Physica D 481 (2025) 134826

Contents lists available at ScienceDirect

Physica D

journal homepage: www.elsevier.com/locate/physd

Communities for the Lagrangian dynamics of the turbulent velocity gradient tensor: A network participation approach

Christopher J. Keylock [a,b,*], Maurizio Carbone [b,c,d]

[a] Loughborough University, School of Architecture, Building and Civil Engineering, Loughborough, LE11 3TU, Leicestershire, UK
[b] University of Bayreuth, Department of Theoretical Physics, Bayreuth, 95440, Germany
[c] Istituto dei Sistemi Complessi, CNR, Via dei Taurini 19, 00185, Rome, Italy
[d] INFN "Tor Vergata", Via della Ricerca Scientifica 1, 00133, Rome, Italy

ARTICLE INFO

Communicated by Dmitry Pelinovsky

Dataset link: https://turbulence.idies.jhu.edu/home

Keywords:
Turbulence
Velocity gradient tensor
Complex networks
Network modularity
Schur decomposition

ABSTRACT

In this paper, we present a network-based framework for analyzing the Lagrangian dynamics of the velocity gradient tensor (VGT). Each node represents a flow state, and link weights correspond to the transition probabilities between states, derived from Direct Numerical Simulation (DNS) of statistically steady, isotropic turbulence. The network provides a compact representation of the VGT's continuum dynamics by discretizing it into a finite set of states. We investigate the optimal variables for this discretization, classifying VGT states into groups that best capture the flow's physics. To this end, we test several classifications based on topology and various properties of the background flow coherent structures. We do this using the notion of "community" or "module", namely clusters of nodes that are optimally distinct while also containing diverse nodal functions. The most effective classification, informed by VGT invariants frequently used in the literature, combines the signs of the principal invariants Q, R, and the discriminant Δ, distinguishing regions of real and complex eigenvalues. We refine this further by incorporating the relative magnitude of the non-normal contributions to enstrophy and straining, derived from the Schur decomposition of the VGT. Accounting for non-normality significantly enhances classification fidelity and underscores the critical role of the unclosed and complex contributions to VGT dynamics from the pressure Hessian and viscous terms. A comparison between the DNS data and an enhanced Gaussian closure model reveals the challenges for conventional modeling approaches in accurately capturing the non-normal contributions to the VGT dynamics.

1. Introduction

The use of symbolic dynamics to characterize fluid flow dates back to Hadamard [1] who studied geodesics of surfaces of negative curvature in terms of a sequence of symbols. Such an approach has been used relatively sparingly since in fluid dynamics, with a few theoretical contributions [2], and then applications in e.g., the geosciences, where it has helped characterize the interaction between large-scale flow structures and natural, mobile boundaries from single-point velocity measurements [3,4]. More generally, in nonlinear time-series analysis, symbolic sequence analysis has been employed to enhance the signal-to-noise ratio [5] and to underpin the construction of discrete, complex network approaches to model continuous systems [6] with applications in a range of fields, such as neuroscience [7], transport planning [8] and electrical power provision [9]. In fluid dynamics, complex network approaches have been relatively under-utilized. However, examples include the investigation of vortex interaction in two-dimensional turbulence [10,11] and the study of time irreversibility in near-wall flows [12]. However, complex networks have typically not been used to study the dynamics of the turbulent velocity gradient tensor (VGT) which is the focus of this work.

An open question in the study of the VGT concerns the best way to classify the flow dynamics into different communities of flow states. Such a question has typically been considered from a continuum perspective, rather than utilizing discrete networks, and has been addressed by defining communities of flow structures based on topological properties of the VGT [13]. Although there have been noteworthy departures from the standard approach such as the three-dimensional formalism due to Lüthi and co-workers [14], typically the approach has been two-dimensional, based on the second and third invariants of the VGT (defined below). Classifying VGT states in this manner in terms of their topology is closely related to the definitions of coherent flow structures in turbulent flows [15–18].

Examples of such classification methods in nonlinear network physics include investigations into cliques [19], communities [20], or

* Corresponding author at: Loughborough University, School of Architecture, Building and Civil Engineering, Loughborough, LE11 3TU, Leicestershire, UK.
E-mail address: c.j.keylock@lboro.ac.uk (C.J. Keylock).

https://doi.org/10.1016/j.physd.2025.134826
Received 2 May 2024; Received in revised form 23 June 2025; Accepted 30 June 2025
Available online 14 July 2025

modules [21]. These entail applying some form of partition to extract a set of discrete and non-overlapping "units" that, when taken together include all nodes in a network (although the definition of clique due to [22] permits these to be nested within a given module or overlap with more than one module). In general, modules provide information about the network structure at an intermediate scale between the study of individual nodes (as exemplified by measures of node centrality [23,24]) and the whole network (such as contrasting the exponential or power-law degree distributions of, respectively, the Erdös-Rényi [25] and Barabási and Albert [26] networks).

Whilst in network physics a search for communities is typically empirical, in fluid dynamics, we can formulate communities *a priori*, based on quantities classically considered in the literature, and then test their effectiveness. This is the approach adopted in this paper: We construct a "baseline" network consisting of 844 nodes, with weights for the edge from node i to j reflecting the probability of the flow transitioning between these two states. We then group the nodes into communities/modules in a number of different ways reflecting the physics embodied in the different classifications. These classifications are then evaluated based on the modularity (how the transition probabilities are grouped within the module) as well as measures of the functionality of nodes in the network (do they act as hubs for redistributing the transition probabilities or not?). We then discuss the fluid mechanics contained within our favored classification.

The paper is structured as follows: We first provide a background on the relevant invariants and properties of the VGT, then use them to define communities in a complex network representing the VGT dynamics and, finally, classify the resulting communities. We build the complex network and analyze the communities using Lagrangian data from a direct numerical simulation of homogeneous, isotropic turbulence (HIT) [27] and compare the results from this to those obtained using an enhanced Gaussian closure (EGC) model for the VGT dynamics due to Wilczek and Meneveau [28].

2. The dynamics of the velocity gradient tensor

Our starting point is the continuity equation and Navier–Stokes momentum equation for an incompressible, Newtonian fluid:

$$\nabla \cdot \boldsymbol{u} = 0, \tag{1}$$

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} = -\nabla p + \nu\nabla^2\boldsymbol{u}, \tag{2}$$

where $\boldsymbol{u}(\boldsymbol{x},t)$ is the three-dimensional velocity vector field, $p(\boldsymbol{x},t)$ is the pressure field divided by the constant density, and ν is the kinematic viscosity. The VGT is defined as $\boldsymbol{A} \equiv \nabla\boldsymbol{u}$ and its evolution equation is obtained by taking the spatial gradient of (2):

$$\frac{\partial \boldsymbol{A}}{\partial t} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{A} = -\left(\boldsymbol{A}^2 - \frac{\mathrm{tr}(\boldsymbol{A}^2)}{3}\boldsymbol{I}\right) - \boldsymbol{H} + \nu\nabla^2\boldsymbol{A}, \tag{3}$$

where $\boldsymbol{I}$ is the identity matrix, tr(…) indicates the matrix trace, and $\boldsymbol{H}$ is the anisotropic pressure Hessian, $\boldsymbol{H} \equiv \boldsymbol{\nabla\nabla}p - 2Q\boldsymbol{I}/3$, with $Q \equiv -\mathrm{tr}(\boldsymbol{A}^2)/2$ the second principal invariant of the VGT.

The relation between the trace of the pressure Hessian and the second principal invariant of the VGT highlights the dynamical significance of the VGT invariants [29,30]. The invariants also play a direct role in parameterizing the single-point statistics of the VGT in isotropic turbulence [31,32] and the classification of small-scale flow states [33]. Following [34], we can use (3) to write down equations for the Lagrangian evolution of the VGT principal invariants, Q and $R \equiv -\mathrm{tr}(\boldsymbol{A}^3)/3$, in terms of each other, formally including the viscous effects and the coupling between the VGT $\boldsymbol{A}$, the deviatoric pressure Hessian $\boldsymbol{H}$ and viscous contributions $\nu\nabla^2\boldsymbol{A}$:

$$\frac{dQ}{dt} = -3R - \nu\,\mathrm{tr}\left(\boldsymbol{A}\nabla^2\boldsymbol{A}\right) + \mathrm{tr}\left(\boldsymbol{A}\boldsymbol{H}\right) \tag{4}$$

$$\frac{dR}{dt} = \frac{2}{3}Q^2 - \nu\,\mathrm{tr}\left(\boldsymbol{A}^2\nabla^2\boldsymbol{A}\right) + \mathrm{tr}\left(\boldsymbol{A}^2\boldsymbol{H}\right), \tag{5}$$

where d/dt is the Lagrangian derivative taken along fluid particle trajectories. Given the principal invariants Q and R, the VGT eigenvalues λ_i follow from the characteristic equation

$$\lambda_i^3 + P\lambda_i^2 + Q\lambda_i + R = 0, \tag{6}$$

where the first invariant, P, is zero in incompressible flow. Neglecting the last two terms on the right-hand side of (4) and (5) yields the restricted Euler model [30,35] in which principal invariants, Q and R evolve independently of the other VGT components. However, the restricted Euler model features a finite-time blowup for almost all initial conditions, which is not observed in numerical simulations of the Navier–Stokes equations. This indicates that the deviatoric pressure Hessian and viscous terms crucially enter the VGT dynamics by affecting the VGT magnitude and alignments. In particular, the effects of the pressure and viscous terms reflect on the statistics of the strain-rate tensor $\boldsymbol{S}$, rotation-rate tensor $\boldsymbol{\Omega}$ and their statistical alignments [36], where

$$\boldsymbol{S} = \frac{1}{2}(\boldsymbol{A} + \boldsymbol{A}^*) \qquad \boldsymbol{\Omega} = \frac{1}{2}(\boldsymbol{A} - \boldsymbol{A}^*) \tag{7}$$

and the asterisk indicates matrix transposition and complex conjugation. Formulating models for the dynamics that correctly capture the complexities introduced by the non-local effects of the pressure Hessian has been the focus of a significant body of work, e.g. [28,37–43], as reviewed in [44,45].

