Data-based approach for time-correlated closures of turbulence models

J. Domingues Lemos [1] and A. A. Mailybaev [2]

[1] *Universidade Federal do Rio de Janeiro, Rua Antônio Barros de Castro, 119 Cidade Universitária, Rio de Janeiro, RJ 21941-853, Brazil*

[2] *Instituto Nacional de Matematica Pura e Aplicada, Estrada Dona Castorina, 110 Jardim Botânico, Rio de Janeiro, RJ 22460-320, Brazil*

(Received 14 August 2023; accepted 4 January 2024; published 1 February 2024)

Developed turbulent motion of fluid still lacks an analytical description despite more than a century of active research. Nowadays, phenomenological ideas are widely used in practical applications, such as small-scale closures for numerical simulations of turbulent flows. In the present paper, we use a shell model of turbulence to construct a closure intended to have a solid theoretical background and to capture intrinsic probabilistic features of turbulence. Shell models of turbulence are dynamical deterministic systems used to model energy cascade and other key aspects of the Navier-Stokes such as intermittency. We rescale the variables of the Sabra model in a way which leads to hidden symmetries and universal distributions. We then use such fine distributions to write closures, i.e., missing expressions for some of the Sabra variables. Our closures rely on approximating probability density functions using a Gaussian mixture model, which makes them probabilistic by nature and allows us to write time-correlated closures. We also provide a framework where other machine learning tools can be employed with reduced black-box aspects.

DOI: 10.1103/PhysRevE.109.025101

I. INTRODUCTION

A fully resolved simulation of a developed turbulent flow, which contemplates all scales of motion up to Kolmogorov's viscous scale, is unfeasible for many real-world flows. In the context of the Navier-Stokes equation, dealing with closure problems is essentially handling the fact that there are more degrees of freedom than there are equations we are able to solve. Many approaches for closure problems in Navier-Stokes turbulence rely on phenomenologial predictions or averaging of fluctuations like, e.g., Reynolds-averaged Navier-Stokes equations and large eddy simulations [1–3]. Some works, such as Ref. [4], put effort into the modeling error at small scales, while Refs. [5,6] go into more general approaches in the proposing of subgrid models. The overall idea is to introduce an effective small scale (cutoff), such that only the larger scales are resolved while smaller scales are modeled by the closure.

Even though approximate schemes may be satisfactory for practical applications, little is clear about a fundamental question: Can one formulate a small-scale closure that provides an accurate description of the flow at all resolved scales? The major obstacle for answering this question in turbulence is the intermittency phenomenon [7] because intermittent velocity fields are not statistically self-similar at small scales. Thus, a proper closure must reproduce multifractal properties of the flow at and beyond a cutoff scale. Another important, though less evident condition, is related to the observations that solutions for the Navier-Stokes equations at very large Reynolds numbers are expected to evolve as a stochastic process triggered by a small-scale noise [8–13], i.e., the solutions are spontaneously stochastic [14–18]. This implies that even tiny small-scale noise cannot be ignored in equations of motion when simulations aim for finite-time predictions, while this aspect becomes less important for long-term statistics. As a consequence, a proper closure must be probabilistic, reproducing the intrinsic stochastic nature of small-scale turbulence. We conclude that the existence of accurate closure depends on the two phenomena: intermittency and spontaneous stochasticity, both still not well understood theoretically.

A similar problem can be formulated for shell models of turbulence [19–21], which are more tractable than the Navier-Stokes equation while preserving its key aspects, such as intermittency, energy cascade, and spontaneous stochasticity. The attempt of formulating an accurate closure for a shell model was done in Ref. [22], where a probabilistic approach was combined with the use of Kolmogorov multipliers. The Kolmogorov multipliers are ratios of velocity fluctuations at adjacent scales, which appear to have a universal (not intermittent) single-time statistics in the inertial interval of developed turbulence [23–25]. As reported in Ref. [22], accuracy of the proposed closures was limited. One possible reason is that multitime multiplier correlations are not taken into account in these closures. The multitime multiplier correlations, however, are again intermittent, i.e., do not possess a universal statistics.

In this paper, we develop the formalism for probabilistic, time-correlated closures in the Sabra shell model of turbulence [26]. This formalism describes the closures which are potentially accurate in both of the key properties: intermittency and stochasticity. To overcome the problem that velocity fluctuations have multifractal statistics across the inertial interval, we apply the space-time rescaling proposed in Refs. [27,28]. This rescaling recovers the hidden scale invariance of the multitime and multiscale statistics. As a consequence, the resulting closure becomes universal, scale independent, and can be properly conditioned to the resolved scales not only at a current time but also to a prehistory of computations.

For a practical realization of our formalism, we employ the probability density estimation with Gaussian mixture models [29]. This approach is convenient for data generation because the estimated density is computed as a weighted sum of Gaussian components, which can easily generate conditional samples of new, unseen instances of data. On the other hand, this method demonstrates limitations in the learning process, especially in cases where the density estimation translates into an optimization problem in a space of high dimension. Numerical tests were performed leading to a satisfactory but not systematically improving performance of the models. These tests indicate that more elaborate tools of machine learning must be combined with our formalism to achieve the accurate closure.

We discuss the Sabra shell model in Sec. II and its rescaled version with the hidden scaling symmetry in Sec. III. Section IV describes the closure problem. In Sec. V, we discuss the process of density estimation and conditional sampling. We present numerical results in Secs. VI for single-time closures and in Sec. VII for closures with time conditioning. We summarize the results and discuss further perspectives in the discussion of Sec. VIII. The Appendix contains supplementary computations and tables.

II. SHELL MODEL

Shell models are infinite-dimensional systems of ordinary differential equations which rely on a discretization of the phase space that considers a sequence of wave numbers as a geometric progression: $|\mathbf{k}| = k_n = k_0\lambda^n$. Typically, $\lambda = 2$ and $k_0 = 1$. The Sabra model [26], which we are considering in this paper, is given by the following set of equations:

$$\frac{du_n}{dt} = i\left(k_{n+1}u_{n+2}u^*_{n+1} - \frac{1}{2}k_n u_{n+1}u^*_{n-1} + \frac{1}{2}k_{n-1}u_{n-1}u_{n-2}\right) - \nu k_n^2 u_n + f_n, \tag{1}$$

where $u_n \in \mathbb{C}$ is a complex-valued variable describing a velocity fluctuation at shell n, ν is the viscosity coefficient, and f_n is the forcing term. In this paper, f_1 is a nonzero constant, and $f_n = 0$ for $n \geqslant 2$. The Reynolds number is defined as $\mathrm{Re} = UL/\nu$, with the integral scale $L = 1/k_0$ and the characteristic large-scale velocity $U = \sqrt{|f_1|/k_0}$.

The system of equations (1) mimics important properties of developed hydrodynamic turbulence such as the energy cascade to small scales and intermittency [7,21]. The regime of fully developed turbulence corresponds to very large Reynolds numbers. In this case, one distinguishes a wide range of scales called the inertial interval, for which both forcing and dissipative terms can be neglected. Within the K41 approximation [7], the inertial interval extends to the wave numbers $1 \ll k/k_0 \ll \mathrm{Re}^{3/4}$; this estimate is rough, but only because of the anomalous corrections [30,31]. The forcing range corresponds to the wave numbers $k_n/k_0 \sim 1$ at which the energy is produced by the work of large-scale forces. This energy is transported through the inertial interval until it dissipates at wave numbers of the dissipation range, $k/k_0 \gtrsim \mathrm{Re}^{3/4}$.

In the context of shell models, the closure problem involves first setting a finite number of resolved scales $n = 1, \ldots, s$ to be computed by solving the equations of motion, while the remaining unresolved scales with $n > s$ are to be provided by the closure model. Considering equation (1) with $n = 1, 2, \ldots, s$ leaves us with missing expressions for shell velocities u_{s+1} and u_{s+2}, which are complex numbers. If the total amount of resolved scales is large enough to cover forcing range, inertial range and dissipation range, then one can set $u_{s+1} = u_{s+2} = 0$. In this situation, the simulation is considered to be fully resolved. The goal of the closure, however, is to set s to a much larger scale (smaller wave number), namely, we consider k_s to be a wave number in the inertial interval. Then the closure consists of writing expressions for the shell speeds u_{s+1} and u_{s+2}, which are the only variables missing in Eq. (1) for $n = s - 1$ and s.

III. RESCALED VARIABLES AND HIDDEN SYMMETRY

The major problem in defining closures for turbulence models is the intermittency. As a consequence, statistics collected in time are not universal with respect to the choice of s, since velocity fluctuations are getting more intermittent with the increase of k_s. Similarly, there is a lack of universality with respect to the forcing as well. It was noted, however, that universal statistics are recovered for single-time ratios of velocities [24,25,32], which are called Kolmogorov multipliers following the ideas of the K62 theory [7,23]. It was shown recently that this universality is a consequence of the hidden scaling symmetry of the equations of motion [27,28], which allows us to recover the full scale invariance in the inertial range by using a properly rescaled velocities and time.

