Complex network approach to turbulent velocity gradient dynamics: High- and low-probability Lagrangian paths

C. J Keylock [1,2,*] and M. Carbone [2,3,4]

[1] *School of Architecture, Building and Civil Engineering, Loughborough University, Epinal Way, Loughborough, Leicestershire LE11 3TU, United Kingdom*

[2] *Theoretical Physics I, University of Bayreuth, 95440 Bayreuth, Germany*

[3] *Istituto dei Sistemi Complessi, CNR, Via dei Taurini 19, 00185 Rome, Italy*

[4] *INFN "Tor Vergata", Via della Ricerca Scientifica 1, 00133 Rome, Italy*

(Received 9 April 2024; accepted 25 April 2025; published 15 May 2025)

Understanding the dynamics of the turbulent velocity gradient tensor (VGT) is essential to gain insights into the Navier-Stokes equations and improve small-scale turbulence modeling. However, characterizing the VGT dynamics conditional on all its relevant invariants in a continuous fashion is extremely difficult. Hence, in this paper, we represent the Lagrangian dynamics using a network where each node represents a unique flow state in the space of the invariants of the VGT. The sign and relative magnitude ranking of the invariants resulting from a Schur decomposition of the VGT determines the discrete flow state vector. Our analysis reveals intriguing features of the resulting network dynamics, such as a grouping of the influential nodes where the eigenvalues of the VGT are real, in the proximity of the right Vieillefosse tail. We relate our complex network approach to the well-established VGT discretization based on the sign of its principal invariants, Q and R, and its discriminant, Δ into six regions representing very different physical and topological states. We study the shortest paths on the network based on the six regions to which their starting and ending nodes belong. This analysis shows that the typically clockwise path on the Q-R diagram arises in a varying manner: sometimes region transitions are based on multiple pathways of more equal probability, while others are focused on specific nodes that are pivotal in these interregion dynamics. Comparisons to an enhanced Gaussian closure model show that the model captures the out-degree distribution for the nodes well. However, it uses a greater number of nodes to transition from real to complex eigenvalues and is far less efficient at generating clockwise motions that are counter to the restricted Euler dynamics.

DOI: 10.1103/PhysRevFluids.10.054606

I. INTRODUCTION

Understanding the properties of turbulent fluid flows at small scales is crucial for gaining insights into the mechanisms by which flows transfer and dissipate energy [1–3] and how this relates to extreme events [4,5]. This is important both for modeling the velocity gradients [6] and for understanding the physical mechanisms that lead to constraints on the smoothness of the

*Contact author: C.J.Keylock@lboro.ac.uk

underpinning Navier-Stokes equations [7–10]

$$\nabla \cdot \boldsymbol{u} = 0, \quad \frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u} \cdot \nabla)\boldsymbol{u} = -\nabla p + \nu \nabla^2 \boldsymbol{u}. \tag{1}$$

Here $\boldsymbol{u}$ is the three-dimensional velocity vector, t is time, p is the pressure field divided by the constant density, and ν is the kinematic viscosity. The smallest scales and extreme events are conveniently addressed from the viewpoint of the velocity gradient tensor (VGT), $\boldsymbol{A} \equiv \nabla \boldsymbol{u}$. The governing equation for the VGT follows by taking the spatial gradient of the Navier-Stokes equations (1)

$$\mathrm{tr}(\boldsymbol{A}) = 0, \quad \frac{\partial \boldsymbol{A}}{\partial t} + (\boldsymbol{u} \cdot \nabla)\boldsymbol{A} = -\left(\boldsymbol{A}^2 - \frac{\mathrm{tr}(\boldsymbol{A}^2)}{3}\boldsymbol{I}\right) - \boldsymbol{H} + \nu \nabla^2 \boldsymbol{A}, \tag{2}$$

where $\boldsymbol{H}$ is the deviatoric part of the pressure Hessian, $\boldsymbol{H} \equiv \nabla\nabla p - \nabla^2 p \boldsymbol{I}/3$, and $\boldsymbol{I}$ is the identity matrix in three dimensions. Equation (2) has been studied extensively, especially concerning its reduced-order modeling [11–15], as reviewed by [6,16]. Reduced-order modeling is crucial when Eq. (2) is interpreted from a Lagrangian perspective, that is, when one considers the dynamic evolution of the VGT along a fluid particle trajectory. Both the pressure Hessian and viscous terms are nonlocal functionals of the velocity gradient [17], and they are not uniquely determined by the knowledge of $\boldsymbol{A}$ at a single point in space. From this single-particle/single-time viewpoint, Eq. (2) is unclosed, and the unclosed/nonlocal terms $\boldsymbol{H}$ and $\nu\nabla^2\boldsymbol{A}$ require modeling.

Of particular importance in this field is the restricted Euler model [18,19], consisting of Eq. (2) without the deviatoric pressure Hessian and viscous contributions. The restricted Euler model provides qualitative insights into the actual turbulent VGT dynamics and also helps introduce several quantities crucial for the VGT dynamics, extensively studied in the literature, such as the VGT principal invariants, the dissipation rate, enstrophy, etc. The restricted Euler model yields two coupled ordinary differential equations for the evolution of the second, $Q = -\mathrm{tr}(\boldsymbol{A}^2)/2$, and third, $R = -\det(\boldsymbol{A})$, principal invariants of the VGT, where $\mathrm{tr}(\ldots)$ is the trace and $\det(\ldots)$ is the determinant:

$$\frac{dQ}{dt} = -3R, \quad \frac{dR}{dt} = \frac{2}{3}Q^2. \tag{3}$$

Solutions to the model (3) follow lines of constant discriminant,

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

and thus, proceed from negative R and positive or negative Q to a state give by positive R and negative Q, with a timescale $\tau = |\Delta|^{-1/6}$. This results in a finite-time blow-up for almost all initial conditions along the Vieillefosse tail [11,18]. Contrarily, this finite-time blow-up is not observed in the numerical solution of the Navier-Stokes equations, in which the phase-space trajectories explore the full Q-R diagram [20,21]. However, the statistical distribution of the phase-space trajectories display signatures of the restricted Euler model, with values skewed towards positive R and negative Q (see Fig. 1). For a Taylor Reynolds number of $\mathrm{Re}_\lambda = 71$, it has been shown that a typical orbit around the Q-R diagram takes approximately 30 Kolmogorov times [22], which for that simulation equated to about three eddy turnover times.

Remarkably, within the restricted Euler model, the evolution of the VGT principal invariants Q and R is independent of the other VGT invariants [22], as in equation (3). Extending the restricted Euler model to the full VGT requires consideration of the deviatoric part of the pressure Hessian $\boldsymbol{H}$ and the viscous effects stemming from $\nu\nabla^2\boldsymbol{A}$ in (2), and how their contributions, conditional on the local VGT state, make the VGT invariants interact with each other. Several studies have proposed phenomenological models and statistical closures for these nonlocal terms (e.g., [21,23–25]) which can qualitatively capture the VGT statistics, and recent machine learning methods have provided new approaches to prediction and modeling [26,27]. These latter methods require relatively

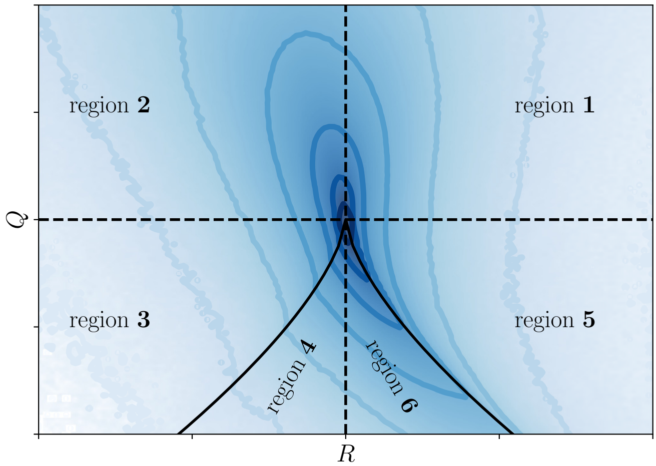

FIG. 1. Contours of the joint distribution function of the VGT principal invariants, Q and R, with the six regions identified based on the signs of Q, R (dashed lines) and Δ (solid lines). The latter lines indicate the Vieillefosse tails with the joint distribution of Q and R biased to the right tail.

advanced computational tools and very large datasets. In this work, we propose a complementary description of the VGT, based on projecting the VGT Lagrangian dynamics on to a complex network. This allows us to gain an intuitive and lower-complexity understanding of small-scale turbulence based on discrete VGT states, crucial for improving, e.g., flow classifications [28]. Furthermore, we can employ the tools from network theory [29] to characterize the VGT dynamics indirectly, by studying the properties of the associated complex network.

II. COMPLEX NETWORK APPROACH TO THE VGT DYNAMICS

Modeling the VGT Lagrangian dynamics remains challenging even in the case of homogeneous, isotropic turbulence (HIT). This is due to the high dimensionality of the problem since the phase space consists of eight continuous variables (the independent components of the VGT) with their intricate interplay and time correlations. To circumvent some of these difficulties, in this paper, we propose a different approach based on projecting the dynamics of the VGT onto a discrete network. We define the network nodes by partitioning the phase space into discrete regions according to the sign and relative magnitudes of relevant VGT invariants. From a statistical viewpoint, classic Lagrangian VGT modeling aims at reproducing the relevant joint distribution functions of the VGT, while the present complex network approach amounts to binning those joint distributions on a coarsely partitioned phase space. This considerably reduces the degrees of freedom and makes it easier to explore the interplay of multiple flow features.

Using complex networks to study nonlinear phenomena is an important strand of nonlinear physics (e.g., [30,31]). In the context of turbulence research, applications of network-based approaches are not as numerous. However, they have been used to study, for example, velocity induction by mutual vortex interaction in 2D turbulence [32,33] coherent structures [34], the intermittency of a transported temperature field [35], time irreversibility [36] and the manner by which network analyses of large-scale mixing and dispersion can be related to Lyapunov exponent characteristics [37]. The majority of previous network-based studies in turbulence science have defined the weighted adjacency matrix in terms of physical-space realizations of the flow. The approach adopted here is to define the nodes based on the VGT phase-space variables and for the adjacency matrix to then represent the transition between states. Specifically, we employ scalar invariants that can be formed from the VGT [38–41] and VGT components obtained from its complex Schur decomposition [42] to define our network nodes, which is a unique aspect of our study. In common with the work of Fernex *et al.* [43], we also form our adjacency/transition matrix such that we only consider transitions between nodes, rather than the probability of remaining at a given node. In contrast to those authors and Iacobello *et al.* [44], both of whom used statistical procedures to obtain the optimal number of nodes for the problem of interest, the number of nodes adopted here is given by the physics of the VGT on the assumption that the signs and relative ranking of the magnitudes of the invariants capture the physical and topological characteristics. For example, we include the sign of Δ from (4) as this indicates whether or not the local flow has a closed contour in the Lagrangian frame and, thus, can be considered to be a coherent flow structure [45] or indicates the existence of a normal enstrophy term [42]. Lagrangian turbulence data have also been used by Perrone *et al.* [46] in their network-based analysis of a turbulent channel flow, although in their case, the nodes were values for the distance from the wall, y^+, discretized into 100 levels rather than formed from a state-space vector.

A. From tensor invariants to a state-space vector

The VGT scalar invariants provide insight into the VGT state since they relate to enstrophy, dissipation, production terms, and flow topology. They are even more relevant in statistically isotropic turbulence in which the VGT single-time statistics can be parameterized in terms of only five independent invariants [39,41]. We can make a variety of choices on the independent invariants and we begin with those complementing the classical Q-R picture (3) [11,13]. A Hermitian or skew-Hermitian decomposition consists in splitting the VGT into its symmetric part, the strain rate $\boldsymbol{S} = (\boldsymbol{A} + \boldsymbol{A}^*)/2$, and its antisymmetric part, the rotation rate $\boldsymbol{\Omega} = (\boldsymbol{A} - \boldsymbol{A}^*)/2$, where the asterisk indicates complex conjugation and matrix transposition. Using this decomposition one can rewrite the principal invariants for incompressible flow as

$$Q = \tfrac{1}{2}(\|\boldsymbol{\Omega}\|^2 - \|\boldsymbol{S}\|^2), \tag{5}$$

$$R = -\det(\boldsymbol{S}) - \mathrm{tr}(\boldsymbol{\Omega}^2\boldsymbol{S}), \tag{6}$$

where $\| \ldots \|$ denotes the Frobenius norm. The second invariant is thus related to the enstrophy and dissipation rate while the third relates to the strain production and enstrophy production. It is clear from (5) and (6) that the signs of the second and third invariants encapsulate distinct physical states. The physical interpretation of Q is straightforward as both the total strain and the enstrophy are non-negative. Hence, for example, Hunt *et al.* [47] used $Q > 0$ as a criterion to delimit coherent structures in the flow. Interpretation of (6) is rather more complex as a given sign for R may arise as a consequence of different signs for the strain production and enstrophy production terms, or a difference in their magnitudes for the same sign. In practice, the latter situation is more probable given that both these production terms are positive on average [48].

