

Quantifying collective motion in living matter with spatial correlation functions

Hongmei Xu* and Yucheng Huo

School of Mechanical and Aerospace Engineering, *Nanyang Technological University*, Singapore 637460, Singapore

K. Jimmy Hsia[†]

School of Mechanical and Aerospace Engineering, *Nanyang Technological University*, Singapore 637460, Singapore
and School of Chemistry, Chemical Engineering and Biotechnology, *Nanyang Technological University*, Singapore 637371, Singapore

Endao Han[‡]

Division of Physics and Applied Physics, School of Physical and Mathematical Sciences,
Nanyang Technological University, Singapore 637371, Singapore



(Received 3 December 2025; accepted 8 March 2026; published 12 August 2026)

Active and living matter exhibits a wide range of complex phenomena, including collective motion, pattern formation, and hierarchical structures. Collective motion in living systems is often quantified using spatial correlation functions; however, competing definitions—especially for velocity fields—yield inconsistent results across studies. Here, we clarify the distinction between “normalize-first” and “average-first” velocity correlations, and establish a robust and practical framework for computing correlation functions and extracting director and velocity correlation lengths, ξ_d and ξ_v , respectively, from a wide range of active cell systems. Using synthetic shear and vortex fields and experimental datasets spanning epithelial monolayers, neural progenitors, and bacterial colonies, we demonstrate how noise and fitting choices affect inferred length scales, and provide recipes that yield consistent outcomes. Finally, we introduce two state diagrams— ξ_v versus ξ_d and ξ_v/ξ_d versus motility anisotropy—that locate systems within different dynamical regimes and track their trajectories during development. These tools standardize correlation analysis and link single-cell motility to emergent colony-scale flows, enabling rigorous, comparable measurements across living systems.

DOI: [10.1103/6zq7-33zc](https://doi.org/10.1103/6zq7-33zc)

I. INTRODUCTION

Collective migration is observed in various biological systems, from a school of fish or a flock of birds at the animal scale to colonies of living cells or bacteria at the cellular scale. The collective migration of living cells is essential in producing various biological structures and performing biological functions, such as wound healing [1–9], biofilm formation [10–16], embryonic development [17], and formation of hierarchical structures [18–20]. These collective behaviors stem from the motility of individual cells and intricate intercellular biophysical and biochemical interactions. Even when these interactions are short-ranged, long-range collective motion can emerge, which has been the focus of active matter research.

Depending on the polarity of individual cells and the interactions between them, the colony can be modeled as polar or

nematic active matter. In a polar system, each particle has a head-tail asymmetry (polarity), and the neighboring particles are forced to align toward the same direction, like magnetic dipoles in a ferromagnetic material [21]. In contrast, a nematic system can be made of either polar or apolar particles. Still, its dynamics are governed by the local orientation of these particles rather than the polarity, which is analogous to a liquid crystal [22,23]. In a colony with high cell density, the cell polarity or orientation can be represented by a polarity or director field, respectively. Furthermore, the colony’s local cell density, velocity, and mechanical forces can all be represented as fields. The spatial and temporal correlations of these fields provide valuable insights into the dynamics of these cell collectives.

Spatial correlation functions are often employed to identify phase transitions in physics. In active matter, the spatial correlation functions of director and velocity fields are used to measure if the system has local, quasi-long-range, or long-range orders [2,24–28]. In living systems composed of cells, it is a valuable tool to characterize and understand collective cell behaviors quantitatively. Figure 1 shows examples of a low-density, less ordered state [panel (a)–(c)], and a high-density, more ordered state [panel (d)–(f)] in a C2C12 cell monolayer. C2C12 is an immortalized mouse myoblast cell line and a standard model for muscle development. During migration, their locomotion is driven by leading-edge actin

*Contact author: hongmei001@e.ntu.edu.sg

†Contact author: kjhsia@ntu.edu.sg

‡Contact author: endao.han@ntu.edu.sg

Published by the American Physical Society under the terms of the *Creative Commons Attribution 4.0 International* license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