The rescaled variables are introduced by fixing a reference shell m and setting the momentary temporal scale (turnover time) for this shell as

$$T_m(t) = \left(k_0^2U^2 + \sum_{n<m} k_n^2|u_n|^2\right)^{-1/2}, \tag{2}$$

where U is the characteristic velocity defined above. The quantity T_m is always positive and the sum in the parenthesis can be seen as accumulated enstrophy of shells up to (but not including) m. We then introduce the rescaled time as

$$\tau = \int_0^t \frac{dt'}{T_m(t')}. \tag{3}$$

The nonlinear change from original time t to rescales time τ acts in such a way that the long stretches of time when shell velocities would be near zero are shrunk and short periods where velocities would rapidly vary are stretched out. This notion becomes clearer when we think of rescaled time increments $d\tau$ as a rescaling of original time increments dt by a turnover time $T_m(t)$. Next, we introduce the relative shell number $N = n - m$ and define rescaled velocities $\mathcal{U}_N$ as

$$\mathcal{U}_N = k_m T_m(t) u_{N+m}(t). \tag{4}$$

We can rewrite the Sabra system from Eq. (1) in its rescaled version, which for the full (forced and viscous) system

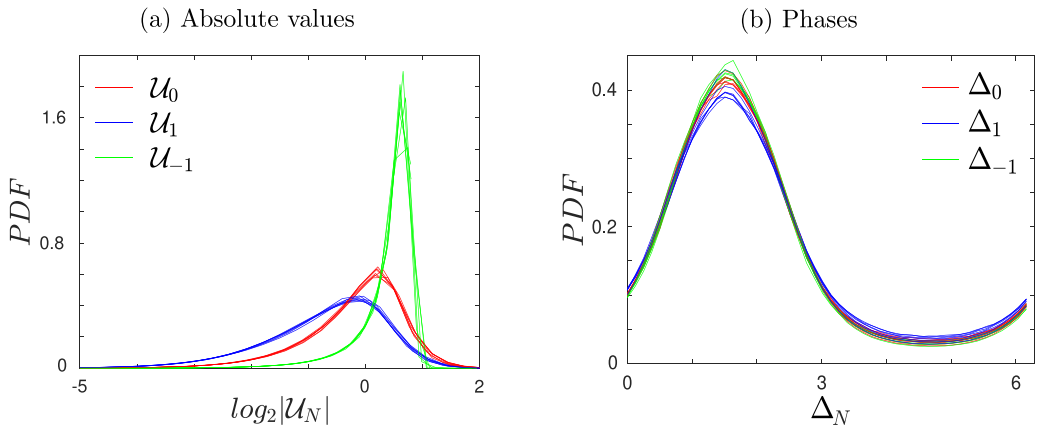

FIG. 1. PDFs of absolute values and phases of $\mathcal{U}_{-1}, \mathcal{U}_0, \mathcal{U}_1$ $m = 8, \ldots, 14$. Statistics are collected in time and rescaled variables are computed from one Sabra simulation.

becomes

$$\frac{d\mathcal{U}_N}{d\tau} = i\Big(k_{N+1}\mathcal{U}_{N+2}\mathcal{U}^*_{N+1} - \frac{1}{2}k_N\mathcal{U}_{N+1}\mathcal{U}^*_{N-1} + \frac{1}{2}k_{N-1}\mathcal{U}_{N-1}\mathcal{U}_{N-2}\Big) + (\xi_{\text{total}} - \nu k^2_{N+m}T_m)\mathcal{U}_N + T^2_m k_m f_{N+m}, \tag{5}$$

where

$$\xi_{\text{total}} = \xi + \xi_\nu + \xi_f, \tag{6}$$

$$\xi = \sum_{N<0} k^3_N \,\text{Im}\left(2\mathcal{U}^*_N\mathcal{U}^*_{N+1}\mathcal{U}_{N+2} - \frac{1}{2}\mathcal{U}^*_{N-1}\mathcal{U}^*_N\mathcal{U}_{N+1} - \frac{1}{4}\mathcal{U}^*_{N-1}\mathcal{U}_N\mathcal{U}^*_{N-2}\right), \tag{7}$$

$$\xi_\nu = \nu T_m k^2_m \sum_{N<0} k^4_N |\mathcal{U}_N|^2, \tag{8}$$

$$\xi_f = -T^2_m \sum_{N<0} k_{N+m}k_N \,\text{Re}\,(\mathcal{U}^*_N f_{N+m}), \tag{9}$$

$$T_m = \frac{1}{k_0 U}\left(1 - \sum_{N<0} k^2_N |\mathcal{U}_N|^2\right)^{1/2}. \tag{10}$$

For the detailed derivation of Eqs. (5)–(10), see Appendix B.

The hidden scale invariance [27,33] is the independence of the statistics of rescaled velocities $\mathcal{U}_N(\tau)$ in the inertial interval with respect to the choice of the reference shell m. Thus, rescaled variables restore the universality broken by the intermittency. This indicates that the presence of intermittent fluctuations, which prevented Sabra's statistics from being universal, has been encoded in the change of variables from Eqs. (3) and (4).

To observe hidden symmetry, we use the data from a simulation of fully resolved Sabra system with 30 shells, while simultaneously solving for $\tau(t)$ using Eq. (3). In this simulation, $u_0 = u_{-1} = 0$, $\nu = 10^{-8}$, $f_1 = 1 + i$ and the initial condition is taken from a stationary state. The results are presented in Fig. 1(a) showing probability density functions (PDFs) of absolute values for $\mathcal{U}_N$ for $n = -1, 0, 1$. Colors indicate different values of N, while a number of the curves of the same color correspond to the reference shell varied in the inertial interval as $m = 8, \ldots, 14$. The collapse of the curves of the same color verifies the hidden scale invariance.

Likewise, we can look at PDFs of the phases of rescaled variables. The simplest nontrivial (non-uniform) PDF corresponds to the phase differences of the form [24,25]

$$\Delta_N = \arg(\mathcal{U}_N) - \arg(\mathcal{U}_{N-1}) - \arg(\mathcal{U}_{N-2}). \tag{11}$$

PDFs for these phases are presented in Fig. 1(b). Again, each of Δ_{-1}, Δ_0, and Δ_1 is represented by a different color, and the curves of the same color collapse, verifying the hidden scale invariance.

IV. CLOSURE FOR A TRUNCATED RESCALED SYSTEM

In the context of the Sabra model, because the coupling of scales is local, writing a closure for a reduced model with s scales translates into writing expressions for the closure variables u_{s+1} and u_{s+2}. Because shell velocities are complex numbers, we need to compute absolute values and phases of each one. When s is a scale of motion in the inertial range, it is unreasonable to set $u_{s+1} = u_{s+2} = 0$ because energy cascade is still at play and the viscous term is overcome by the nonlinear term and is, therefore, not strong enough to dissipate any energy.

In terms of rescaled variables, we need to first choose a reference shell. In this paper, we are committing to $m = s + 1$, because this sets the new local temporal scale at shell $s + 1$, which is one of the closure variables. With this setting, the closure variables for the rescaled system are $\mathcal{U}_0$ and $\mathcal{U}_1$; see Eq. (4).

There are two important aspects of the closures presented in this paper. The first one is that they are probabilistic. Instead of fixing estimates for absolute values and phases of the closure variables, we are estimating the PDFs of these quantities, like the ones presented in Figs. 1(a) and 1(b), and sampling new instances of data from the density estimation. The new samples are used to evolve the reduced models in time.

