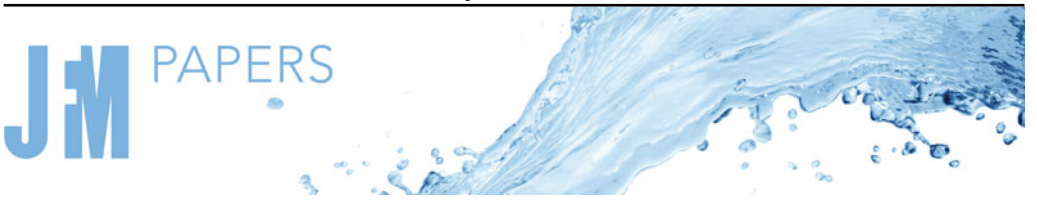
J. Fluid Mech. (2026), *vol.* 1027, A39, doi:10.1017/jfm.2025.11087



Fluctuations around turbulence models

Flavio Tuteri[1], **Alexandros Alexakis**[1] and **Sergio Chibbaro**[2]

[1]Laboratoire de Physique de l'Ecole normale supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université Paris Cité, F-75005 Paris, France

[2]Laboratoire Interdisciplinaire des Sciences du Numérique, LISN, CNRS, CentraleSupélec, Inria, Université Paris-Saclay, F-91405 Orsay, France

Corresponding author: Flavio Tuteri, flavio.tuteri@phys.ens.psl.eu

(Received 25 July 2025; revised 29 October 2025; accepted 15 December 2025)

Numerical simulations of turbulent flows at realistic Reynolds numbers generally rely on filtering out small scales from the Navier–Stokes equations and modelling their impact through the subgrid-scale stress tensor τ_{ij}. Traditional models approximate τ_{ij} solely as a function of the filtered velocity gradient, leading to deterministic subgrid-scale closures. However, small-scale fluctuations can locally exhibit instantaneous values whose deviation from the mean can have a significant influence on the flow dynamics. In this work, we investigate these effects by employing direct numerical simulations combined with Gaussian filtering to quantify subgrid-scale effects and evaluating the local energy flux in both space and time. The mean performance of the canonical Clark model is assessed by conditioning the energy flux distributions on the invariants of the filtered velocity gradient tensor, Q and R. The Clark model captures to a good degree the mean energy flux. However, the fluctuations around these mean values for given (Q, R) are of the order of the mean, displaying fat-tailed distributions. To be more precise, we examine the joint distributions of true energy flux and the predictions from both the Clark and the Smagorinsky models. This approach mirrors the strategy adopted in early stochastic subgrid-scale models. Clear non-Gaussian characteristics emerge from the obtained distributions, particularly through the appearance of heavy tails. The mean, the variance, the skewness and the flatness of these distributions are quantified. Our results emphasise that fluctuations are an integral component of the small-scale feedback onto the large-scale dynamics and should be incorporated into subgrid-scale modelling through an appropriate stochastic framework.

Key words: isotropic turbulence, turbulence modelling, turbulence simulation

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/jfm.2025.11087 Published online by Cambridge University Press

1. Introduction

Turbulent flows are the fundamental basis of many applications in engineering (Pope 2001), geophysics (Thorpe 2007; Gill 2016) and astrophysics (Biskamp 2003) among others. Despite the considerable improvement in computational power, performing direct numerical simulations (DNS) that resolve all active scales in realistic turbulent fluids remains out of reach for most industrial and natural flows (Moin & Mahesh 1998; Pope 2001). This is related to the fact that the number of degrees of freedom grows fast with the size of the system. To circumvent this prohibitive cost, large-eddy simulations (LES) are widely adopted in engineering and planetary atmospheric models. The key idea is to take the evolution only down to a specified filtering scale, with the influence of smaller, unresolved subfilter motions taken into account through explicit modelling (Smagorinsky 1963; Deardorff 1970; Germano 1992; Meneveau & Katz 2000; Sagaut 2005). Because the equations remain unclosed, due to the nonlinear Navier–Stokes operator coupling all scales, empirical models grounded in turbulence theory are employed. Many of these models account for the influence of the small-scale motions through the subgrid-scale stress tensor τ_{ij}, which captures the action of unresolved scales on the resolved ones. Typically, τ_{ij} is expressed as a function of the large-scale velocity gradients, consistent with the symmetries of the flow, in order to provide a deterministic model closure for turbulence (Meneveau & Katz 2000). The models by Smagorinsky (1963) and Clark, Ferziger & Reynolds (1979) are some of the most well-known ones. Being quasi-empirical, these models require extensive testing and validation through experiments and simulations (Piomelli *et al.* 1991; Domaradzki, Liu & Brachet 1993; Härtel *et al.* 1994; Liu, Meneveau & Katz 1994). These models have been found to reasonably reproduce even some intermittency measures Linkmann, Buzzicotti & Biferale (2024) and statistical symmetries Magacho *et al.* (2024), however, more severe tests reveal often significant discrepancies between modelled and measured stress components (Meneveau & Katz 2000; Moser, Haering & Yalla 2021). In particular, looking at the full turbulent cascade process, the eddy-viscosity approximation may lead to oversimplifying the picture with large errors in fluctuations. In previous works (Borue & Orszag 1998; Tsinober 2009; Alexakis & Biferale 2018; Alexakis & Chibbaro 2020), the local energy flux transferred from unfiltered to filtered scales was investigated. The mean was found to be consistent with the predictions of the simple Smagorinsky model, however, significant deviations were observed. It is possible that, since these deviations originate from the small, unresolved scales, their impact on the resolved scales could be captured using stochastic modelling. While stochastic subgrid models have been proposed in previous studies (Mason & Thomson 1992; Gicquel *et al.* 2002; Carati, Winckelmans & Jeanmart 2007; Marstorp, Brethouwer & Johansson 2007), the statistical properties of the associated noise terms were typically prescribed in an *ad hoc* fashion, and notably are not meant to capture extreme events. The need for an accurate stochastic description of the small scale effects is emphasised by recent results proposing that the Navier–Stokes equations can develop randomness at large scales in a finite time (Eyink & Bandak 2020; Thalabard, Bec & Mailybaev 2020; Bandak *et al.* 2024).

In the present work, we aim to quantify the fluctuations around the mean subgrid-scale energy flux and compare these with the predictions of classical models. To this end, we employ data from DNS of the Navier–Stokes equations and follow two complementary approaches: (i) the analysis of several conditional means, as suggested in previous works (Buaria, Pumir & Bodenschatz 2020; Johnson & Wilczek 2024); notably we shall analyse probabilities conditioned on QR, the invariants of the velocity gradient tensor, which have not been investigated so far. (ii) The joint probability density function (PDF) of

https://doi.org/10.1017/jfm.2025.11087 Published online by Cambridge University Press

model predictions and ground truth, as in Alexakis & Chibbaro (2020), to examine the distribution of subgrid-scale fluxes conditioned on a given model. The ultimate goal is to guide and constrain the development of improved stochastic subgrid-scale models for turbulent flow simulations. To this end, we restrict our analysis of fluctuations to the simplest, and most used deterministic closures, namely the Smagorinsky and Clark models. Although it should be possible in principle to refine the accuracy of such closures by including higher-order velocity gradients in the expansion of the subgrid-scale stress tensor (Carati, Winckelmans & Jeanmart 2001; Eyink 2006; Chang, Yuan & Wang 2022), we deliberately adopt simple forms in order to emphasise the stochastic component, and to provide insights for further practical models. Moreover, with regard to possible future machine learning approaches, the key point is to provide a well-conditioned framework for learning. Whether more accurate deterministic closures are required to achieve this is beyond the scope of the present work.

