, t_n) \mathcal{G}^{(v)}(k, \Delta t)^{-2} k^{-\beta} F(-k, t_n). \quad (\text{B14b})$$

Notice that Eq. (B14b) is the discretized version of Eq. (17). The const in Eq. (B14a) is related to the contribution of $\ln J$ to \mathcal{S} , which is constant in the case of the explicit-time schemes as introduced in Sec. VIII with fixed initial boundary conditions, e.g., here $v(x, t_0) = 0$, and open final conditions [52].

In general, any constant contribution to the action can be removed, as the sampling is left unaffected, since the HMC considers the differences of the Hamiltonian. On the other hand, as stated in Sec. VI A in the case of periodic boundary conditions in time $v(x, t + T) = v(x, t)$, the contribution $\ln J$ to \mathcal{S} is field dependent [52] and therefore, in principle, it has to be evaluated during the course of the simulation.

- [1] K. R. Sreenivasan and R. A. Antonia, *Annu. Rev. Fluid Mech.* **29**, 435 (1997).
- [2] U. Frisch, *Turbulence: The Legacy of A. N. Kolmogorov* (Cambridge University Press, Cambridge, 1995).
- [3] R. H. Kraichnan, *Phys. Fluids* **1**, 358 (1958).
- [4] G. Rosen, *Phys. Fluids* **3**, 519 (1960).
- [5] G. Rosen, *Phys. Fluids* **3**, 525 (1960).
- [6] H. W. Wyld, Jr., *Ann. Phys. (NY)* **14**, 143 (1961).
- [7] D. Forster, D. R. Nelson, and M. J. Stephen, *Phys. Rev. Lett.* **36**, 867 (1976).
- [8] V. Yakhot and S. A. Orszag, *J. Sci. Comput.* **1**, 3 (1986).
- [9] A. M. Polyakov, *Phys. Rev. E* **52**, 6183 (1995).
- [10] L. Adzhemyan, N. Antonov, M. Kompaniets, and A. Vasil'ev, *Int. J. Mod. Phys. B* **17**, 2137 (2003).
- [11] H.-K. Janssen, *Z. Phys. B* **23**, 377 (1976).
- [12] C. de Dominicis, *J. Phys. Colloq.* **37**, C1-247 (1976).
- [13] P. C. Martin, E. D. Siggia, and H. A. Rose, *Phys. Rev. A* **8**, 423 (1973).
- [14] I. Hosokawa, *J. Phys. Soc. Jpn.* **21**, 538 (1966).
- [15] I. Hosokawa, *J. Phys. Soc. Jpn.* **25**, 271 (1968).
- [16] G. Rosen, *J. Math. Phys.* **24**, 193 (1983).
- [17] W. D. Thacker, *J. Math. Phys.* **38**, 300 (1997).
- [18] V. Gurarie and A. A. Migdal, *Phys. Rev. E* **54**, 4908 (1996).
- [19] E. Balkovsky, G. Falkovich, I. Kolokolov, and V. Lebedev, *Phys. Rev. Lett.* **78**, 1452 (1997).
- [20] J. D. Cole, *Q. Appl. Math.* **9**, 225 (1951).
- [21] E. Hopf, *Commun. Pure Appl. Math.* **3**, 201 (1950).
- [22] F. Bouchet and E. Simonnet, *Phys. Rev. Lett.* **102**, 094504 (2009).
- [23] F. Bouchet, J. Laurie, and O. Zaboronski, *J. Phys.: Conf. Ser.* **318**, 022041 (2011).
- [24] C.-E. Bréhier, M. Gazeau, L. Goudenège, T. Lelièvre, and M. Rousset, *Ann. Appl. Probab.* **26**, 3559 (2016).
- [25] J. Rolland, F. Bouchet, and E. Simonnet, *J. Stat. Phys.* **162**, 277 (2016).
- [26] G. Ferré and H. Touchette, *J. Stat. Phys.* **172**, 1525 (2018).
- [27] C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler, *J. Chem. Phys.* **108**, 1964 (1998).
- [28] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, *Annu. Rev. Phys. Chem.* **53**, 291 (2002).