The second one is the fact that the new instances of data for absolute values and phases can be sampled conditioned to the system's prehistory, which should take into account the fact that shell velocities are time-correlated. This raises the

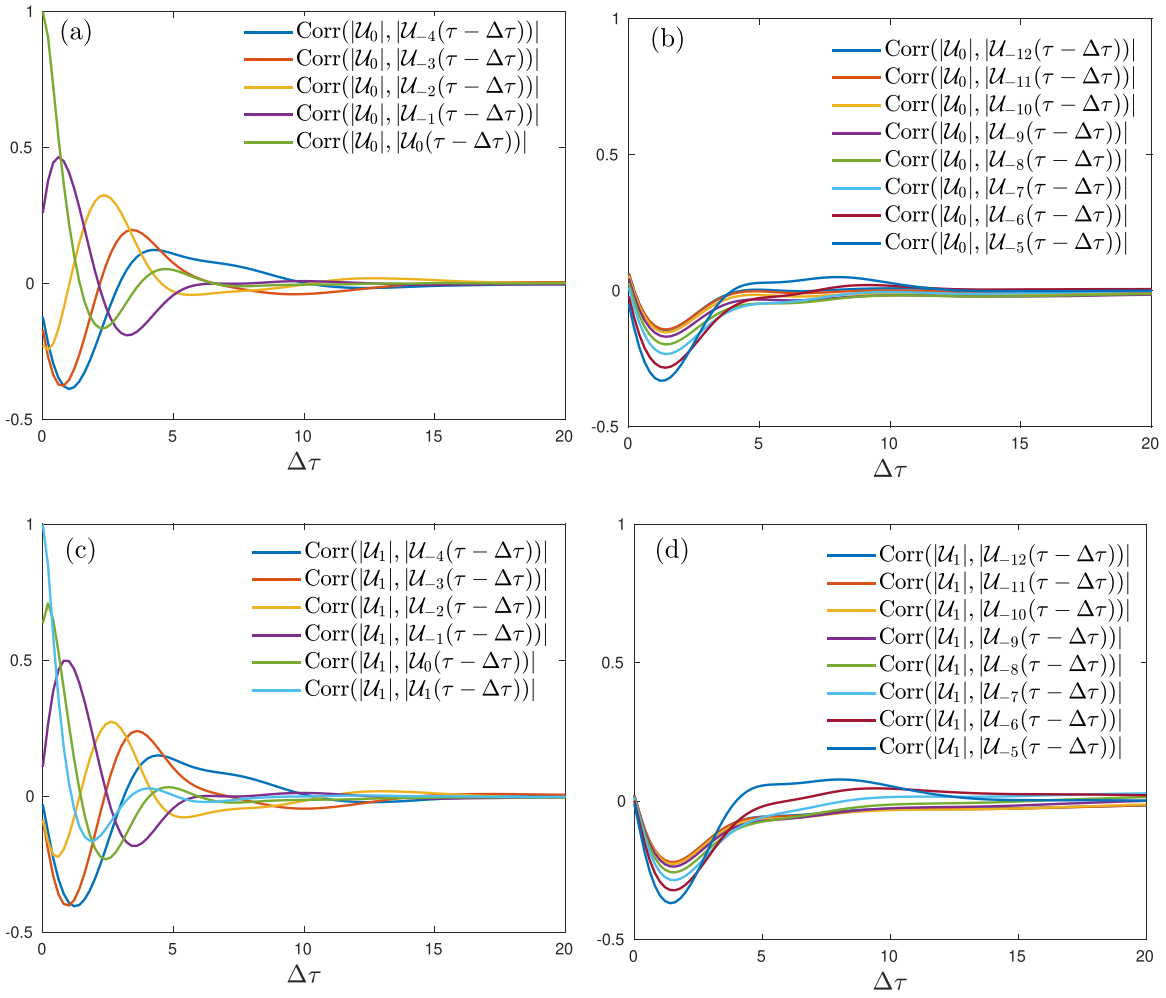

FIG. 2. Figures 2(a) and 2(b) show correlation coefficients between $\mathcal{U}_0$ and other delayed shells, and Figs. 2(c) and 2(d) show correlation coefficients between $\mathcal{U}_1$ and other delayed shells for different values of $\Delta\tau$. Here, $s = 12$ and $m = s + 1$.

question of how long into the pre-history we should be looking. It is important to recall that the representation in rescaled variables $\mathcal{U}_N(\tau)$ has universal multitime statistics [27], with the exceptional role played by the rescaling of time. This property is crucial for the formulation of prehistory-dependent closures, which is the major advance of our formulation compared to the previous ideas based on Kolmogorov multipliers [22,24,25].

In Fig. 2, we plot correlation coefficients between $\mathcal{U}_0$ and all other previous shells for different time delays $\Delta\tau$, as well as correlation coefficients between $\mathcal{U}_1$ and all other previous shells, also $\Delta\tau$ delayed. What we can see is that correlation decays as we progress into the past, as well as between different scales as they grow further apart in space. This indicates that the bulk of relevant information may be captured by conditioning the sampling of new instances of data to neighboring shells delayed some $\Delta\tau \in [0, 10]$ approximately. Such conditioning can be explicitly made in each time step and, moreover, it is possible to condition to more than one value at a time. We stress that maximum correlations are generally attained at considerable (order-one) delays. This highlights the importance of prehistory for the closures.

In the rescaled reduced system when the shell s corresponds to a scale of motion in the inertial range, the viscous terms are negligible. Recalling that we committed to $m = s + 1$, the rescaled reduced system is given by

$$\begin{aligned}\frac{d\mathcal{U}_N}{d\tau} =& i\Big(k_{N+1}\mathcal{U}_{N+2}\mathcal{U}^*_{N+1} - \frac{1}{2}k_N\mathcal{U}_{N+1}\mathcal{U}^*_{N-1} \\ &+ \frac{1}{2}k_{N-1}\mathcal{U}_{N-1}\mathcal{U}_{N-2}\Big) \\ &+ (\xi + \xi_f)\mathcal{U}_N + T_m^2 k_m f_{N+m}, \\ N =& -s, \dots, -1, \end{aligned} \tag{12}$$

where ξ, ξ_f, and T_m are given by Eqs. (7), (9), and (10). These equations can be solved numerically for $N = -s, \dots, -1$ as long as the closure provides values of $\mathcal{U}_0$ and $\mathcal{U}_1$ at each time step. In our closures, instead of the complex values $\mathcal{U}_0$ and $\mathcal{U}_1$, we find it more convenient to use their absolute values and phases. Furthermore, instead of the phases of $\mathcal{U}_0$ and $\mathcal{U}_1$, one can reformulate the closure in terms of phase differences Δ_0 and Δ_1 given by Eq. (11) [22].

V. GAUSSIAN MIXTURE MODELS

Before we start writing closures, we need to discuss how the density estimation is performed. We are employing a well-established method called the Gaussian mixture models

(GMMs), which relies solely on data from the target distribution. In our case, the target distribution consists of absolute values of $\mathcal{U}_0$ and $\mathcal{U}_1$, as well as the associated multipliers' phases and any other values to which we wish to condition them. The GMM works on the assumption that the distribution can be approximated as a weighted sum of Gaussian components. Given a PDF p and a positive amount of Gaussian components K, we want to find means μ_k, variances Σ_k, and positive weights π_k such that

$$p(\mathbf{z}) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\mathbf{z}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k). \tag{13}$$

Here, $\mathcal{N}(\mathbf{z}|\mu, \Sigma)$ denotes the normal distribution of mean μ and variance Σ, while $\mathbf{z} \in \mathbb{R}^d$ for integer $d \geqslant 1$. The parameters are computed as an iterative process using an expectation-maximization algorithm, which is based on maximum likelihood estimates. We initialize with random guesses. This process and algorithm have been detailed, for example, in Ref. [29].

The amount of Gaussian components K is a parameter that has to be decided beforehand. Theoretical results guarantee that we can find a suitable approximation [34], given enough data and small overlapping of the Gaussian components [35,36], which means there are systematical ways of choosing this parameter. However, some knowledge of what the data set looks like is surely beneficial to this process.

What we saw in this paper, though, was a relevant struggle on the method's performance when dealing with high-dimensional data. This relates closely to the fact that Euclidean distances between data points increase exponentially with the dimension, which turns the data set into a sparse one [37,38]. This can be mitigated by increasing the amount of data samples fed to the density estimator, but only up to a certain point, limited by finite data and RAM memory.

The choice of GMM for the density estimation relates heavily to the sampling process. To evolve reduced models in time, we need new instances of data from the densities being approximated, and Gaussian generators are widely available. Once we are in possession of an estimate, to obtain a new datum we can use the weights π_k to choose one of the Gaussian components and then use any in-built Gaussian generator with mean μ_k and variance Σ_k to generate a new sample.

If we want to sample conditionally, once the Gaussian component is selected, it is enough to recalculate the mean and the variance of the selected component since any slice of a Gaussian distribution is another Gaussian distribution [39]. Given a concatenated random variable $Y = [y_1, y_2]$, $Y \sim \mathcal{N}(\mu, \Sigma)$, we write

$$\mu = [\mu_1, \mu_2], \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}, \tag{14}$$

with compatible dimensions. Then, $p(y_1|y_2) = \mathcal{N}(\bar{\mu}, \bar{\Sigma})$ with [39]

$$\bar{\mu} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(y_2 - \mu_2), \quad \bar{\Sigma} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}. \tag{15}$$

This process for generating a new data sample happens at each time step in the process of evolving the reduced model in time. We intend on sampling absolute values and multiplier's phases for $\mathcal{U}_0$ and $\mathcal{U}_1$, as presented in Figs. 1(a) and 1(b), while conditioning them to the prehistory of shell velocities. Once sampled, these values are used in Eq. (12) to close the system.

VI. SINGLE-TIME CLOSURES

As we described in Sec. III, the universal closure of the shell model can be formulated in terms of rescaled variables. These variables define scale invariant statistics for closure variables $\mathcal{U}_0$ and $\mathcal{U}_1$, namely, PDFs of module and phase of each one. Now we apply the method described in Sec. V to estimate such densities and generate new data under the same distributions to evolve reduced models in time. The reduced models we simulate were discussed in Sec. IV and are given by Eq. (12).