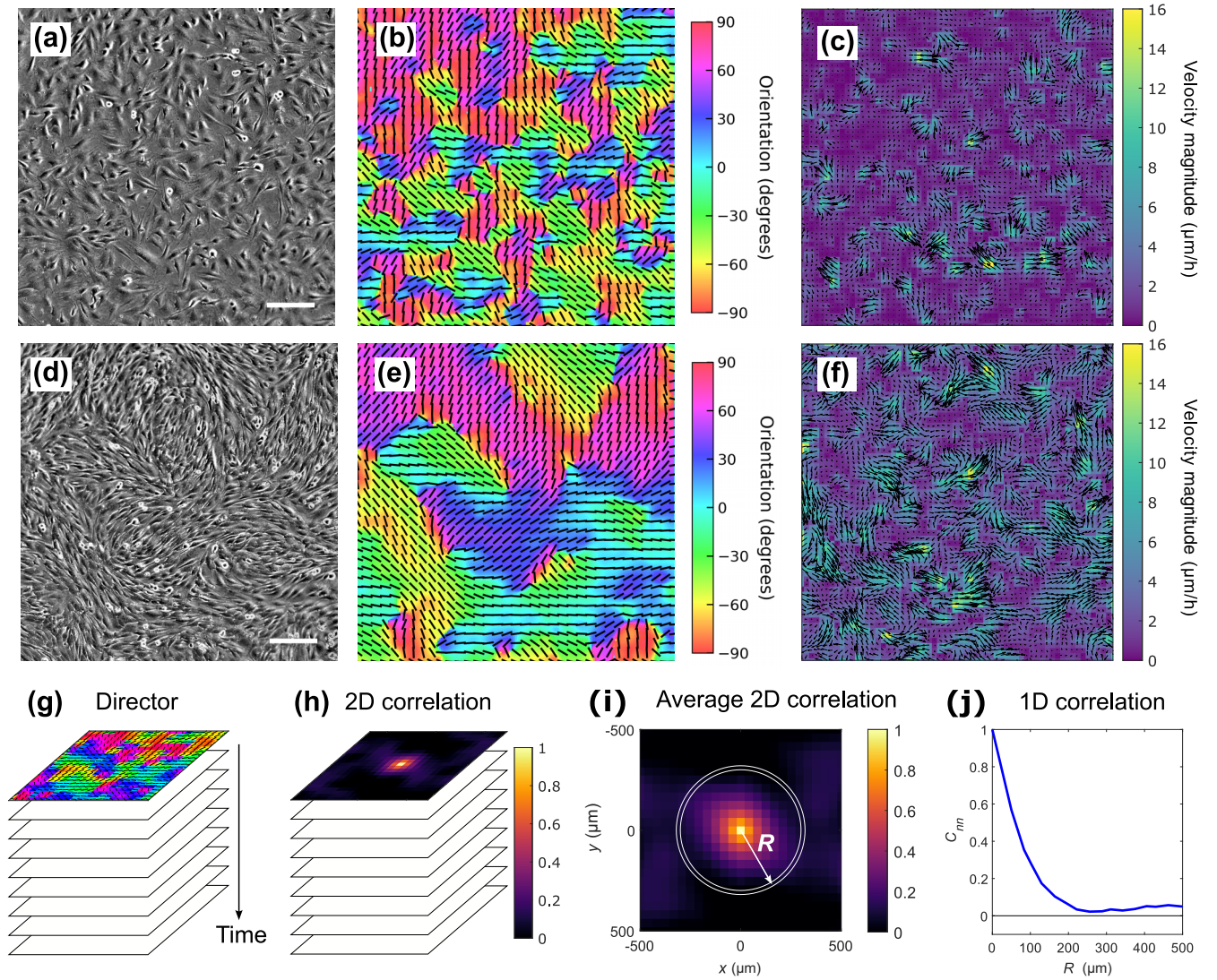


FIG. 1. C2C12 cell monolayers with lower (top row) or higher (bottom row) nematic order. Scale bars: 200 μm. (a) A bright-field image of a C2C12 cell monolayer with low cell density. (b) Cell orientation (director) field of panel (a). (c) Velocity field of panel (a). (d) A bright-field image of a C2C12 cell monolayer with high cell density. (e) Cell orientation (director) field of panel (d), which shows longer-range nematic order compared to panel (b), with larger regions of aligned cells. (f) Velocity field of panel (d). Larger groups of cells move collectively in panel (f) compared to panel (c). (g) A time series of director fields. (h) A series of two-dimensional (2D) correlation maps obtained from the director fields in panel (g). (i) 2D correlation map averaged over time. At each radius R , we averaged this 2D map azimuthally to obtain the one-dimensional (1D) correlation curve $C_{\hat{n}\hat{n}}(R)$ in panel (j).

protrusion and adhesion, followed by cell-body translocation and rear retraction. The colony behaves as an active nematic system, in which half-integer topological defects [10,25,29–37] are observed. The increase in order is evident in the director field, which shows fewer defects and larger regions with cell alignment, and in the velocity field, where larger groups of collectively moving cells are observed. To characterize the ordered state of the cell monolayer, we use the spatial correlation function to determine the extent of correlation in cell orientation or velocity.