Here, we explicitly separate the normal part of $\boldsymbol{A}$, related to its eigenvalues λ_i, and the non-normal part of $\boldsymbol{A}$ by using a complex Schur transform [46]. This transform imposes a unitary form for the rotation vectors and moves the non-normality, $\boldsymbol{N}$, into the upper-triangular part of the central matrix of the transform, $\boldsymbol{T}$, [47]

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{T}\boldsymbol{U}^*, \tag{8}$$

where $\boldsymbol{T} = \boldsymbol{\Lambda} + \boldsymbol{N}$, $\boldsymbol{\Lambda}$ is a diagonal matrix of eigenvalues such that $\Lambda_{i,i} = \lambda_i$, and $\boldsymbol{N}$ is strictly upper-triangular. The normal and non-normal contributions to the dynamics may be isolated according to, $\boldsymbol{B} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^*$ and $\boldsymbol{C} = \boldsymbol{U}\boldsymbol{N}\boldsymbol{U}^*$. The principal invariants Q and R can be written purely in terms of strain and rotation components of $\boldsymbol{B}$, while the enstrophy, total straining and production terms all involve contributions from $\boldsymbol{C}$ [47]. Hence, for the second-order invariants we have

$$Q = \frac{1}{2}\left(\|\boldsymbol{\Omega}\|^2 - \|\boldsymbol{S}\|^2\right), \tag{9a}$$

$$\|\boldsymbol{\Omega}\|^2 = \|\boldsymbol{\Omega}_B\|^2 + \|\boldsymbol{\Omega}_C\|^2, \tag{9b}$$

$$\|\boldsymbol{S}\|^2 = \|\boldsymbol{S}_B\|^2 + \|\boldsymbol{\Omega}_C\|^2, \tag{9c}$$

where $\|\ldots\|$ is the Frobenius norm. The equivalent expressions for the third-order invariants are

$$R = -\det(\boldsymbol{S}) - \mathrm{tr}\left(\boldsymbol{\Omega}^2\boldsymbol{S}\right), \tag{10a}$$

$$-\det(\boldsymbol{S}) = -\det\left(\boldsymbol{S}_B\right) + \mathrm{tr}\left(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B\right) - \det\left(\boldsymbol{S}_C\right), \tag{10b}$$

$$\mathrm{tr}\left(\boldsymbol{\Omega}^2\boldsymbol{S}\right) = \mathrm{tr}\left(\boldsymbol{\Omega}_B^2\boldsymbol{S}_B\right) + \mathrm{tr}\left(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B\right) - \det\left(\boldsymbol{S}_C\right), \tag{10c}$$

where det(…) is the determinant. This decomposition has been adopted to gain an insight into several properties of turbulence including the reason for the preferred alignment between the vorticity vector and the eigenvector corresponding to the intermediate eigenvalue of the strain rate tensor [47], the physics of the flow when non-normality is maximal [48], and to characterize the role of in-rushing sweeps in boundary layers [49]. In addition, in spatially developing flows, it has been shown that $\|\boldsymbol{\Omega}_C\|^2$ is crucial for the dynamics before turbulence is fully established (i.e. regions where pressure Hessian contributions typically dominate the kinetic energy budget) [50].

In this study, we follow [51] and formulate a classification for the VGT state by making use of the terms in (9)–(10) together with the discriminant,

$$\Delta = Q^3 + \frac{27}{4}R^2, \tag{11}$$

Table 1
The various components of the state space vector used to define each node in the network.

Flow properties	Manner implemented in the network
1. Regions of the $Q-R$ diagram	Signs of Q(9a), R(10a), and Δ(11)
2. Constituents of R	Signs of $-\det(\boldsymbol{S})$ and $\mathrm{tr}\left(\boldsymbol{\Omega}^2\boldsymbol{S}\right)$(10a)
3. Non-normal constituents of the production terms	Signs of $-\det\left(\boldsymbol{S}_C\right)$ and $\mathrm{tr}\left(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B\right)$ ((10b),(10c))[a]
4. Relative magnitudes of the second-order terms	Signs of ξ and χ ((23),(24))[b]
5. Relative magnitudes of the third-order terms	Twenty-four state variable based on all possible rank-orderings of $\lvert\det\left(\boldsymbol{S}_B\right)\rvert$, $\lvert\mathrm{tr}\left(\boldsymbol{\Omega}_B^2\boldsymbol{S}_B\right)\rvert$, $\lvert\det\left(\boldsymbol{S}_C\right)\rvert$, $\lvert\mathrm{tr}\left(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B\right)\rvert$ ((10b),(10c))

[a] The signs of $-\det\left(\boldsymbol{S}_B\right)$ and $\mathrm{tr}\left(\boldsymbol{\Omega}_B^2\boldsymbol{S}_B\right)$ are opposite with the former given by the sign of R so do not need to be included.
[b] The sign of Q is already included in the first row of the table.

which separates states of $\boldsymbol{A}$ featuring real and complex eigenvalues. Hence, we construct a network where each node is a unique flow state, characterized by the signs and relative magnitudes of the invariants ((9)–(10)–(11)), as detailed in Table 1. The various permutations given in Table 1 result in 848 possible flow states as the physics constrains each term in this table from being treated independently. A simple example of this is that if $\Delta < 0$ then $Q < 0$, and another is that if $\|\boldsymbol{S}_B\|^2 > \|\boldsymbol{\Omega}_C\|^2$ then $\lvert\det\left(\boldsymbol{S}_B\right)\rvert > \lvert\det\left(\boldsymbol{S}_C\right)\rvert$. Given this "full" network, we define classifications based on grouping nodes into different modules or communities. We can then make use of network community extraction techniques to determine the best classification for the Lagrangian dynamics of the VGT. In practice, we reduce the 848 nodal states to 844 nodes as four states require such particular alignment properties that they very rarely occur. This is described in more detail below.

3. Networks and network analysis terminology

A network or graph, $\mathcal{G}$, without self-loops, comprises a set of N nodes (vertices), edges between nodes and weights associated with each edge. The N by N adjacency matrix, $\boldsymbol{M}$, contains the connectivity information between the nodes and consists of zeros along the primary diagonal and, for the unweighted case, a one if there is a direct connection between nodes i and j and a zero if no connection exists. For the weighted case, the weight w_{ij} is proportional to the probability of transitioning from node i to j. In our case, this is the probability of the flow changing from a state described by node i to a state described by node j. An undirected network is one where $w_{ij} = w_{ji}$ while a directed case will lack this symmetry. For example, in our case, the restricted Euler model implies that the transition from a state given by $Q > 0$, $R < 0$ to one given by $Q > 0$, $R > 0$ is much more likely than the reverse case, reflecting the preferential clockwise trajectories in the Q-R plane [35].

3.1. Modularity

The notion of a community or "module" within a network can be formalized to enable a quantitative analysis of intermediate scale structure between that of the individual vertex and the network as a whole [21]. An important principle to find optimal partitions has been to maximize the modularity [52], which for a given partition into N_M distinct modules is given by:

$$M \equiv \sum_{k=1}^{N_M}\left[\frac{S_k}{S_W} - \frac{\mathcal{D}_k}{2S_W}\right]. \tag{12}$$

Here N_M is the number of modules,

$$S_W = \sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij} \tag{13}$$

is the sum of all weights in the network, S_k is the sum of all weights over all nodes in module k

$$S_k = \sum_{i\in\mathcal{M}_k}\sum_{j\in\mathcal{M}_k} w_{ij}, \tag{14}$$

where $\mathcal{M}_k$ is the set of all node indices in module k, and

$$\mathcal{D}_k = \sum_{i\in\mathcal{M}_k} D_i, \tag{15}$$

is the module degree, given by the sum of the degrees $D_i = \sum_{j=1}^{N} w_{ij}$ of all the nodes in module k. While D_i measures how well node i is connected to the rest of the network, $\mathcal{D}_k$ measures the connection strength of module k towards both itself and the rest of the network, and S_k quantifies how strongly nodes within module k are connected amongst themselves. Based on this, maximum modularity M in (12) implies that nodes within a module are minimally connected to the rest of the network. Furthermore, the second term within the bracket in Eq. (12) prevents a maximization of S_k that leads to a single module corresponding to the whole network. In this paper, we adopted an alternative approach to this maximization problem. We first note that $\sum S_k \leq S_W$ for a fully connected network and, thus, that $p_W = (\sum_k S_k)/S_W$ is the effectiveness of the modules for incorporating all the edges in the network and that our normalization to $S_W = 1$ allows us to simplify to $p_W = \sum_k S_k$. In addition, it is desirable that S_k is similar in value for all k, i.e. each module is of a similar "size". It follows that the best classification will have a high p_W and maximize an entropy based on the set of values for S_k. These two statements may be combined through the conditional entropy:

$$\mathcal{E} = -p_W\sum_{k=1}^{N_M} p_M(k)\log_2 p_M(k), \tag{16}$$

where

$$p_M(k) = \frac{S_k}{\sum_{k'} S_{k'}}. \tag{17}$$

This formalism intrinsically prevents an optimum corresponding to the whole network because $N_M = 1$ gives $\mathcal{E} = 0$. While it is the case that, *ceteris paribus*, $\mathcal{E}$ grows with N_M it will increase less steeply if further division into new modules does not result in significant entropy increase. Hence, it is useful to characterize this rate of change of $\mathcal{E}$ with N_M. The rate of entropy change, $\delta\mathcal{E}$ is given by

$$\delta\mathcal{E} = \frac{\mathcal{E}}{\log_2(N_M)}, \tag{18}$$

and we use $\delta\mathcal{E}$ as well as $\mathcal{E}$ to evaluate our different classifications from the direct numerical simulation and the enhanced Gaussian closure.