Below we formulate different closures given by expressions for $\mathcal{U}_0$ and $\mathcal{U}_1$. Then, we evolve the reduced models and compare statistics to the baseline fully resolved Sabra model. For the comparison, we look at PDFs of energy flux Π_n from shell n to $n+1$, which is given by

$$\Pi_n = \mathrm{Im}\left(k_{n+1}u_{n+2}u_{n+1}^*u_n^* + \tfrac{1}{2}k_n u_{n+1}u_n^*u_{n-1}^*\right), \tag{16}$$

as well as moments of order p defined as

$$S_p(k_n) = \langle |u_n|^p \rangle, \tag{17}$$

where the averaging is with respect to time. Also, we compare normalized PDFs of real parts of velocities. The summary of all closures, including other choices for slightly different closures not reported in the main text, is given in Appendix A. In this section, we consider a simpler class of single-time closures, while the more elaborated closures with multi-time conditioning are studied in the next section.

It is important to note that while the reduced models and closures are written for rescaled variables, the simulations are all run for the original velocities in rescaled time. In each time step, we switch to rescaled variables to compute the closure variables and then switch back to original variables to evolve it in time. This was a friendlier alternative to computing rescaled variables in rescaled time, both numerically and for comparison purposes, and it is done noting that

$$\frac{du_n}{d\tau} = \frac{du_n}{dt}\frac{dt}{d\tau} = \frac{du_n}{dt}T_m. \tag{18}$$

Each simulation was run until the rescaled time $\tau = 30\,000$ with time step 10^{-4} using an Adams-Bashforth scheme [40].

In the simplest closure, called MJoint, we consider the absolute values of the rescaled variables $|\mathcal{U}_0|$ and $|\mathcal{U}_1|$ only, while their phases will be fixed by the maximum energy flux condition. For determining distribution of these absolute values, we feed to the GMM a data set consisting of pairs $(\log_2|\mathcal{U}_0|, \log_2|\mathcal{U}_1|) = (z_0, z_1)$ using three two-dimensional Gaussian components. Then, once we have an estimation g for the density of the random vector $\mathbf{z} = (z_0, z_1)$, for each time step we can sample one $\mathbf{z}$ and use it to evolve the reduced model in time. This closure is written as

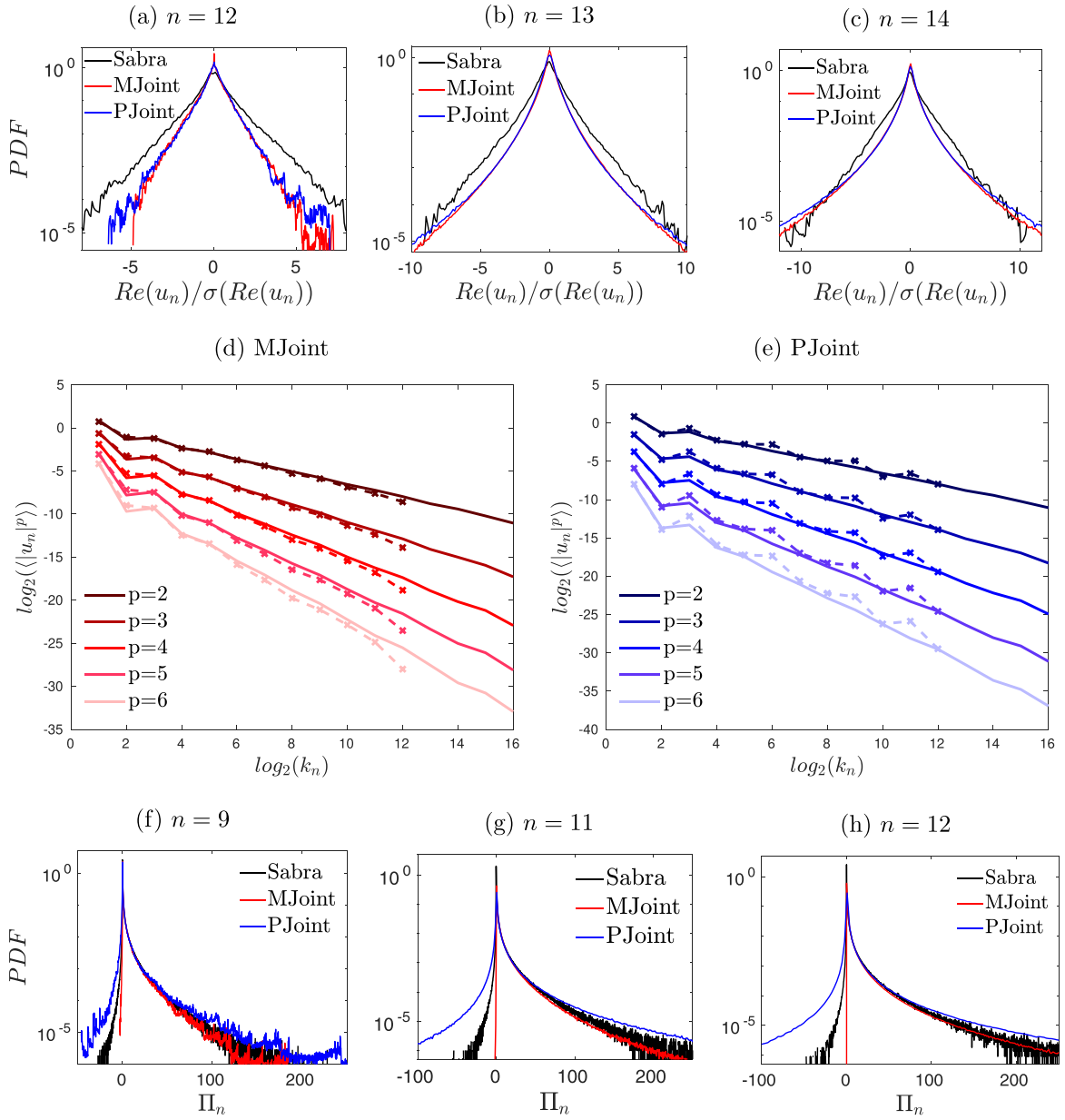

FIG. 3. Cutoff shell is $s = 12$. Figures 3(a)–3(c) present normalized PDFs of real parts for the closure with modules modeling (MJoint) and the closure with modules and phases modeling (PJoint), while Figs. 3(f)–3(h) show energy flux PDFs. All are compared to Sabra statistics in black. Figures 3(d) and 3(e) show moments of orders 2 up to 6 for the closure with module modeling and with phase modeling, respectively. Both are compared to Sabra moments in solid lines.

follows:

$$\mathcal{U}_0 = 2^{z_0} e^{i(\frac{\pi}{2}+\alpha_{-1}+\alpha_{-2})}, \tag{19}$$

$$\mathcal{U}_1 = 2^{z_1} e^{i(\frac{\pi}{2}+\alpha_0+\alpha_{-1})}, \tag{20}$$

$$\mathbf{z} = (z_0, z_1) \sim g(\mathbf{z}). \tag{21}$$

where $\alpha_N = \arg(\mathcal{U}_N)$. This closure only takes into account the modules and fixes multipliers phases in $\pi/2$, which is the value of highest dissipation [22].

An analogous but more elaborate closure, called PJoint, includes both absolute values and phases of the rescaled variables $|\mathcal{U}_0|$ and $|\mathcal{U}_1|$. In this case, we feed the density estimator data of the form $(\log_2|\mathcal{U}_0|, \Delta_0, \log_2|\mathcal{U}_1|, \Delta_1) = (z_0, z_1, z_2, z_3)$, using six Gaussian components to compute the approximation. Then the closure is formulated as

$$\mathcal{U}_0 = 2^{z_0} e^{i(z_1+\alpha_{-1}+\alpha_{-2})}, \tag{22}$$

$$\mathcal{U}_1 = 2^{z_2} e^{i(z_3+\alpha_0+\alpha_{-1})}, \tag{23}$$

$$\mathbf{z} = (z_0, z_1, z_2, z_3) \sim g(\mathbf{z}). \tag{24}$$

For a cutoff shell of $s = 12$ we can see in Figs. 3(a)–3(c) the PDFs for real parts of velocities in shells 12 through 14. Figure 3(d) shows moments of order 2 up to 6 of the closure with module modeling, given by Eqs. (19)–(21). Figure 3(e) shows moments of order 2 up to 6 for the closure with modules

and phases, given by Eqs. (22)–(24). In Figs. 3(f)–3(h), we see energy flux PDFs for shells 9 through 12.