Consideration of the above points is why we formulate our state-space vector in terms of the signs and relative magnitudes of the various invariants considered. However, rather than restricting ourselves to the quantities in (4)–(6), we pursue a more detailed approach introduced by Keylock [42]. This permits us to discriminate between the dynamics that are related to the eigenvalues of the VGT and those related to the departure of the eigenvectors from a unitary form. Hence, we express the VGT, $\boldsymbol{A}$, in terms of the complex Schur decomposition

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{T}\boldsymbol{U}^*, \quad \boldsymbol{T} = \boldsymbol{L} + \boldsymbol{N}, \tag{7}$$

where $\boldsymbol{U}$ is a unitary rotations matrix, and the Schur matrix, $\boldsymbol{T}$ may be separated into a diagonal matrix of eigenvalues, $\boldsymbol{L}$, and the strictly upper-triangular non-normality matrix, $\boldsymbol{N}$ [49]. By defining the normal matrix, $\boldsymbol{B}$, and the non-normal matrix, $\boldsymbol{C}$, and decomposing them into their Hermitian and anti-Hermitian parts

$$\boldsymbol{B} = \boldsymbol{U}\boldsymbol{L}\boldsymbol{U}^* = \boldsymbol{S}_B + \boldsymbol{\Omega}_B, \quad \boldsymbol{C} = \boldsymbol{U}\boldsymbol{N}\boldsymbol{U}^* = \boldsymbol{S}_C + \boldsymbol{\Omega}_C, \tag{8}$$

we write the second-order invariants in terms of the normal enstrophy, normal strain, and non-normality

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

$$\|\boldsymbol{S}\|^2 = \|\boldsymbol{S}_B\|^2 + \|\boldsymbol{\Omega}_C\|^2. \tag{10}$$

Here $\boldsymbol{S}_B = (\boldsymbol{B} + \boldsymbol{B}^*)/2$ denotes the Hermitian part of $\boldsymbol{B}$ while $\boldsymbol{\Omega}_B = (\boldsymbol{B} - \boldsymbol{B}^*)/2$ denotes its anti-Hermitian part (with the same decomposition for $\boldsymbol{C}$). Thus, it now follows from (5), (9) and (10) that there are three second-order terms that are relevant to the VGT dynamics, and $\|\boldsymbol{\Omega}_C\|^2$ cancels out when forming Q. Hence, while the sign of Q establishes the relative magnitudes of $\|\boldsymbol{\Omega}_B\|^2$ and $\|\boldsymbol{S}_B\|^2$, we need to consider the magnitude of $\|\boldsymbol{\Omega}_C\|^2$ relative to the other two terms to obtain a more complete description of the flow using our state-space vector.

Using (8) we can rewrite the strain production and enstrophy production as

$$-\det(\boldsymbol{S}) = -\det(\boldsymbol{S}_B) - \det(\boldsymbol{S}_C) + \operatorname{tr}\big(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B\big), \tag{11}$$

$$\operatorname{tr}(\boldsymbol{\Omega}^2\boldsymbol{S}) = \operatorname{tr}\big(\boldsymbol{\Omega}_B^2\boldsymbol{S}_B\big) - \det(\boldsymbol{S}_C) + \operatorname{tr}\big(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B\big), \tag{12}$$

where the three terms on the right-hand side of (11) represent the normal strain production, $-\det(\boldsymbol{S}_B)$, non-normal production, $-\det(\boldsymbol{S}_C)$, and interaction production, $\operatorname{tr}(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B)$; the additional term on the right-hand side of (12) is the normal enstrophy production, $\operatorname{tr}(\boldsymbol{\Omega}_B^2\boldsymbol{S}_B)$. Therefore, rather than just considering the signs and relative magnitudes of the two terms on the left-hand side of (11) and (12) as would be standard, we also consider the signs and magnitudes of the terms on the right-hand sides of these equations.

From the above discussion, we use the invariants and associated quantities described to define the discrete state-space vector $\overrightarrow{v}(t)$. This consists of three key components, which are displayed on separate lines of (13): on the first line, the signs of seven terms (128 possible states); the relative magnitudes of the three second order invariant terms on the second line; and the relative magnitudes of the four third order invariant terms on the third line:

$$\begin{aligned}\overrightarrow{v}(t) = \operatorname{sgn}\big\{&\big[Q; R; \Delta; -\det(\boldsymbol{S}); \operatorname{tr}(\boldsymbol{\Omega}^2\boldsymbol{S}); -\det(\boldsymbol{S}_C); \operatorname{tr}\big(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B\big)\big];\\ &\operatorname{rank}\big[\|\boldsymbol{\Omega}_B\|^2, \|\boldsymbol{\Omega}_C\|^2, \|\boldsymbol{S}_B\|^2\big];\\ &\operatorname{rank}\big[-\det(\boldsymbol{S}_B), -\det(\boldsymbol{S}_C), \operatorname{tr}\big(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B\big), \operatorname{tr}\big(\boldsymbol{\Omega}_B^2\boldsymbol{S}_B\big)\big]\big\},\end{aligned} \tag{13}$$

where sgn[...] indicates all the possible sign combinations of the arguments and rank[...] indicates descending ordering of the arguments. Semicolons are used to delimit the separate components of the state-space vector such that the first seven elements (on the first line) each has two possible states, the eighth element (on the second line) has six possible states, and the ninth (on the third line) has 24 possible states. Notice that the signs of $-\det(\boldsymbol{S}_B)$ and $\operatorname{tr}(\boldsymbol{\Omega}_B^2\boldsymbol{S}_B)$ do not need to be included in the definition of (13) since they are equal to, and opposite to, the sign of R, respectively, where $\|\boldsymbol{\Omega}_B\|^2 \neq 0$ (the case of $\|\boldsymbol{\Omega}_B\|^2 = 0$ and, thus, $\operatorname{tr}(\boldsymbol{\Omega}_B^2\boldsymbol{S}_B) = 0$, is dealt with by the negative sign for Δ).

Each node in the complex network is formed based on its own unique combination of signs/rankings of the terms in (13). For example, the Q-R decomposition, classically used to characterize the VGT and that we adopt for much of the paper, may be defined in terms of the signs of Q, R, and Δ, yielding the six regions sketched in Fig. 1. The total number of possible states in (13) is reduced by physical or mathematical constraints from the $2^7 \times 6 \times 24 = 18\,432$ states that appear to arise. For example, as already mentioned, when $\Delta < 0$ then $\|\boldsymbol{\Omega}_B\|^2 = 0$ and is therefore ranked the smallest of the terms given in the second row of (13). It then follows that $\operatorname{tr}(\boldsymbol{\Omega}_B^2\boldsymbol{S}_B) = 0$ and is ranked the smallest of the production terms in the third row of (13). Moreover, if $R > 0$, then $-\det(\boldsymbol{S}_B) > 0$ and, thus, from Eqs. (11) and (12) if the non-normal production and interaction production are positive then so must be $-\det(\boldsymbol{S})$. In addition, if $|\det(\boldsymbol{S}_C)| > |\operatorname{tr}(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B)|$

and $\det(\boldsymbol{S}_B) > 0$ then $-\det(\boldsymbol{S})$ will still be positive even if $\mathrm{tr}(\boldsymbol{\Omega}_C^2 \boldsymbol{S}_B) < 0$, etc. Taking into account all these constraints gives $N = 848$ possible states for the VGT based on (13), which form the nodes for our networks. However, in practice, four nodes that were visited extremely infrequently because of their unique physical characteristics were excluded from our analysis as detailed below.

B. Constructing the complex network

A network (graph) $\mathcal{G}$, without self-loops, comprises a set of N nodes (vertices), $\mathcal{V}$, edges $\mathcal{E}$, and weights, $\mathcal{W}$. The possible configurations of the discrete state $\overrightarrow{v}$ define the nodes in this study, while the weights in our adjacency matrix follow from the probability of transitioning from one node to any of the others. These are determined based on Lagrangian tracking within a flow simulation. We consider time realizations of the VGT, $\boldsymbol{A}(t)$, thus of the discretized state vector $\overrightarrow{v}(t)$, along an ensemble of fluid particle trajectories. Then we count the number of transitions from node i at some time, t, to node j at a later time $t + Z\delta t$, where Z is a positive integer and δt is the sampling step of the Lagrangian trajectories. With this Lagrangian approach the network edge weights act as a lens for the dynamics. We adopted a sub-Kolmogorov time resolution to provide as much detail as possible as described in greater detail in the Methods section below.

The weights may be considered via the $N \times N$ adjacency matrix $\boldsymbol{M}$ [29]

$$m_{ij} = \begin{cases} w_{ij} & \text{if } (i, j) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases}, \tag{14}$$

where the w_{ij} are the observed probabilities of changes in the state-space vector and, thus, from node i to node j, with $\sum_i \sum_j m_{ij} = 1$. In the absence of symmetry for these weights the network is *directed* and the out-degree, $D_O(i) \equiv \sum_j m_{ij}$ will not equate to the in-degree, $D_I(j) \equiv \sum_i m_{ij}$. To capture this asymmetry certain metrics can be applied separately to both the out-degree and in-degree. For example, the out-degree eigenvector centrality [50], C_{eO}, is based on the right eigenvector for the largest eigenvalue of $\boldsymbol{M}$ while the in-degree variant, C_{eI}, uses the equivalent left eigenvector.

Given the importance of the principal invariants for studying the VGT dynamics [18,19,48,51,52], we also focus explicitly on the transitions between regions of the Q-R plane that are defined based on the signs of Q, R, and Δ as shown in Fig. 1. Thus, generalizing our notation from (14), we consider the adjacency matrix weights as $m_{i,\alpha;j,\beta}$, where the (i, j) subscripts label each node and $(\alpha, \beta) \in \{1, \dots, 6\}$ indicate that node i belongs to region α and node j belongs to region β. To focus on the flux around the Q-R diagram and to exclude the dynamics within a region (the *intraregion dynamics*), we then have the alternate adjacency matrix, $\boldsymbol{M}^X$, with weights

$$m^X_{i,\alpha;j,\beta} = \begin{cases} m_{i,j} & \text{if } \alpha \neq \beta \\ 0 & \text{otherwise} \end{cases}, \tag{15}$$

where the "X" superscript highlights that this is the adjacency matrix for edges that exist between nodes in different regions of the Q-R diagram, i.e., $\alpha \neq \beta$.

III. METHODS

We employ a direct numerical simulation of HIT at a Taylor Reynolds number of $\mathrm{Re}_\lambda = 433$ in a 1024^3 computational domain with Lagrangian tracking of fluid particles from the Johns Hopkins Turbulence Database [53,54]. Within the computational domain, we chose 27 initial seeding locations along each axis and we track the 27^3 fluid particles for 240 Kolmogorov times, τ_η with a temporal sampling step of $\delta t = \tau_\eta/20$, using cubic Hermitian polynomials to interpolate in time between adjacent time slices. In constructing the network, when the flow packet transitions from the state indexed by i, to that indexed by j, we increment the value of w_{ij} by one. The typical time that the flow remained at a given node before transitioning to a new flow state on each trajectory was 0.195 τ_η (3.9 time samples). This quantity was normally distributed with 99% of trajectories having a typical time between transitions greater than 2.8 time samples. Hence, our intention in this study is to study the dynamics with the transitions well resolved. Future work using our framework could examine how these vary at coarse grainings corresponding to the Kolmorogov and Taylor timescales. After tracking all particles for the entire time series duration we normalize the w_{ij} by $\sum_{i,j} w_{ij}$ such that the adjacency matrix, $\boldsymbol{M}$ contains normalized weights summing to unity that obey the law of total probability. While typically in the literature the w_{ij} are left un-normalized, in this paper we make use of adjacency matrices consisting of all edges and those merely for region exit pathways (with the "X" superscript), as well as simulation and model results as explained below. Hence, the normalization is needed to compare results from these matrices, particularly given our definition of exit node importance that is introduced in Sec. IV D. This normalization has no impact on the calculation of the transition probability vector used in random walk analysis, which for node i uses a normalization to obtain an appropriate conditional probability, for the connection from i to j,

$$p(j|i) = \begin{cases} \frac{m_{ij}}{D_O(i)} & \text{if } (i,j) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases}. \tag{16}$$

A random walk approach underpins our analysis of commute distances on the network in Sec. IV E.

The transitions between flow states we analyze differ from those evaluated directly from the Lagrangian dynamics of the VGT since we exclude self-loops, as is typical in network analysis. Hence, the residence time in a flow state i before the transition to a distinct node j is not incorporated into our approach. Thus, our approach is similar to symbolic sequence analyses of turbulence time series [55] and if we had four flow states and measured the flow state at six regularly sampled points in time to obtain $\{1, 2, 2, 3, 4, 4\}$ on one occasion and $\{1, 1, 1, 2, 3, 4\}$ on another, both would give the same directed network with the same edge weights.

The results from direct numerical simulation are denoted by "DNS" below. As noted in the introduction, a number of models have been formulated that attempt to characterize the Lagrangian dynamics of the VGT. In this study, we compare the DNS results to the Enhanced Gaussian Closure (EGC) model of Wilczek and Meneveau [24], with results sampled from this model at a time resolution that matches those described for the DNS. Hence, in addition to examining the VGT dynamics from DNS data, we can also determine the extent to which these features can be captured by the EGC model. In the following, quantities relative to the EGC model are denoted by a tilde.

The EGC model consists of a Langevin equation describing a set of fluid particles that share the same velocity gradient configuration [24]

$$\begin{aligned} d\tilde{\mathbf{A}} = \big[&- \big(\widetilde{\boldsymbol{A}}^2 - \tfrac{1}{3}\mathrm{tr}(\widetilde{\boldsymbol{A}}^2)\boldsymbol{I}\big) - \pi_1\big(\widetilde{\boldsymbol{S}}^2 - \tfrac{1}{3}\mathrm{tr}(\widetilde{\boldsymbol{S}}^2)\boldsymbol{I}\big) - \pi_2\big(\widetilde{\boldsymbol{\Omega}}^2 - \tfrac{1}{3}\mathrm{tr}(\widetilde{\boldsymbol{\Omega}}^2)\boldsymbol{I}\big) \\ &- \pi_3(\widetilde{\boldsymbol{S}}\widetilde{\boldsymbol{\Omega}} - \widetilde{\boldsymbol{\Omega}}\widetilde{\boldsymbol{S}}) + \delta\widetilde{\boldsymbol{A}}\big]dt + d\boldsymbol{\mathcal{F}}, \end{aligned} \tag{17}$$

where $\tilde{\mathbf{A}}$, $\widetilde{\boldsymbol{S}}$ and $\widetilde{\boldsymbol{\Omega}}$ are realizations of the VGT, strain rate and rotation rate fields, respectively, and $\boldsymbol{\mathcal{F}}$ is a white-in-time Gaussian forcing with zero mean and isotropic statistics. The model coefficients π_i are fitted from the DNS data in order to approximate the average of the pressure Hessian conditional on the local velocity gradient, while the Gaussian closure hypothesis prescribes the tensorial functional form [24]. Similarly, the damping coefficient δ follows from approximating the average velocity gradient Laplacian conditional on the local velocity gradient itself. All model coefficients are constant and their values are in good agreement among different DNS and experimental datasets, up to a Reynolds number dependency [56]. Using trajectories sampled from the EGC model (17), we computed an adjacency matrix over the set of 848 discrete states defined through (13), analogous to the adjacency matrix constructed for the DNS (14).