3.2. Participation

Optimizing the modularity does not distinguish between different forms of nodal function and the extent to which these are represented within a given module. This can be investigated using the participation coefficient, which for node i is defined as [21]

$$\rho_i = 1 - \sum_{k=1}^{N_M}\left(\frac{\mathcal{D}_{ik}}{D_i}\right)^2, \tag{19}$$

where

$$\mathcal{D}_{ik} = \sum_{j\in\mathcal{M}_k} w_{ij}. \tag{20}$$

Hence, if a vertex is only connected to other nodes in its own module, $i \in \mathcal{M}_k$, then $\rho_i = 0$, while even connectivity to all modules gives $\rho_i \to 1$. Again, a probabilistic interpretation of ρ_i is possible if we define $p_D(i,k) = \mathcal{D}_{ik}/D_i$. As noted by [53] this means that (19) may be interpreted as equivalent to a Gini coefficient, which is derived from the Shannon entropy by expanding the logarithm and truncating to the leading term.

Table 2
Classification of node function based on z_i and ρ_i following [21] A non-hub has $z < 2.5$ while $z_i \geq 2.5$ corresponds to a hub.

Role	Hub/non-hub	Description
Ultra-peripheral	non-hub	$\rho_i \sim 0$
Peripheral	non-hub	$\rho_i \leq 0.625$
Connector	non-hub	$0.625 < \rho_i \leq 0.8$
Kinless	non-hub	$\rho_i > 0.8$
Provincial	hub	Node with a large degree and $\rho_i \sim 0.3$
Connector	hub	Node with a large degree and $0.3 < \rho_i \leq 0.75$
Kinless	hub	$\rho_I > 0.75$

It was proposed by [21] to complement ρ_i with a metric for within-module connectivity based on the statistical notion of a z-score. With D_{ik}^{+} the variant of D_{ik} that is restricted/conditioned to nodes in the same module, i.e. $D_{ik}^{+} = \sum_{j\in\mathcal{M}_k}\left(W_{ij} \mid i \in \mathcal{M}_k\right)$, then

$$z_i = \frac{D_{ik}^{+} - \overline{D}_{ik}^{+}}{\sigma_D(k)}, \tag{21}$$

where the overbar indicates the mean over all nodes in module k, and the standard deviation of this set of nodes is given by $\sigma_D(k)$. Based on a threshold applied to z it was then suggested that $z_i \geq 2.5$ was a *hub* node and $z_i < 2.5$ was a non-hub. That is, the degree of within module connectivity exhibited by a hub would have an approximate probability of occurrence of 1% if the D_{ik}^{+} values were normally distributed. One reason the classifications studied in this paper have no more than $N_M = 64$ modules is that, on average, there are $844/64 = 13$ nodes per classification so that the mean and standard deviation in (21) can be meaningfully estimated.

Combining this threshold with ρ_i resulted in the seven-state classification given in Table 2. Note that, for a directed network the participation coefficient, ρ_i will potentially differ depending upon if one is considering in-degree or out-degree characteristics. Thus, actually, ρ_i in (19) is an abbreviation for the out-degree participation, ρ_i^O, and the in-degree participation, ρ_i^I would require D_{ik} to be replaced by $D_{kj} = \sum_{i\in\mathcal{M}_k} w_{ij}$.

4. Numerical method and closure model approximation

4.1. Numerical approach

To populate the adjacency matrix for our network, we studied Lagrangian trajectories in a direct numerical simulation (DNS) of homogeneous isotropic turbulence (HIT) [27]. We released 27^3 tracers into the simulation with an initial separation of two Taylor microscales. Tracers were followed for 240 Kolmogorov times, τ_η, at a resolution of $0.05\tau_\eta$. Transition probabilities were determined by counting every time the flow changed from node i, that is, a state defined by the signs and relative magnitudes of the velocity gradient invariants in Table 1, to another node j, featuring different signs and relative magnitudes in Table 1. Consequently, our analysis focuses on the probability of changes in flow states rather than the duration of occupation of a state. In the Appendix, we show that we have sampled a sufficient number of trajectories for our results to have converged, based on the residual error between the out-degree and in-degree and from bootstrapping subsets of our database.

When examining the results from the numerical simulation, four of the 848 flow states required a strong anti-alignment between the normal straining and the non-normality such that, while $\|\boldsymbol{\Omega}_C\|^2 < \|\boldsymbol{S}_B\|^2$, one has that $|\det(\boldsymbol{S}_C)| > |\mathrm{tr}(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B)|$. This situation was so infrequent that these four nodes were visited at least five times less frequently than the fifth least-frequented node. This may be seen in the degree distribution in Fig. A.6b-d in the Appendix. Hence, we excluded these four infrequent nodes from further consideration, and the network analyzed consists of $N = 844$ nodes.

Table 3
Topological flow states in the Q-R plane following [54].

Sign(Δ)	Sign(R)	Topology
−	−	stable-node/saddle/saddle (SN/S/S)
−	+	unstable-node/saddle/ saddle (UN/S/S)
+	−	stable-focus/stretching (SF/S)
+	+	unstable-focus/ contracting (UF/C)

4.2. Enhanced Gaussian closure (EGC) modeling

We complement the DNS network analysis with data from a reduced-order velocity gradient model. Developing stochastic models for the Lagrangian dynamics of the VGT is important as their verisimilitude encapsulates our understanding of the Navier–Stokes equations whilst also providing a potentially powerful method in applied computational fluid dynamics. In this study, we adopt the Enhanced Gaussian Closure (EGC) model of Wilczek and Meneveau [28] to provide a comparative dimension to our study. This is based on a Langevin equation describing a set of fluid particles with the same velocity gradient characteristics:

$$d\widetilde{\boldsymbol{A}} = \Big[-\Big(\widetilde{\boldsymbol{A}}^2 - \tfrac{1}{3}\mathrm{tr}(\widetilde{\boldsymbol{A}}^2)\boldsymbol{I}\Big) - \pi_1\Big(\widetilde{\boldsymbol{S}}^2 - \tfrac{1}{3}\mathrm{tr}(\widetilde{\boldsymbol{S}}^2)\boldsymbol{I}\Big) - \pi_2\Big(\widetilde{\boldsymbol{\Omega}}^2 - \tfrac{1}{3}\mathrm{tr}(\widetilde{\boldsymbol{\Omega}}^2)\boldsymbol{I}\Big) - \pi_3\Big(\widetilde{\boldsymbol{S}}\widetilde{\boldsymbol{\Omega}} - \widetilde{\boldsymbol{\Omega}}\widetilde{\boldsymbol{S}}\Big) + \delta\widetilde{\boldsymbol{A}}\Big]\,dt + d\boldsymbol{\mathcal{F}}, \tag{22}$$

where $\widetilde{\boldsymbol{A}}$, $\widetilde{\boldsymbol{S}}$ and $\widetilde{\boldsymbol{\Omega}}$ are realizations of the VGT, strain rate and rotation rate fields, respectively. The stochastic nature of the model is dictated by $\boldsymbol{\mathcal{F}}$, which is a Gaussian forcing with zero mean and isotropic statistics that acts as a white noise in the time domain. The model coefficients π_i are fitted from DNS data to approximate the average behavior of the pressure Hessian conditional on the local velocity gradient, while the tensorial functional form is prescribed by a Gaussian closure hypothesis [28]. The damping coefficient, δ, arises from the approximation for the average velocity gradient Laplacian conditional on the local velocity gradient itself. The model coefficients are fixed as constants up to a Reynolds number dependency for π_3, and these values are in good agreement among different DNS and experimental datasets. Using trajectories sampled from the EGC model in (22), we computed the values for the edges of an adjacency matrix based on the same 848 discrete states (and, thus, nodes) used for the DNS, as in Table 1, at the same temporal resolution and Reynolds number. This allows a comparison of the nodal properties and community structures of the two networks.

5. *A priori* network classifications

Rather than adopting an empirical approach to community network classification, our approach is to test existing physics-based classifications of the VGT dynamics, as well as extensions of them based on the terms in (9a)–(10c) as listed in Table 1. In this section, we outline the basis for the ten primary classifications examined in this paper.

From the restricted Euler model given by the simplification of (4) and (5) that arises from disregarding the last two terms on the right-hand side of each of these equations, it follows that the dynamics on the Q-R plane [30,55] are the logical starting point for looking to define communities. This approach increased in popularity significantly following the work of [13,54] (see [44] for a review), as it provided a topological classification of four different flow configurations based on the sign of R and the sign of the discriminant function, Δ in (11). These four topologies are defined in Table 3 and whilst this provides a means to classify the flow into different, non-overlapping modules/communities, it is not the only approach possible on the Q-R plane. For example, a positive value of Q has been used extensively as a criterion for coherent flow structure identification [15,16], which follows from the physical interpretation of $Q > 0$ as the excess enstrophy (9a). Hence, an alternative definition of modules/communities to that in Table 3 would be one based on the signs of Q and R. Thus, we have:

C1 A classification into $N_M = 4$ modules based on the sign of R and the sign of the discriminant function, Δ [13];

C2 A classification into $N_M = 4$ modules based on the sign of R and the sign of Q.

A superposition of **C1** and **C2** leads to a classification of the VGT in terms of six regions and it has been shown that there are distinct behaviors in each of these regions even though they are not all topologically distinct [47]. Thus, our third classification is a hybrid of the first two:

C3 A classification into $N_M = 6$ modules based on the signs of R, Q and Δ.

A departure from working in the Q-R plane was introduced by [14] who separated R into the terms on the right-hand side of (10a). This leads to

C4 A classification into $N_M = 8$ modules based on the sign of Q and then the signs of the strain production, $-\det(\boldsymbol{S})$ and the enstrophy production, $\mathrm{tr}(\boldsymbol{\Omega}^2\boldsymbol{S})$.