Both closures reproduce statistics to some degree of fidelity, especially moments. It is important to note that the closure without any phase modeling does not present negative flux of energy in the cutoff shell, which is due to the fixing of the phases at a value that strictly dissipates energy. On the other hand, including phases in the modeling introduces negative flux with higher probability than required.

VII. CLOSURES WITH TIME CONDITIONING

Using the conditional sampling method we introduced in Sec. V, we can write closures that take the prehistory of the reduced model into account. Here the use of rescaled variables is crucial because only in these variables can one recover the universal multitime statistics. To that end, we train the density estimator on a data set containing the modules and phases of the closure variables, which are to be sampled, and all the variables in which we wish to condition. Both closures in this section use six Gaussian components in the density estimation.

Given that the correlation between shell velocities decays in time and in space, the first closure in this section will sample only modules of the closure variables conditioned to the modules of the last three shells of the reduced model. These are the ones representing the most significant correlation in space. Different values of delays $\Delta\tau$ yield different results in the conditioning process due to the fact that when the delay is either too small or too long, the correlation is either too high or essentially nonexistent. Recalling Fig. 2, we remark that different values of $\Delta\tau$ were initially tested to select a suitable delay and are reported in Ref. [41]. The delay selected for our simulations below is $\Delta\tau = 2.4$. Then the closure, called 3Most, can be written as

$$\mathcal{U}_0 = 2^{z_0} e^{i(\frac{\pi}{2}+\alpha_{-1}+\alpha_{-2})}, \tag{25}$$

$$\mathcal{U}_1 = 2^{z_1} e^{i(\frac{\pi}{2}+\alpha_0+\alpha_{-1})}, \tag{26}$$

$$\mathbf{z} = (z_0, z_1) \sim g(\mathbf{z}|\log_2|\mathcal{U}_{-3}(\tau-\Delta\tau)|, \log_2|\mathcal{U}_{-2}(\tau-\Delta\tau)|, \times \log_2|\mathcal{U}_{-1}(\tau-\Delta\tau)|), \tag{27}$$

The last expression signifies that the closed variables are selected randomly with a given distribution, which is conditioned on the three larger-scale absolute velocities taken at time $t-\Delta\tau$.

Another type of conditioning can be performed by taking into account information of the closure variables themselves, allowing them to evolve conditioned to their own prehistory. In this self-conditioning closure, we are also sampling (and conditioning to) multipliers phases and the delay used in the conditioning is also $\Delta\tau = 2.4$. We can write this closure, called self, as

$$\mathcal{U}_0 = 2^{z_0} e^{i(z_1+\alpha_{-1}+\alpha_{-2})}, \tag{28}$$

$$\mathcal{U}_1 = 2^{z_2} e^{i(z_3+\alpha_0+\alpha_{-1})}, \tag{29}$$

$$\mathbf{z} = (z_0, z_1, z_2, z_3) \sim g(\mathbf{z}|\log_2|\mathcal{U}_0(\tau-\Delta\tau)|, \Delta_0, \times \log_2|\mathcal{U}_1(\tau-\Delta\tau)|, \Delta_1). \tag{30}$$

It is important to mention that, for both closures, the expressions above make sense for $\tau \geqslant \Delta\tau$. In the initial stage of simulation, when $0 < \tau < \Delta\tau$, one can employ some deterministic closure like the one described in Ref. [22].

It is relevant to note that even with time conditioning, omitting the modeling of phases still produces drastic differences in statistics. The lack of backscattering in energy flux PDFs is, again, characteristic of the fixing phases to strictly dissipating values for the model (25)–(27). On the other hand, the inclusion of phases in the model (28)–(30) introduces negative energy flux events, but in an exaggerated manner.

The crucial achievement in the formulation of our approach is its universality with respect to the cutoff scale s, because the statistics of the rescaled variables are universal through the inertial interval. Once we change the cutoff shell s, we just need to change the reference shell m accordingly. We refer to Fig. 1 to recall that the densities for $\mathcal{U}_0$ and $\mathcal{U}_1$ do not change when we change the reference shell. To better see this, we will use the density estimations used in the simulations with the cutoff $s = 12$ to run another simulations with cutoff shell $s = 9$. The results are presented in Fig. 5, which correspond to the history conditioned closure in Eqs. (25)–(27), as well as the closure in Eqs. (28)–(30). Comparison of Figs. 4 and 5 shows that the performance of these closures is indeed independent of the cutoff scale s. One can see from Figs. 4(d) and 5(d) that the behavior of closures near the cutoff have similar details, e.g., the same bump in the graphs of the moments at shells 9 and 6, respectively. Such a detailed coincidence was related in Ref. [31] to the extended hidden scale invariance of closures in shell models.

In summary, we observe that all proposed closures lead to a satisfactory performance per se, but the addition of the extra information such as conditioning on extra variables and previous times, do not lead to a noticeable improvement: some of the characteristics improve while others get worse. We discuss now the possible causes for such a behavior and the ways how it can be improved.

VIII. DISCUSSION

In this paper, we addressed the question of whether it is possible and how to formulate a small-scale closure for the shell model of turbulence in such a way that it accurately reproduces the dynamics at all resolved scales of motion. The problem of formulating this closure is related to two important properties of turbulent solutions: intermittency [7] and spontaneous stochasticity [14]. We argue that the proper closure that satisfies these properties must be probabilistic and time-correlated to prehistory. Working with the Sabra model, we apply a space-time rescaling that reveals the hidden scaling symmetry and, hence, restores universality of multiscale and multitime correlations. We then use a density estimation process to approximate the density of rescaled closure variables and apply this approximation to (conditionally) sample new instances of data from this density. Lastly, use these new instances to evolve reduced models in time.

The performance of the closures tested numerically in this paper varied, but in general managed to recover at least some statistics of Sabra model with reasonable accuracy. Even

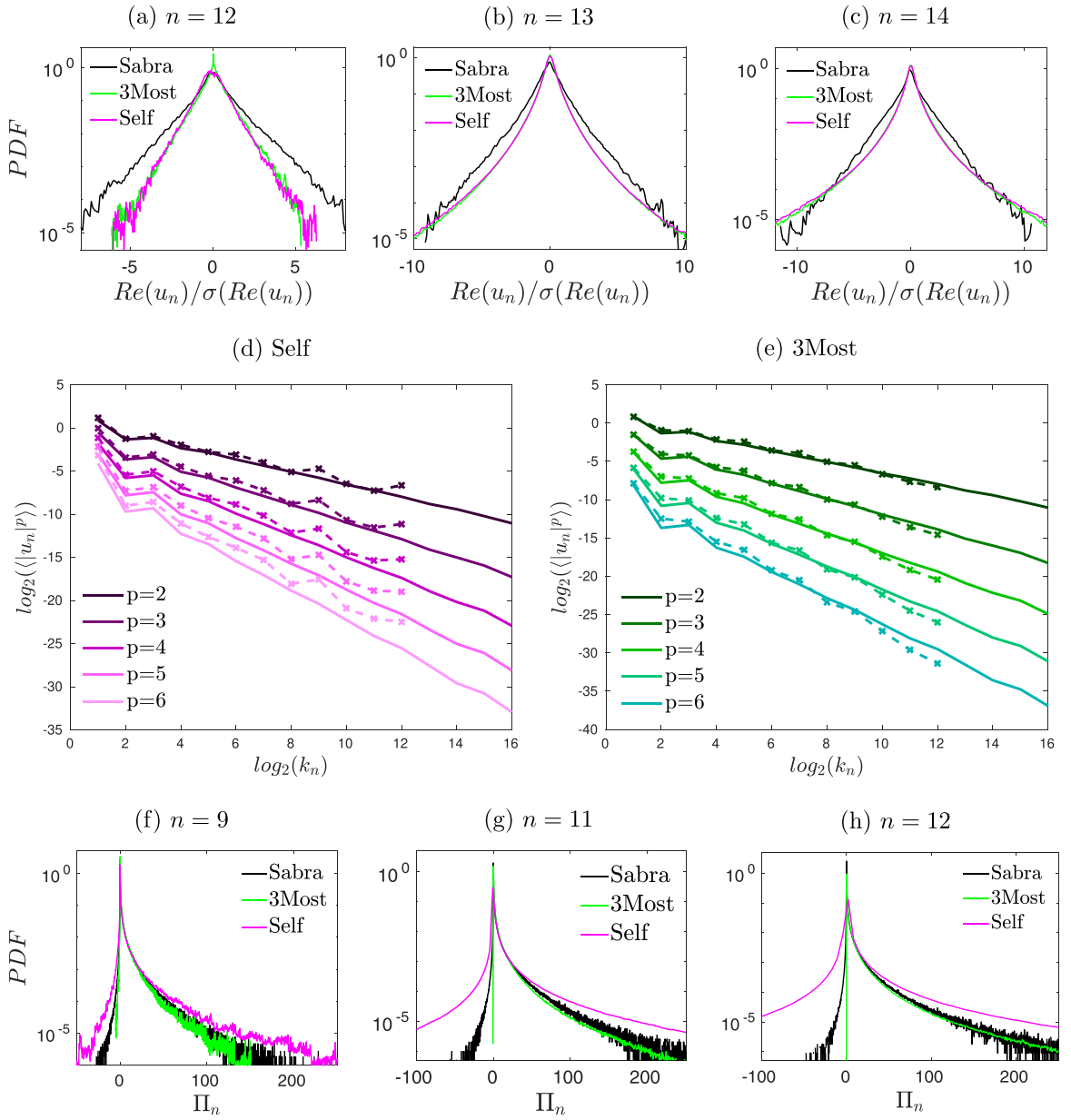

FIG. 4. Cutoff shell is $s = 12$. Figures 4(a)–4(c) present normalized PDFs of real parts for the closure with time conditioning to the three closes shells (3Most) and the closure with time conditioning to the closure variables (self), while Figs. 4(f)–4(h) show energy flux PDFs. All are compared to Sabra statistics in black. Figures 4(d) and 4(e) show moments of orders 2 up to 6. Both are compared to Sabra moments in solid lines.