Network metrics

Classic network analysis commonly draws upon the notion of *centrality* to evaluate the role played by a particular node. We employ the eigenvector and betweenness centralities to characterize the complex network and then focus in greater detail on the properties of the shortest paths and random walks on the network. To capture the extent to which important nodes are influencing other important nodes we make use of the out-degree eigenvector centrality, C_{eO} [50], and our metric of asymmetry between the out-degree and in-degree, C_{eI}, centralities is

$$\kappa_e(i) = \frac{C_{eO}(i) - C_{eI}(i)}{C_{eO}(i) + C_{eI}(i)}, \tag{18}$$

where C_{eO} employs the right eigenvectors of $\boldsymbol{M}$ and C_{eI} the left eigenvectors, and as these are equivalent for an undirected network, one would obtain $\kappa_e = 0$ in that case. The Perron-Frobenius theorem ensures that the largest eigenvalue of $\boldsymbol{M}$, $\lambda_{1,1}$, is positive and, thus, all the elements of the first column vector of the eigenvector matrix, which defines C_{eO} (or C_{eI}), are also positive. We adopt the standard normalization of the eigenvectors such that $\sum_{i=1}^{N} C_{eO}(i) = 1$. The inclusion of an "X" superscript with C_{eO}, κ_e etc., indicates that the underpinning adjacency matrix is $\boldsymbol{M}^X$ rather than $\boldsymbol{M}$.

Further insight is provided by the shortest paths on the network between vertices i and j. Noting that our network is strongly connected, there is a path from any i to any j, and the shortest path is the maximum of the sum of the m_{ij} over all pathways from i to j (or, it is the minimum in the cost/resistance network defined where the nonzero values in the adjacency matrix are converted to $1/m_{ij}$). The betweenness centrality for vertex k is the fraction of shortest paths between the i and j that pass through k [57,58]. With $\phi(i, j)$ the total number of shortest paths and $\phi(i, j|k)$ those passing through node k, the betweenness centrality is

$$C_b(k) = \frac{\phi(i, j|k)}{(N-1)(N-2)\phi(i, j)}, \tag{19}$$

where $\phi(i, j|k) = 0$ if $k \in \{i, j\}$. Hence, $C_b(k)$ is the extent to which node k controls, on average, the pair-wise interaction between other nodes. To consider the structure of the network in greater detail, we focus explicitly on the shortest paths and describe these both in terms of the nodes and the regions of the *Q-R* diagram they pass through. A shortest path starting from node i in region α, ending at node j in region β and passing through a node k in region $\gamma \neq \alpha$ can be written as the set of the vertices in the order they are visited

$$\mathcal{S}(i, \alpha; j, \beta) = \{v_{i,\alpha}, v_{k_1,\gamma_1}, \dots, v_{k_\xi,\gamma_\xi}, \dots, v_{k_T,\gamma_T}, v_{j,\beta}\}, \tag{20}$$

where vertices v_{k_1} to v_{k_T} are the intermediary nodes that are visited, and these are in intermediary regions, γ_ξ, which may equate to α or β depending on the nature of the path. Hence, the length of the shortest path may be expressed in terms of the number of edges traversed, $\ell_S(i, j) = T + 1$. With definition (20) an intraregion shortest path has $\alpha = \beta$ and an interregion shortest path has $\alpha \neq \beta$. In the majority of cases, therefore, the γ_ξ for an intraregion shortest path are all equal to α (and β), while an interregion shortest path between nonadjacent *Q-R* regions (see Fig. 1) will necessarily have at least one $\gamma_\xi \neq \{\alpha, \beta\}$.

Other measures of nodal separation for directed networks include the hitting and commute distances [59]. The former is the average distance between nodes i and j based on a random walk on the network, while the latter is the average distance from node i to j and then back to i. Hence, the hitting and commute distances move away from a focus on shortest paths to consider all possible pathways on the network. To parallel our approach to the length of shortest paths, the length of a commuting random walk is expressed as the average total number of nodes visited, $\overline{\ell_C}(i, j) = \overline{T} + 1$, with the bar denoting the average over a large number of random walks (while a shortest path is a single path realization on the network by definition).

Having outlined some of the underpinning physics of the VGT dynamics, defined our discrete network of flow states and described the primary methods from the complex network literature we will adopt, we can now analyze how the VGT dynamics impact upon the associated complex network structure, and how the network metrics highlight aspects of the VGT dynamics.

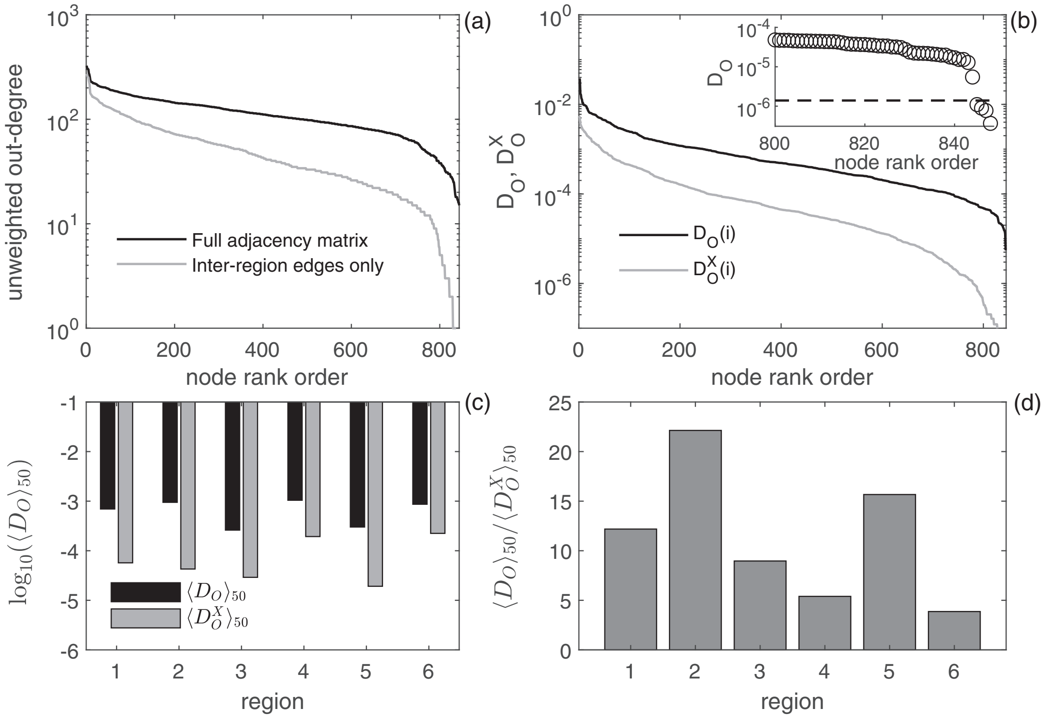

FIG. 2. The out-degree distributions, D_O and D_O^X for the the full weighted networks are given in (b), while out-degree distributions for the unweighted networks where nonzero entries of $\boldsymbol{M}$ are replaced by ones are given in (a). The median values of D_O and D_O^X for each Q-R region are reported on a log scale in (c), and the ratio of these values is in (d). Here and throughout $\langle \ldots \rangle_{50}$ indicates the median, and the "X" superscript indicates that the analysis is based on the region exit node network $\boldsymbol{M}^X$. The inset to panel (b) shows the smallest 48 nodes and the clear separation for the smallest four nodes beneath the dashed line is readily apparent.

IV. RESULTS

As noted above, in total, all combinations of signs and relative magnitudes of the terms in the state vector (13) gave $N = 848$ nodes. However, four of those nodes were visited very infrequently as can be seen in the inset to Fig. 2(b). These nodes were visited at least five times less frequently than the fifth smallest node; the horizontal dashed line in this panel is for $D_O = 1/(847 \times 848)$ and these four nodes are beneath this line. That is, their out-degree (sum of the m_{ij} over 847 nodes) is less than the expected value for an individual m_{ij}. As they were visited so infrequently, these four vertices were subsequently excluded from the network to give $N = 844$ nodes that were used in analysis. Inspection of the state-space vector for these four deleted nodes showed they had common properties, indicating flow states that rarely occur in HIT. All arose where $Q > 0$ and all had a relative ranking of the second-order terms, i.e., the second row of (13), such that $\|\boldsymbol{\Omega}_B\|^2 > \|\boldsymbol{\Omega}_C\|^2 > \|\boldsymbol{S}_B\|^2$. They then formed two pairs based on the signs and relative magnitudes of the third-order terms, where the members of each pair differed by the sign of R (i.e., each pair had one node in region $\alpha = 1$ and one in region $\alpha = 2$). The properties of these four vertices are given in Table I. The reason why these flow states are so rare can be discerned from the final column of this table. The interaction production is the smallest magnitude of the four production terms for all of these cases yet, given that $\|\boldsymbol{\Omega}_C\|^2 > \|\boldsymbol{S}_B\|^2$, one would anticipate that the normal strain production would be the smallest magnitude term. That this is not the case indicates that the required antialignment between the non-normality and the normal straining is a rare occurrence in HIT, explaining the factor of five difference in the values of D_O for these four nodes and the fifth smallest.

TABLE I. Characteristics of the four vertices that were excluded from the adjacency matrix due to the very low probability of them being visited. They form two pairs based on the relative magnitudes of the production terms (final column), with each member of the pair a mirror image of the other based on the region of Q-R space, and the signs of the production terms.

Pair	Region	Signs for $-\det(\boldsymbol{S})$, $\mathrm{tr}(\boldsymbol{\Omega}^2\boldsymbol{S})$	Signs for $-\det(\boldsymbol{S}_C)$, $\mathrm{tr}\big(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B\big)$	Rank order of production term magnitudes $\lvert\det(\boldsymbol{S}_B)\rvert$, $\big\lvert\mathrm{tr}\big(\boldsymbol{\Omega}_B^2\boldsymbol{S}_B\big)\big\rvert$, $\lvert\det(\boldsymbol{S}_C)\rvert$, $\big\lvert\mathrm{tr}\big(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B\big)\big\rvert$
A	1	(+,−)	(+,−)	3,2,1,4
	2	(−,+)	(−,+)	3,2,1,4
B	1	(+,+)	(+,+)	3,1,2,4
	2	(−,−)	(−,−)	3,1,2,4

A. Degree distribution

The degree distribution of a network provides a signature of the network's structure with, for example, the preferential attachment model leading to a power-law degree distribution and a small-world network [60]. Such analyses are typically undertaken for an unweighted network such that the degree distribution reflects the network topology rather than both the topology and the strength of the edges. Hence, in Fig. 2(a) we first show the equivalent unweighted degree distributions for the full adjacency matrix $\boldsymbol{M}$ and that based on the edges that connect nodes in different regions, $\boldsymbol{M}^X$. The unweighted equivalent adjacency matrices were obtained by replacing each nonzero value in $\boldsymbol{M}$ and $\boldsymbol{M}^X$ with 1 and then suitably renormalizing to $\sum_i \sum_j m_{ij} = 1$. These results may be compared to the distributions of D_O and D_O^X in panel (b). In all cases, the distribution is exponential from approximately the one-hundredth ranked node to about node 750, before a steep drop in the degree. This is particularly large for D_O^X and its equivalent in Fig. 2(a) because a number of nodes are primarily involved in adjusting the dynamics of the tensor within a region, α, with constant Q, R and Δ. The impact of the weights is to steepen the slope of the exponential part of the distribution, rather than to change the qualitative nature of the overall topology from exponential to, for example, a power law.

Region-based median values of D_O and D_O^X are given in Fig. 2(c) on a logarithmic scale, and the ratio between these values is given in Fig. 2(d) to highlight the differences between the two forms of adjacency matrix. Regions 4 and 6 (where $\Delta < 0$), as well as region 2 ($Q > 0$, $R < 0$), have the largest average values of D_O but while this pattern persists for D_O^X for the $\Delta < 0$ regions, the average value for region 2 declines by a factor of 20, as shown in Fig. 2(d). Given a typical clockwise motion around the $Q - R$ diagram (e.g., see Fig. 6 of [61] for an example of this for the same simulation), this large value for region $\alpha = 2$ in Fig. 2(d) implies a relatively infrequent transition from region 2 to region 1. That is, where $Q > 0$ and $R < 0$ (hence, the normal enstrophy production is positive), many of the changes in the dynamics are about reconfiguring the properties of the vortex in a manner that does not impact the signs of the normal enstrophy production and normal strain production.

The second largest value in Fig. 2(d) is for region $\alpha = 5$, also implying significant flow reconfiguration within that region. Given the typical clockwise Lagrangian path on the Q-R plane, and the number of nodes in region 5 and 6 (Table II), there is a reduction from 239 nodes in region 5 to 60 nodes in region 6. This is driven by $\|\boldsymbol{\Omega_B}\|^2 = 0$ when $\Delta \leqslant 0$. So the flow needs to reconfigure internally within region 5 from vertices that represent a relatively high ranking for the normal enstrophy or the magnitude of its production in the second and third rows of (13), to those where these terms are sufficiently small to be connected to nodes in region 6 where these terms are zero.