2. Theoretical background

The stage is set by the incompressible Navier–Stokes equations, governing the evolution of u, divergence-free velocity field,

$$\left[\partial_t - \nu\partial_{jj}\right]u_i + \partial_j\left(u_j u_i\right) + \partial_i p = f_i. \tag{2.1}$$

Here, ν is the kinematic viscosity and p is the pressure computed through incompressibility. An external forcing term is added on the right-hand side, acting at the largest scales of the system. This forcing is statistically homogeneous and injects energy into the flow at a mean rate ϵ. The domain is a triply periodic cubic box of size $2\pi L$, which ensures the absence of boundary-layer effects. We focus on the high Reynolds number regime, with $Re \equiv \epsilon^{1/3}L^{4/3}/\nu$, so that a wide separation of scales emerges between the forcing scale L and the viscous dissipation scale $\eta \equiv \nu^{3/4}\epsilon^{-1/4}$.

2.1. *Filtering approach*

A typical starting point for turbulence modelling is to separate the components describing variations at large and small scales, although their dynamics remain irrevocably coupled because of the nonlinear term. Following Germano (1992) we work with the coarse graining at scale l of $\boldsymbol{u}(\boldsymbol{x}, t)$, the solution of the three-dimensional incompressible Navier–Stokes problem (2.1), obtained by suppressing high-frequency fluctuations through local averaging

$$\overline{\boldsymbol{u}}^l(\boldsymbol{x}, t) = \int G^l(\boldsymbol{r})\boldsymbol{u}(\boldsymbol{x} - \boldsymbol{r}, t)\mathrm{d}\boldsymbol{r}, \tag{2.2}$$

where $G^l(\boldsymbol{r}) = l^{-3}G(\boldsymbol{r}/l)$ is a localised smooth filtering function such that $\int G(\boldsymbol{r})\mathrm{d}\boldsymbol{r} = 1$. In particular, we use the Gaussian filter, given in Fourier space by

$$\widehat{G}(\boldsymbol{k}) = \exp\left[-\frac{k^2}{2}\right]. \tag{2.3}$$

The dynamics of the resulting large-scale field is not closed because of the nonlinearity

$$\left[\partial_t - \nu\partial_{jj}\right]\overline{u_i}^l + \left[\delta_{ki} - \partial_{ki}\nabla^{-2}\right]\partial_j\left(\overline{u_j}^l\overline{u_k}^l\right) = \overline{f_i}^l - \partial_j\tau^l(u_i, u_j). \tag{2.4}$$

On the left-hand side, the standard Navier–Stokes operator acts on the coarse-grained field, while the right-hand side depends on the velocity field at all scales through the subgrid-scale tensor

$$\tau^l(u_i, u_j) = \overline{u_i u_j}^l - \overline{u_i}^l\overline{u_j}^l =: \tau_{ij}^l, \tag{2.5}$$

https://doi.org/10.1017/jfm.2025.11087 Published online by Cambridge University Press

the stress exerted on scales larger than l by the chaotic fluctuations at smaller scales. This tensor depends on both the filtered velocity field and the underlying small-scale fluctuations, and it determines the Galilean-invariant local energy transfer from the large-scale to the small-scale component

$$\Pi^l(\boldsymbol{x}, t) = -\tau_{ij}^l \partial_j \overline{u_i}^l. \tag{2.6}$$

We note that the value of local energy flux here depends on the choice of filter and technical differences between filters have been noted (Alexakis & Chibbaro 2020). An alternative (filter-independent) definition of the energy flux within the structure function framework has been explored (Yao *et al.* 2024). Here, we will refrain from examining further filters or alternative definitions and we will limit ourselves to the Gaussian filter. A model of turbulence can be obtained via the closure of (2.4) by expressing the unknown right-hand side in terms of the resolved large-scale field. Traditionally, the subgrid-scale stress tensor τ is modelled as a function of the first-order gradients of the filtered velocity. For example, the model by Clark *et al.* (1979)

$$\tau_{ij}^l \sim l^2 \partial_k \overline{u}_i^l \partial_k \overline{u}_j^l, \tag{2.7}$$

which may be obtained as a first term of an expansion of (2.5) in l (Borue & Orszag 1998; Carati *et al.* 2001; Eyink 2006; Johnson 2021). Another example is the traditional eddy-viscosity model by Smagorinsky (1963)

$$\tau_{ij}^l - \frac{\delta_{ij}}{3}\mathrm{Tr}[\tau^l] \sim -l^2 \sqrt{\overline{S}_{kh}^l \overline{S}_{hk}^l}\, \overline{S}_{ij}^l, \tag{2.8}$$

where S is the symmetric part of the velocity gradient tensor, the strain rate. Such models lead to a deterministic evolution of the filtered field.

2.2. *Velocity gradient invariants*

The tensor $\partial_i u_j = S_{ij} + \Omega_{ij}$ (where S_{ij} is the symmetric and Ω_{ij} the antisymmetric part of $\boldsymbol{\nabla u}$) encapsulates the spatial variations of the velocity field around a given point and offers key insights into the local flow topology, intermittency and the mechanisms driving the energy cascade. For isotropic turbulence, τ should depend only on the properties of $\boldsymbol{\nabla u}$ and higher-order derivatives that are invariant under rotations (Carati *et al.* 2001; Eyink 2006). In models, however, only information from the gradient tensor $\boldsymbol{\nabla u}$ is usually kept. In total, the gradient tensor is fully characterised by 5 scalar-invariant quantities, that are independent of the orientation of the coordinate system. This reflects the fact that the tensor $\boldsymbol{\nabla u}$ possesses five effective degrees of freedom, obtained by starting from its nine components and accounting for the incompressibility constraint and the assumption of isotropy (Meneveau 2011; Johnson & Wilczek 2024). To provide a synthetic description of the local fluid configurations without referring directly to the tensor components, all 5 invariants are in principle needed. However, here, we limit ourselves to Q and R, the invariants under reference frame rotations and reflections that characterise the spectrum of $\boldsymbol{\nabla u}$. In general, for a traceless real matrix M we will denote

$$Q_M = -\frac{1}{2}\mathrm{Tr}(M^2), \quad R_M = -\frac{1}{3}\mathrm{Tr}(M^3). \tag{2.9}$$