Considering our expansion of R in (10b) and (10c) it is clear that **C4** permits the combined effect of the non-normal production, $\det(\boldsymbol{S}_C)$, and the interaction production, $\mathrm{tr}(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B)$, to be included in analysis (these terms cancel when considering R instead). However, because these two terms feature in both (10b) and (10c) there is an innate correlation between these variables. While the signs of the normal strain production, $\det(\boldsymbol{S}_B)$, and the normal enstrophy production, $\mathrm{tr}(\boldsymbol{\Omega}_B^2\boldsymbol{S}_B)$ are necessarily opposite where $\Delta > 0$, (with $\mathrm{sgn}[-\det(\mathbf{S}_B)] = \mathrm{sgn}[R]$), there is no similar constraint on the signs for the other two terms. Thus, one may extend **C4** to form

C5 A classification with $N_M = 16$ modules based on the signs of Q and R (with the latter setting the signs for $-\det(\mathbf{S}_B)$ and $\mathrm{tr}(\boldsymbol{\Omega}_B^2\mathbf{S}_B)$), and then the signs of $-\det(\mathbf{S}_C)$ and $\mathrm{tr}(\boldsymbol{\Omega}_C^2\mathbf{S}_B)$.

Our final general classification type extends the logic [14] applied to R to also consider Q. From (9b) we have the non-normality, $\|\boldsymbol{\Omega}_C\|^2$, in addition to the normal enstrophy, $\|\boldsymbol{\Omega}_B\|^2$ and the normal straining, $\|\boldsymbol{S}_B\|^2$. As all these terms are non-negative, it only makes sense to consider the sign of their differences as in (9a). Hence, with Q retained to contrast the magnitudes of the normal enstrophy and normal strain, we may define

$$\chi = \|\boldsymbol{\Omega}_C\|^2 - \|\boldsymbol{\Omega}_B\|^2 \tag{23}$$

$$\xi = \|\boldsymbol{\Omega}_C\|^2 - \|\mathbf{S}_B\|^2, \tag{24}$$

and there will be six possible values for the signs of Q, χ and ξ when taken together (because if $Q > 0$ and $\chi > 0$, necessarily $\xi > 0$; likewise if $Q < 0$ and $\xi > 0$, necessarily $\chi > 0$). If, in addition, we follow **C3** and incorporate the sign of Δ into our analysis then there is a further constraint that restricts the number of combinations of signs that can be observed because $\|\boldsymbol{\Omega}_B\|^2 = 0$ when $\Delta < 0$, which imposes $\chi < 0$. Hence, our sixth classification type is

C6 A classification with $N_M = 64$ modules based on the signs of Q, χ, ξ, Δ, R, $-\det(\mathbf{S}_C)$ and $\mathrm{tr}(\boldsymbol{\Omega}_C^2\mathbf{S}_B)$.

The introduction of the variables ξ and χ means that we may also extend the classifications **C2** to **C5** (all of which involve Q) by introducing these two additional variables, which increases the number of modules by a factor of 3 where $\Delta > 0$ and by a factor of 2 where $\Delta < 0$. Thus, in addition to **C1** to **C6** we have four additional categories, **C2b**-**C5b**, which extend **C2**-**C5** by including ξ and χ.

6. Analysis

6.1. Out-degree

Fig. 1 shows the out-degree for each of the 848 nodes of our network for the DNS results and the EGC model [28]. There is, in general, a good level of agreement between the two, with the four nodes that were excluded from analysis due to their very low frequency of occurrence found to also be very infrequent in the model (highlighted in the figure by the dotted-dashed box). The EGC model over-represents the connectivity of the most visited nodes with $Q > 0$ (regions 1 and 2 shown as circles and squares) and under-represents it for the most visited nodes with $Q < 0, \Delta < 0$ (regions 4 and 6 shown as triangles). This is in agreement with the thin right Vieillefosse tail produced by the EGC model [28]. While the model behavior is broadly correct in terms of nodal out-degree, it is not obvious from Fig. 1 how any difference in the connections between specific nodal pairs impacts the community structure of the resulting network, which is what we examine next through the modularity enstrophy.

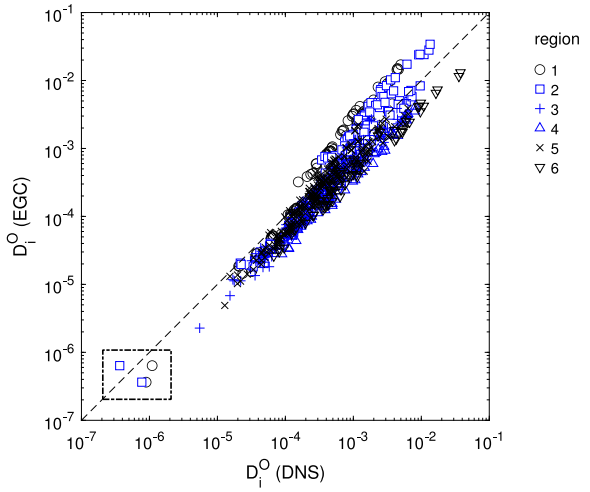

Fig. 1. A scatterplot of the nodal out-degree for the DNS data and the values obtained from the EGC model of [28]. The dashed line is the line of perfect agreement and the box highlights the four nodes that arise very infrequently. Nodes in the six regions of the $Q - R$ diagram, adopted in classification **C3**, are shown by the different symbols: circles/squares are $Q > 0$, crosses/pluses are $Q < 0$; $\Delta > 0$ and triangles are $\Delta < 0$. The black color indicates $R > 0$ and blue $R < 0$.

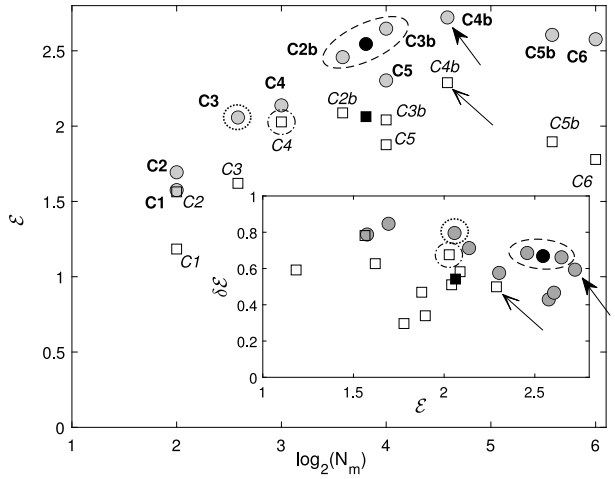

Fig. 2. The main panel shows values of the modularity entropy, $\mathcal{E}$ for the various classifications. The DNS results are given by gray circles and the EGC model as white squares. The inset shows $\mathcal{E}$ vs $\delta\mathcal{E}$. The black symbols are for a reduced variant of **C3b** discussed later in the paper. Additional symbols highlight features discussed in the text.

6.2. Modularity entropy

The main panel of Fig. 2 shows the conditional entropy, $\mathcal{E}$ from (16) as a function of the number of modules N_M, which is displayed on

Table 4
Node types and their stability based on the probability of changing function over the different classifications, as explained in the text, and the associated properties of D_i^O for each type. Values in brackets are from the EGC model.

Type	No. of nodes	$\langle \log_{10}(D_i^O) \rangle$	$\sigma\left[\log_{10}(D_i^O)\right]$	Nodes in common
Stable hub	40 (42)	−2.14 (−2.14)	0.26 (0.35)	40
Non-stable hub	148 (137)	−2.79 (−2.74)	0.31 (0.31)	134
Stable non-hub	346 (346)	−3.82 (−4.09)	0.44 (0.49)	346
Semi-stable 1 non-hub	63 (65)	−3.38 (−3.55)	0.23 (0.26)	63
Semi-stable 2 non-hub	55 (54)	−3.09 (3.25)	0.32 (0.34)	53
Non-stable non-hub	186 (197)	−3.24 (−3.41)	0.35 (0.40)	183
	838 (841)			

a base-2 logarithmic scale, as many of our classifications lead to N_M being an integer power of two. Results are shown for both the DNS and the EGC model, and in general, the EGC model struggles to replicate the observed behavior of the DNS. The entropy $\mathcal{E}$ increases markedly from **C2** to **C3** (the $Q-R$ plane in six regions) for the DNS, increases slightly from **C3** to **C4** and then increases once more for **C2b** as the non-normality $\|\boldsymbol{\Omega}_C\|^2$ is included explicitly. The importance of this term is seen by comparing **C3b** and **C5**, both of which have $N_M = 16$: $\mathcal{E}$ is 15% greater for the former, which includes (23) and (24). This change is significant relative to the 11% change from **C2b** at $N_M = 12$ to the maximum value of $\mathcal{E} = 2.72$, attained by **C4b** with $N_M = 24$. Values for $\mathcal{E}$ then decline for larger N_M.

Although it is also the case that **C4b** maximizes $\mathcal{E}$ for the EGC simulation, most of the values have plateaued by **C4**, with the gain from **C4** to **C4b** only 13% compared to 27% for the DNS, and with no other result showing more than a 3% increase over that for **C4**. In other words, the EGC model cannot capture the non-normal dynamics very effectively. The entropy values for the EGC model are systematically lower than the values for the DNS. This indicates that for a given classification, there is more anisotropy in intra-module nodal behavior in the EGC than in the DNS.

The inset panel in Fig. 2 shows the entropy variation $\delta\mathcal{E}$ against $\mathcal{E}$. This indicates that the most effective classifications are **C4b** with the highest value for $\mathcal{E}$ (solid arrowhead) but also the three values highlighted by the dashed ellipse (**C2b**, **C3b** and the reduced version of **C3b**). At a lower value for $\mathcal{E}$, classification **C3** (dotted circle) is also effective owing to its small N_M. For the EGC model, **C4b** (open arrowhead) and **C4** (dot-dashed circle) are the more effective classifications. Hence, for the EGC model, the key factor is to disaggregate strain production and enstrophy production within R (the key component of the **C4** classifications) rather than capture the non-normal processes, and the consequence is less effective modularity.

6.3. Nodal functionality

Table 4 classifies the network nodes into hub or non-hub classes and then further sub-divides them based on how consistent this was over the ten classifications. The mean and standard deviation of the out-degree is also stated for each case (the out-degree is approximately normally distributed after a logarithmic transformation). Three hundred and eighty six vertices did not alter their classification from hub or non-hub over all ten classifications for the DNS. These are denoted as "stable" in Table 4 and it is clear that the EGC model does an excellent job at identifying these nodes. The results in brackets in the table are from the EGC simulation, and the right-hand column states the number of nodes that are the same for the DNS and the EGC. Hence, all stable hubs and stable non-hub nodes for the DNS are identified in the EGC (although the value in brackets in the second column shows that the EGC finds two additional nodes to be stable hubs).