B. Centrality metrics

The strongly connected nature of $\boldsymbol{M}$ means that the Perron-Frobenius theorem holds allowing for the existence of C_{eO} and C_{eI}. To ensure that this was also the case for $\boldsymbol{M}^X$ it was necessary to remove

TABLE II. Median values of the betweenness centralities, C_b and C_b^X for the nodes belonging to each Q-R region, α (see Fig. 1). The number of nodes in a region is N_α, while the values in italics for N_α are the adjusted number of nodes per region for $\boldsymbol{M}^X$. The values in brackets for the betweenness centralities' columns indicate the proportion of nodes in the region over which the median is calculated (i.e., excluding the nodes that do not lie on any shortest path).

Region α	N_α	$\langle C_b\rangle_{50}$ ($\times 10^{-3}$)	$\langle C_b^X\rangle_{50}$ ($\times 10^{-3}$)
1	123	6.2 (0.85)	6.5 (0.72)
2	123	9.0 (0.83)	6.2 (0.72)
3	239, *232*	3.1 (0.86)	4.7 (0.46)
4	60	4.0 (0.95)	9.4 (0.92)
5	239, *236*	2.4 (0.89)	5.5 (0.46)
6	60	7.4 (0.93)	8.6 (0.90)

seven nodes from region 3 and three from region 5 where either the out-degree, $D_O(i) = \sum_j m_{ij}^X$ or the in-degree, $D_I(j) = \sum_i m_{ij}^X$ was zero (because these vertices were only connected to other nodes in the same region). Summary values for the centrality statistics for both $\boldsymbol{M}$ and $\boldsymbol{M}^X$ are given in Fig. 3 and Table II. The results are reported as a function of the region of the Q-R diagram, labeled by α, as shown in Fig. 1, with N_α in Table II indicating the number of nodes in region α. Figure 3(a) shows that 98% of $\sum C_{eO}$ arises in regions 5 and 6 (those that form the right Vieillefosse tail [18]). In contrast, the results for $\sum C_{eO}^X$ in panel (b) are a strong function of Q, with values less than 2% for regions 1 and 2 where $Q > 0$, between 12% and 20% for regions 3 and 5 ($Q < 0$, $\Delta > 0$) and at least 30% for regions 4 and 6 ($Q < 0$, $\Delta < 0$). Thus, the overwhelming tendency is for nodes with high out-degree centrality to group in the regions adjacent to the right Vieillefosse tail. However, the difference in the results between $\sum C_{eO}$ and $\sum C_{eO}^X$ indicates that a high proportion of these nodal connections are within-region, and when these are excluded, there is greater similarity in network structure for the pairs of regions with a different sign of R but the same signs for Q and Δ.

The results for the median values of κ_e and κ_e^X, shown in Figs. 3(c) and 3(d), also display a marked contrast. While the sign of $\langle\kappa_e\rangle_{50}$ is given by the sign of R (nodes for regions $\alpha \in \{1, 5, 6\}$ with $R > 0$ have greater out-degree than in-degree eigenvector centrality), $\langle\kappa_e^X\rangle_{50}$ is only positive in region 6. Hence, when the focus is on the transitions between regions of the Q-R diagram, nodes that are strong donors of probability flux only arise where $R > 0$ and $\Delta < 0$. Figure 3(d) also shows that region 2 exhibits a preferential occurrence of the nodes that have a strong influence on receipt vertices. This reinforces the result in Fig. 2(d) that region 2 is focused on intraregion flow reconfiguration, helping to explain its enhanced probability of occurrence in general relative to region 1 as seen in Fig. 1 and as can be shown from statistical testing using random matrix approaches [62]. This result indicates the potential of network-based analyses for revealing aspects of the VGT dynamics that may be harder to discern using alternative methods.

Table II shows the values for the median betweenness centralities $\langle C_b\rangle_{50}$ and $\langle C_b^X\rangle_{50}$ accompanied by the proportion of nodes in the region that lie on a shortest path (stated in brackets). The value for N_α decreases for regions 3 and 5 when $\boldsymbol{M}^X$ is analyzed as there were seven (respectively, three) nodes for which all the edges were to nodes within the same region. The median values quoted were determined over the nodes lying on shortest paths since, in the case of $\langle C_b^X\rangle_{50}$ for regions 3 and 5, with only 46% of nodes lying on a shortest path, the median would go to zero. In other words, the low median C_b values for region 3 and 5 would be even lower if one determined the median over all nodes in the region. As with the eigenvector centrality, region 6 plays an important role for both networks, although in contrast to the eigenvector centrality results, the $Q > 0$ regions (1 and 2) are much more important, with region 2 having the largest value for $\langle C_b\rangle_{50}$. Thus, from its eigenvector centrality, region 2 has a relatively low influence of its dynamically important vertices compared to

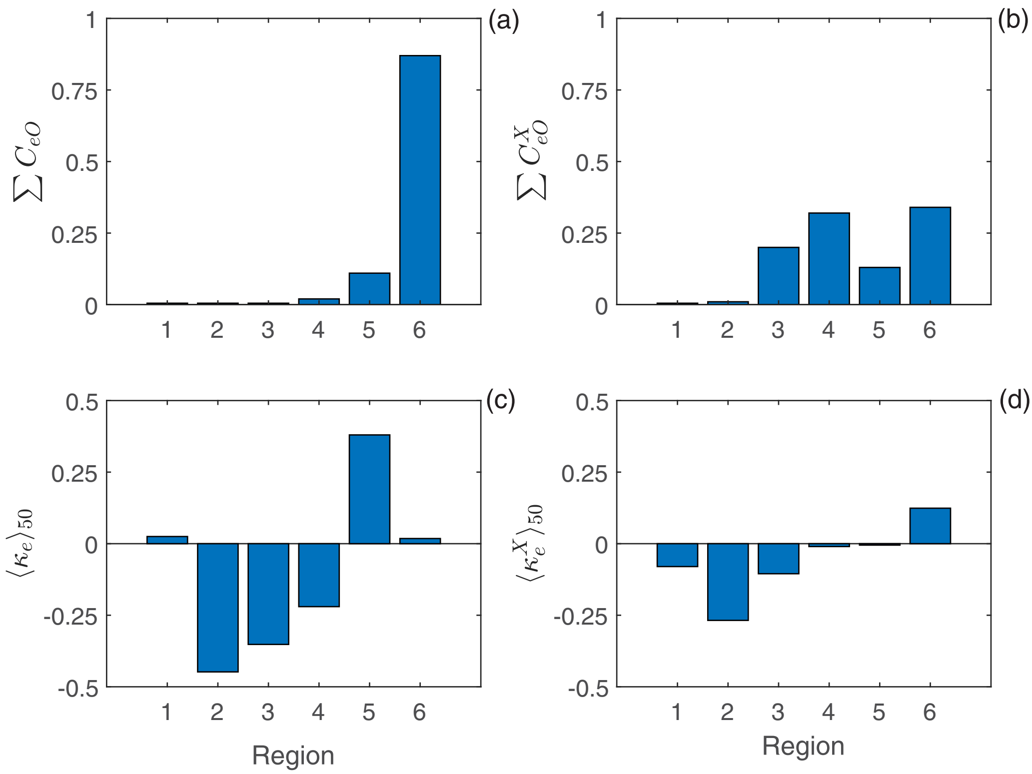

FIG. 3. Summary values for the out-degree eigenvector centrality, C_{eO} (top row), and the out-degree/in-degree asymmetry metric, κ_e (bottom row), for the full adjacency matrix, $\boldsymbol{M}$ (left column), and that defined over the edges that exist between regions of the Q-R diagram, $\boldsymbol{M}^X$ (right column). Results are given for each of the six regions of the Q-R diagram, defined in Fig. 1. The median is shown by $\langle \ldots \rangle_{50}$.

region 6, but from Table II these feature on a large proportion of shortest paths. Table II also shows that $\langle C_b^X \rangle_{50}$ is a decreasing function of N_α. This is the expected pattern since the presence of more vertices in a given region implies fewer shortest paths through a given node on average. The fact that the results for $\langle C_b \rangle_{50}$ depart from this simple pattern highlights that in region 2 there are intranode shortest paths that contribute significantly to the betweenness centrality without directly impacting the Lagrangian circulation around the Q-R diagram. This result is in agreement with the eigenvector centrality ratios in Fig. 2(d).

C. Intraregion shortest paths

Looking at the shortest paths in greater detail, Fig. 4 compares distributions of the length of the shortest path $\ell_S(i, j)$ for different Q-R regions, distinguishing between intra- and interregion shortest paths (the left- and right-hand panels for each region, respectively). The ordering of the panels by region reflects the structure in Fig. 1. The mean shortest path lengths stated in the left-hand panels of each region in Fig. 4 indicate that for the intraregion shortest paths, the path length is primarily a function of N_α with a secondary dependence such that greater values arise for $R > 0$ if $\Delta > 0$. This is similar to the results for $\langle C_b^X \rangle_{50}$ in Table II. The percentage stated in each of the left-hand panels is the proportion of the intraregion shortest paths that include nodes in other regions. That is, a shortest path (20) with $\alpha = \beta$ but some of the γ_ξ values not equal to α. Thus, region 2 is the most self-contained as fewer than 1% of intraregion shortest paths involve vertices outside this region as implied by the earlier analysis of the contrasting nature of C_{eO} and C_b for this region. This leads to

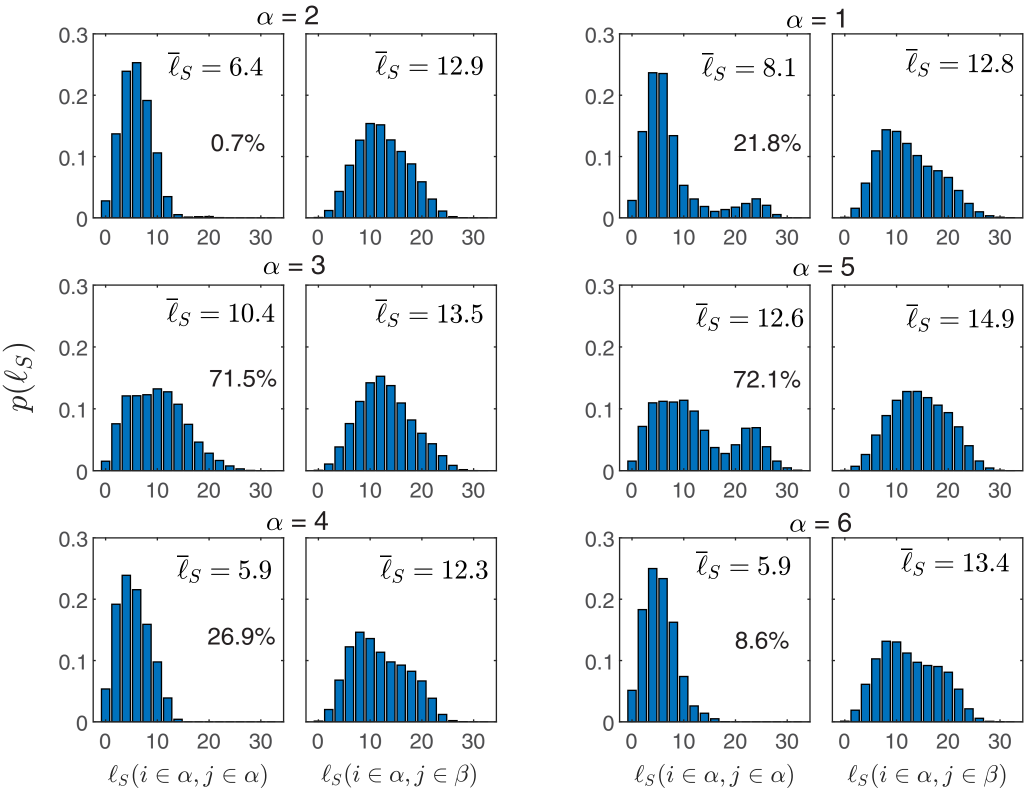

FIG. 4. Probability of the shortest path length $\ell_S(i, j)$ for each *Q*-*R* region, α. The intraregion shortest paths, reported in the left-hand panel of each region, start from node i belonging to region α and end at node j also belonging to region α. The interregion shortest paths, displayed on the right-hand panel of each region, start from node i belonging to region α and end at node j belonging to region $\beta \neq \alpha$. We quote the mean shortest path length $\overline{\ell}_S$ in each panel. The percentage of intraregion shortest paths that explore other regions, i.e., from region α to region $\gamma \neq \alpha$ and back to α, is quoted in each left-hand panel.

a small mean shortest path length $\ell_S(i, j) = 6.4$ for $\alpha = 2$, consistent with the interpretation of the structure of region 2 as involving significant flow reconfiguration within the region.

However, the intraregion shortest paths for regions $\alpha = 4$ and $\alpha = 6$ also have low mean lengths despite a significant proportion of trajectories visiting other intermediate regions $\gamma \neq \alpha$. Investigating this, it was found that such shortest paths only involve the other region with $\Delta < 0$. Hence, in this particular sense, $\Delta < 0$ is more relevant for the VGT dynamics than the sign of R combined with the sign of Δ, as the high amount of interchange between regions 4 and 6 implies they are almost acting as one. It may also be noted that the 26.9% of intraregion shortest paths of region 4 with $\gamma \neq \alpha$ *all* transition via region 6 and thus have a preferential direction opposite to the average clockwise Lagrangian circulation around the *Q*-*R* diagram (although in agreement with the paths for the restricted Euler model of [13]).

While the maximum length of the intraregion shortest paths starting from regions 2, 4, and 6 is short, the length distribution for region 3 displays a significant right tail, such that the intraregion shortest path lengths are similar to the interregion shortest path lengths. For example, 4% of the intraregion shortest paths for region $\alpha = 3$ have length $\ell_S(i, j) > 20$ vertices, and for the interregion shortest paths this is 10%. A related effect is seen for regions 1 and 5 (i.e., $R > 0$, $\Delta > 0$) as the intraregion shortest paths' length distributions have a clear bimodality. In both of these cases, the left mode exhibits a similar shape to the equivalent region at $R < 0$, while the right mode has a

TABLE III. Relative frequency of joint states of the signs of the non-normal production, $-\det(\boldsymbol{S}_C)$, and interaction production, $\mathrm{tr}(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B)$ for long intraregion shortest paths for region 5. These are divided into the vertices that are exit nodes (third column) and returning entrance nodes (fourth column) for region 5. We also show the relative frequencies of these states for all nodes in region 5 and all nodes in the overall network.