From now on, we simply refer to Q and R as invariants of the velocity gradient $\boldsymbol{\nabla u}$ (Chong, Perry & Cantwell 1990)

$$Q = Q_{\boldsymbol{\nabla u}} = -\frac{1}{2}\partial_i u_j \partial_j u_i, \quad R = R_{\boldsymbol{\nabla u}} = -\frac{1}{3}\partial_i u_j \partial_j u_k \partial_k u_i. \tag{2.10}$$

https://doi.org/10.1017/jfm.2025.11087 Published online by Cambridge University Press

Combining the multiscale description, the invariants of velocity gradient and strain rate are sufficient to express the energy fluxes predicted by the introduced models (2.7) and (2.8)

$$\Pi^l_{Clark} = l^2\big[4R_{\overline{S}^l} - R_{\overline{\nabla \boldsymbol{u}}^l}\big], \quad \Pi^l_{Smag.} = l^2\big[-Q_{\overline{S}^l}\big]^{3/2}. \tag{2.11}$$

A striking difference between the models is that Smagorinsky's prediction, unlike Clark's, is non-negative. It is important to note that knowledge of the two invariants of the sole velocity gradient tensor does not uniquely determine those of its symmetric part S, and therefore model predictions (2.11) are not uniquely defined in the (Q, R) space. Yet, to link the energy flux to the invariants Q, R has proven useful to get insights into the dynamical process (Johnson & Wilczek 2024), and also for the construction of LES models (Borue & Orszag 1998). Indeed, Q and R provide information on the local geometry of the flow: depending on their sign, the eigenvalue problem of $\boldsymbol{\nabla u}$ admits either two complex conjugate roots and one real root, positive or negative, or three real roots. Taking into account the incompressibility constraint, this leads to four distinct regions in the (Q, R) plane, where the upper-right quadrant is associated with vortex compression. For more details we refer readers to the literature (Chong *et al.* 1990; Meneveau 2011; Johnson & Wilczek 2024).

3. Results

The present study relies on a numerical solution of the Navier–Stokes equations (2.1). The pseudo-spectral code GHOST (Mininni *et al.* 2011) was employed, using a fourth-order Runge–Kutta scheme and adapting the time step according to the Courant–Friedrichs–Lewy condition. The code considers a periodic cubic box of size 2π with resolution 1024 in each direction. Viscosity was set to $\nu = 5 \times 10^{-4}$. In order to fix the mean energy injection rate ϵ, we imposed the forcing (as, for example, in Alexakis & Chibbaro 2020)

$$\widehat{\boldsymbol{f}}(\boldsymbol{k}, t) = \left[\frac{\epsilon}{\sum_{p\leqslant 2} |\widehat{\boldsymbol{u}}(\boldsymbol{p}, t)|^2} + i\varphi_{\boldsymbol{k}}\right] \widehat{\boldsymbol{u}}(\boldsymbol{k}, t), \tag{3.1}$$

where the phases $\varphi_{\boldsymbol{k}}$ are randomly chosen at initialisation and kept fixed in time. In particular, we put $\epsilon = 1$. After a transient, a steady state was reached, statistically consistent with the classical isotropic turbulence. Figure 1 shows, on the left, the energy spectrum, which provides a Kolmogorov–Obukhov $5/3$ spectrum over almost two decades. On the right, it displays the mean energy transfer Π across scale k, normalised by the mean energy injection rate. The actual inertial range of scales is identified by the region where the normalised transfer is close to 1. The blue line represents the spatial and temporal average of (2.6) using the sharp spectral filter (i.e. the Fourier formulation), while the red triangles correspond to the estimation based on Gaussian filtering. We select the wavenumber 16 as a representative point for the inertial range.

3.1. *Clark vs DNS*

3.1.1. *The QR diagrams*

To compare *a priori* predictions given by models of turbulence we propose to look at the conditional mean $\langle \Pi | Q, R \rangle$. This observable has been first used in a recent work (Vela-Martín & Jiménez 2021), yet in a different context. In figure 2 we consider Q and R normalised by the mean dissipation at the corresponding scale, and showing the mean flux for a given (Q, R) with the colour scale: positive values in red indicate a mean energy

https://doi.org/10.1017/jfm.2025.11087 Published online by Cambridge University Press

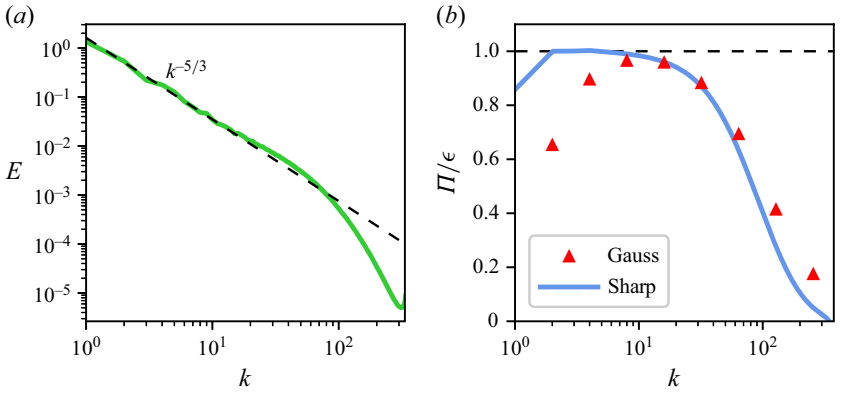


Figure 1. Kolmogorov energy spectrum $E(k)$ in (a) and fraction of the mean energy transfer across scales relative to the mean energy injection rate (b).

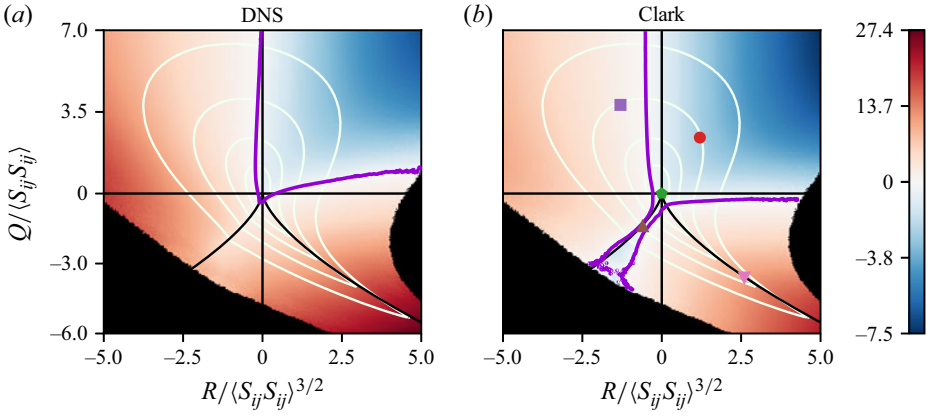


Figure 2. Colour scale represents the mean local energy flux conditioned on (Q, R) configurations at filtering scale $q = 16$ (corresponding to $l = 1/q$). Positive fluxes (forward cascade) are shown in red, and negative fluxes (backscatter) in blue. Black regions correspond to events with probability less than 10^{-8}. The (a) displays DNS results, while the (b) shows the Clark model prediction. White contours indicate isolines of the probability in the (Q, R) plane at levels 10^{-4}, 10^{-5}, 10^{-6} and 10^{-7}, moving outward from the origin. The purple curve denotes the isoline of zero flux. Coloured dots in the (b) indicate the (Q, R) locations used for the conditional analysis in figure 3.