though we succeeded in including the modeling of phases in our closures, we still detect persistent difficulties in obtaining a tight fit of PDFs of real parts and energy flux through the final shells. However, the fact that these closures work on data alone, without any direct adjustment or learning towards phenomenological predictions, can be seen as a success.

Our closures performed similarly to the ones reported in Ref. [22]. Other works, such as Ref. [42], present quite sharp numerical results in their closure models, which are built using a deep neural network. It provides great insight and a powerful working tool, albeit lacking the information on small-scale universality and intrinsic stochasticity of the dynamics.

We note that the limited performance of our models can be related to the choice of the GMM density estimator, which has difficulties in reproducing the statistics at high dimensions. As presented in Appendix A, a wide variety of different quantities with different delays in time conditioning was tested, and those were reported in detail in Ref. [41]. The density estimation problems for closures lower in Table I are to be solved in considerably high dimension, which is surely a struggle for GMM. This cannot be ignored as a source of inaccuracy in numerical results. Other approaches or time conditioning on different quantities could be an important direction to follow. For example, one could improve from GMM to any other machine-learning tool that scales better as the dimension of the optimization problem grows [42,43].

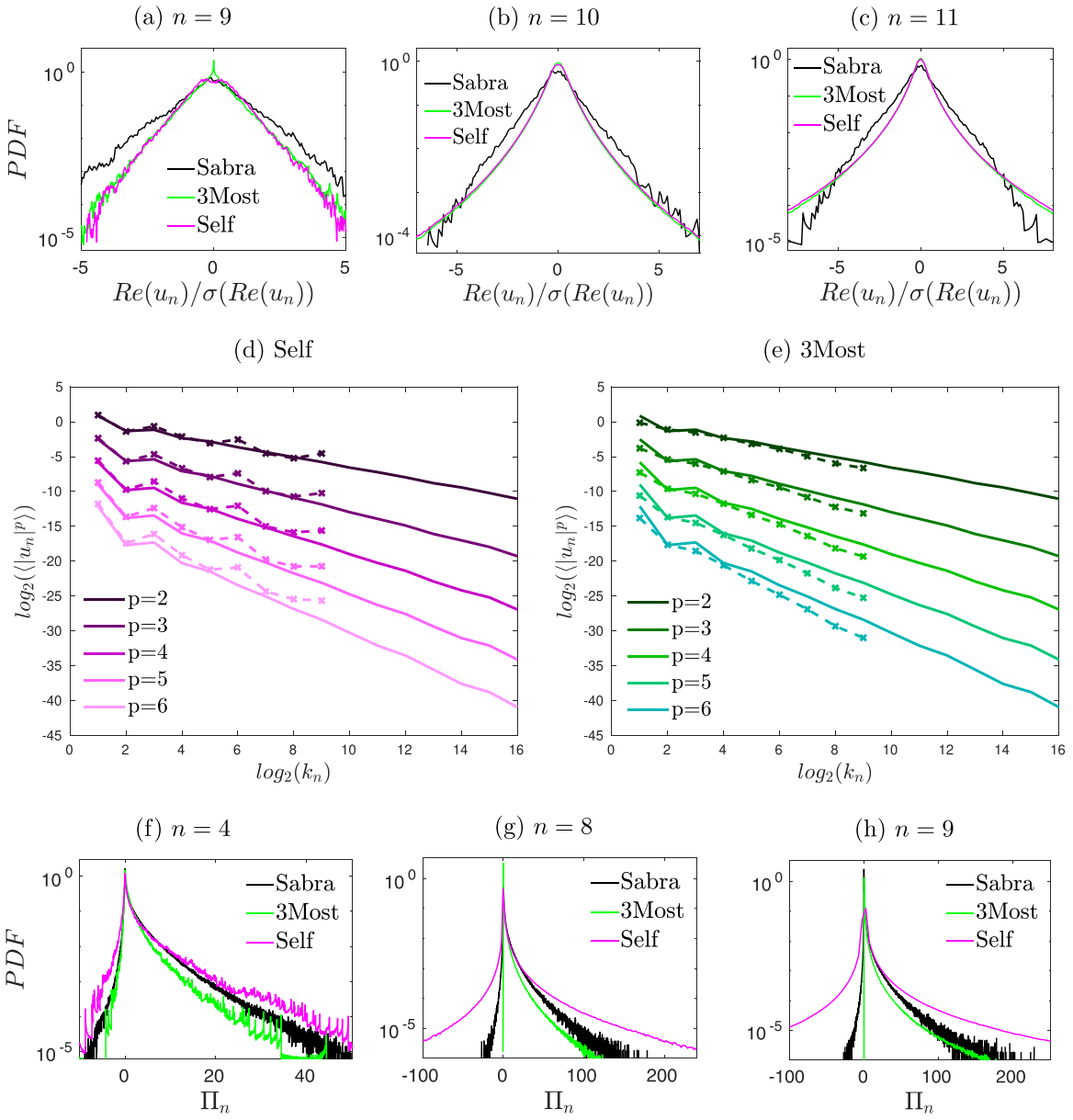

FIG. 5. Cutoff shell is $s = 9$. Figures 3(a)–3(c) present normalized PDFs of real parts for the closure with time conditioning to the three closes shells (3Most) and the closure with time conditioning to the closure variables (self), while Figs. 5(f)–3(h) show energy flux PDFs. All are compared to Sabra statistics in black. Figures 5(d) and 5(e) show moments of orders 2 up to 6. Both are compared to Sabra moments in solid lines.

The main contribution of this paper is developing a systematic framework for the closures based on the hidden scaling symmetry. The latter provides a universal scale-invariant statistical description both at present and past times, which is not possible in the standard formulation due to the intermittency [22]. This framework opens possibilities for employing machine-learning tools in closure problems for turbulence. We stress that the strategy we used is not limited to shell models but can potentially be generalized to the Navier-Stokes turbulence. There, a similar rescaling procedure was proposed, leading to the hidden scale invariance [28,33]. Such an extension would face natural challenges due to the much larger amount of information and the additional complexity that the rescaled representation introduces. In this respect, the similarity of the hidden symmetry formalism of Navier-Stokes turbulence to shell models provides a useful playground for new ideas.

ACKNOWLEDGMENTS

This work was supported by CNPq Grant 308721/2021-7 and FAPERJ Grants No. E-26/201.054/2022 and No. E-26/210.874/2014.

APPENDIX A: LIST OF ALL CLOSURES CONSIDERED

A wide variety of combinations between how many and in which shells we could conditionally sample the closure variables have been tested. Table I shows all closures tested, with the ones reported in Sec. VI in bold letters. The first

TABLE I. All closures written for this paper, with the ones in bold being the ones presented here.