Sign of $-\det(\boldsymbol{S}_C)$	Sign of $\mathrm{tr}\left(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B\right)$	Exit nodes	Entrance nodes	All of region 5	All nodes
$-$	$-$	36%	56%	11%	10%
$-$	$+$	34%	9%	34%	33%
$+$	$-$	15%	31%	13%	14%
$+$	$+$	15%	4%	42%	43%

value of $\ell_S(i, j) \sim 24$. The large volume of the upper mode for region 5 results in the probability of $\ell_S(i, j) > 20$ being greater for the intraregion shortest paths (22%) than the interregion shortest paths (17%). It was found that this upper mode of long shortest paths is driven by cases that undertake a full orbit of the Q-R diagram to reach another flow state in the same region. Therefore, in regions 1 and 5 there is a set of vertices characterized by a relatively low probability of transitioning to other parts of the same region directly.

Further inspection of the results highlights a difference between regions 1 and 5 concerning the origin of these orbiting intraregion shortest paths. For region 1, the intraregion shortest paths with $\ell_S(i, j) > 20$ exhibit a major change in either the non-normality or its production along the path; i.e., one or both of the following apply:

(1) They start from a node i where $\|\boldsymbol{\Omega}_C\|^2$ is ranked third in magnitude of the second-order terms [see the discrete state definition (13)] and terminate at at node j where this term is ranked first; or

(2) At i, $-|\det(\boldsymbol{S}_C)|$ and $|\mathrm{tr}(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B)|$ are ranked third and fourth. This means the two normal production terms are ranked first and second and from the definition of region 1 that must mean $|\det(\boldsymbol{S}_B)|$ is second. However, at j the normal strain production is now ranked fourth. Hence, the production terms involving the non-normal tensor increase in relative prominence, an effect related to, but distinct from the gain in non-normality in point 1.

For all these long intraregion paths in region 1, the normal enstrophy production magnitude, $|\mathrm{tr}(\boldsymbol{\Omega}_B^2\boldsymbol{S}_B)|$ was ranked first. Hence, increasing $\|\boldsymbol{\Omega}_C\|^2$ within region 1 is difficult, and an orbit of the Q-R diagram is required to increase tensor non-normality significantly. As is shown later, this is important for transitioning to region 5 as the most efficient such pathway has $\|\boldsymbol{\Omega}_C\|^2 > \|\boldsymbol{\Omega}_B\|^2 > \|\boldsymbol{S}_B\|^2$.

The mechanism for the bimodality is different in region 5, where $Q < 0$ and, therefore the ordering of the normal enstrophy and normal straining swaps compared to region 1. Bimodality is less about the change in relative magnitude of $\|\boldsymbol{\Omega}_C\|^2$ and more about the signs of the production terms. This is shown in Table III, which looks at the signs of the non-normal production and the interaction production for the nodes that are the last in region 5 for a long intraregion shortest path (an exit node) compared to the entrance node where it returns to region 5. The entrance nodes (fourth column) for the long shortest paths preferentially feature a negative interaction production, $\mathrm{tr}(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B) < 0$ ($56 + 31 = 87\%$ of cases compared to $11 + 13 = 24\%$ in general for nodes in region 5 (fifth column) and $10 + 14 = 24\%$ for the network as a whole given in the sixth column). Given the average positivity of $\mathrm{tr}(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B)$ [42], these intraregion shortest paths are long because they have to orbit the Q-R diagram to gain this relatively rare negative sign for the interaction production.

Although the mechanism explaining the long intraregion paths associated with regions 1 and 5 differs, it is noteworthy that the bimodality arises in these two regions and its explanation involves either the magnitude of the non-normality or the signs of the non-normal and interaction production. It is these two regions of the Q-R diagram where the deviatoric or anisotropic pressure Hessian contribution to (2) results in complex dynamics. This is seen in conditional trajectory analyses [14] that separate contributions from the restricted Euler terms (3), viscous terms and the deviatoric pressure Hessian. An example of such work can be seen in Fig. 3 of [63]. Hence, our network analysis highlights new aspects of such results and shows that certain nodal states within a region are hard to access or escape from without a very long pathway being taken.

D. Interregion shortest paths

Except for the long intraregion shortest paths discussed above, circulation around the *Q-R* diagram is due to the interregion shortest paths, i.e., those where $\alpha \neq \beta$ for the path $\mathcal{S}(i, \alpha; j, \beta)$ defined in (20). Here we focus on the pair of vertices crossed at the transition from region α to the intermediate region γ, namely, how the shortest path exits a region and enters another. The definition in (20) gives a pair of vertices, $\{v_{k_\xi,\alpha}, v_{k_{\xi+1},\gamma}\}$ at the boundary between a starting region α and the next region visited, γ. For every interregion shortest path we can extract the edge weight, $m_{k_\xi,k_{\xi+1}}$, from $\boldsymbol{M}$. We can then determine the number of shortest paths that transition from region α to γ via this pair of vertices, $n^{\mathcal{S}}_{k_\xi,\alpha}$, leading to a definition of the *exit node importance* as

$$I_{k_\xi,\alpha;k_{\xi+1},\gamma} = n^{\mathcal{S}}_{k_\xi,\alpha} m_{k_\xi,k_{\xi+1}}. \tag{21}$$

This quantity has some similarity to the nodal fairness (the reciprocal of the closeness or status [64]). However, while these metrics for node i are based on the sums of the shortest path distances between vertices i and j for all j, here we only make use of $m_{k_\xi,k_{\xi+1}}$, the segment of the shortest path that exits region α to enter region $\gamma \neq \alpha$. To move from the vertex-pair quantity (21) to a summary measure for a given region, we define the total region exit node importance, $I_{\alpha,\gamma}$, as the sum of the $I_{k_\xi,\alpha;k_{\xi+1},\gamma}$ over all the exit nodes from α to γ:

$$I_{\alpha,\gamma} = \sum_{k_\xi,k_{\xi+1}} I_{k_\xi,\alpha;k_{\xi+1},\gamma}. \tag{22}$$

Figure 5 shows some statistics for the interregion connectivity for both the DNS and the EGC model (results in brackets) superimposed on to *Q-R* diagrams. Figure 5(a) shows the number of nodes that act as exit nodes for each region, the total number of nodes in a region and then distinguishes between exit nodes that connect in counterclockwise and clockwise directions. Thus, in region 1, “78/123: 12, 66” shows there are 123 nodes in total of which 78 act as exit nodes, 12 in the counterclockwise direction. Values for $I_{\alpha,\gamma}$ for clockwise paths are given in Fig. 5(b) and for counterclockwise paths are in Fig. 5(c). The primary conclusion from Fig. 5(a) is that the EGC model does a good job at obtaining the correct number of exit nodes for each region apart from those lying immediately above the discriminant function (regions 3 and 5), where the DNS uses fewer nodes to connect to adjacent regions. This is not too surprising as trajectories for the restricted Euler solution follow lines of constant Δ [13] and therefore never cross $\Delta = 0$, meaning that this is where the model is most reliant on its representation of the non-normal dynamics. A comparison of Figs. 5(b) and 5(c) shows that for the DNS the clockwise importance is always greater than the counterclockwise value between the same pair of regions. However, this is not true for the EGC where, for example, $I_{1,2} = 41.25$, but $I_{2,1} = 22.12$. In addition, there are some major differences in the magnitudes of the $I_{\alpha,\gamma}$ and this is particularly true for clockwise pathways involving regions with real eigenvalues. For $I_{5,6}$ the DNS value is 9.3 times greater than the EGC, for $I_{6,4}$ it is 75 times greater and for $I_{4,3}$ it is 50 times greater. Hence, the well-known property of Lagrangian pathways to move in a clockwise direction around the *Q-R* diagram causes difficulties for the EGC model where these trajectories are counter to the direction imposed by the restricted Euler solution.

Whilst the clockwise path is always of greater importance for the DNS than the equivalent counterclockwise path, when comparing paths out of a region, the exception to clockwise dominance is region 1, where $I_{1,5} = 7.46$ and $I_{1,2} = 25.49$. This is in agreement with previous observations that the deviatoric part of the pressure Hessian generates probability fluxes that differ significantly from the clockwise restricted Euler component in region 1 [23,31,63]. By far the largest total exit importance value is $I_{5,6} = 150.2$, i.e., across the right Vieillefosse tail. The eigenvector centrality

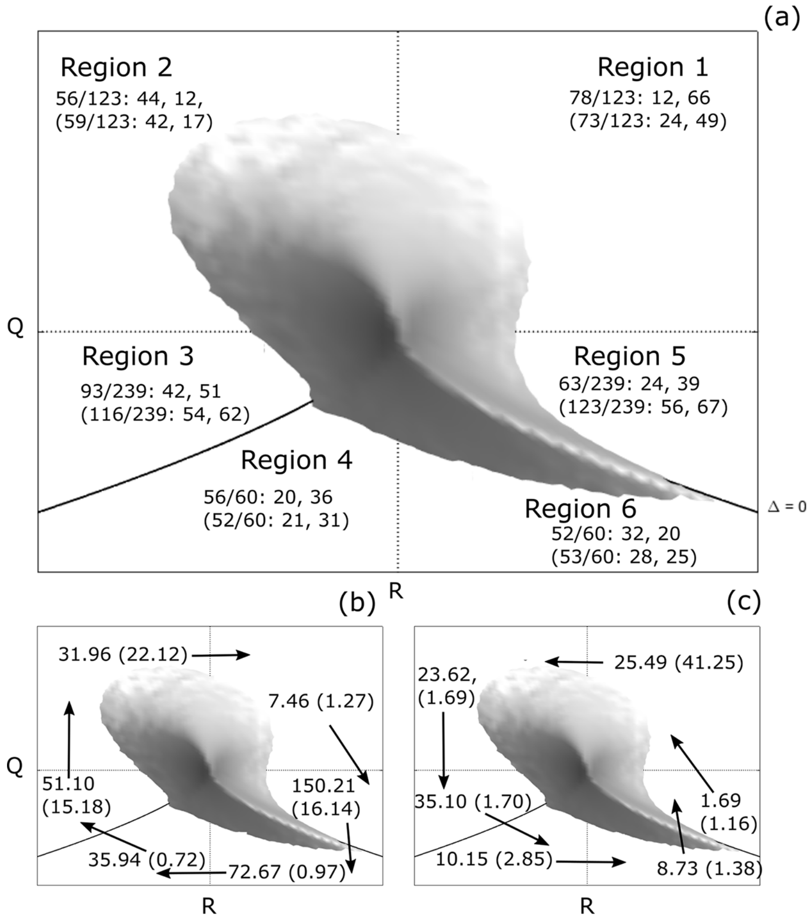

FIG. 5. Three sketches of the joint distribution for Q and R showing the six regions, α, and then region-averaged statistics for interregion shortest paths. Results from the enhanced Gaussian closure model are given in brackets. The main panel (a) gives the number of nodes from the total in that region involved in interregion connectivity, followed by the number of these involved with counterclockwise and then clockwise pathways. The two lower panels give the total exit node importance for each region for clockwise (b) and counterclockwise (c) connections.

results in Fig. 3 also draw particular attention to the vertices in the regions adjacent to the right Vieillefosse tail as being critical for the VGT dynamics.

Moving beyond the region summary values in Fig. 5, we plot the individual exit node importance $I_{k_\xi,\alpha;k_{\xi+1},\gamma}$ in Fig. 6 to examine the vertices through which the transitions between Q-R regions take place. Figure 6(a) shows the results for the DNS and Fig. 6(b) the EGC. The different lines and symbols indicate the different starting region α and intermediately adjacent region γ with clockwise paths in blue and counterclockwise in red. The horizontal lines dashed in Fig. 6(a) and dot-dashed in Fig. 6(b) indicate 75% of the overall total exit importance $I_{\alpha,\gamma}$. The mechanism by which $I_{1,2} > I_{1,5}$ for the DNS (as discussed above) is apparent from this figure. While there are more vertices that serve as exit nodes in the clockwise direction from region 1 to 5 (this is the largest number for any region at 66), rather than the counterclockwise direction from region 1 to 2 (13 nodes), the greatest value for $I_{k_\xi,1;k_{\xi+1},5}$ is just 1.64 while for $\alpha = 1, \gamma = 2$ there are two instances above the threshold given by the dashed line with $I_{k_\xi,1;k_{\xi+1},2} = 13.36$ and $I_{k_\xi,1;k_{\xi+1},2} = 4.80$. Given that the deviatoric pressure Hessian drives the counterclockwise trajectories, these results