- [29] C. Giardina, J. Kurchan, V. Lecomte, and J. Tailleur, *J. Stat. Phys.* **145**, 787 (2011).
- [30] G. Torrie and J. Valleau, *J. Comput. Phys.* **23**, 187 (1977).
- [31] M. R. Shirts and J. D. Chodera, *J. Chem. Phys.* **129**, 124105 (2008).
- [32] M. Karplus and J. A. McCammon, *Nat. Struct. Mol. Biol.* **9**, 646 (2002).
- [33] F. Ragone, J. Wouters, and F. Bouchet, *Proc. Natl. Acad. Sci. USA* **115**, 24 (2018).
- [34] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, *Phys. Lett. B* **195**, 216 (1987).
- [35] J. C. Sexton and D. H. Weingarten, *Nucl. Phys. B* **380**, 665 (1992).
- [36] S. Gottlieb, W. Liu, D. Toussaint, R. L. Renken, and R. L. Sugar, *Phys. Rev. D* **35**, 2531 (1987).
- [37] J. M. Burgers, *The Nonlinear Diffusion Equation: Asymptotic Solutions and Statistical Problems* (Springer Netherlands, Houten, 1974).
- [38] J. Bec and K. Khanin, *Phys. Rep.* **447**, 1 (2007).
- [39] P. Düben, D. Homeier, K. Jansen, D. Mesterházy, G. Münster, and C. Urbach, *Europhys. Lett.* **84**, 40002 (2008).
- [40] D. Mesterházy and K. Jansen, *New J. Phys.* **13**, 103028 (2011).
- [41] D. Mesterházy, L. Biferale, K. Jansen, and R. Tripiccione, in *Proceedings of the 31st International Symposium on Lattice Field Theory (Lattice 2013)*, Mainz, 2013 (Sissa Medialab, Trieste, 2014), p. 054.
- [42] S. L. Adler, *Phys. Rev. D* **23**, 2901 (1981).
- [43] S. L. Adler, *Phys. Rev. D* **37**, 458 (1988).
- [44] E. Medina, T. Hwa, M. Kardar, and Y.-C. Zhang, *Phys. Rev. A* **39**, 3053 (1989).
- [45] A. Chekhlov and V. Yakhot, *Phys. Rev. E* **51**, R2739 (1995).
- [46] F. Hayot and C. Jayaprakash, *Phys. Rev. E* **54**, 4681 (1996).
- [47] M. K. Verma, *Physica A* **277**, 359 (2000).
- [48] R. Phythian, *J. Phys. A: Math. Gen.* **10**, 777 (1977).
- [49] F. Langouche, D. Roekaerts, and E. Tirapegui, *Physica A* **95**, 252 (1979).
- [50] R. V. Jensen, *J. Stat. Phys.* **25**, 183 (1981).
- [51] L. Onsager and S. Machlup, *Phys. Rev.* **91**, 1505 (1953).
- [52] H. Nakazato, K. Okano, L. Schulke, and Y. Yamanaka, *Nucl. Phys. B* **346**, 611 (1990).
- [53] A. D. Kennedy, in *Proceedings of the workshop "Perspectives in Lattice QCD"* (World Scientific, Singapore, Hackensack, NJ, 2007), pp. 15–81.
- [54] M. Luscher, [arXiv:1002.4232](https://arxiv.org/abs/1002.4232).
- [55] R. M. Neal, [arXiv:1206.1901](https://arxiv.org/abs/1206.1901).
- [56] G. G. Batrouni, G. R. Katz, A. S. Kronfeld, G. P. Lepage, B. Svetitsky, and K. G. Wilson, *Phys. Rev. D* **32**, 2736 (1985).
- [57] C. T. H. Davies, G. G. Batrouni, G. R. Katz, A. S. Kronfeld, G. P. Lepage, P. Rossi, B. Svetitsky, and K. G. Wilson, *Phys. Rev. D* **41**, 1953 (1990).
- [58] S. Catterall and S. Karamov, *Phys. Lett. B* **528**, 301 (2002).
- [59] P. Gerhold, Upper and lower Higgs boson mass bounds from a chirally invariant lattice Higgs-Yukawa model, Ph.D. thesis, Humboldt University of Berlin, 2009.