Previous work has discussed extensively how to draw information from the correlation functions measured in active matter or living systems [38]. However, in practice, measurements obtained from cell colonies, especially cell velocities, have large fluctuations due to the intrinsic

stochasticity of living cells and their complex interactions. At the same time, due to the trade-off between the spatial resolution and field of view in microscopy, the experimental system size is often limited. Consequently, it is often challenging to interpret the physical meanings of the experimentally measured correlation functions compared to those obtained from numerical simulations. For example, it is useful to measure if the correlation function follows an exponential decay or power-law decay in the long-range limit, or if the long-range correlation decays to zero or a finite value, which distinguishes the quasi- and true-long-range orders. But such fittings are nontrivial due to the reasons stated above. Furthermore, various ways to calculate the velocity correlation function have been reported in different research. Some of these definitions cannot properly reflect the features in

the velocity field and may lead to misinterpretations of the results.

This paper seeks a practical approach to understanding collective cell motion with the help of spatial correlations. First, we address the ambiguities in the definitions of the correlation function by reviewing the existing ones in the literature, highlighting the inconsistencies that may confound result interpretation, and proposing our suggestions. Second, we investigate the correlation functions of two idealized flow patterns in cell migration: shear flows and isolated vortices. Third, we study the effects of fluctuations on the measured correlation functions. The fluctuations include intrinsic noise due to the stochastic cell activity in self-propulsion, and extrinsic noise, which is introduced by image taking and processing. Then, we compare and evaluate multiple methods to extract the correlation length from the correlation function. Finally, we propose two simple diagrams comparing director and velocity correlation lengths, and linking them with cell-motility anisotropy across multiple biological systems, which help categorize different regimes of collective behaviors and track transitions between states. Overall, our goal is to improve clarity in interpreting spatial correlation measurements, propose a standardized definition to address existing inconsistencies, and offer valuable data analysis tools for the broader biological and bioengineering communities.

II. SPATIAL CORRELATION FUNCTIONS FOR DIRECTOR AND VELOCITY FIELDS

Active matter can be nematic or polar, depending on the symmetry of its constituent particles and particle-particle interactions. Polar particles have a head-tail asymmetry, while nematic particles are rodlike but have two identical ends. Polar active systems generate integer (± 1) defects, while nematic systems generate half-integer ($\pm 1/2$) defects. In this article, we discuss only the latter. Note that a nematic system can be made of either polar or apolar particles. For example, each swimming or gliding bacterium has an instantaneous direction of motion, thus polar, but the overall dynamics of their colonies indicate that they are nematic systems, where cell orientation is more important. The orientation of each cell is represented by a director $\hat{n} = (\cos \theta, \sin \theta)$ along its long axis, where $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ is called the orientation angle [see Figs. 1(b), 1(e), and 8(a)]. In a cell layer, the directors can be represented by a field $\hat{n}(x, y)$. The velocity field $\mathbf{v}(x, y)$ is another important vector field that indicates the local cell velocity [see Figs. 1(c), 1(f), and 8(c)]. Since $\hat{n}(x, y)$ and $\mathbf{v}(x, y)$ can be easily extracted from cell micrographs, they are widely used to test theoretical predictions and probe the dynamics of the living system. A common quantitative tool is the correlation function; however, the literature contains conflicting definitions and calculation methods.

We now examine the existing definitions of spatial correlation functions for $\hat{n}(x, y)$ and $\mathbf{v}(x, y)$, which are defined differently due to their different symmetries. First, we look at the correlation between a pair of directors or velocities. For directors, we define the pairwise nematic correlation as $2 \cos^2 \Delta \theta - 1$, where $\Delta \theta$ ($< \pi$) is the angle difference between these two directors. Consequently, a pair of directors has a correlation of $+1$ when they are parallel, -1 when they

are perpendicular, and 0 when the angle between them is $\pi/4$. For a velocity pair, the correlation can be written as $\frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{|\mathbf{v}_1||\mathbf{v}_2|}$. With this definition, two velocity vectors pointing in the same direction yield a correlation $+1$, those pointing in opposite directions yield a correlation -1 , and the perpendicular ones yield zero correlation regardless of their magnitudes [Fig. 8(d) illustrates these examples]. Additionally, suppose that one velocity vector changes direction randomly while the other remains constant, the temporally averaged correlation is also zero.