transfer to smaller scales, while negative values in blue correspond to a net transfer toward larger scales. White contours indicate the isolines of the probability density in the (Q, R) space, corresponding to levels 10^{-4}, 10^{-5}, 10^{-6} and 10^{-7}, starting from the curve closest to the origin, which corresponds to the maximum of the distribution. We have chosen the observation window such that the probability isoline corresponding to 10^{-7} is largely included, and for events with probability below 10^{-8} we have not shown the mean energy flux (black regions) because of the lack of examples to compute the averages. The violet curve marks the isoline where the mean flux vanishes. The left panel presents the results obtained from (2.6), whereas the right panel shows the prediction of the Clark model (2.7), calculated using the complete DNS fields. First, it is noteworthy that, in the DNS, the backward cascade (backscatter) is confined to the upper-right quadrant, associated with vortex compression configurations. Then, the Clark model successfully captures the

https://doi.org/10.1017/jfm.2025.11087 Published online by Cambridge University Press

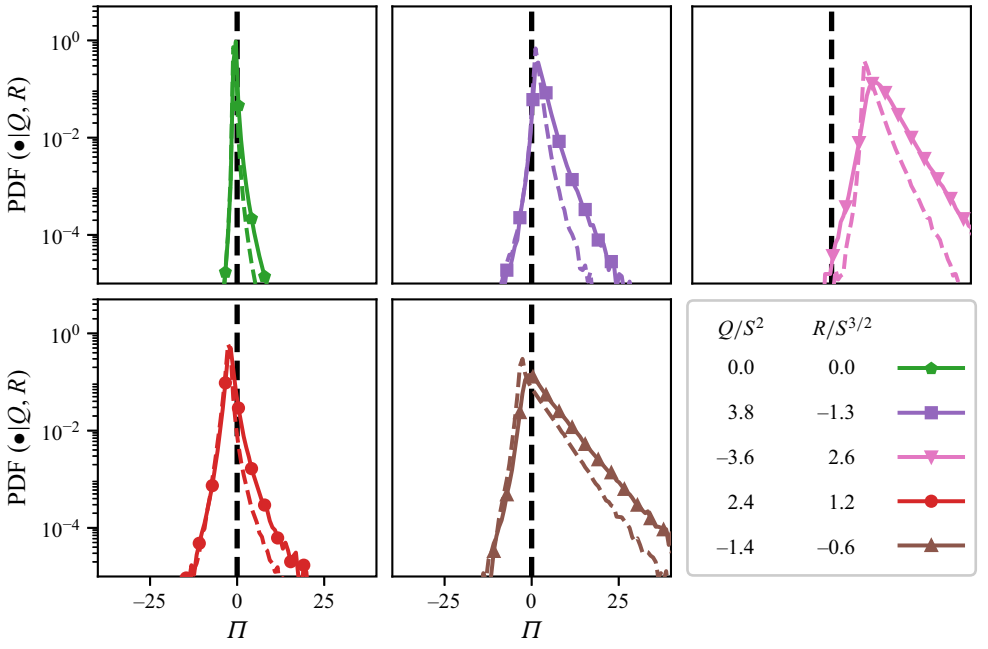


Figure 3. Comparison of local energy flux PDFs conditioned on (Q, R) configurations, between ground-truth DNS (solid lines) and Clark model predictions (dashed lines). The corresponding (Q, R) locations are indicated in figure 2.

qualitative distribution of the mean energy flux across the (Q, R) diagram. However, quantitatively, it underestimates the right value. It is possible to see that a region with small backscatter flow is observable in the lower-left quadrant, as highlighted by the violet curve. We can discuss a little more the error made by the Clark model from a physical point of view. As discussed in previous works (Eyink 2006), the turbulent subgrid stress may be written as an infinite series of terms in which the Clark model represents the first term. In particular, for a Gaussian filter as the one used here, it may be written (Johnson 2021)

$$\tau_{ij}^{l} \sim l^2 \partial_k \overline{u}_i^{l} \partial_k \overline{u}_j^{l} + \int_0^{l^2} \mathrm{d}\theta \left(\overline{\overline{\nabla_i \boldsymbol{u}_k}^{\sqrt{\theta}} \overline{\nabla_j \boldsymbol{u}_k}^{\sqrt{\theta}}}^{\phi} - \overline{\overline{\nabla_i \boldsymbol{u}_k}^{\sqrt{\theta}}}^{\phi} \overline{\overline{\nabla_j \boldsymbol{u}_k}^{\sqrt{\theta}}}^{\phi} \right), \tag{3.2}$$

where $\phi = \sqrt{l^2 - \theta}$. The first term is scale local and corresponds to the Clark model, which therefore neglects the higher-order non-local terms involving scales smaller than the filter one. The difference between the Clark model and DNS in the energy flux is related to this term. We emphasise that the discussion is made at the filter scale $q = 16$, that is in the inertial range. We can see that the local-in-scale model globally underestimates the mean. In figure 3, we move beyond the analysis of the mean and examine the distribution of the Clark prediction versus the ground-truth energy flux conditioned on five representative points in the (Q, R)-space: the origin and one point in each of the four quadrants, highlighted in figure 2. This assessment is severe, but it permits us to fully recognise the capability to reproduce the cascade process by a model. This type of conditioning is particularly interesting since it corresponds to imposing a constraint on the large-scale flow topology. The solid lines show the conditioned PDF for the measured energy flux directly from the DNS, while the dashed lines give the conditioned PDFs for the predicted energy

https://doi.org/10.1017/jfm.2025.11087 Published online by Cambridge University Press

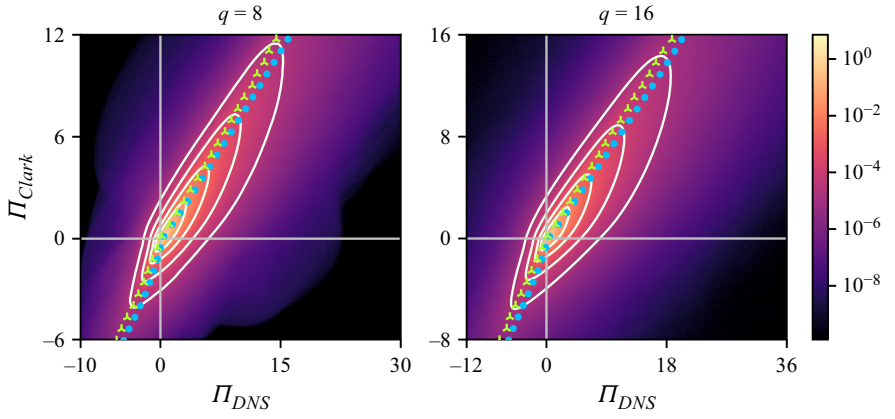


Figure 4. Colour scale represents the joint PDFs of (2.6) and (2.7). White contours indicate probability isolines at levels 10^{-4}, 10^{-5}, 10^{-6} and 10^{-7}, progressing outward from the origin. Green upward tripod markers denote the peaks of the conditional PDFs obtained by fixing the model prediction (horizontal cuts), while blue circles represent the corresponding mean values.