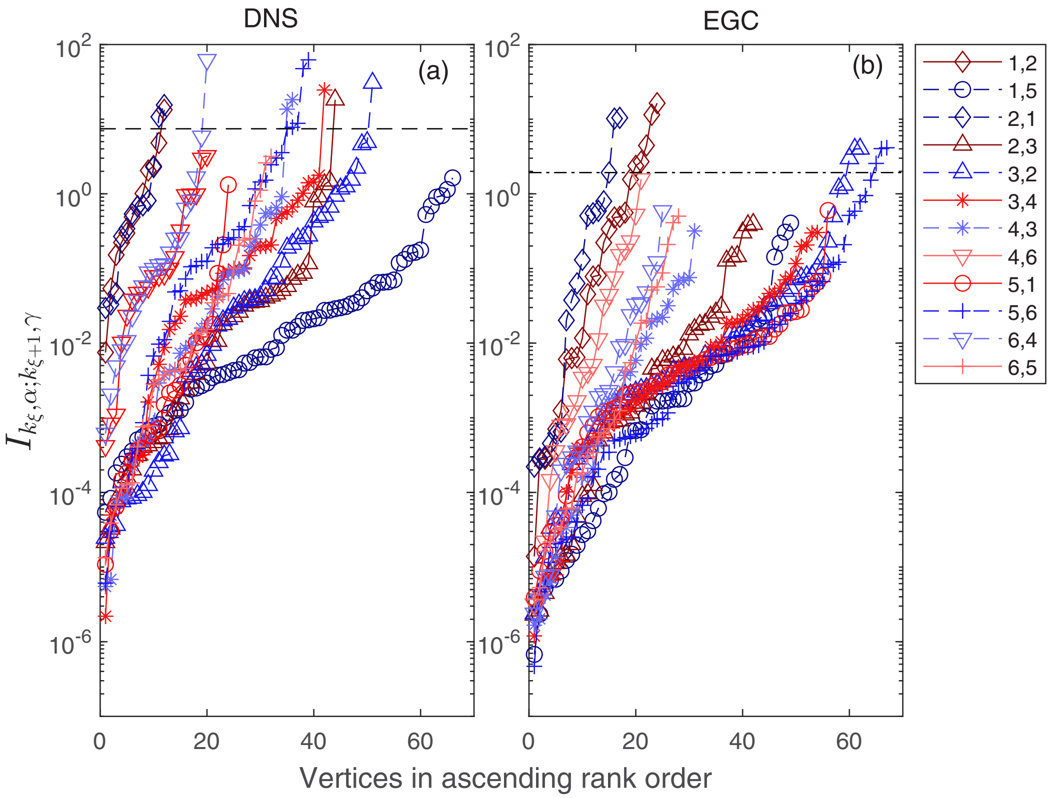

FIG. 6. Exit importance $I_{k_\xi,\alpha;k_{\xi+1},\gamma}$ for each exit node for the interregion shortest paths for the DNS (a) and the EGC (b). Each line shows the exit importance for different combinations of Q-R regions, labeled by α and γ, as given in the legend, with vertices placed in ascending rank order. The sum of the values on each curve for fixed α and γ gives $I_{\alpha,\gamma}$ reported in Figs. 5(b) and 5(c). The nodes placed above the horizontal, dashed line in both panels account for more than 75% of $\sum_{\alpha,\gamma} I_{\alpha,\gamma}$ over all nodes. Darkening colors correspond to increasing Q for the exit region, α, with blue indicating clockwise and red counterclockwise paths.

show that in the direction of the restricted Euler model we have a large number of low activity paths (the action is relatively diffusive), while the deviatoric pressure Hessian opens up a small number of select paths that are very active. The two nodes with large values for $I_{k_\xi,1;k_{\xi+1},2}$ are unusual in that they involve a negative value for $\mathrm{tr}(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B)$ and the same ordering of the production magnitudes with the non-normal production (typically the smallest term) largest in magnitude: $|\det(\boldsymbol{S}_C)| > |\mathrm{tr}(\boldsymbol{\Omega}_B^2\boldsymbol{S}_B)| > |\mathrm{tr}(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B)| > |\det(\boldsymbol{S}_B)|$. This latter ranking arises for these two nodes despite $\|\boldsymbol{\Omega}_B\|^2 > \|\boldsymbol{\Omega}_C\|^2$, meaning that either the normal straining is particularly small in magnitude or there is very weak alignment between the normal straining eigenvectors and the normal vorticity. This particular result highlights some of the advantages of using a discrete network to study the dynamics as while previous work has illustrated the complexity of Lagrangian paths exiting region 1 [23,63], our results regarding the range of ways that this can occur in one direction relative to more selective paths in the counterdirection highlight the utility of the network approach.

A comparison between Figs. 6(a) and 6(b) shows that the EGC model is able to broadly capture the difference in nature between the $I_{k_\xi,1;k_{\xi+1},2}$ and $I_{k_\xi,1;k_{\xi+1},5}$ pathways, although the latter makes use of an insufficient number of vertices. However, when looking in greater detail, the vertices with large values for $I_{k_\xi,1;k_{\xi+1},2}$ are rather different in nature in the EGC case, with $|\mathrm{tr}(\boldsymbol{\Omega}_B^2\boldsymbol{S}_B)| > |\det(\boldsymbol{S}_C)| > |\mathrm{tr}(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B)| > |\det(\boldsymbol{S}_B)|$, i.e., the normal enstrophy production is greater in magnitude that the non-normal production; the model finds it difficult to place sufficient emphasis on the purely non-normal term.

Figure 6(a) also shows that there are three particularly large values for $I_{k_\xi,\alpha;k_{\xi+1},\gamma}$ and these occur in regions 5 and 6, driving the large values for $I_{5,6}$ and $I_{6,4}$ in Fig. 5. Concerning the discrete states for these nodes, with reference to (13), the two nodes in region 5 with high exit importance that facilitate the transition from complex to real eigenvalues are very similar: they have an intermediate ranking for $\|\boldsymbol{\Omega}_C\|^2$, positive signs of $-\det(\boldsymbol{S})$ and $\mathrm{tr}(\boldsymbol{\Omega}^2\boldsymbol{S})$, a positive sign for $\mathrm{tr}(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B)$, and the same ranking for the production terms. Hence, they differ only in terms of the sign of $-\det(\boldsymbol{S}_C)$. The particularly important node for $I_{6,4}$ (acting counter to the restricted Euler dynamics) would be expected to have significant contributions from the non-normal terms. This is also borne out as it has $\|\boldsymbol{\Omega}_C\|^2$ as the largest in magnitude of the three second-order terms with, of the production terms, the interaction production largest in magnitude, followed by the non-normal production. Having $\|\boldsymbol{\Omega}_C\|^2 > \|\boldsymbol{S}_B\|^2$ but $\mathrm{tr}(\boldsymbol{\Omega}_C^2\boldsymbol{S}_B) > -\det(\boldsymbol{S}_C)$ implies that the normal strain and the non-normality are well aligned, indicating the importance of strain and vortical alignments [48] in addition to the magnitude of the non-normality.

E. Commute distances

The dynamics over the Q-R plane can be characterized using shortest paths, as in the previous sections or, as noted in the introduction, using random walks over the network that include less frequented pathways. This requires the adoption of the conditional transition probability vector for evaluating the transitions from node i to j given in (16). To this end, we focused on the average commute distances, i.e., we determined the mean length $\overline{\ell}_C(i,j)$ of the random walks from node i back to i via node j, for all i and j. The averaging was undertaken over at least 20 random walk realizations for a given (i,j) node pair, with a median of 3800 realizations. This high variability is a consequence of the four orders of magnitude of variation in out-degree seen in Fig. 2(b). To focus on paths that have traversed a large portion of the Q-R diagram, we consider the intermediate node j to be in the regions nonadjacent to the region α to which the starting or arrival node i belongs. We then define $\ell_C(i,j_{\min})$ as the mean commute distance for node i in α to i via an intermediate node $j_{\min}$ in region γ that gave the shortest mean commute of all the j in region γ:

$$\ell_C(i,j_{\min}) = \min_j(\langle\overline{\ell}_C(i,\alpha;j,\gamma)), \tag{23}$$

Consequently, $\ell_C(i,j_{\min})$ measures how well node i is connected to region γ from a commute distance perspective.

We examine the characteristics of the $j_{\min}$ nodes (dark gray bars) relative to all nodes in the same region (light gray bars) in Fig. 7. The regions of concern are indicated by the value for α for node i (each row of panels) and then the value for γ for the $j_{\min}$ nodes (dark gray bars) or all j (light gray bars) stated in the left-hand panel of each pair. Results are provided for two choices of γ: the regions that are three (left pair of columns) and four (right pair of columns) regions apart from α based on a clockwise trajectory around the Q-R diagram. We focus on three aspects of the flow state vector for each vertex: the relative ranking of the magnitude of $\|\boldsymbol{\Omega}_C\|^2$ (compared to $\|\boldsymbol{\Omega}_B\|^2$ and $\|\boldsymbol{S}_B\|^2$), i.e., the second row of (13), in the left-hand panels of each pair, and the signs of the non-normal production and the interaction production, i.e., the final two components of the first row of (13), in the right-hand panels of each pair. These are the same terms considered earlier to explain the bimodality of intraregion shortest paths for regions 1 and 5.

For commutes from the $Q>0$ regions $\alpha\in\{1,2\}$ via vertices within $\Delta<0$ regions $\gamma\in\{4,6\}$ (or *vice versa*), and particularly for $\alpha=2$, there is a clear preference for $\|\boldsymbol{\Omega}_C\|^2$ to be the largest term in magnitude at the $j_{\min}$ node. Based on a chi-squared test, all of the differences in $\|\boldsymbol{\Omega}_C\|^2$ between the $j_{\min}$ nodes and all the nodes in γ (i.e., between the dark gray and light gray bars) for commutes between these types of regions were statistically significant apart from the $\alpha=6,\gamma=1$ case in the bottom right panels of this figure. Thus, non-normality plays an important role in facilitating short commutes around Q-R space despite not appearing in either Q or R. This suggests that it may

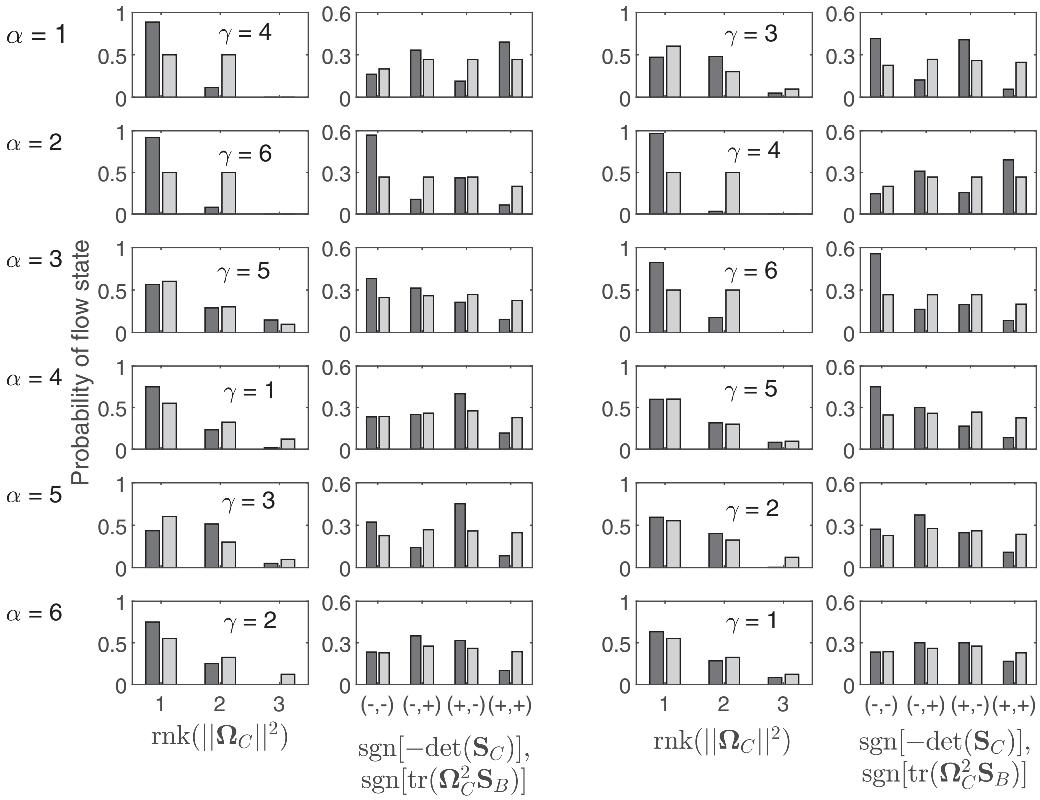

FIG. 7. The probability of different flow states for nodes, $j_{\min}$, which form the shortest mean commutes, $\ell_C(i, j_{\min})$ from region α to region γ (dark gray) compared to all the vertices in that region γ (light gray). For a region α containing the node i we show results for two regions γ as indicated in the left-hand panel of each pair, which are three and four regions apart from the starting region α based on a clockwise motion around the Q-R diagram. For each pair of panels, the left-hand panel shows the ranking of the non-normality, $\|\boldsymbol{\Omega}_C\|^2$, while the right-hand panel shows the signs of the non-normal production and interaction production.

be advantageous in future model formulations to consider this component separately rather than integrating it into the conventional enstrophy and straining terms seen on the left-hand sides of (9) and (10).

The two regions where $Q < 0$ but $\Delta > 0$, i.e., regions 3 and 5 both exhibit significant differences concerning the signs of the non-normal production and the interaction production whether they are a source node, α, or an intermediate node, γ. In particular, where $\alpha \in \{1, 5\}$ and $\gamma = 3$, such results are driven by $\mathrm{tr}(\boldsymbol{\Omega}_C^2 \boldsymbol{S}_B) < 0$ state. In contrast, where $\gamma = 5$, which is the case for $\alpha \in \{3, 4\}$, the $j_{\min}$ cases are driven by $-\det(\boldsymbol{S}_C) < 0$. From Table III we have that intraregion shortest paths which traverse the Q-R space are associated with negative values for non-normal production and, in particular, interaction production. The fact that the shortest commute distances are also aided by negative non-normal and interaction production further highlights the importance of these flow characteristics. These are relatively unusual flow characteristics given that the general positivity of the interaction production is what drives the typically positive signs for both the enstrophy production and strain production on the left-hand side of (11) and (12) [42], as seen in the last column of Table III.