Based on the correlation of director or velocity pairs, spatial correlation functions are defined for the fields $\hat{n}(x, y)$ and $\mathbf{v}(x, y)$. For $\hat{n}(x, y)$, the spatial correlation function $C_{\hat{n}\hat{n}}(R)$ is [10,24,39–41]

$$C_{\hat{n}\hat{n}}(R) = \langle 2 \cos^2 [\theta(\mathbf{r}_j) - \theta(\mathbf{r}_j + \mathbf{R})] \rangle_j - 1 \rangle_\varphi = \langle \hat{n}_{2\theta}(\mathbf{r}_j) \cdot \hat{n}_{2\theta}(\mathbf{r}_j + \mathbf{R}) \rangle_j \rangle_\varphi, \quad (1)$$

where $\langle \cdot \rangle_j$ denotes an average over all sampled positions $\mathbf{r}_j$ in the field of view (e.g., grid points or cell centers), $\langle \cdot \rangle_\varphi$ denotes an average over all azimuthal directions of $\mathbf{R}$ at fixed $R = |\mathbf{R}|$, and $\hat{n}_{2\theta} = (\cos 2\theta, \sin 2\theta)$. In Figs. 1(g)–1(j), we illustrate how a 1D correlation function is obtained from a time series of director fields. For each frame [Fig. 1(g)], we first compute the corresponding 2D spatial correlation map [Fig. 1(h)]. We then average these 2D maps over time to obtain a temporally averaged correlation map. Finally, we perform azimuthal averaging of this map over all directions at fixed separation R [Fig. 1(i)], yielding the 1D correlation curve $C_{\hat{n}\hat{n}}(R)$ shown in Fig. 1(j).

The definition of $C_{\hat{n}\hat{n}}(R)$ is generally agreed upon. In contrast, the methods to calculate the velocity correlation function vary in the literature and thus require careful investigation. In general, there are two ways to calculate the correlation function, depending on how the normalization is implemented. One approach is the correlation of normalized velocities, referred to as the “normalize-first” method, denoted by $C_{\hat{v}\hat{v}}(R)$:

$$C_{\hat{v}\hat{v}}(R) = \left\langle \left\langle \frac{\delta \mathbf{v}(\mathbf{r}_j)}{|\delta \mathbf{v}(\mathbf{r}_j)|} \cdot \frac{\delta \mathbf{v}(\mathbf{r}_j + \mathbf{R})}{|\delta \mathbf{v}(\mathbf{r}_j + \mathbf{R})|} \right\rangle_j \right\rangle_\varphi, \quad (2)$$

where $\delta \mathbf{v}(\mathbf{r}_j) = \mathbf{v}(\mathbf{r}_j) - \mathbf{v}_B$ is the cell velocity with the velocity of the background drift in the frame $\mathbf{v}_B$ removed. The velocities are normalized to unit vectors before calculating the spatial and angular averages [42–45]. The other approach is referred to as the “average-first” method, denoted by $C_{vv}(R)$:

$$C_{vv}(R) = \left\langle \frac{\langle \delta \mathbf{v}(\mathbf{r}_j) \cdot \delta \mathbf{v}(\mathbf{r}_j + \mathbf{R}) \rangle_j}{\langle |\delta \mathbf{v}|^2 \rangle} \right\rangle_\varphi, \quad (3)$$

where the spatial average is calculated before being normalized by the variance of velocity $\langle |\delta \mathbf{v}|^2 \rangle$. Specifically, there are four numerical methods to compute $\langle |\delta \mathbf{v}|^2 \rangle$ within $C_{vv}(R)$, as summarized in Table I. However, their results differ only slightly. In the following sections, we compare $C_{\hat{v}\hat{v}}(R)$ and $C_{vv}(R)$ to identify which one better captures the features of velocity fields in two-dimensional living systems.