Semi-stable type 1 nodes were those that were the same for all classifiers except for one of the two that have the greatest number of modules (either **C5b** or **C6**). Semi-stable type 2 nodes were those that were the same for all classifiers except both of **C5b** and **C6**; all of these cases were non-hub nodes. The non-stable cases were those that were classified as hubs or non-hubs the majority of times, but the classifications that gave alternate results were not restricted to the high N_M cases of **C5b** and **C6**. Six nodes were classified as hub or non-hub an equal number of times for the DNS and are therefore excluded from this table, which is why the number of nodes sums to 838. For the EGC, three nodes were excluded on this basis.

Based on a t-test, all of the mean values for the DNS results quoted in the table are significantly different at the 1% significance level except those for the two semi-stable non-hub node groupings ($p = 0.03$) as might be expected given the closely related definition of these two groupings. As there is close to a factor of 50 difference in the mean out-degree for the stable hubs compared to the stable non-hubs and given that a large degree forms part of the description of a hub in Table 2, then Table 4 indicates that we can robustly discriminate between the stable hubs and stable non-hubs. Hence, we can legitimately examine the different fluid mechanics of stable hubs and non-hubs.

6.4. Fluid mechanical properties of stable hubs and non-hubs

Fig. 3 illustrates how the sign and ranking of the VGT invariants varies for the stable hub nodes (black bars), stable non-hubs (gray bars) and for all the nodes (white bars). All the terms used in classifications **C1** to **C3b** feature in panels (a) to (c), while the signs of the production terms, which feature in the remaining classifications are given in panels (d) and (e). Note that for compactness we state the rank order of $\|\boldsymbol{\Omega}_C\|^2$ relative to $\|\boldsymbol{\Omega}_B\|^2$ and $\|\boldsymbol{S}_B\|^2$, rather than using ξ and χ explicitly. Also, we do not give the signs of $-\det\left(\boldsymbol{S}_B\right)$ and $\mathrm{tr}\left(\boldsymbol{\Omega}_B^2\boldsymbol{S}_B\right)$ as these are opposite to one another for $\Delta > 0$, with the sign of the former given by the sign of R. Note that the sign of R is stated on the abscissa for panel Fig. 3a and is used to separate the flow states considered in panels (b) and (c).

Panel (a) focuses on the $Q > 0$ cases, and among these enstrophy-dominated nodes the stable non-hubs (gray) show very similar relative frequencies to all nodes (white). In contrast, stable hub nodes (black) have very distinct behavior, which is not symmetric in R, with strong non-normality when $R < 0$ (i.e. $\mathrm{rnk}\left(\|\boldsymbol{\Omega}_C\|^2\right) = 1$ is much more frequent) and lower non-normality where $R > 0$. This distinction highlights why classification **C3b** out-performs **C3** as it can capture the role played by the non-normality.

Panels (b) and (c) show that stable hubs rarely occur where $Q < 0$ and $\Delta > 0$, and not at all if, in addition, $R < 0$ as shown in panel (b). Given the typical clockwise Lagrangian path around the Q-R plane, $Q < 0$, $\Delta > 0$, $R < 0$ is where vorticity is first established (as the flow crosses the zero-discriminant line $\Delta = 0$ at $R < 0$, and the eigenvalues of $\boldsymbol{A}$ transition from all being real to one real value and a conjugate pair, resulting in $\|\boldsymbol{\Omega}_B\|^2 > 0$). This asymmetry in the VGT behavior concerning R explains why it is important for a classification to capture its sign. Panels (b) and (c) also show that where the eigenvalues of $\boldsymbol{A}$ are real ($\Delta < 0$), stable hubs are preferentially associated with $\|\boldsymbol{\Omega}_B\|^2 < \|\boldsymbol{\Omega}_C\|^2 < \|\boldsymbol{S}_B\|^2$, i.e. the $\Omega_C(2)$ state in the x-axis label.

These results highlight why neither **C1** nor **C2** is sufficient for effective classification of the VGT dynamics since different signs of Q, R as well as Δ correspond to dynamically different configurations of the VGT. Let us highlight two examples of this in the following using the

results in Fig. 3: Panel (a) shows that where $\Delta > 0$ and $\|\boldsymbol{\Omega}_B\|^2 > \|\boldsymbol{S}_B\|^2$ (i.e. $Q > 0$) the preference is for hubs to arise when non-normality is larger than the normal straining. However, classification **C1** groups the $Q > 0$ cases from panel (a) with the $\Delta > 0$ cases in panels (b) and (c), and hubs do not arise when $R < 0$, $Q < 0$ and $\Delta > 0$ conflating dynamically distinct states. configurations. Panel (b) shows that stable hubs only arise significantly where normal enstrophy is zero and non-normality is small relative to normal straining (i.e., $R < 0$, $\Delta < 0$ and $\Omega_C(2)$ pertains). A similar tendency is seen in panel (c) for $R > 0$. However, classification **C2** conflates the distinct Δ_+ and Δ_- cases, indicating why **C3** is preferable to **C1** and **C2** as also seen in Fig. 2.

While the occurrence of stable hubs displays marked asymmetries concerning the signs of the VGT invariants (particularly the sign of R), stable non-hubs show a more symmetric behavior. Panels (b) and (c) in Fig. 3 indicate there is a preferred state for the stable non-hubs, which is approximately symmetric in R, and arises where $\Delta > 0$ and $\|\boldsymbol{\Omega}_C\|^2$ is the largest magnitude term. Hence, in two of the six regions defined by **C3** where hubs arise only rarely (and only at all for $R > 0$), non-hubs occur preferentially. Hence, again, including the signs of Q and Δ in the classification of the $Q - R$ plane (i.e. **C3**) is effective in separating out distinct nodal function. However, because $\|\boldsymbol{\Omega}_C\|^2$ would also appear to be relevant, and as this is only represented explicitly in **C2b**-**C5b** and **C6**, **C3b** is an effective extension to **C3**.

The last two panels in Fig. 3, (d) and (e), focus on the signs of the production terms. Given that the signs of the normal enstrophy production and normal strain production are opposite (where the former is non-zero) with the sign of the normal strain production equal to the sign of R [47], we focus on the classical strain and enstrophy production in (d) and the non-normal and interaction production in (e). The results between these panels are correlated due to the decompositions (10b) and (10c), and show that stable hubs arise preferentially when all production terms are positive, while the occurrence of stable non-hubs is no different to the nodes in general (the gray bars and white bars are similar throughout these panels). The favored positivity of the two terms shown in panel (d) is a well-known feature of small-scale turbulence. However, for HIT in general, [47] showed that while the interaction production was strongly biased towards positive values, the probability of positive values for the non-normal production was approximately 50%. Hence, the bias seen on the right-hand side of panel (e) is an interesting new result that shows how contributions to the production budget from the deviatoric pressure Hessian and viscous terms need to act to produce flow states that are effective in providing interconnections to other states.

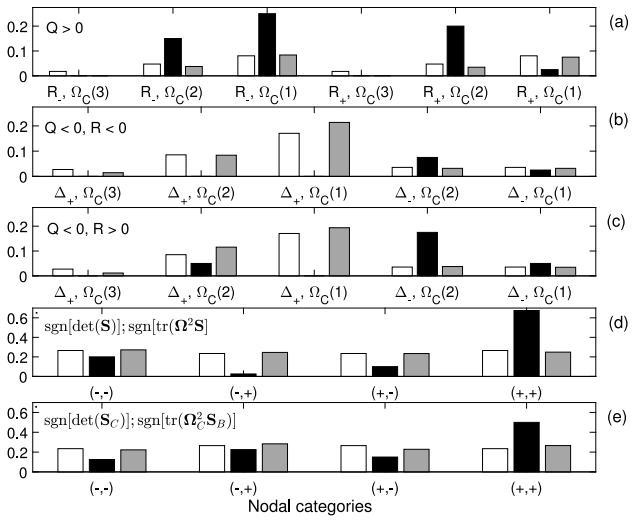

Fig. 3. The proportion of all nodes (white), stable hubs (black) and stable non-hubs (gray) that arise for different combinations of the terms used to analyze the velocity gradient tensor. Panels (a) to (c) consider the signs of Q, R, and Δ indicated by $\pm$ subscripts, and the rank order of $\|\boldsymbol{\Omega}_C\|^2$ relative to the normal enstrophy and normal straining when considered in descending order of magnitude. Hence, "$\Omega_C(1)$" for $Q > 0$ means that $\|\boldsymbol{\Omega}_C\|^2 > \|\boldsymbol{\Omega}_B\|^2 > \|\boldsymbol{S}_B\|^2$. Panels (d) and (e) give the signs of the classic production terms and the non-normal production and interaction production, respectively. The order in which these terms are considered is indicated in the top-left of each panel and then the relevant sign combinations are provided on the horizontal axis.

7. Discussion

The modularity entropy results in Fig. 2 suggest that the best classifiers of the VGT Lagrangian dynamics are **C2b**, **C3b**, and **C4b**. The difference in the values of $\mathcal{E}$ for these classifications relative to **C3**, and **C4** indicates that explicitly representing the non-normality via (23) and (24) is a more important way to develop the classification from the signs of Q and R (the definition of **C2**) than including the sign of Δ or the signs of the production terms featured in (10a), as arises for **C3** and **C4**, respectively.

7.1. A potential simplification based on nodal functionality

As shown in Fig. 3b, when it comes to examining nodal functionality, there is no clear distinction between $R < 0, \Delta < 0$ and $R > 0, \Delta < 0$, with stable hubs occurring at an approximately equal frequency in these two regions. This then raises the possibility that **C3b** can be enhanced by simplifying from six to five regions in the Q-R plane by merging the regions where $\Delta < 0$. This gives a new classification with $N_M = 14$ communities rather than $N_M = 16$ shown as the black, solid symbols in Fig. 2 and referred to as **C3c**. However, as the main panel of this figure shows, $\mathcal{E}$ is lower for the new case, and the inset shows that the reduction in N_m is not sufficient to increase $\delta\mathcal{E}$.