- [60] L. Biferale, R. Grauer, K. Jansen, G. Margazoglou, D. Mesterházy, T. Rosenow, and R. Tripiccione (unpublished).
- [61] D. J. Higham, *SIAM Rev.* **43**, 525 (2001).
- [62] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer, Berlin, 1992).
- [63] C. Canuto, M. Hussaini, A. Quarteroni, and T. Zang, *Spectral Methods in Fluid Dynamics* (Springer, New York, 1988).
- [64] U. Wolff, *Comput. Phys. Commun.* **156**, 143 (2004).
- [65] H. Meyer, H. Simma, R. Sommer, M. D. Morte, O. Witzel, and U. Wolff, *Comput. Phys. Commun.* **176**, 91 (2007).
- [66] H. Müller-Krumbhaar and K. Binder, *J. Stat. Phys.* **8**, 1 (1973).
- [67] N. Madras and A. D. Sokal, *J. Stat. Phys.* **50**, 109 (1988).
- [68] A. M. Ferrenberg, D. P. Landau, and K. Binder, *J. Stat. Phys.* **63**, 867 (1991).
- [69] A. M. Ferrenberg and R. H. Swendsen, *Phys. Rev. Lett.* **61**, 2635 (1988).
- [70] A. Pelissetto and F. Ricci-Tersenghi, in *Large Deviations in Physics: The Legacy of the Law of Large Numbers*, edited by A. Vulpiani, F. Cecconi, M. Cencini, A. Puglisi, and D. Vergni, Lecture Notes in Physics Vol. 885 (Springer, Berlin, 2014).
- [71] A. M. Ferrenberg, D. P. Landau, and R. H. Swendsen, *Phys. Rev. E* **51**, 5092 (1995).

- [72] A. M. Ferrenberg and R. H. Swendsen, *Phys. Rev. Lett.* **63**, 1195 (1989).
- [73] U. Ray, G. K.-L. Chan, and D. T. Limmer, *J. Chem. Phys.* **148**, 124120 (2018).
- [74] T. R. Gingrich and P. L. Geissler, *J. Chem. Phys.* **142**, 234104 (2015).
- [75] G. Falkovich, I. Kolokolov, V. Lebedev, and A. Migdal, *Phys. Rev. E* **54**, 4896 (1996).
- [76] M. Chertkov, *Phys. Rev. E* **55**, 2722 (1997).
- [77] I. Daumont, T. Dombre, and J.-L. Gilson, *Phys. Rev. E* **62**, 3592 (2000).
- [78] L. Biferale, I. Daumont, T. Dombre, and A. Lanotte, *Phys. Rev. E* **60**, R6299 (1999).
- [79] J. Laurie and F. Bouchet, *New J. Phys.* **17**, 015009 (2015).
- [80] E. R. Weeks, Y. Tian, J. S. Urbach, K. Ide, H. L. Swinney, and M. Ghil, *Science* **278**, 1598 (1997).
- [81] M. J. Schmeits and H. A. Dijkstra, *J. Phys. Oceanogr.* **31**, 3435 (2001).
- [82] T. Grafke, R. Grauer, and T. Schäfer, *J. Phys. A: Math. Theor.* **48**, 333001 (2015).
- [83] J. Hubbard, *Phys. Rev. Lett.* **3**, 77 (1959).
- [84] R. L. Stratonovich, *Sov. Phys. Dokl.* **2**, 416 (1957).
- [85] A. I. Chernykh and M. G. Stepanov, *Phys. Rev. E* **64**, 026306 (2001).
- [86] T. Grafke, R. Grauer, and T. Schäfer, *J. Phys. A* **46**, 062002 (2013).
- [87] T. Grafke, R. Grauer, T. Schäfer, and E. Vanden-Eijnden, *Europhys. Lett.* **109**, 34003 (2015).
- [88] M. V. Feigel'man, *Zh. Eksp. Teor. Fiz.* **79**, 1095 (1980) [*Sov. Phys. JETP* **52**, 555 (1980)].
- [89] S. A. Orszag, *J. Atmos. Sci.* **28**, 1074 (1971).
- [90] D. R. Durran, *Numerical Methods for Fluid Dynamics* (Springer, New York, 2010).