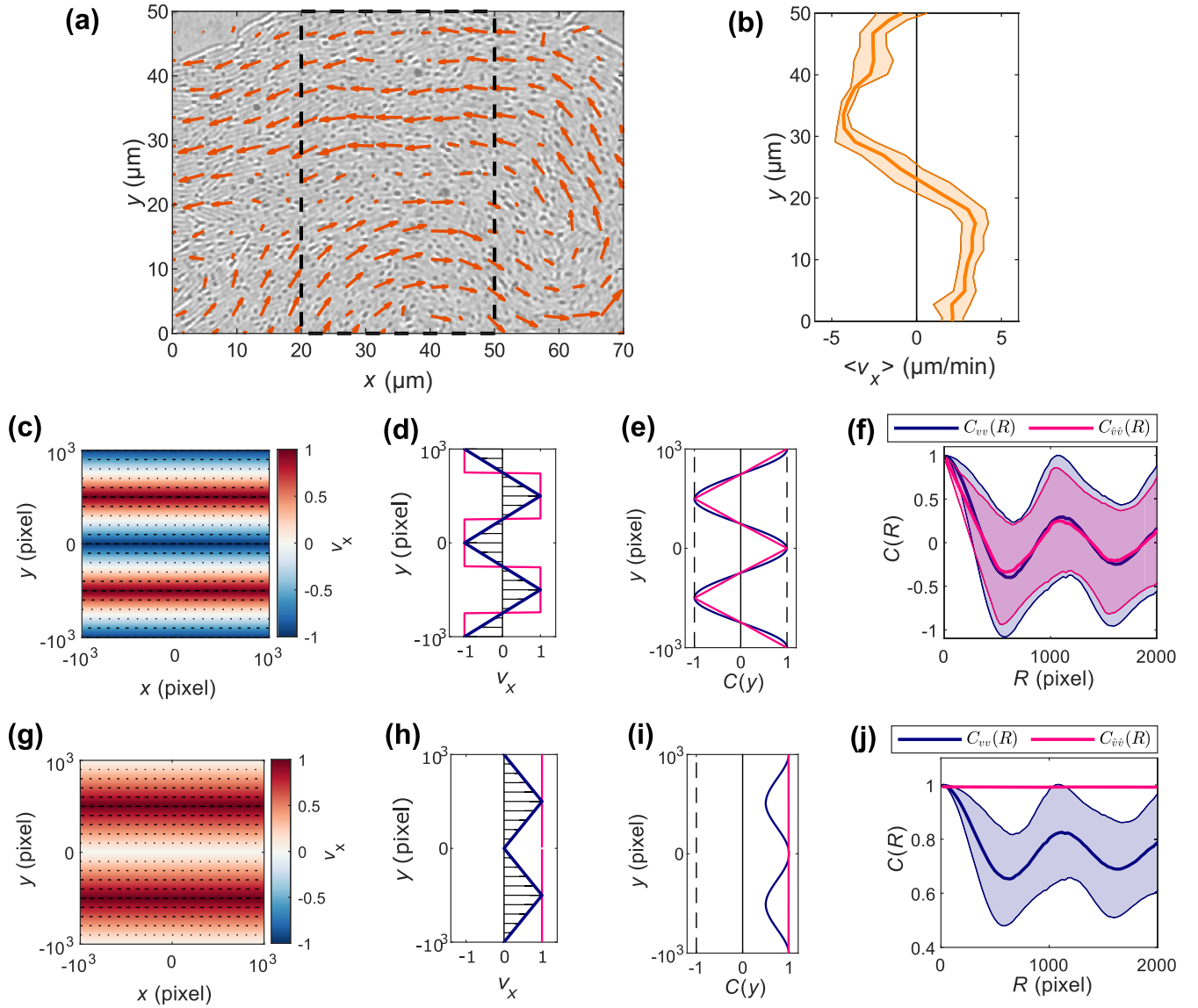


FIG. 2. Spatial correlations of the velocity field in a *M. xanthus* monolayer and simplified periodic shear flows. (a) An exemplary bright field image of *M. xanthus* monolayer on a solid surface and the corresponding velocity field labeled with orange arrows. (b) Horizontal component of the cell velocity averaged in the x direction $\langle v_x \rangle$ within the black dashed rectangle in panel (a). The thick orange curve shows the mean, and the lighter orange shade shows one standard deviation. (c) Velocity field of an artificial periodic shear flow with positive and negative velocities. (d) One-dimensional profile of the horizontal flow $v_x(y)$ in panel (c), shown by the blue lines. The magenta lines show the normalized horizontal velocity profile $\hat{v}_x(y)$. (e) Correlation functions $C_{v_x v_x}(R)$ (blue) and $C_{\hat{v}_x \hat{v}_x}(R)$ (magenta) for $v_x(y)$ in panel (d). (f) Azimuthally averaged correlation functions $C_{vv}(R)$ (blue) and $C_{\hat{v}\hat{v}}(R)$ (magenta) for the 2D flow in panel (c). (g) Velocity field of an artificial periodic shear flow with only positive velocities. (h) 1D profile of the horizontal flow $v_x(y)$ in panel (g) (blue) and the normalized flow $\hat{v}_x(y)$. Symbols identical to panel (d). (i) $C_{v_x v_x}(R)$ (blue) and $C_{\hat{v}_x \hat{v}_x}(R)$ (magenta) for $v_x(y)$ in panel (h). (j) Azimuthally averaged $C_{vv}(R)$ (blue) and $C_{\hat{v}\hat{v}}(R)$ (magenta) for the 2D flow in panel (g).

III. SPATIAL CORRELATION FUNCTIONS OF SHEAR FLOWS

We first examine the velocity correlation functions of a simple flow pattern—shear flow. Figure 2(a) shows an exemplary shear flow observed in a thin colony of a rod-shaped bacterium *Myxococcus xanthus* moving on a solid surface. *Myxococcus xanthus* is a social bacterium with rich multicellular behaviors. It moves on solid surfaces using two distinct motility systems: gliding and twitching (driven by type IV

pili). Interestingly, it also periodically reverses its direction of motion, controlled by an internal Frz-based oscillator sensitive to environmental cues such as nutrient level [46]. In the upper half of the field of view, the bacteria move to the left, while in the lower half, they move to the right. This creates an approximately linear shear flow in the middle, at the communal boundary of the two cell streams, as shown in Fig. 2(b), where we calculated the average horizontal velocity $\langle v_x \rangle$ at different y positions within the dashed rectangle in Fig. 2(a).