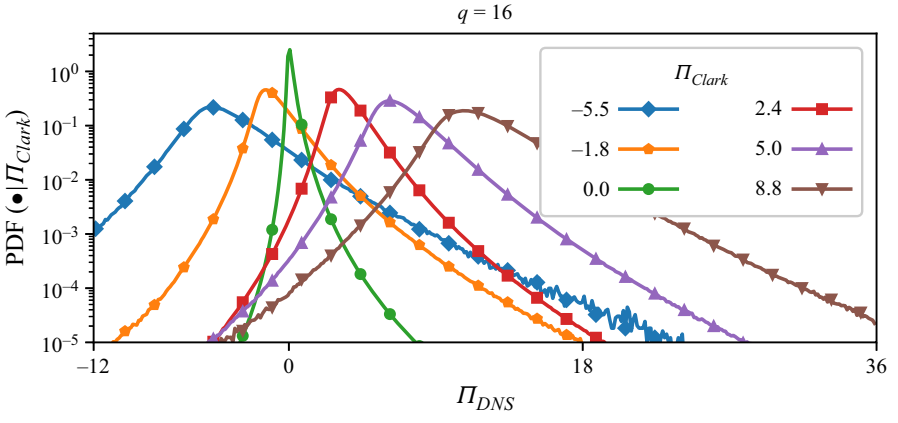


Figure 5. The PDFs of the (2.6) energy flux, conditioned on values of the Clark estimation (2.7).

flux from the Clark model. We recall that Q, R do not uniquely determine the energy flux in the Clark model and that, for the same Q, R, a distribution of energy fluxes is observed. Globally speaking, the Clark model makes a decent job overall. The essential features of the flux are captured, and in particular we observe a good agreement in the left/backscatter tail. Nonetheless, the skewness is not correct and the right/positive tail is systematically underestimated. The shape of the PDF appears similar, but a more careful look at the right tails shows that the scaling is also different. Overall, stronger fluctuations are observed in the DNS than the Clark model predictions. Considering the decomposition in scales (3.2), the non-local terms appear important in the positive tails of the flux distributions.

3.1.2. *Joint PDFs*

To provide a more precise characterisation of fluctuations, we introduce the joint probability density of the local energy flux (2.6) and the Clark model prediction (2.7). This joint PDF serves as the appropriate observable to assess the fluctuations around the model: for a fixed value of the Clark prediction, the complete flux exhibits a distribution of

https://doi.org/10.1017/jfm.2025.11087 Published online by Cambridge University Press

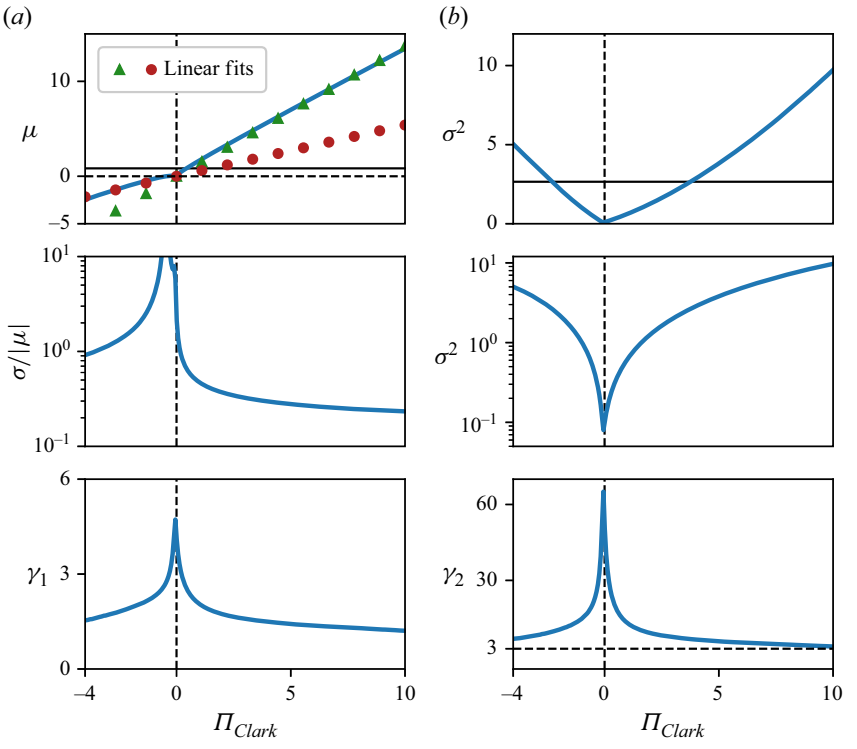


Figure 6. First and third rows: statistical cumulants (mean, variance, skewness and flatness) of the DNS density, computed conditionally on a given value of the Clark flux. Horizontal solid lines indicate the unconditional values. Middle row: the conditional variance shown in logarithmic scale (*b*, to highlight variability) and the relative statistical error $\sigma/|\mu|$ (*a*). In the top-left panel (mean), two linear fits with different slopes, respectively for positive and negative values, are indicated by green triangles and red circles.

values, reflecting the influence of the chaotic small-scale dynamics. Our primary objective is to investigate the significance of these fluctuations. In figure 4 we present this joint PDF. The white contours have the same meaning as in the previous plot. Green markers indicate the maxima of the conditional PDFs for fixed model values (horizontal cuts), while blue circles represent the corresponding mean values. The discrepancies between the mode and the mean provide a first indication of non-Gaussian behaviour. Furthermore, to a good approximation the means lie on a linear curve, although the slope differs slightly between the negative and positive ranges, as quantified in figure 6. These evidences corroborate the previous conclusion that the Clark model performs well: it is possible to adjust a multiplicative constant (or, more precisely, two constants depending on whether the cascade is forward or backward) in front of (2.7) so that the model matches the average of the DNS configurations. In the inertial range, the linear behaviour with slope of order one shows that the contributions of the non-local terms are negligible with respect to the mean flux (Johnson 2021). Examples of selected conditional distributions, given the model prediction (i.e. the resolved field), are presented in figure 5. These correspond to horizontal cuts in figure 4. Large fluctuations with a variance of the order of the mean are observed. Then, the PDFs are characterised by pronounced non-Gaussian fat tails for each value of the flux. These heavy-tailed statistics indicate the occurrence of rare but intense deviations from the average behaviour, highlighting the intermittent and strongly nonlinear nature of the system. The observed behaviour suggests that the tails could follow

https://doi.org/10.1017/jfm.2025.11087 Published online by Cambridge University Press