Fig. 4 compares the participation values of all nodes for classifications **C3b** and **C3c**. For all nodes where there is a significant departure from equality, **C3c** induces a loss of nodal diversity (a decline in ρ_i^O). The consequence is that stable hubs (triangles) or stable non-hubs (circles) acting as connectors in **C3b** are peripheral under **C3c**. This can be seen in more detail by looking at the two hubs that undergo the greatest change, which are labeled (i) and (ii) in Fig. 4. These are very similar in terms of the fluid dynamics they represent: Both correspond to $\Delta < 0$, $R < 0$, with positive signs of $-\det(\boldsymbol{S})$, $\mathrm{tr}(\boldsymbol{\Omega}^2\mathbf{S})$, $-\det(\boldsymbol{S}_C)$ and $\mathrm{tr}(\boldsymbol{\Omega}_C^2\mathbf{S}_B)$, with $|\mathrm{tr}(\boldsymbol{\Omega}_C^2\mathbf{S}_B)|$ the largest in magnitude of the production terms. There were only four stable hubs in the $\Delta < 0$, $R < 0$ region and (i) and (ii) were much more strongly connected to nodes with a different sign of R (approximately 25% of their out-degree compared to just 4% for the other two stable hubs in this region). Hence, collapsing the $\Delta < 0$ region into one grouping results in a large drop in ρ_i^O for these two nodes, as their different function compared to the other two nodes is lost in **C3c** when the sign of R is removed from the classification for the $\Delta < 0$ regions. Hence, merging nodes with all real eigenvalues into one classification is not an effective approach, despite the implication of the results in Fig. 3b. The restricted Euler solution to the dynamics for $\Delta > 0$ gives a movement from $R < 0$ to $R > 0$ [34], while the contribution from the deviatoric pressure Hessian is in the counter direction [56]. Incorporating the non-normality provides a means to include dynamic processes that cannot be written in terms of the eigenvalues of the VGT and, hence, having extended the **C3** classification to **C3b** to include such processes, **C3c** reduces the nodal participation for the two nodal hubs that are crucial for capturing the exchanges that arise in both directions where $\Delta < 0$.

7.2. Incorporating non-normality into the representation of the second and third invariants

Adopting an entropy form for the network modularity, $\mathcal{E}$, given by (16), and combining this with a consideration of nodal participation (19)–(21), allows us to identify effective classifications for the

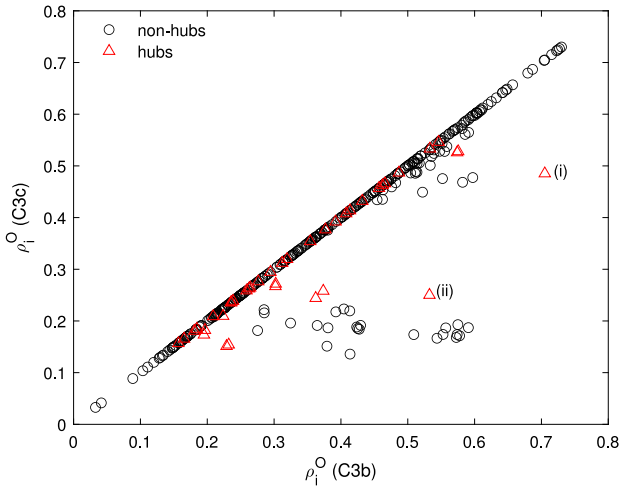

Fig. 4. Values of the participation ρ_i^O for classifications **C3b** and the reduced variant, **C3c**, with stable hubs shown as triangles and stable non-hubs as circles. The labels highlight two individual hubs discused in the text.

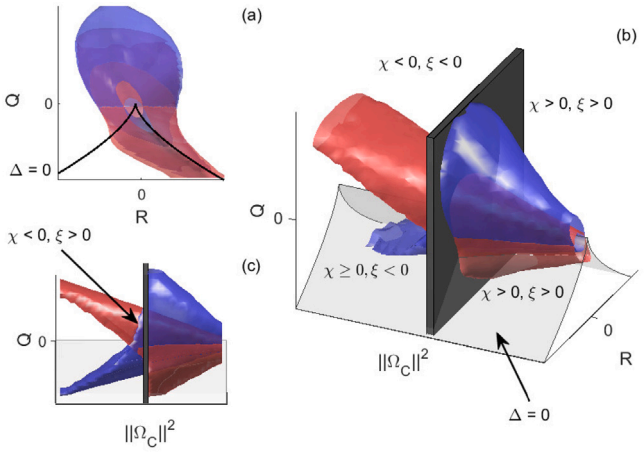

Fig. 5. A schematic of the extension of the $Q-R$ diagram into three dimensions based on the values for $\|\boldsymbol{\Omega}_C\|^2$ and thus, the signs of χ and ξ. Panel (a) presents the standard view of the Q and R axes, while (b) and (c) include the new axis. The line representing $\Delta = 0$ in (a) is a triangular-like surface in (b). The red and blue areas indicate where χ and ξ, respectively, are preferentially expressed. The vertical dark gray slice delimits the $\chi > 0, \xi > 0$ regions from the others. The arrow in panel (c) highlights a small region associated with $\chi < 0, \xi > 0$.

Lagrangian dynamics of the velocity gradient tensor (VGT). The key term to enhance the initial classifications is the non-normality, which is implemented using (23) and (24). The main panel of Fig. 2 shows that **C4b** yields the highest value for $\mathcal{E}$, but makes use of 24 different modules, compared to 12 and 16 for **C2b** and **C3b**. Incorporating this into the evaluation through (18) as shown in the inset of Fig. 2 indicates that **C2b** and **C3b** are also effective. These classifications are initially based on the signs of Q and R and, in the case of the latter, Δ. These terms are illustrated schematically in Fig. 5a, which provides a conventional view of the Q-R diagram. The non-normality is then provided by the additional axis in Fig. 5b,c. To the right of the vertical dark gray slice is where both ξ and χ are positive, i.e. $\|\boldsymbol{\Omega}_C\|^2$ is greater than either the normal enstrophy, $\|\boldsymbol{\Omega}_B\|^2$, or the normal total strain, $\|\boldsymbol{S}_B\|^2$. As the sign of Q dictates the relative magnitude of the two normal terms, we see on the right-hand side of panels (b) and (c) that the areas where χ is preferentially expressed (in red) are those where $Q < 0$, while ξ (in blue) is associated with $Q > 0$. This is not surprising as, for example, given $\|\boldsymbol{\Omega}_C\|^2$ is largest in magnitude, $\chi > 0$ will arise more readily where $\|\boldsymbol{\Omega}_B\|^2 < \|\boldsymbol{S}_B\|^2$.

There is a small region colored blue and highlighted by the arrow in panel (b) where $Q > 0$, $\chi < 0$ and $\xi > 0$, i.e. it is rare for $\|\boldsymbol{\Omega}_B\|^2 > \|\boldsymbol{\Omega}_C\|^2 > \|\boldsymbol{S}_B\|^2$ and this arises where $\|\boldsymbol{\Omega}_C\|^2$ is still relatively large in magnitude so that it exceeds a significantly non-zero normal total strain rate. Otherwise, the results to the left of the vertical dark gray slice separate into two much larger regions. In the red region, χ occurs preferentially and this is where $Q > 0$, $\chi < 0$ and $\xi < 0$, i.e. $\|\boldsymbol{\Omega}_B\|^2 > \|\boldsymbol{S}_B\|^2 > \|\boldsymbol{\Omega}_C\|^2$ and it is the negative values for χ that dominate. The $Q < 0$, $\chi < 0$ and $\xi > 0$ case tends to form the right Vieillefosse tail where the non-normality is low and is shown in blue. This arises close to or within the $\Delta < 0$ region where all the VGT eigenvalues are real. This imposes $\|\boldsymbol{\Omega}_B\|^2 = 0$ and, therefore $\xi > 0$ to high probability. However, as one moves to smaller values for $\|\boldsymbol{\Omega}_C\|^2$ in panels (b) and (c), $\chi > 0$ is less important than the large magnitude of the normal straining, which is why this region is colored blue as it is dominated by $\xi < 0$.

8. Conclusion

Applying methods from complex networks' theory [21,52] to a network encoding the VGT dynamics has shed light on the intricate time evolution and preferential flow state configurations of the VGT. In this paper, we have focused on extracting communities of VGT states and studying their interaction, as well as the extent to which different nodes within a community act as hubs or non-hubs, and how those hubs or non-hubs occur when certain fluid dynamical constraints are met. In particular, Fig. 3a shows that for $Q > 0$ the stable hub nodes require $\|\boldsymbol{\Omega}_C\|^2$ to be greater than at least one of $\|\boldsymbol{\Omega}_B\|^2$ or $\|\boldsymbol{S}_B\|^2$ and that stable hubs occur more frequently where the non-normality is largest in magnitude, and $R < 0$. Such regions are shown towards the right-hand side of Fig. 5b,c and arise preferentially where $Q > 0$ (highlighted in blue). However, conditional trajectory analysis indicates that the region affected directly by the deviatoric pressure Hessian in the most complex manner is the adjacent $Q > 0$, $R > 0$ region [39,56]. From this region, a significant proportion of trajectories exit to the $Q > 0$, $R < 0$ region rather than $Q < 0$, $\Delta > 0$, $R > 0$ as anticipated from the restricted Euler model [34]. Given the restricted Euler model is based on the normal dynamics, these trajectories in the counter-direction will preferentially be those with higher non-normality. That hubs arise preferentially where non-normality (and the production of non-normality) is large highlights the significance of the dynamical contribution from the deviatoric pressure Hessian for establishing flow states that are critical for the Lagrangian dynamics.