Inspired by this type of cell flow, we compared the two approaches to calculating the velocity correlation functions, $C_{\hat{v}\hat{v}}(R)$ and $C_{vv}(R)$, considering two idealized quasi-1D periodic shear flows, as shown in Figs. 2(c)–2(j). In Fig. 2(c), the velocities are all horizontal and only vary in the y direction. The horizontal component $v_x(y)$ is a triangular wave [Fig. 2(d), dark blue line] with a spatial period L :

$$v_x(y) = 1 - 4 \left| \text{frac}\left(\frac{y}{L}\right) - \frac{1}{2} \right|, \quad (4)$$

where $\text{frac}(y/L)$ denotes the fractional part of y/L , in the range $[0, 1)$. When “normalized,” the original triangular velocity profile becomes a square profile

$$\hat{v}_x(y) = \frac{v_x(y)}{|v_x(y)|} = -\text{sgn}\left[\cos\left(\frac{2\pi}{L}y\right)\right], \quad (5)$$

where sgn denotes the sign of the following function [Fig. 2(d), magenta line].

The corresponding correlation functions are both periodic, so we calculated both $C_{v_x v_x}(y)$ and $C_{\hat{v}_x \hat{v}_x}(y)$ in the range $-L/2 \leq y \leq L/2$, and got

$$C_{v_x v_x}(y) = 1 - 24 \left(\frac{|y|}{L}\right)^2 + 32 \left(\frac{|y|}{L}\right)^3 \quad (6)$$

and

$$C_{\hat{v}_x \hat{v}_x}(y) = 1 - \left(\frac{4|y|}{L}\right), \quad (7)$$

respectively. Their peak-to-peak distances both reflect the wavelength of the original flow [Fig. 2(e)]. However, because the normalize-first approach changes the velocity profile from a triangular wave v_x to a square wave $\hat{v}_x$, the corresponding correlation functions $C_{v_x v_x}(y)$ [dark blue line in Fig. 2(e)] and $C_{\hat{v}_x \hat{v}_x}(y)$ [magenta line in Fig. 2(e)] differ in shape. We further calculated the azimuthally averaged correlation functions of the 2D flow in Fig. 2(c). Both $C_{\hat{v}\hat{v}}(R)$ and $C_{vv}(R)$ keep their periodicity, but the amplitude of the oscillation decreases as R increases. At the same time, $C_{\hat{v}\hat{v}}(R)$ and $C_{vv}(R)$ remain different as in Fig. 2(e), although the difference is marginal.

In contrast, when the flow becomes non-negative [Figs. 2(g) and 2(h)], we have

$$v_x(y) = 1 - 2 \left| \text{frac}\left(\frac{y}{L}\right) - \frac{1}{2} \right| \quad (8)$$

and

$$\hat{v}_x(y) = \frac{v_x(y)}{|v_x(y)|} = 1. \quad (9)$$

Now $C_{\hat{v}\hat{v}}(R)$ and $C_{vv}(R)$ show significant difference to the point that $C_{\hat{v}\hat{v}}(R)$ does not reflect any features of the flow. The normalize-first approach leads to a uniform flow, which, in turn, results in a flat correlation— $C_{\hat{v}_x \hat{v}_x}(y) = C_{\hat{v}\hat{v}}(R) = 1$ throughout y and R , respectively, as Figs. 2(i) and 2(j) show. Conversely, $C_{v_x v_x}(y)$ and $C_{vv}(R)$ still capture the periodicity of $v_x(y)$, although they no longer oscillate around zero. For example,

within $-L/2 \leq y \leq L/2$,

$$C_{v_x v_x}(y) = 1 - 6 \left(\frac{|y|}{L}\right)^2 + 8 \left(\frac{|y|}{L}\right)^3. \quad (10)$$

This example demonstrates that the average-first approach ($C_{vv}(R)$) is a more reasonable option compared to the normalize-first approach ($C_{\hat{v}\hat{v}}(R)$).