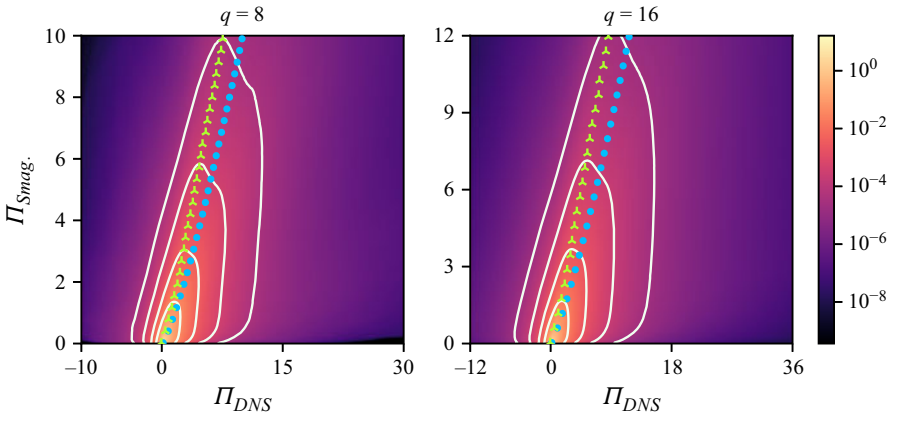


Figure 7. Colour scale represents the joint PDFs of (2.6) and (2.8). White contours indicate probability isolines at levels 10^{-4}, 10^{-5}, 10^{-6} and 10^{-7}, progressing outward from the origin. Green upward tripod markers denote the peaks of the conditional PDFs obtained by fixing the model prediction (horizontal cuts), while blue circles represent the corresponding mean values.

a distribution intermediate between exponential and power-law decay. Yet, it appears that there is not a universal shape for the PDF, and we have not been able to rescale toward a sole master curve. To quantify these observations, we compute the first four cumulants of the conditional PDFs, varying the value of the model. These are mean and variance

$$\mu = \int \Pi \, \mathrm{d}\mathbb{P}[\Pi|\Pi_{model}], \quad \sigma^2 = \int (\Pi - \mu)^2 \, \mathrm{d}\mathbb{P}[\Pi|\Pi_{model}], \tag{3.3}$$

skewness and flatness

$$\gamma_1 = \frac{1}{\sigma^3} \int (\Pi - \mu)^3 \, \mathrm{d}\mathbb{P}[\Pi|\Pi_{model}], \quad \gamma_2 = \frac{1}{\sigma^4} \int (\Pi - \mu)^4 \, \mathrm{d}\mathbb{P}[\Pi|\Pi_{model}], \tag{3.4}$$

where Π denotes the DNS value (2.6). As a first non-trivial result, we note that, on average, the DNS flux has the same sign as the Clark model prediction. This allows us to fit two linear slopes, depending on the sign, both constrained to pass through the (0, 0) flux point, yielding 1.4 ± 0.2 for the forward cascade and 0.5 ± 0.2 for the backward case. Continuing the analysis, and interpreting the variance as a measure of statistical uncertainty, we observe that extreme events (i.e. those with large absolute values) are associated with greater variability. Interestingly, the ratio $\sigma/|\mu|$ appears to decrease as one moves away from zero. The skewness, however, clearly indicates that the distribution remains not symmetric. Gaussian flatness is recovered for extreme Clark model predictions, whereas fat-tailed behaviour dominates near the most probable events.

3.2. *Smagorinsky vs DNS*

For completeness we perform a parallel analysis for the classical eddy-viscosity model of Smagorinsky, described by expression (2.8). Unlike the Clark model, the Smagorinsky formulation predicts only non-negative energy flux, thereby completely failing to account for the backscatter. In figure 7, we present the joint PDF of the ground truth and model prediction: a multiplicative constant can be fitted to align either the maximum probability or the mean value, as shown in figure 9. As in figure 5, figure 8 displays the conditional PDFs of the DNS energy flux for fixed values of the Smagorinsky output. These distributions exhibit prominent heavy tails, characteristic of rare but intense events. A fit of the conditional distributions reveals that the tails decay more slowly than

https://doi.org/10.1017/jfm.2025.11087 Published online by Cambridge University Press

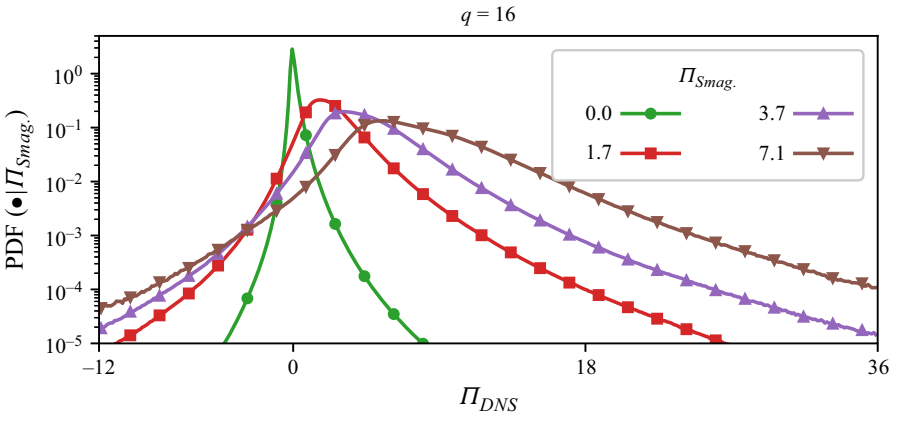


Figure 8. The PDFs of the (2.6) energy flux, conditioned on Smagorinsky estimations (2.8).