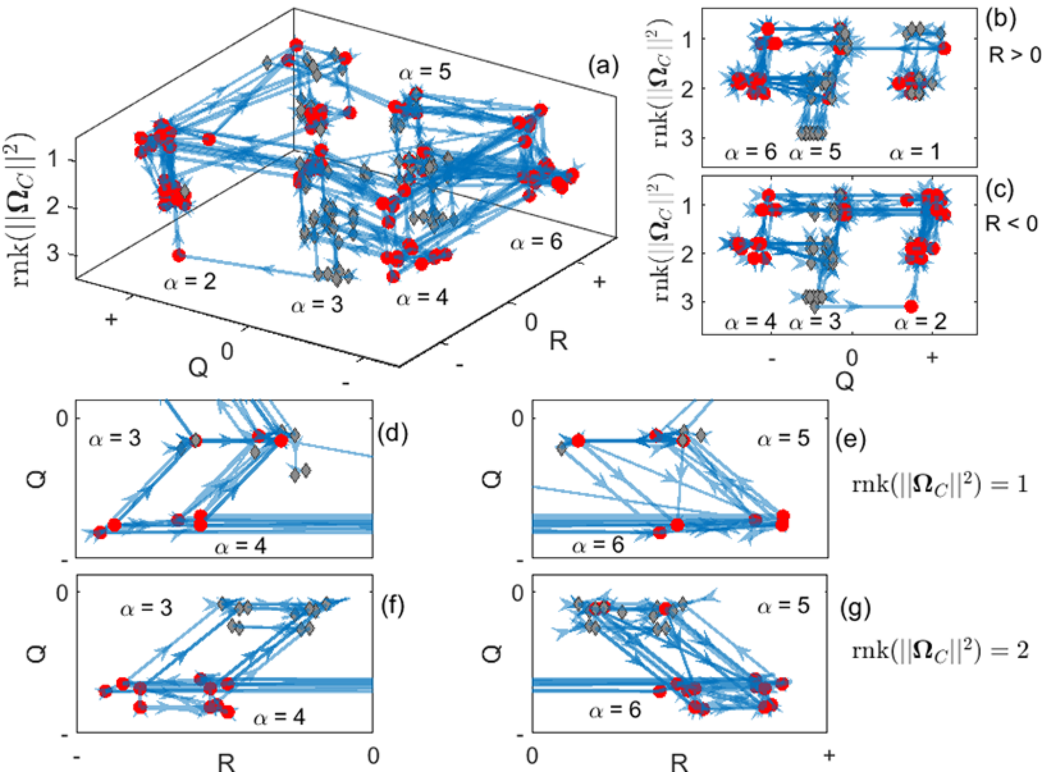

FIG. 8. A visualization of the subnetwork constructed from $\boldsymbol{M}$ using the 178 nodes with highest out-degree is shown in full in panel (a). Each node is placed in its appropriate Q-R region in the horizontal plane, with the vertical axis indicating the ranking of non-normality, $\|\mathbf{\Omega}_C\|^2$, compared to the enstrophy and total straining [see definition (13)]. The red circle nodes are those in the top 10% based on their out-degree, while the gray diamonds are the others that meet the additional selection criterion (see text). We show the edges with a weight exceeding a given threshold, and the arrowhead indicates the net direction if the flux in both directions exceeds this threshold. Panels (b)–(g) emphasize subcomponents of the main panel. Nodes belonging to the same region, α, are displaced from one another horizontally and vertically to aid visualization.

V. INTERPRETATION WITH RESPECT TO THE VISUALIZED NETWORK STRUCTURE

Figure 8 illustrates a network extracted from $\boldsymbol{M}$ and consisting of 178 nodes, selected based on their out-degree: the largest 10% (84 nodes) were initially selected (shown as red circles); we then extracted the next largest nodes based on their out-degree from each region (Fig. 1) until the total number of nodes selected for a region, α, equaled or first exceeded $0.2N_\alpha$. The 178 nodes extracted (21% of the total) accounted for 56% of the total out-degree over all nodes. Edges are included in Fig. 8 where $m_{i,j} > 75/N(N-1)$ and the line thickness is proportional to the logarithm of m_{ij}. Further simplification for visualization purposes is achieved by only showing a single edge between vertices; if both m_{ij} and m_{ji} exceed the threshold, only the larger is shown as indicated by the direction of the arrow. The vertical axis in Fig. 8 shows the relative ranking of $\|\mathbf{\Omega}_C\|^2$ compared to the other two terms on the right-hand side of (9) and (10), while the horizontal plane represents, schematically, the traditional Q-R plane. Because $\|\mathbf{\Omega}_B\|^2 = 0$ in regions 4 and 6 there are no nodes at $\mathrm{rank}(\|\mathbf{\Omega}_C\|^2) = 3$ for these two regions. No such constraint arises for region 1, but Figs. 8(a) and 8(b) show that there are no selected nodes in this region for $\mathrm{rank}(\|\mathbf{\Omega}_C\|^2) = 3$ with sufficiently large out-degree to be shown. Note this relates directly to our previous statements about the complex behavior of trajectories in region 1 driven by the deviatoric pressure Hessian contribution to the VGT dynamics [Fig. 5(c) and [23]]. The earlier result concerning the need in region 1 for nodes to gain $\|\boldsymbol{\Omega}_C\|^2$ to connect to major pathways is supported both by the links from region 1 to region 5 that arise only where rank($\|\boldsymbol{\Omega}_C\|^2$) = 1 and the absence of featured vertices and, thus, visible pathways at rank($\|\boldsymbol{\Omega}_C\|^2$) = 3. Hence, to gain non-normality and exit region 1 from a major pathway, the flow undertakes a circuit around the $Q - R$ diagram taking low probability pathways that are not shown in Fig. 8.

That relatively few major edges connect region 1 to region 5 reflects the results in Fig. 6(a) where the stronger pathways were counterclockwise (from region 1 to 2). These results are of relevance to turbulence modeling because the restricted Euler model implies a direct connection between the $Q > 0$, enstrophy-dominated region 1 and the focus for preferential dissipation along the Vieillefosse tail at the boundary between regions 5 and 6. With dissipation in classical closures, such as the Smagorinsky subgrid-scale closure in large eddy simulation [65] related directly to $\|\boldsymbol{S}\|^2$, the shape of the Q-R diagram (Fig. 1) means that dissipation will arise preferentially in regions 5 and 6. From the perspective of the resticted Euler model, the impossibility of Lagrangian trajectories moving from $\Delta > 0$ to $\Delta < 0$ means that region 5 acts as the focus of dissipation for all of regions 3, 2, and 1. Our results highlight the importance of the non-normality for mediating these dynamics and, thus, the important role played by the deviatoric pressure Hessian in turbulence dissipation. The nonlocal nature of this term [17] therefore highlights the potential role for nonlocal closure models.

Figures 8(a) and 8(c) show that one node in region 2 ($Q > 0$, $R < 0$) has both a minimal ranking for the non-normality, rank($\|\boldsymbol{\Omega}_C\|^2$) = 3, and large out-degree (within the largest 10%). In contrast, none of the nodes with rank($\|\boldsymbol{\Omega}_C\|^2$) = 3 in regions 3 or region 5 have as large an out-degree (all are shown as gray diamonds). Therefore, this node in region 2 plays a unique role as it provides a path to move clockwise from region 3 to region 2 with low non-normality. Given that $\|\boldsymbol{\Omega}_B\|^2 = 0$ in regions 4 and 6, as the VGT moves clockwise and crosses the left Vieillefosse tail in to region 3 its eigenvalues become complex and the VGT gains some $\|\boldsymbol{\Omega}_B\|^2$. In region 3, the normal enstrophy production is positive since $R < 0$ and, consequently, $\|\boldsymbol{\Omega}_B\|^2$ increases rapidly and overtakes the ranking of the non-normality so that a number of nodes are shown at rank($\|\boldsymbol{\Omega}_C\|^2$) = 3 in Fig. 8(c). There is then a single path from these nodes to the one described in region 2. Given that in region 2, by definition, $\|\boldsymbol{\Omega}_B\|^2 > \|\boldsymbol{S}_B\|^2$ then for this node, $\|\boldsymbol{\Omega}_B\|^2 > \|\boldsymbol{S}_B\|^2 > \|\boldsymbol{\Omega}_C\|^2$, meaning that this is a coherent flow structure pathway in the sense of [66,67] as the structures have a strong solid-body rotation and low non-normality. This "coherent flow structure" pathway is rather novel as the majority of the major connections between regions 3 and 2 arise where $\|\boldsymbol{\Omega}_C\|^2 > \|\boldsymbol{\Omega}_B\|^2 > \|\boldsymbol{S}_B\|^2$.

The densest connections are between regions 5 and 6 as anticipated from the results in Fig. 5(a), and a comparison of Figs. 8(e) and 8(g) shows that typically these arise where rank($\|\boldsymbol{\Omega}_C\|^2$) = 2 rather than rank($\|\boldsymbol{\Omega}_C\|^2$) = 1. Figures 8(e) and 8(g) include edges pointing in both directions. This indicates that the VGT state can oscillate between these two Q-R regions, driving the large eigenvector centralities seen in Fig. 3(a) for these regions and the well-known concentration of the VGT state along the Vieillefosse tail [11,13,68]. Finally, Figs. 8(b) and 8(c) indicate that the major transfers between regions where $Q < 0$ [regions 5 and 6 in Fig. 8(b) and 3 and 4 in Fig. 8(c)] take place at a constant ranking of the non-normality. Hence, regions 3 and 5 are crucial in providing "vertical" pathways for re-organizing the importance of the non-normality relative to the normal straining, which is greater than the normal enstrophy in these regions by definition. Hence, regions 3 and 5 have the smallest values for the median betweenness centrality $\langle C_b^X \rangle_{50}$ in Table II as they are more focused on internal flow reconfiguration rather than transitions to other regions.

VI. SUMMARY AND CONCLUSION

Representing the Lagrangian dynamics of the velocity gradient tensor as a complex network has allowed us to explore the VGT time evolution conditional on several discrete variables, something challenging to achieve in a statistically robust manner in a continuous framework. The complex network associated with the VGT dynamics comprised $N = 848$ nodes, representing different states of the flow based on the signs and relative magnitudes of different invariants of the velocity gradient tensor. In practice, this was reduced to $N = 844$ nodes because four states were very rarely frequented as seen in the inset to Fig. 2(b), with the properties of these four nodes given in Table I. The nodes are defined in (13) and include terms up to third order as well as their Schur-decomposed versions [42], and the discriminant function, Δ, which is a combination of two such terms as given in (4). An advantage of defining the nodes in terms of the mathematical properties of the tensor, rather than the flow physics is that it provides a certain universality, permitting an analysis of the characteristics of different flows, or the same flow at different Reynolds numbers, or in our case, a direct numerical simulation (DNS) due to [53] and the enhanced Gaussia closure (EGC) model of [24]. This is because, with a constant nodal structure, differences between flows will be seen in the edge weights, which capture the probability of transitioning between the discrete flow states, following the Lagrangian evolution of the flow along fluid particle trajectories. The performance of the EGC model is compared to the DNS in Figs. 5 and 6 and is found to perform well, although it struggles to correctly model the trajectories in the counter direction to the restricted Euler model where the eigenvalues of the VGT are real and where the trajectories transition from real to complex eigenvalues.

We have analyzed the network using a variety of well-established metrics of network structure [29,30] as well as variants we have formulated to focus on the transitions between nodes belonging to different regions of the Q-R plane. We have also considered the shortest paths on the network, conditional on their starting and arrival region, discerning between those starting and ending in the same Q-R region (intraregion shortest paths) and those originating and ending in distinct Q-R regions (interregion shortest paths). There is a preferential grouping of nodes with high eigenvector centrality on the two regions adjacent to the right Vieillefosse tail [13,15]. This is particularly the case for region 6, immediately below the Vieillefosse tail (where $Q < 0$, $R > 0$, $\Delta < 0$) as seen in Fig. 3(a). Such a strong bias towards this region is reduced when the network is constructed in terms of edges between regions only Fig. 3(b), indicating it is connections within region 6 that drive the strong grouping of nodes with high connectivity.

In addition to gaining an understanding of the VGT dynamics using network-related metrics, the reduced number of degrees of freedom in the network provide a means to gain new insights into the VGT dynamics more generally. For example, we found that regions 1 and 5 were particularly unusual in terms of having long paths for some of their intraregion shortest paths (Fig. 4). The reason for this was that gaining sufficient non-normal characteristics to access other nodes within this region requires the flow to take less frequented paths that resulted in an orbit the Q-R diagram. The driving mechanism for this differed for the two regions: In region 1 ($Q > 0$, $R > 0$) these trajectories arose for flow states that needed to increase the non-normality, $\|\boldsymbol{\Omega}_C\|^2$, from low to high values [see Fig. 8(b)], while in region 5 these long intraregion shortest paths were associated with developing negative production of the non-normality by the action of the normal straining (Table III). Regions 1 and 5 are the Q-R regions where the deviatoric pressure Hessian introduces convoluted probability fluxes, very different from the underlying restricted Euler part [19,23,63] and these dynamical complexities are clearly captured by our analysis, but we have been able to go further using the discrete approach to highlight particular flow states that drive the restricted Euler and deviatoric pressure Hessian components of the fluxes.

Additional emphasis on the non-normality comes from the random walk commute distances from node i and back to i via node j. The nodes $j_{\min}$, responsible for the shortest commutes preferentially have the non-normal rotation rate $\boldsymbol{\Omega}_C$ larger in magnitude than either the normal strain rate, $\boldsymbol{S}_B$, or the normal rotation rate, $\boldsymbol{\Omega}_B$. Such vertices with large non-normality form the upper layer in the reduced network shown in Fig. 8, and the edges between such nodes are of particular importance for the transitions from the strain-dominated, $Q < 0$, to entrophy dominated, $Q > 0$, regions and vice versa [see Figs. 8(b) and 8(c)]. The non-normality $\|\boldsymbol{\Omega}_C\|^2$ does not feature explicitly in the expression of the VGT principal invariants as seen from Eqs. (9) and (12), demonstrating the importance of extending the discrete and qualitative description of the VGT beyond the Q-R diagram. A previous approach in this direction due to [69] considered Q together with the strain production

(11) and enstrophy production (12) separately. The Schur decomposition of the VGT enriches this by including the separate effects of normal enstrophy production, $\mathrm{tr}(\boldsymbol{\Omega}_B^2 \boldsymbol{S}_B)$, normal strain production, $-\det(\boldsymbol{S}_B)$, non-normal production, $-\det(\boldsymbol{S}_C)$ and interaction production, $\mathrm{tr}(\boldsymbol{\Omega}_C^2 \boldsymbol{S}_B)$. It is the positivity of the last term that drives the general positivity of the classical enstrophy production and strain production [42]. Hence, that negative values for this term are important for the long intraregion shortest paths starting from region 5 as shown in Table III and the minimum commute distances from regions 1 and 5 to region 3 (Fig. 7) highlights the distinct role of a component of the production budget that is not isolated in conventional analyses that utilize the left-hand sides of (11) and (12).