IV. SPATIAL CORRELATION FUNCTIONS OF ISOLATED VORTICES

Vortex is another typical flow pattern in active and living matter, especially in the regime of active turbulence, where rich vortices are created [Fig. 3(a)] [47]. To study how the spatial correlation function $C_{vv}(R)$ reveals features in such flows, we consider an idealized velocity field around an isolated vortex, based on the standard Lamb-Oseen vortex [Fig. 3(b)]. The tangential velocity $v_t(r)$ is

$$v_t(r) = \begin{cases} v_0 \frac{r}{r_0}, & \text{for } 0 \leq r \leq r_0, \\ v_0 e^{-(r-r_0)/r_c}, & \text{for } r > r_0, \end{cases} \quad (11)$$

where r is the distance from the vortex center, r_0 the radius at which the tangential velocity peaks at v_0 , and r_c the characteristic decay length. This velocity field has a rotational rigid body core, where $v_t(r)$ increases linearly with r , followed by an exponentially decaying tail [Figs. 3(c) and 3(f)].

To evaluate how well the features in the correlation function $C_{vv}(R)$ capture those in the flow, we vary r_0 [Fig. 3(c)] and r_c [Fig. 3(f)] independently. Figures 3(d) and 3(g) show the corresponding correlation functions. In general, each curve starts from $C_{vv}(R=0) = 1$, decays as R increases, and then reaches a trough where $C_{vv} < 0$. Beyond the minimum of the trough, the curve increases, gradually levels off, and approaches zero. The width of the decaying section increases with either r_0 or r_c . At the same time, the trough deepens as r_0 increases, but becomes shallower and broader as r_c increases. The same features are reflected in the 2D color maps shown in Figs. 3(e) and 3(h), as r_0 and r_c change, respectively. Furthermore, we use two features in $C_{vv}(R)$ to quantify the characteristic length scales ξ_v of the vortex: the decaying length by fitting the curve with an exponential function

$$C_{vv}(R) = e^{-R/\xi_v^{\text{exp}}} \quad (12)$$

and the location of the minimum in the trough ξ_v^{min} , which are both indicated in Figs. 3(e) and 3(h). ξ_v^{exp} increases linearly with r_0 or r_c . In comparison, ξ_v^{min} shows a good linear trend only when r_0 and r_c are above certain thresholds. Especially, ξ_v^{exp} is approximately independent of r_0 when $r_0 < r_c$. As a result, compared to ξ_v^{min} , ξ_v^{exp} represents the characteristic size of an idealized vortex flow more consistently, especially at small r_0 .

In contrast, the correlation function $C_{\hat{v}\hat{v}}(R)$ with the normalize-first approach exhibits a continuous and monotonically decreasing function throughout the range of R as shown in Fig. 14. After normalization, the isolated vortex loses its characteristic length scales, so the shape of $C_{\hat{v}\hat{v}}(R)$ is only controlled by the system size.