A comparison between our results and those of the enhanced Gaussian closure (EGC) model [28] shows that the latter does very well at correctly predicting the out-degree of a given flow state (Fig. 1). However, the modularity entropy results in Fig. 2 show that the EGC model does not see the same gain in modularity entropy that occurs for the direct numerical simulation when the non-normality is incorporated into the nodal classifications. This means that current state-of-the-art modeling approaches can be enhanced by explicitly including the non-normality in reduced-order models that develop further the formalism seen in (22). Based on the results in this paper, formulating such an approach for stochastic closures is something we believe is an important direction for future work.

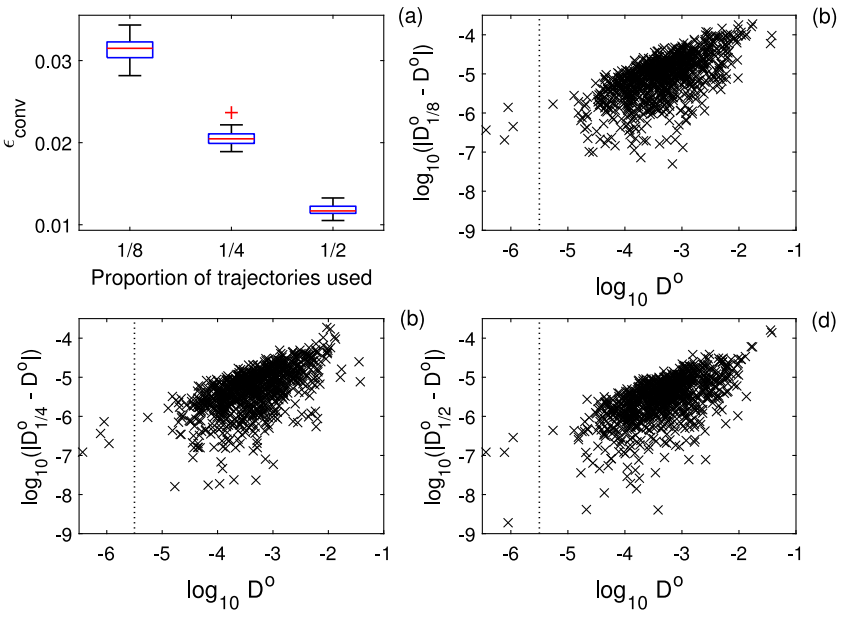

Fig. A.6. Values for the bootstrapped mean convergence error, ϵ_{conv} as a function of ℓ, the proportion of trajectories used in the bootstrap are given by the boxplots in panel (a). Each box is based on 50 realizations with the median shown by the central line in the box, the upper and lower quartiles by the upper and lower edges of the box, the whiskers extending up to 1.5 times the quartile deviation and outliers shown with a "+". Panels (b)–(d) then show the dependence of the absolute error on the out-degree for each choice of ℓ for one of the fifty realizations used to form the boxes in panel (a). The vertical dotted lines separate the four excluded rarely occurring nodes from the other 844 nodes.

CRediT authorship contribution statement

Christopher J. Keylock: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Maurizio Carbone:** Writing – review & editing, Methodology, Investigation.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Chris Keylock reports financial support was provided by The Leverhulme Trust. The other author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors gratefully acknowledge the Leverhulme Trust International Fellowship 2023-014 that permitted CK's visit to Bayreuth and enabled this research. They are extremely grateful to Michael Wilczek for hosting CK in Bayreuth.

Appendix. Sampling convergence

For a directed network, the out-degree, $D_i^{out} = \sum_j w_{ij}$ for any given node may not equate to the in-degree, $D_j^{in} = \sum_i w_{ij}$. However, in our case, where we have a circulation, even though our network is directed, it should be the case that $\sum_j D_j^{in} = \sum_i D_i^{out}$ up to precision errors resulting from the finite nature of the simulation and the evaluation of the flow states. With a normalization of $\boldsymbol{M}$ such that $\sum_j D_j^{in} = \sum_j D_i^{out} = 1$, then the expected degree for the i'th node is $\langle D_i^{in} \rangle = \langle D_i^{out} \rangle = 1/N$. In our case, the median error for the DNS was $|D_j^{in} - D_i^{out}|_{DNS} = 1.32 \times 10^{-6}$, or approximately 1000 times smaller than $\langle D_i^{out} \rangle$. For the enhanced Gaussian closure, the value was roughly an order of magnitude smaller, $|D_j^{in} - D_i^{out}|_{EGC} = 1.04 \times 10^{-7}$.

Considering the out-degree for the DNS in greater detail, we undertook a bootstrap of our results to determine the sensitivity of the network weights to the number of trajectories employed in the analysis. On average, there was $0.19\tau_\eta$ (or 3.79 samples) between state transitions, indicating that our analysis captured approximately 25 million transitions and that the temporal sampling is sufficiently resolved to (correctly) introduce temporal autocorrelation into individual Lagrangian trajectories. Because the initial spatial separation of each trajectory was designed to reduce spatial autocorrelation between trajectories, the bootstrapping was undertaken by selecting subsets of trajectories rather than perturbing particular trajectories. To form the bootstrap, on fifty separate occasions, we randomly sampled (without replacement) $\ell \in \{1/8, 1/4, 1/2\}$ of the 27^3 trajectories and, for each case formed an adjacency matrix. Because values for the D_i^{out} vary over several orders of magnitude and the error for node i is a strong function of D_i^{out}, we made use of a relative convergence error term

$$\epsilon_{conv} = \left\langle \frac{|D^{out} - D^{out}_\ell|}{D^{out}} \right\rangle, \tag{A.1}$$

where the averaging indicated by the angled braces is undertaken over all nodes, and the explicit dependence on node i has been dropped from the notation for compactness. Fig. A.6a shows that the typical estimation error for the out-degree reduced from over 3% when $1/8$ of the trajectories were used to less than 1.5% when $1/2$ of the trajectories were used. Panel (a) summarizes the results for 50 random realizations corresponding to different ℓ, while the remaining three panels display the out-degree of all nodes in one of the fifty realizations, with ℓ changing for each panel. These panels separate the terms on the numerator and denominator of (A.1) and adopt a log scale for visibility. It has already been noted that four nodes were excluded from further analysis in this paper, and these occur to the left of the vertical dotted lines in Fig. A.6b-d.

Fig. A.7 examines the convergence error for the modularity entropy, $\mathcal{E}$. Eight of the eleven cases from Fig. 2 are considered (the cases excluded are **C1** and **C2**, which had much lower values for $\mathcal{E}$ and the additional classification, **C3c** introduced in the discussion section). The upper panel shows that even when using $\ell = 1/8$ of the original trajectories, the errors are small relative to the degree of variation in $\mathcal{E}$ between classifications such that all are clearly distinguished. Defining $\epsilon_\mathcal{E}$ as the relative error in the estimate of ϵ, i.e.

$$\epsilon_\mathcal{E} = \frac{\mathcal{E}_\ell - \mathcal{E}}{\mathcal{E}}, \tag{A.2}$$

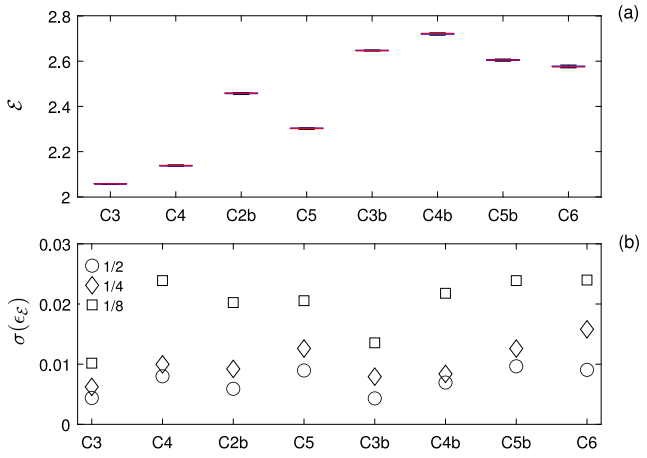

Fig. A.7. Boxplots of the bootstrapped values for the modularity entropy, $\mathcal{E}$ for eight classifications from Fig. 2 using one eighth of the 27^3 trajectories are shown in panel (a). Panel (b) gives the standard error of the bootstrapped terms relative to the value used in this study for one half (circles), one quarter (diamonds) and one eighth (squares) of the 27^3 trajectories.

then the standard error of this quantity is given in Fig. A.7b for $\ell \in \{1/8, 1/4, 1/2\}$. The standard error, $\sigma(\epsilon_{\mathcal{E}})$, for $\ell = 1/2$ remains below 1% for all classifications, while for $\ell = 1/4$ it averages 1% and for $\ell = 1/8$ it averages just under 2%. There is a general tendency for the error to increase with N_m (from left to right on the x-axis), although sampling would appear to be sufficient to ensure that this effect is weak. Hence, our use of approximately 25 million changes in flow state occurring on just under 20 000 Lagrangian tracks would appear to be sufficient to obtain converged results.

Data availability

The data used are already in the public domain. We deliberately used a public simulation database to aid reproducibility (https://turbulence.idies.jhu.edu/home)

References

[1] J. Hadamard, Les surfaces a courbures opposées e leur liques géodesiques, J. Math. Pures Appl 4 (1898) 27–73.

[2] R. Bowen, Symbolic dynamics for hyperbolic flows, Am. J. Math 95 (1973) 429–460.

[3] A.D. Kirkbride, R. Ferguson, Turbulent flow structure in a gravel-bed river: Markov chain analysis of the fluctuating velocity profile, Earth Surf. Proc. Landforms 20 (1995) 721–733, http://dx.doi.org/10.1002/esp.3290200804.

[4] C.J. Keylock, S.N. Lane, K.S. Richards, Quadrant/octant sequencing and the role of coherent structures in bed load sediment entrainment, J. Geophys. Res. 119 (2014) 264–286, http://dx.doi.org/10.1002/2012JF002698.

[5] Beim Graben, Estimating and improving the signal-to-noise ratio of time series by symbolic dynamics, Phys. Rev. E 64 (2001).