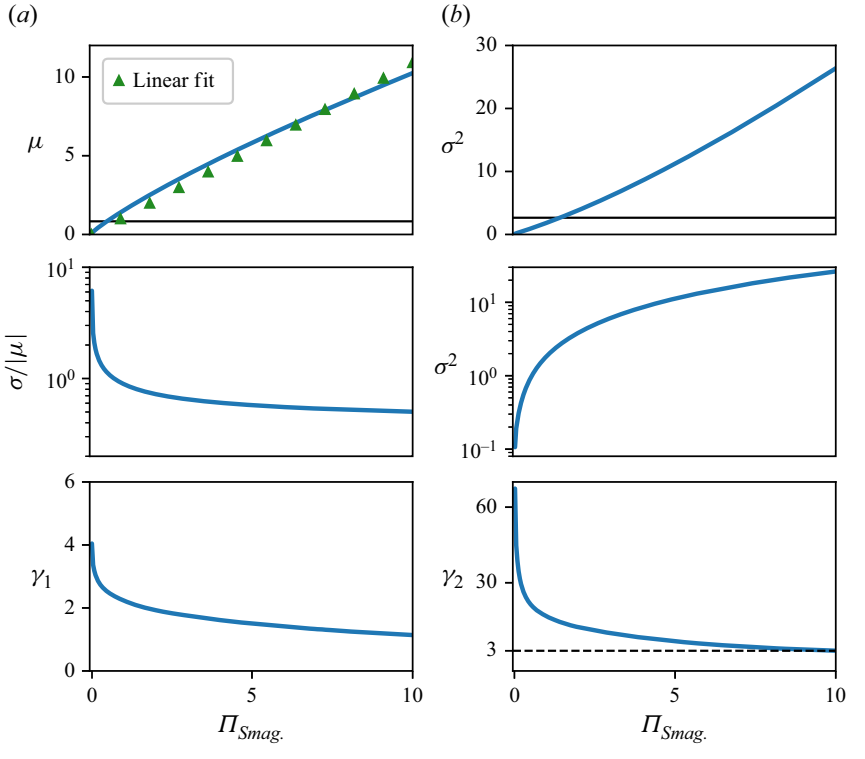


Figure 9. First and third rows: statistical cumulants (mean, variance, skewness and flatness) of the DNS density, computed conditionally on a given value of the Smagorinsky prediction. Horizontal solid lines indicate the unconditional values. Middle row: the conditional variance shown in logarithmic scale (*b*, to highlight variability) and the relative statistical error $\sigma/|\mu|$ (*a*). In the top-left panel (mean), a linear fit constrained to the (0, 0) flux point is indicated by green triangles.

exponentially, supporting the idea that rare events dominate the flux statistics and cannot be neglected in closure strategies. To further quantify these features, we look at the first four statistical cumulants of the conditional distributions, shown in figure 9. As with the Clark model, the conditional mean has a linear trend versus the model prediction: a linear

https://doi.org/10.1017/jfm.2025.11087 Published online by Cambridge University Press

fit gives 1.1 ± 0.4, where the uncertainty corresponds to the root-mean-square deviation. Yet the variance increases with increasing model flux much more for the Smagorinsky case, indicating greater uncertainty for more intense events. However, in that regime, the relative uncertainty $\sigma/|\mu|$ decreases. The skewness exhibits a pronounced positive trend, highlighting the strong asymmetry of the distributions. As for the flatness, values close to the Gaussian reference (i.e. 3) are approached at large predicted values, while extreme flatness, indicative of intermittent fat-tailed fluctuations, dominates in a region of width ϵ around $\Pi_{Smag.} = 0$, where ϵ denotes the mean energy injection rate. These findings illustrate that the behaviour near the Smagorinsky prediction is not Gaussian.

4. Discussion

We have investigated different conditional statistics of the filtered turbulent velocity field in three dimensions. The analysis of the mean energy flux in the *QR* diagram shows unambiguously that the backscatter inverse cascade is correlated solely with the vortex compression, while positive flux is due to all quadrants. Results highlight the role of the non-local terms, which are found to be responsible for an increase of the positive extreme events of the energy flux. With respect to the LES framework, we have found that the Clark model reproduces satisfactorily the mean energy flux and provides an approximate reconstruction of the fluctuations around the mean, but it remains insufficient to fully capture the complex statistical structure of the energy flux in turbulent flows. While the Clark model can be considered a good first-order approximation, we have highlighted the presence of heavy-tailed distributions and marked asymmetries, hallmarks of the intermittent nature of the modelled unresolved scales.

These findings underscore the limitations of classical deterministic closures based solely on filtered quantities and point toward the need for stochastic components to better account for the influence of unresolved scales. The strong skewness and enhanced flatness observed at moderate model outputs demonstrate that rare events and intermittent bursts remain statistically significant and cannot be neglected. In this context, fluctuations in the subgrid stress tensor cannot be modelled by simple Gaussian noise. Starting from available stochastic modelling for LES (Gicquel *et al.* 2002), our results provide a quantitative foundation for the development of stochastic subgrid-scale models that can incorporate both the mean behaviour and the statistical variability of energy transfer.

We need to remark that, in this work, we have focused on stationary statistics to construct time-independent PDFs. An important direction for future research is to investigate temporal correlations, with the aim of constructing fully dynamical stochastic processes for the subgrid-scale energy flux. Temporal correlations indeed constitute a critical point. For instance, in the limiting case of a delta-correlated-in-time closure, a dependence on the integration time step arises: if the random variable is sampled much more frequently than the eddy turnover time of the LES field, the deterministic dynamics effectively perceives only its mean. As discussed in previous works on stochastic modelling (Dreeben & Pope 1998; Minier, Chibbaro & Pope 2014; Innocenti, Marchioli & Chibbaro 2016), physical and numerical constraints on the time step may be needed. We regard this as an important question to be addressed in future research.

We conclude by outlining the perspective of a data-driven approach to capture fluctuations around simple deterministic models. The main present challenge is the curse of dimensionality in the three-dimensional setting: collecting sufficient data to train a generative model for full three-dimensional Eulerian fields is practically impossible, and convergence of the training is also difficult to achieve. A more tractable strategy, in our view, is to reproduce the fluctuations around a first-order approximation expressed

https://doi.org/10.1017/jfm.2025.11087 Published online by Cambridge University Press

in terms of the velocity gradient, rather than attempting to model the entire subgrid-scale stress tensor. The neglected contributions can be written in terms of higher-order gradients, and locality allows the training to be well conditioned when restricted to a neighbourhood of data, thereby improving stability. Unfortunately, the powerful neural networks underlying state-of-the-art generative models, such as diffusion models, remain essentially uninterpretable from a physical point of view, leaving them as promising but opaque tools.

Acknowledgements. This work received financial support from the CNRS through the MITI interdisciplinary initiatives, under its exploratory research program. S.C. acknowledges the funding by ANR SPEED project ANR-20-CE23-0025-01. It was granted access to the HPC resources of GENCI-TGCC & GENCI- CINES (Project No. A0170506421). Finally, we would like to thank L. Biferale for useful discussions.

Declaration of interests. The authors report no conflict of interest.

References

ALEXAKIS, A. & BIFERALE, L. 2018 Cascades and transitions in turbulent flows. *Phys. Rep.* **767**, 1–101.

ALEXAKIS, A. & CHIBBARO, S. 2020 On the local energy flux of turbulent flows. *Phys. Rev. Fluids* **5** (9), 094604.

BANDAK, D., MAILYBAEV, A.A., EYINK, G.L. & GOLDENFELD, N. 2024 Spontaneous stochasticity amplifies even thermal noise to the largest scales of turbulence in a few eddy turnover times. *Phys. Rev. Lett.* **132** (10), 104002.

BISKAMP, D. 2003 *Magnetohydrodynamic Turbulence*. Cambridge University Press.

BORUE, V. & ORSZAG, S.A. 1998 Local energy flux and subgrid-scale statistics in three-dimensional turbulence. *J. Fluid Mech.* **366**, 1–31.

BUARIA, D., PUMIR, A. & BODENSCHATZ, E. 2020 Self-attenuation of extreme events in Navier–Stokes turbulence. *Nat. Commun.* **11** (1), 5852.

CARATI, D., WINCKELMANS, G. & JEANMART, H. 2007 A stochastic subgrid model with application to turbulent flow and scalar transport. *Phys. Fluids* **19** (3), 035107.