Translating an infinite number of continuous flow states into a network of a finite number of discrete states provides several advantages to tackling the multiscale complexity of turbulence [16]. It constitutes a viable approach for elucidating the physics of the VGT through the application of network theory, also opening up the potential to use discrete control formulations for, e.g., turbulent drag reduction. It also provides a means to test the effectiveness of current models for capturing the relevant dynamics. Our results in Figs. 5 and 6 show that such models capture the general characteristics of the dynamics, but struggle to place sufficient emphasis on the most important nodes. In regions 3 and 5 (where $Q < 0$ and $\Delta > 0$) this arises because too many nodes are involved with interregion movements, while the other major source of discrepancy is in regions 5 and 6 ($Q < 0$, $R > 0$) where dynamics arise that are inadmissible in the restricted Euler model and therefore are stringent tests of modeling capability: the crossing the Vieillefosse tail from region 5 to 6, which involves a move from complex to real eigenvalues for the VGT; and the path from region 6 to region 4 ($R < 0$) with all real eigenvalues that is counter to the direction inferred from the restricted Euler dynamics. Our results are highly suggestive that an improvement to such models could result from explicit rather than implicit modeling of the non-normal dynamical quantities we have incorporated explicitly into our definition of the nodes of our network.

ACKNOWLEDGMENTS

The authors thank M. Wilczek for fruitful discussions about this work and for hosting C.K. in Bayreuth. C.K. acknowledges financial support from Leverhulme Trust International Fellowship 2023-014 that permitted the visit to Bayreuth that enabled this research.

DATA AVAILABILITY

The data used in this study were extracted from the Johns Hopkins Turbulence database, available in Ref. [70].

[1] J. C. Vassilicos, Dissipation in turbulent flows, Annu. Rev. Fluid Mech. **47**, 95 (2015).

[2] M. Carbone and A. D. Bragg, Is vortex stretching the main cause of the turbulent energy cascade? J. Fluid Mech. **883**, R2 (2020).

[3] P. L. Johnson, Energy transfer from large to small scales in turbulence by multiscale nonlinear strain and vorticity interactions, Phys. Rev. Lett. **124**, 104501 (2020).

[4] D. Buaria, A. Pumir, E. Bodenschatz, and P. K. Yeung, Extreme velocity gradients in turbulent flows, New J. Phys. **21**, 043004 (2019).

[5] A. Vela-Martín, The energy cascade as the origin of intense events in small-scale turbulence, J. Fluid Mech. **937**, A13 (2022).

[6] C. Meneveau, Lagrangian dynamics and models of the velocity gradient tensor in turbulent flows, Annu. Rev. Fluid Mech. **43**, 219 (2011).

[7] U. Frisch, *Turbulence* (Cambridge University Press, Cambridge, 1995).

[8] K. Ohkitani, Study of the Navier–Stokes regularity problem with critical norms, Fluid Dyn. Res. **48**, 021401 (2016).
[9] T. D. Drivas and T. M. Elgindi, Singularity formation in the incompressible Euler equation in finite and infinite time, EMS Surv. Math. **10**, 1 (2023).
[10] E. Miller, Finite-time blowup for a Navier-Stokes model equation for the self-amplification of strain, Analysis & PDE **16**, 997 (2023).
[11] P. Vieillefosse, Internal motion of a small element of fluid in an inviscid flow, Physica A **125**, 150 (1984).
[12] A. Tsinober, L. Shtilman, and H. Vaisburd, A study of properties of vortex stretching and enstrophy generation in numerical and laboratory turbulence, Fluid Dyn. Res. **21**, 477 (1997).
[13] B. J. Cantwell, On the behavior of velocity gradient tensor invariants in direct numerical simulations of turbulence, Phys. Fluids **5**, 2008 (1993).
[14] A. Ooi, J. Martin, J. Soria, and 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**, 141 (1999).
[15] J. M. Chacin and B. J. Cantwell, Dynamics of a low Reynolds number turbulent boundary layer, J. Fluid Mech. **404**, 87 (2000).
[16] P. L. Johnson and M. Wilczek, Multiscale velocity gradients in turbulence, Annu. Rev. Fluid Mech. **56**, 463 (2024).
[17] K. Ohkitani and S. Kishiba, Nonlocal nature of vortex stretching in an inviscid fluid, Phys. Fluids **7**, 411 (1995).
[18] P. Vieillefosse, Local interaction between vorticity and shear in a perfect incompressible fluid, J. Phys. France **43**, 837 (1982).
[19] B. J. Cantwell, Exact solution of a restricted Euler equation for the velocity gradient tensor, Phys. Fluids A **4**, 782 (1992).
[20] J. H. Chen, M. S. Chong, J. Soria, R. Sondergaard, A. E. Perry, M. M. Rogers, R. D. Moser, and B. J. Cantwell, A study of the topology of dissipating motions in direct numerical simulations of time-developing compressible and incompressible mixing layers Phys. Fluids **6**, 871 (1994).
[21] S. S. Girimaji and S. B. Pope, A diffusion model for velocity gradients in turbulence, Phys. Fluids **2**, 242 (1990).
[22] J. Martín, A. Ooi, M. S. Chong, and J. Soria, Dynamics of the velocity gradient tensor invariants in isotropic turbulence, Phys. Fluids **10**, 2336 (1998).
[23] L. Chevillard, C. Meneveau, L. Biferale, and 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**, 101504 (2008).
[24] M. Wilczek and C. Meneveau, Pressure Hessian and viscous contributions to velocity gradient statistics based on Gaussian random fields, J. Fluid Mech. **756**, 191 (2014).
[25] P. L. Johnson and C. Meneveau, A closure for Lagrangian velocity gradient evolution in turbulence using recent-deformation mapping of initially Gaussian fields, J. Fluid Mech. **804**, 387 (2016).
[26] D. Buaria and K. R. Sreenivasan, Forecasting small-scale dynamics of fluid turbulence using deep neural networks, Proc. Natl. Acad. Sci. USA **120**, e2305765120 (2023).
[27] M. Carbone, V. J. Peterhans, A. S. Ecker, and M. Wilczek, Tailor-designed models for the turbulent velocity gradient through normalizing flow, Phys. Rev. Lett. **133**, 184001 (2024).
[28] B. Sharma, R. Das, and S. S. Girimaji, Local vortex line topology and geometry in turbulence, J. Fluid Mech. **924**, A13 (2021).
[29] V. Latora, V. Nicosia, and G. Russo, *Complex Networks: Principles, Methods and Applications* (Cambridge University Press, New York, 2017).
[30] M. E. J. Newman, The structural and function of complex networks, SIAM Rev. **45**, 167 (2003).
[31] X. Xiao, H. Chen, and P. Bogdan, Deciphering the generating rules and functionalities of complex networks, Sci. Rep. **11**, 22964 (2021).
[32] A. G. Nair and K. Taira, Network-theoretic approach to sparsified discrete vortex dynamics, J. Fluid Mech. **768**, 549 (2015).
[33] K. Taira, A. G. Nair, and S. L. Brunton, Network structure of two-dimensional decaying isotropic turbulence, J. Fluid Mech. **795**, R2 (2016).

[34] S. Scarsoglio, G. Iacobello, and L. Ridolfi, Complex networks unveiling spatial patterns in turbulence, Int. J. Bifurcation Chaos **26**, 1650223 (2016).
[35] S. Chowdhuri, G. Iacobello, and T. Banerjee, Visibility network analysis of large-scale intermittency in convective surface layer turbulence, J. Fluid Mech. **925**, A38 (2021).
[36] G. Iacobello, S. Chowdhuri, L. Ridolfi, L. Rondoni, and S. Scarsoglio, Coherent structures at the origin of time irreversibility in wall turbulence, Commun. Phys. **6**, 91 (2023).
[37] E. Ser-Giacomi, V. Rossi, C. López, and E. Hernandez-Garcia, Flow networks: A characterization of geophysical fluid transport, Chaos **25**, 036404 (2015).
[38] S. B. Pope, A more general effective-viscosity hypothesis, J. Fluid Mech. **72**, 331 (1975).
[39] M. Itskov, *Tensor Algebra and Tensor Analysis for Engineers* (Springer International Publishing, Cham, 2015).
[40] L. A. Leppin and M. Wilczek, Capturing velocity gradients and particle rotation rates in turbulence, Phys. Rev. Lett. **125**, 224501 (2020).
[41] M. Carbone and M. Wilczek, Asymptotic predictions on the velocity gradient statistics in low-Reynolds number random flows: Onset of skewness, intermittency and alignments, arXiv:2305.03454 (2023).
[42] C. J. Keylock, The Schur decomposition of the velocity gradient tensor for turbulent flows, J. Fluid Mech. **848**, 876 (2018).
[43] D. Fernex, B. R. Noack, and R. Semaan, Cluster-based network modeling—From snapshots to complex dynamical systems, Sci. Adv. **7**, eabf5006 (2021).
[44] G. Iacobello, F. Kaiser, and D. E. Rival, Load estimation in unsteady flows from sparse pressure measurements: Application of transition networks to experimental data, Phys. Fluids **34**, 025105 (2022).
[45] P. Chakraborty, S. Balachandar, and R. J. Adrian, On the relationships between local vortex identification schemes, J. Fluid Mech. **535**, 189 (2005).
[46] D. Perrone, J. G. M. Kuerten, L. Ridolfi, and S. Scarsoglio, Wall-induced anisotropy effects on turbulent mixing in channel flow: A network-based analysis, Phys. Rev. E **102**, 043109 (2020).
[47] J. C. R. Hunt, A. A. Wray, and P. Moin, *Eddies, stream, and convergence zones in turbulent flows*, Technical Report CTR-S88 (Center for Turbulence Research, Stanford University, 1988), p. 193.
[48] R. Betchov, An inequality concerning the production of vorticity in isotropic turbulence, J. Fluid Mech. **1**, 497 (1956).
[49] P. Henrici, Bounds for iterates, inverses, spectral variation and fields of values of non-normal matrices, Numer. Math. **4**, 24 (1962).
[50] P. Bonacich, Factoring and weighting approaches to status scores and clique identification, J. Math. Sociol. **2**, 113 (1972).
[51] K. K. Nomura and G. K. Post, The structure and dynamics of vorticity and rate of strain in incompressible homogeneous turbulence, J. Fluid Mech. **377**, 65 (1998).
[52] M. Carbone and M. Wilczek, Only two Betchov homogeneity constraints exist for isotropic turbulence, J. Fluid Mech. **948**, R2 (2022).
[53] Y. Li, E. Perlman, M. Wan, Y. Yang, R. Burns, C. Meneveau, S. Chen, A. Szalay, and G. Eyink, A public turbulence database cluster and applications to study Lagrangian evolution of velocity increments in turbulence, J. Turbul. **9**, N31 (2008).
[54] H. Yu, K. Kanov, E. Perlman, J. Graham, E. Frederix, R. Burns, A. Szalay, G. Eyink, and C. Meneveau, Studying Lagrangian dynamics of turbulence using on-demand fluid particle tracking in a public turbulence database, J. Turbul. **13**, N12 (2012).
[55] C. J. Keylock, S. N. Lane, and K. S. Richards, Quadrant/octant sequencing and the role of coherent structures in bed load sediment entrainment, JGR Earth Surface **119**, 264 (2014).
[56] J. Lawson and J. Dawson, On velocity gradient dynamics and turbulent structure, J. Fluid Mech. **780**, 60 (2015).
[57] L. C. Freeman, A set of measures of centrality based upon betweenness, Sociometry **40**, 35 (1977).
[58] S. P. Borgatti and M. G. Everett, A graph-theoretic perspective on centrality, Social Netw. **28**, 466 (2006).
[59] D. Boley, G. Ranjan, and Z.-L. Zhang, Commute times for a directed graph using an asymmetric Laplacian, Linear Algebra Appl. **435**, 224 (2011).
[60] A.-L. Barabási and R. Albert, Emergence of scaling in random networks, Science **286**, 509 (1999).

[61] 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**, 283 (2019).
[62] C. J. Keylock, Synthetic velocity gradient tensors and the identification of significant aspects of the structure of turbulence, Phys. Rev. Fluids **2**, 084607 (2017).
[63] Y. Zhou, K. Nagata, Y. Sakai, Y. Ito, and T. Hayase, On the evolution of the invariants of the velocity gradient tensor in single square-grid-generated turbulence, Phys. Fluids **27**, 075107 (2015).
[64] F. Harary, Status and contrastatus, Sociometry **22**, 23 (1959).
[65] J. Smagorinsky, General circulation experiments with the primitive equations, Mon. Weather Rev. **91**, 99 (1963).
[66] M. S. Chong, A. E. Perry, and B. J. Cantwell, A general classification of three-dimensional flow fields, Phys. Fluids A **2**, 765 (1990).
[67] H. Xu, X. S. Cai, and C. Liu, Liutex (vortex) core definition and automatic identification for turbulence vortex structures, J. Hydrodyn. **31**, 857 (2019).
[68] S. Laizet, J. Nedic, and J. C. Vassilicos, Influence of the spatial resolution on fine-scale features in DNS of turbulence generated by a single square grid, Int. J. Comput. Fluid Dyn. **29**, 286 (2015).
[69] B. Lüthi, M. Holzner, and A. Tsinober, Expanding the Q-R space to three dimensions, J. Fluid Mech. **641**, 497 (2009).
[70] https://turbulence.idies.jhu.edu/home.