	$\lvert\mathcal{U}_0\rvert$	$\lvert\mathcal{U}_1\rvert$	Δ_0	Δ_1	Conditioning
Kolmogorov	$\lvert\mathcal{U}_{-1}\rvert\lambda^{-1/3}$	$\lvert\mathcal{U}_0\rvert\lambda^{-1/3}$	$\pi/2$	$\pi/2$	×
Half closure	2^{z_0}	$\lvert\mathcal{U}_0\rvert\lambda^{-1/3}$	$\pi/2$	$\pi/2$	×
Joint	2^{z_0}	2^{z_1}	$\pi/2$	$\pi/2$	×
Simple cond	2^{z_0}	$\lvert\mathcal{U}_0\rvert\lambda^{-1/3}$	$\pi/2$	$\pi/2$	$\lvert\mathcal{U}_{-1}\rvert$ at $\tau-\Delta\tau$
Joint cond	2^{z_0}	2^{z_1}	$\pi/2$	$\pi/2$	$\lvert\mathcal{U}_{-1}\rvert$ at $\tau-\Delta\tau$
3-Most	2^{z_0}	2^{z_1}	$\pi/2$	$\pi/2$	$\lvert\mathcal{U}_{-3}\rvert$, $\lvert\mathcal{U}_{-2}\rvert$, $\lvert\mathcal{U}_{-1}\rvert$ at $\tau-\Delta\tau$
3-Most 9	2^{z_0}	2^{z_1}	$\pi/2$	$\pi/2$	$\lvert\mathcal{U}_{-3}\rvert$, $\lvert\mathcal{U}_{-2}\rvert$, $\lvert\mathcal{U}_{-1}\rvert$ at $\tau-\Delta\tau$
Long	2^{z_0}	2^{z_1}	$\pi/2$	$\pi/2$	$\lvert\mathcal{U}_{-s}\rvert,\ldots,\lvert\mathcal{U}_{-1}\rvert$ at $\tau-\Delta\tau$
Joint phases	2^{z_0}	2^{z_2}	z_1	z_3	×
Joint cond phases	2^{z_0}	2^{z_2}	z_1	z_3	$\lvert\mathcal{U}_{-1}\rvert$, $\Delta-1$ at $\tau-\Delta\tau$
Self	2^{z_0}	2^{z_2}	z_1	z_3	$\lvert\mathcal{U}_0\rvert$, $\lvert\mathcal{U}_1\rvert$, Δ_0, Δ_1 at $\tau-\Delta\tau$
Self 9	2^{z_0}	2^{z_2}	z_1	z_3	$\lvert\mathcal{U}_0\rvert$, $\lvert\mathcal{U}_1\rvert$, Δ_0, Δ_1 at $\tau-\Delta\tau$
Global	2^{z_0}	2^{z_2}	z_1	z_3	$\lvert\mathcal{U}_{-2}\rvert$, $\lvert\mathcal{U}_{-1}\rvert$, Δ_{-2}, Δ_{-1} at τ $\lvert\mathcal{U}_{-2}\rvert$, $\lvert\mathcal{U}_{-1}\rvert$, $\lvert\mathcal{U}_0\rvert$, $\lvert\mathcal{U}_1\rvert$, Δ_{-2}, Δ_{-1}, Δ_0, Δ_1 at $\tau-\Delta\tau$
Two times	2^{z_0}	2^{z_2}	z_1	z_3	$\lvert\mathcal{U}_{-2}\rvert$, $\lvert\mathcal{U}_{-1}\rvert$, Δ_{-2}, Δ_{-1} at τ $\lvert\mathcal{U}_{-2}\rvert$, $\lvert\mathcal{U}_{-1}\rvert$, $\lvert\mathcal{U}_0\rvert$, $\lvert\mathcal{U}_1\rvert$, Δ_{-2}, Δ_{-1}, Δ_0, Δ_1 at $\tau-\Delta\tau$ $\lvert\mathcal{U}_{-2}\rvert$, $\lvert\mathcal{U}_{-1}\rvert$, $\lvert\mathcal{U}_0\rvert$, $\lvert\mathcal{U}_1\rvert$, Δ_{-2}, Δ_{-1}, Δ_0, Δ_1 at $\tau-2\Delta\tau$

line in the table refers to Kolmogorov's closure [22]. It is the deterministic closure given by the following expressions:

$$\mathcal{U}_0 = |\mathcal{U}_{-1}|\lambda^{-1/3}e^{i(\frac{\pi}{2}+\alpha_{-1}+\alpha_{-2})}, \tag{A1}$$

$$\mathcal{U}_1 = |\mathcal{U}_0|\lambda^{-1/3}e^{i(\frac{\pi}{2}+\alpha_0+\alpha_{-1})}. \tag{A2}$$

This closure is also used to cover the initial time steps in simulations that require time conditioning. For more details about the performance of closures from Table I we refer to the Ph.D. thesis in Ref. [41]. Table II shows the closures reported in Secs. VI and VII to clear the fact that they can be grouped in more than one way, referring to whether they include time conditioning or not and whether they include phase modeling or not.

APPENDIX B: THE RESCALED SYSTEM

Here we derive the equations of motion in terms of intrinsic time τ for the rescaled variables $\mathcal{U}_N$. From the definition of τ in Eq. (3), we have

$$\frac{d\mathcal{U}_N}{d\tau} = \frac{d\mathcal{U}_N}{dt}\frac{dt}{d\tau}. \tag{B1}$$

The first factor can be computed deriving Eq. (4) with respect to t, finding

$$\frac{d\mathcal{U}_N}{dt} = \frac{dT_m}{dt}k_m u_{N+m} + T_m k_m \frac{du_{N+m}}{dt}. \tag{B2}$$

TABLE II. The closures we presented in this paper.

	Includes phases	Only modules
Conditioning	Self, Self 9	3-Most, 3-Most 9
No conditioning	Joint phases	Joint

and the second can be computed rewriting Eq. (3) as

$$\frac{dt}{d\tau} = T_m. \tag{B3}$$

Equation (B1) then becomes

$$\frac{d\mathcal{U}_N}{d\tau} = \left(\frac{dT_m}{dt}k_m u_{N+m} + T_m k_m \frac{du_{N+m}}{dt}\right)T_m. \tag{B4}$$

We compute the derivative of T_m with respect to t as

$$\frac{dT_m}{dt} = -\frac{1}{2}\left(k_0^2U^2 + \sum_{n<m}k_n^2|u_n|^2\right)^{-3/2}\sum_{n<m}k_n^2 2\,\mathrm{Re}\left(u_n^*\frac{du_n}{dt}\right). \tag{B5}$$

Using Eqs. (2) and (1) in the above expression, we find

$$\frac{dT_m}{dt} = -T_m^3\sum_{n<m}k_n^2\,\mathrm{Re}\left(u_n^*\left[i\left(k_{n+1}u_{n+2}u_{n+1}^* - \frac{1}{2}k_n u_{n+1}u_{n-1}^* + \frac{1}{2}k_{n-1}u_{n-1}u_{n-2}\right) - \nu k_n^2 u_n + f_n\right]\right), \tag{B6}$$

which is the same as

$$\frac{dT_m}{dt} = -\sum_{n<m}\mathrm{Re}\left(i\left(k_n T_m u_n^* k_n T_m u_{n+2}k_{n+1}T_m u_{n+1}^* - \frac{1}{2}k_n T_m u_n^* k_n T_m u_{n+1}k_n T_m u_{n-1}^* + \frac{1}{2}k_n T_m u_n^* k_{n-1}T_m u_{n-1}k_n T_m u_{n-2}\right) - \nu k_n^4 T_m^3 u_n^* u_n + k_n^2 T_m^3 u_n^* f_n\right). \tag{B7}$$

A direct manipulation, noting that $k_n = k_{n-m}k_m$, $k_1 = 2$, $k_{-1} = 1/2$, rewriting the indexes as $N = n - m$ and recalling that $n < m$ is the same as $N < 0$, yields

$$\frac{dT_m}{dt} = -\sum_{N<0} \mathrm{Re}\left(i\left(k_N^3 2\mathcal{U}_N^*\mathcal{U}_{N+2}\mathcal{U}_{N+1}^* - \frac{1}{2}k_N^3\mathcal{U}_N^*\mathcal{U}_{N+1}\mathcal{U}_{N-1}^* + \frac{1}{4}k_N^3\mathcal{U}_N^*\mathcal{U}_{N-1}\mathcal{U}_{N-2}\right) - \nu k_{N+m}^2 k_N^2 T_m|\mathcal{U}_N|^2 + k_{N+m}k_N T_m^2\mathcal{U}_N^* f_{N+m}\right), \quad \text{(B8)}$$

which can be further reduced to the form

$$\frac{dT_m}{dt} = \sum_{N<0} k_N^3 \,\mathrm{Im}\left(2\mathcal{U}_N^*\mathcal{U}_{N+1}^*\mathcal{U}_{N+2} - \frac{1}{2}\mathcal{U}_{N-1}^*\mathcal{U}_N^*\mathcal{U}_{N+1} - \frac{1}{4}\mathcal{U}_{N-1}^*\mathcal{U}_N\mathcal{U}_{N-2}^*\right) + \nu T_m \sum_{N<0} k_{N+m}^2 k_N^2|\mathcal{U}_N|^2 - T_m^2\sum_{N<0} k_{N+m}k_N\,\mathrm{Re}\,(\mathcal{U}_N^* f_{N+m}). \quad \text{(B9)}$$