CARATI, D., WINCKELMANS, G.S. & JEANMART, H. 2001 On the modelling of the subgrid-scale and filtered-scale stress tensors in large-eddy simulation. *J. Fluid Mech.* **441**, 119–138.

CHANG, N., YUAN, Z. & WANG, J. 2022 The effect of sub-filter scale dynamics in large eddy simulation of turbulence. *Phys. Fluids* **34** (9), 095122.

CHONG, M.S., PERRY, A.E. & CANTWELL, B.J. 1990 A general classification of three-dimensional flow fields. *Phys. Fluids A: Fluid Dyn.* **2** (5), 765–777.

CLARK, R.A., FERZIGER, J.H. & REYNOLDS, W.C. 1979 Evaluation of subgrid-scale models using an accurately simulated turbulent flow. *J. Fluid Mech.* **91** (1), 1–16.

DEARDORFF, J.W. 1970 A numerical study of three-dimensional turbulent channel flow at large Reynolds numbers. *J. Fluid Mech.* **41** (2), 453–480.

DOMARADZKI, J.A., LIU, W. & BRACHET, M.E. 1993 An analysis of subgrid-scale interactions in numerically simulated isotropic turbulence. *Phys. Fluids A: Fluid Dyn.* **5** (7), 1747–1759.

DREEBEN, T.D. & POPE, S.B. 1998 Probability density function/Monte Carlo simulation of near-wall turbulent flows. *J. Fluid Mech.* **357**, 141–166.

EYINK, G.L. 2006 Multi-scale gradient expansion of the turbulent stress tensor. *J. Fluid Mech.* **549**, 159–190.

EYINK, G.L. & BANDAK, D. 2020 Renormalization group approach to spontaneous stochasticity. *Phys. Rev. Res.* **2** (4), 043161.

GERMANO, M. 1992 Turbulence: the filtering approach. *J. Fluid Mech.* **238**, 325–336.

GICQUEL, L.Y.M., GIVI, P., JABERI, F.A. & POPE, S.B. 2002 Velocity filtered density function for large eddy simulation of turbulent flows. *Phys. Fluids* **14** (3), 1196–1213.

GILL, A.E. 2016 *Atmosphere—Ocean Dynamics*. Elsevier.

HÄRTEL, C., KLEISER, L., UNGER, F. & FRIEDRICH, R. 1994 Subgrid-scale energy transfer in the near-wall region of turbulent flows. *Phys. Fluids* **6** (9), 3130–3143.

INNOCENTI, A., MARCHIOLI, C. & CHIBBARO, S. 2016 Lagrangian filtered density function for LES-based stochastic modelling of turbulent particle-laden flows. *Phys. Fluids* **28** (11), 115106.

JOHNSON, P.L. 2021 On the role of vorticity stretching and strain self-amplification in the turbulence energy cascade. *J. Fluid Mech.* **922**, A3.

JOHNSON, P.L. & WILCZEK, M. 2024 Multiscale velocity gradients in turbulence. *Annu. Rev. Fluid Mech.* **56**, 463–490.

https://doi.org/10.1017/jfm.2025.11087 Published online by Cambridge University Press

LINKMANN, M., BUZZICOTTI, M. & BIFERALE, L. 2024 Multiscale properties of large eddy simulations: correlations between resolved-scale velocity-field increments and subgrid-scale quantities. *Phys. Fluids* **36**, 085115.

LIU, S., MENEVEAU, C. & KATZ, J. 1994 On the properties of similarity subgrid-scale models as deduced from measurements in a turbulent jet. *J. Fluid Mech.* **275**, 83–119.

MAGACHO, B., THALABARD, S., BUZZICOTTI, M., BONACCORSO, F., BIFERALE, L. & MAILYBAEV, A.A. 2024 Scale invariance of intermittency in LES turbulence. *J. Fluid Mech.* **987**, A35.

MARSTORP, L., BRETHOUWER, G. & JOHANSSON, A.V. 2007 A stochastic subgrid model with application to turbulent flow and scalar mixing. *Phys. Fluids* **19** (3), 035102.

MASON, P.J. & THOMSON, D.J. 1992 Stochastic backscatter in large-eddy simulations of boundary layers. *J. Fluid Mech.* **242**, 51–78.

MENEVEAU, C. 2011 Lagrangian dynamics and models of the velocity gradient tensor in turbulent flows. *Annu. Rev. Fluid Mech.* **43**, 219–245.

MENEVEAU, C. & KATZ, J. 2000 Scale-invariance and turbulence models for large-eddy simulation. *Annu. Rev. Fluid Mech.* **32** (1), 1–32.

MINIER, J.-P., CHIBBARO, S. & POPE, S.B. 2014 Guidelines for the formulation of Lagrangian stochastic models for particle simulations of single-phase and dispersed two-phase turbulent flows. *Phys. Fluids* **26** (11), 113303.

MININNI, P.D., ROSENBERG, D., REDDY, R. & POUQUET, A. 2011 A hybrid MPI–OpenMP scheme for scalable parallel pseudospectral computations for fluid turbulence. *Parallel Comput.* **37** (6–7), 316–326.

MOIN, P. & MAHESH, K. 1998 Direct numerical simulation: a tool in turbulence research. *Annu. Rev. Fluid Mech.* **30** (1), 539–578.

MOSER, R.D., HAERING, S.W. & YALLA, G.R. 2021 Statistical properties of subgrid-scale turbulence models. *Annu. Rev. Fluid Mech.* **53** (1), 255–286.

PIOMELLI, U., CABOT, W.H., MOIN, P. & LEE, S. 1991 Subgrid-scale backscatter in turbulent and transitional flows. *Phys. Fluids A: Fluid Dyn.* **3** (7), 1766–1771.

POPE, S.B. 2001 Turbulent flows. *Meas. Sci. Technol.* **12** (11), 2020–2021.

SAGAUT, P. 2005 *Large Eddy Simulation for Incompressible Flows: an Introduction*. Springer Science & Business Media.

SMAGORINSKY, J. 1963 General circulation experiments with the primitive equations: I. the basic experiment. *Mon. Weath. Rev.* **91** (3), 99–164.

THALABARD, S., BEC, J. & MAILYBAEV, A.A. 2020 From the butterfly effect to spontaneous stochasticity in singular shear flows. *Commun. Phys.* **3** (1), 122.

THORPE, S.A. 2007 *An Introduction to Ocean Turbulence*, vol. 10. Cambridge University Press Cambridge.

TSINOBER, A. 2009 *An Informal Conceptual Introduction to Turbulence*. Springer.

VELA-MARTÍN, A. & JIMÉNEZ, J. 2021 Entropy, irreversibility and cascades in the inertial range of isotropic turbulence. *J. Fluid Mech.* **915**, A36.

YAO, H., SCHNAUBELT, M., SZALAY, A.S., ZAKI, T.A. & MENEVEAU, C. 2024 Comparing local energy cascade rates in isotropic turbulence using structure–function and filtering formulations. *J. Fluid Mech.* **980**, A42.

https://doi.org/10.1017/jfm.2025.11087 Published online by Cambridge University Press