[6] M. McCullough, M. Small, T. Stemler, H.C. Iu, Time lagged ordinal partition networks for capturing dynamics of continuous dynamical systems, Chaos 25 (2015).

[7] Q.K. Telesford, S.L. Simpson, J.H. Burdette, S. Hayasaka, P.J. Laurienti, The brain as a complex system: Using network science as a tool for understanding the brain, Brain Connect. 1 (2011) 295–308, http://dx.doi.org/10.1089/brain.2011.0055.

[8] J. Wang, H. Mo, F. Wang, F. Jin, Exploring the network structure and nodal centrality of China's air transport network: A complex network approach, J. Transp. Geog 19 (2011) 712–721, http://dx.doi.org/10.1016/j.jtrangeo.2010.08.012.

[9] P. Crucitti, V. Latora, M. Marchiori, Model for cascading failures in complex networks, Phys. Rev. E 69 (4) (2004).

[10] A.G. Nair, K. Taira, Network-theoretic approach to sparsified discrete vortex dynamics, J. Fluid Mech. 768 (2015) 549–571.

[11] K. Taira, A.G. Nair, S.L. Brunton, Network structure of two-dimensional decaying isotropic turbulence, J. Fluid Mech. 795 (2016) http://dx.doi.org/10.1017/jfm.2016.235.

[12] G. Iacobello, S. Chowdhuri, L. Ridolfi, L. Rondoni, S. Scarsoglio, Coherent structures at the origin of time irreversibility in wall turbulence, Commun. Phys. 6 (2023) 91, http://dx.doi.org/10.1038/s42005-023-01215-y.

[13] A.E. Perry, M.S. Chong, A study of eddying motions and flow patterns using critical point concepts, Ann. Rev. Fluid Mech 19 (1987) 125–155.

[14] B. Lüthi, M. Holzner, A. Tsinober, Expanding the Q-R space to three dimensions, J. Fluid Mech. 641 (2009) 497–507.

[15] J.C.R. Hunt, A.A. Wray, P. Moin, Eddies, Stream, and Convergence Zones in Turbulent Flows. Technical Report CTR-S88. Center for Turbulence Research, Stanford University, 1988.

[16] Y. Dubief, F. Delcayre, On coherent-vortex identification in turbulence, J. Turbul. 1 (2000).

[17] P. Chakraborty, S. Balachandar, R.J. Adrian, On the relationships between local vortex identification schemes, J. Fluid Mech. 535 (2005) 189–214.

[18] H. Xu, X.S. Cai, C. Liu, Liutex (vortex) core definition and automatic identification for turbulence vortex structures, J. Hydrodyn. 31 (2019) 857–863.

[19] B. Bollobás, P. Erdös, Cliques in random graphs, Math. Proc. Camb. Phil. Soc 80 (1976) 419–427.

[20] S. Fortunato, Community detection in graphs, Phys. Rep. 486 (2010) 75–174.

[21] R. Guimerà, L.A. Nunes Amaral, Cartography of complex networks: modules and universal roles, J. Stat. Mech. (2005) http://dx.doi.org/10.1088/1742-5468/2005/02/P02001.

[22] G. Palla, I. Derényi, I. Farkas, T. Vicsek, Uncovering the overlapping community structure of complex networks in nature and society, Nature 435 (2005) 814–818, http://dx.doi.org/10.1038/nature03607.

[23] A. Bavelas, Communication patterns in task-oriented groups, J. Acoust. Soc. Am. 22 (1950) 725–730.

[24] L.C. Freeman, A set of measures of centrality based upon betweenness, Sociometry 40 (1977) 35–41.

[25] P. Erdös, A. Rényi, On random graphs. i, Publ. Math. 6, 290–297.

[26] A.L. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (1999) 509–512.

[27] Y. Li, E. Perlman, M. Wan, Y. Yang, R. Burns, S. C.MeneveauChen, A. Szalay, G. Eyink, A public turbulence database cluster and applications to study Lagrangian evolution of velocity increments in turbulence, J. Turbul. 9 (2008).

[28] M. Wilczek, C. Meneveau, Pressure hessian and viscous contributions to velocity gradient statistics based on Gaussian random fields, J. Fluid Mech. 756 (2014) 191–225.

[29] R. Betchov, An inequality concerning the production of vorticity in isotropic turbulence, J. Fluid Mech. 1 (1956) 497–504.

[30] P. Vieillefosse, Local interaction between vorticity and shear in a perfect incompressible fluid, J. Phys. Fr. 43 (1982) 837–842, http://dx.doi.org/10.1051/jphys:01982004306083700.

[31] M. Itskov, Tensor Algebra and Tensor Analysis for Engineers, Springer International Publishing, Switzerland, 2015, http://dx.doi.org/10.1007/978-3-319-16342-0.

[32] M. Carbone, M. Wilczek, Asymptotic predictions on the velocity gradient statistics in low-Reynolds number random flows: Onset of skewness, intermittency and alignments, 2023, arXiv:2305.03454.

[33] B. Sharma, R. Das, S.S. Girimaji, Local vortex line topology and geometry in turbulence, J. Fluid Mech. 924 (A13) (2021) http://dx.doi.org/10.1017/jfm.2021.613.

[34] B.J. Cantwell, On the behavior of velocity gradient tensor invariants in direct numerical simulations of turbulence phys, Fluids A 5 (1993) 2008–2013, http://dx.doi.org/10.1063/1.858828.

[35] B.J. Cantwell, Exact solution of a restricted Euler equation for the velocity gradient tensor, Phys. Fluids A 4 (1992) 782–793.

[36] A. Tsinober, L. Shtilman, H. Vaisburd, A study of properties of vortex stretching and enstrophy generation in numerical and laboratory turbulence, Fluid Dyn. Res. 21 (1997) 477–494.

[37] S.S. Girimaji, S.B. Pope, A diffusion model for velocity gradients in turbulence, Phys. Fluids 2 (1990) 242–256.

[38] J. Martiín, A. Ooi, M.S. Chong, J Soria, Dynamics of the velocity gradient tensor invariants in isotropic turbulence, Phys. Fluids 10 (1998) http://dx.doi.org/10.1063/1.869752.

[39] Chevillard L., C. Meneveau, L. Biferale, F. Toschi, Modeling the pressure hessian and viscous Laplacian in turbulence: comparisons with direct numerical simulation and implications on velocity gradient dynamics, Phys. Fluids 20 (2008) 101504.

[40] P.L. Johnson, C. Meneveau, A closure for Lagrangian velocity gradient evolution in turbulence using recent-deformation mapping of initially Gaussian fields, J. Fluid Mech. 804 (2016) 387–419.

[41] D. Buaria, K.R. Sreenivasan, Forecasting small-scale dynamics of fluid turbulence using deep neural networks, in: Proc. Natl. Acad. Sci. U.S.A, vol. 120, 2023.

[42] R. Das, S.S. Girimaji, Data-driven model for lagrangian evolution of velocity gradients in incompressible turbulent flows, J. Fluid Mech. 984 (A39) (2024) http://dx.doi.org/10.1017/jfm.2024.235.

[43] M. Carbone, V.J. Peterhans, A.S. Ecker, M. Wilczek, Tailor-designed models for the turbulent velocity gradient through normalizing flow, Phys. Rev. Lett. 133 (2024) 184001.

[44] C. Meneveau, Lagrangian dynamics and models of the velocity gradient tensor in turbulent flows, Ann. Rev. Fluid Mech 43 (2011) 219–245.
[45] P.L. Johnson, M. Wilczek, Multiscale velocity gradients in turbulence, Ann. Rev. Fluid Mech 56 (2024) 463–490, http://dx.doi.org/10.1146/annurev-fluid-121021-031431.
[46] I. Schur, On the characteristic roots of a linear substitution with an application to the theory of integral equations, Math. Ann. 66 (1909) 488–510.
[47] C.J. Keylock, The Schur decomposition of the velocity gradient tensor for turbulent flows, J. Fluid Mech. 848 (2018) 876–904.
[48] C.J. Keylock, Turbulence at the Lee bound: maximally non-normal vortex filaments and the decay of a local dissipation rate, J. Fluid Mech. 881 (2019) 283–312, http://dx.doi.org/10.1017/jfm.2019.779.
[49] C.J. Keylock, Studying turbulence structure near the wall in hydrodynamic flows: An approach based on the Schur decomposition of the velocity gradient tensor, J. Hydrodyn. 34 (2022) 806–825, http://dx.doi.org/10.1007/s42241-022-0068-6.
[50] P. Beaumard, O.R.H. Buxton, C.J. Keylock, The importance of nonnormal contributions to velocity gradient tensor dynamics for spatially developing, inhomogeneous, in: Turbulent Flows, 2019.
[51] C.J. Keylock, M. Carbone, Complex network approach to the turbulent velocity gradient dynamics: High- and low-probability Lagrangian paths, Phys. Rev. Fluids 10 (2025) http://dx.doi.org/10.1103/PhysRevFluids.10.054606.
[52] M.E.J. Newman, M. Girvan, Finding and evaluating community structure in networks, Phys. Rev. E 69 (2004) http://dx.doi.org/10.1103/PhysRevE.69.026113.
[53] P. Cajic, D. Agius, O.M. Cliff, J.M. Shine, J.T. Lizier, B.D. Fulcher, On the information-theoretic formulation of network participation, J. Phys.: Complex. (2024).
[54] A. Ooi, J. Martin, J. Soria, M.S. Chong, A study of the evolution and characteristics of the invariants of the velocity-gradient tensor in isotropic turbulence, J. Fluid Mech. 381 (1999) 141–174.
[55] P. Vieillefosse, Internal motion of a small element of fluid in an inviscid flow, Phys. A 125 (1984) 150–162.
[56] K. Zhou, Y. Sakai, Y. Ito, T. Hayase, On the evolution of the invariants of the velocity gradient tensor in single square-grid-generated turbulence, Phys. Fluids 27 (2015) http://dx.doi.org/10.1063/1.4926472.