Splitting the sum above in several pieces, we write

$$\frac{dT_m}{dt} = \xi + \xi_\nu + \xi_f = \xi_{\text{total}}, \quad \text{(6)}$$

$$\xi = \sum_{N<0} k_N^3 \,\mathrm{Im}\left(2\mathcal{U}_N^*\mathcal{U}_{N+1}^*\mathcal{U}_{N+2} - \frac{1}{2}\mathcal{U}_{N-1}^*\mathcal{U}_N^*\mathcal{U}_{N+1} - \frac{1}{4}\mathcal{U}_{N-1}^*\mathcal{U}_N\mathcal{U}_{N-2}^*\right), \quad \text{(7)}$$

$$\xi_\nu = \nu T_m k_m^2 \sum_{N<0} k_N^4|\mathcal{U}_N|^2, \quad \text{(8)}$$

$$\xi_f = -T_m^2\sum_{N<0} k_{N+m}k_N\,\mathrm{Re}\,(\mathcal{U}_N^* f_{N+m}). \quad \text{(9)}$$

We can now go back to computing the main system, plugging Eqs. (6)–(9) back into Eq. (B4). This yields

$$\frac{d\mathcal{U}_N}{d\tau} = \left(\xi_{\text{total}}k_m u_{N+m} + T_m k_m\frac{du_{N+m}}{dt}\right)T_m. \quad \text{(B10)}$$

We then use Eq. (1), with $b = -1/2$ and $c = 1/2$, to write

$$\frac{d\mathcal{U}_N}{d\tau} = \xi_{\text{total}}T_m k_m u_{N+m} + T_m^2 k_m\left(i\left(k_{N+m+1}u_{N+m+2}u_{N+m+1}^* - \frac{1}{2}k_{N+m}u_{N+m+1}u_{N+m-1}^* + \frac{1}{2}k_{N+m-1}u_{N+m-1}u_{N+m-2}\right) - \nu k_{N+m}^2 u_{N+m} + f_{N+m}\right), \quad \text{(B11)}$$

which becomes, doing the change of variables from Eq. (4) again:

$$\frac{d\mathcal{U}_N}{d\tau} = \xi_{\text{total}}\mathcal{U}_N + i\left(k_{N+1}\mathcal{U}_{N+2}\mathcal{U}_{N+1}^* - \frac{1}{2}k_N\mathcal{U}_{N+1}\mathcal{U}_{N-1}^* + \frac{1}{2}k_{N-1}\mathcal{U}_{N-1}\mathcal{U}_{N-2}\right) - \nu k_{N+m}^2 T_m\mathcal{U}_N + T_m^2 k_m f_{N+m}. \quad \text{(B12)}$$

Reorganizing the terms, we have

$$\frac{d\mathcal{U}_N}{d\tau} = i\left(k_{N+1}\mathcal{U}_{N+2}\mathcal{U}_{N+1}^* - \frac{1}{2}k_N\mathcal{U}_{N+1}\mathcal{U}_{N-1}^* + \frac{1}{2}k_{N-1}\mathcal{U}_{N-1}\mathcal{U}_{N-2}\right) + \left(\xi_{\text{total}} - \nu k_{N+m}^2 T_m\right)\mathcal{U}_N + T_m^2 k_m f_{N+m}. \quad \text{(5)}$$

To write this system with no dependence on t, we must also write T_m as a function of $\mathcal{U}$ instead of u. Recall the expression for T_m in Eq. (2), expand the sum, and apply the change of variables in Eq. (4):

$$T_m = \left(k_0^2U^2 + \frac{k_1^2}{k_m^2T_m^2}|\mathcal{U}_{1-m}|^2 + \frac{k_2^2}{k_m^2T_m^2}|\mathcal{U}_{2-m}|^2 + \cdots + \frac{k_{m-1}^2}{k_m^2T_m^2}|\mathcal{U}_{-1}|^2\right)^{-1/2}. \quad \text{(B13)}$$

Regrouping the sum,

$$\frac{1}{T_m^2} = k_0^2U^2 + \frac{1}{T_m^2}\sum_{N<0} k_N^2|\mathcal{U}_N|^2. \quad \text{(B14)}$$

It then follows that

$$T_m = \frac{1}{k_0U}\left(1 - \sum_{N<0} k_N^2|\mathcal{U}_N|^2\right)^{1/2}. \quad \text{(10)}$$

[1] O. Reynolds, Philos. Trans. R. Soc. London A, 123 (1895).
[2] J. Smagorinsky, Monthly Weather Review **91**, 99 (1963).
[3] S. B. Pope, *Turbulent Flows* (Cambridge University Press, Cambridge, 2000).
[4] N. Chen, A. Farhat, and E. Lunasin, Physica D **443**, 133552 (2023).
[5] F. J. Alexander, G. Johnson, G. L. Eyink, and I. G. Kevrekidis, Phys. Rev. E **77**, 026701 (2008).
[6] Y. Li, L. Chevillard, G. Eyink, and C. Meneveau, Phys. Rev. E **79**, 016305 (2009).
[7] U. Frisch, *Turbulence: The Legacy of A. N. Kolmogorov* (Cambridge University Press, Cambridge, 1995).
[8] E. N. Lorenz, Tellus A **21**, 289 (1969).
[9] C. E. Leith and R. H. Kraichnan, J. Atmos. Sci. **29**, 1041 (1972).
[10] D. Ruelle, Phys. Lett. A **72**, 81 (1979).
[11] G. L. Eyink, J. Stat. Phys. **83**, 955 (1996).
[12] G. Boffetta, M. Cencini, M. Falcioni, and A. Vulpiani, Phys. Rep. **356**, 367 (2002).
[13] T. N. Palmer, A. Döring, and G. Seregin, Nonlinearity **27**, R123 (2014).

[14] A. A. Mailybaev, Nonlinearity **29**, 2238 (2016).
[15] A. A. Mailybaev, Nonlinearity **30**, 2466 (2017).
[16] L. Biferale, G. Boffetta, A. A. Mailybaev, and A. Scagliarini, Phys. Rev. Fluids **3**, 092601(R) (2018).
[17] S. Thalabard, J. Bec, and A. Mailybaev, Commun. Phys. **3**, 122 (2020).
[18] D. Bandak, A. A. Mailybaev, G. L. Eyink, and N. Goldenfeld, arXiv:2401.13881 [Phys. Rev. Lett. (to be published)].
[19] E. B. Gledzer, Soviet Physics Doklady **18**, 216 (1973).
[20] M. Yamada and K. Ohkitani, Phys. Rev. Lett. **60**, 983 (1988).
[21] L. Biferale, Annu. Rev. Fluid Mech. **35**, 441 (2003).
[22] L. Biferale, A. A. Mailybaev, and G. Parisi, Phys. Rev. E **95**, 043108 (2017).
[23] A. N. Kolmogorov, J. Fluid Mech. **13**, 82 (1962).
[24] R. Benzi, L. Biferale, and G. Parisi, Physica D **65**, 163 (1993).
[25] G. Eyink, S. Chen, and Q. Chen, J. Stat. Phys. **113**, 719 (2003).
[26] V. S. L'vov, E. Podivilov, A. Pomyalov, I. Procaccia, and D. Vandembroucq, Phys. Rev. E **58**, 1811 (1998).
[27] A. A. Mailybaev, Phys. Rev. Fluids **6**, L012601 (2021).
[28] A. A. Mailybaev and S. Thalabard, Philos. Trans. R. Soc. A. **380**, 20210098 (2022).
[29] C. M. Bishop, *Pattern Recognition and Machine Learning*, 1st ed., Information Science and Statistics (Springer, New York, 2006).
[30] U. Frisch and M. Vergassola, Europhys. Lett. **14**, 439 (1991).
[31] A. A. Mailybaev, Phys. Rev. Fluids **8**, 054605 (2023).
[32] Q. Chen, S. Chen, G. L. Eyink, and K. R. Sreenivasan, Phys. Rev. Lett. **90**, 254501 (2003).
[33] A. A. Mailybaev, Nonlinearity **35**, 3630 (2022).
[34] A. P. Dempster, N. M. Laird, and D. B. Rubin, J. R. Stat. Soc. Series B: **39**, 1 (1977).
[35] L. Xu and M. I. Jordan, Neural Comput. **8**, 129 (1996).
[36] J. V. Psutka and J. Psutka, Pattern Recognition **91**, 25 (2019).
[37] R. Bellman, *Dynamic Programming* (Princeton University Press, Princeton, 1972).
[38] C. C. Aggarwal, A. Hinneburg, and D. A. Keim, in *Database Theory — ICDT 2001*, edited by J. Van den Bussche and V. Vianu (Springer, Berlin, 2001), pp. 420–434.
[39] M. L. Eaton, *Multivariate Statistics: A Vector Space Approach* (Institute of Mathematical Statistic, Beachwood, 2007).
[40] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems* (Springer, Berlin, 1996).
[41] J. D. Lemos, Data-based approach for time-correlated closures of turbulence models, Ph.D. thesis, National Institute for Pure and Applied Mathematics, 2022.
[42] G. Ortali, A. Corbetta, G. Rozza, and F. Toschi, Phys. Rev. Fluids **7**, L082401 (2022).
[43] S. Hochreiter and J. Schmidhuber, Neural Comput. **9**, 1735 (1997).