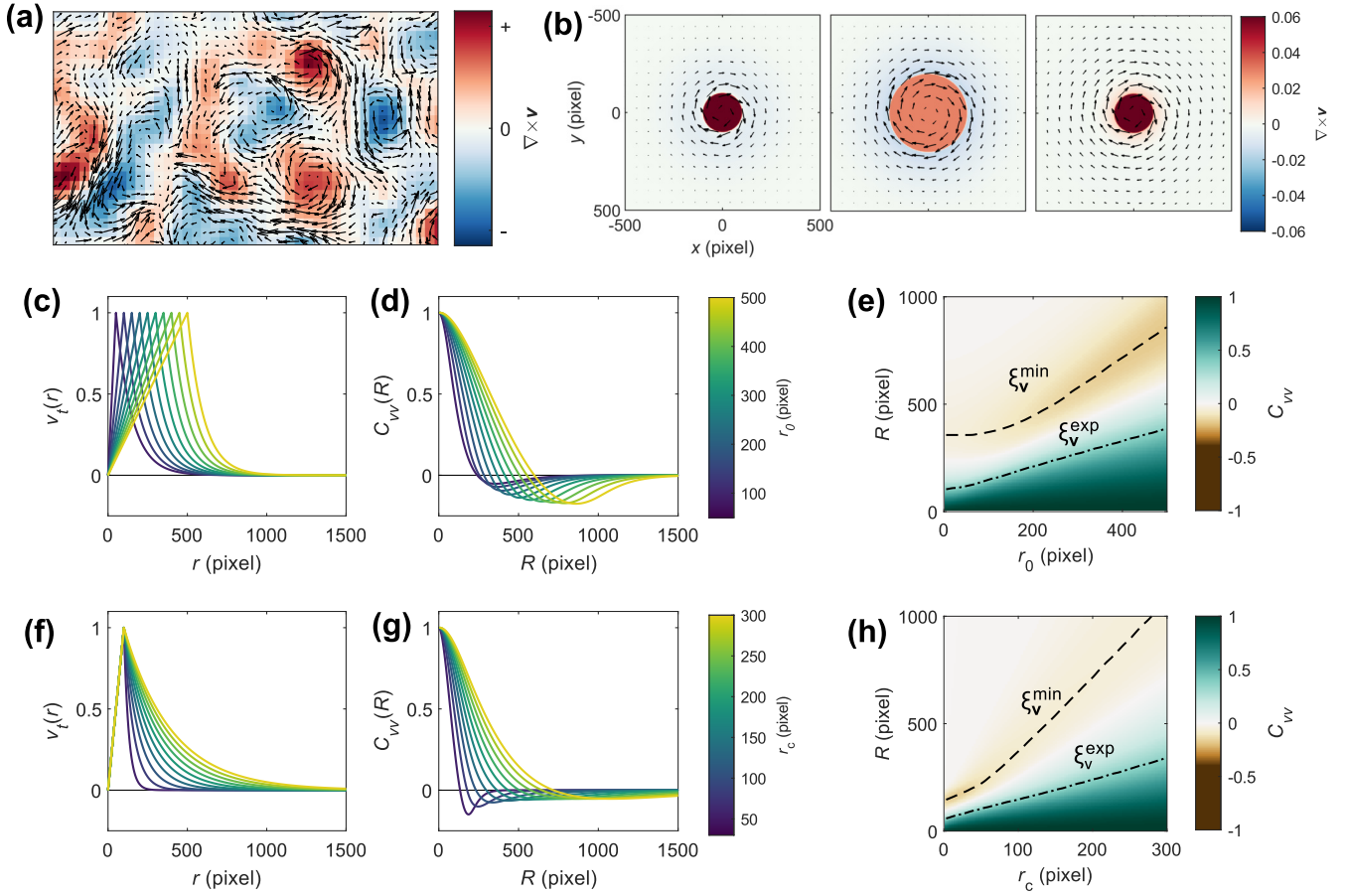


FIG. 3. Correlation functions $C_{vv}(R)$ of idealized isolate vortices. (a) An exemplary velocity field of *B. subtilis* swarms. The black arrows indicate local velocity, and the color map shows the vorticity $\nabla \times \mathbf{v}$. The data are from Ref. [47]. (b) Velocity fields of three idealized isolated vortices. They all follow Eq. (11) with the maximum tangential velocity $v_0 = 1$. From left to right, the length scales of the rigid body core r_0 and the exponential decay r_c are $r_0 = 100, r_c = 100$; $r_0 = 200, r_c = 100$; and $r_0 = 100, r_c = 200$, respectively. (c) Tangential velocity profiles $v_t(r)$ of vortices with $r_c = 100$ and r_0 ranging from 50 to 500, where r represents the radial distance from the center. (d) Correlation functions $C_{vv}(R)$ corresponding to the velocity profiles in panel (c). (e) A 2D color map showing how $C_{vv}(R)$ changes with r_0 . The dashed curve labels the minima in $C_{vv}(R)$, and the dot-dashed curve labels the characteristic length ξ_v^{exp} obtained from exponential fitting [Eq. (12)]. Tangential velocity profiles $v_t(r)$ of vortices with $r_0 = 100$ and r_c ranging from 30 to 300 [panel (f)], the corresponding correlation functions $C_{vv}(R)$ [panel (g)], and 2D color map $C_{vv}(r_c, R)$ [panel (h)].

V. EFFECTS OF NOISE ON CORRELATION

Based on the idealized vortex flow, we move on to study the effects of noise in the velocity fields of cell migration. This noise can be intrinsic or extrinsic. The *intrinsic noise* reflects natural fluctuations in living systems, including those caused by individuality in cell motility and stochasticity in cell-cell and cell-substrate interactions. For example, the active force generated by each cell could vary significantly compared to active nonliving systems such as microtubules, where the activity is controlled globally by the ATP concentration [29,32]. On the other hand, all cell velocity measurements and analytical tools are limited by instrumentation resolution and numerical accuracy, which heavily depend on image quality and the algorithm used for analysis. These *extrinsic noises* do not reflect any intrinsic properties of the cells but are unavoidable.

The intrinsic noise affects how cells move, so we assume that it manipulates the direction angle ψ and speed v of cell

velocity $\mathbf{v} = v(\cos\psi\hat{x} + \sin\psi\hat{y})$ following

$$\begin{aligned}\psi^N &= \psi + \epsilon_\psi, & \epsilon_\psi &\sim \mathcal{N}(0, \sigma_\psi), \\ v^N &= v(1 + \epsilon_m), & \epsilon_m &\sim \mathcal{N}(0, \sigma_m),\end{aligned}\quad (13)$$

where ψ^N and v^N represent the direction and magnitude of the cell velocity upon adding the noise, respectively. The noise terms ϵ_ψ and ϵ_m each follow a Gaussian distribution $\mathcal{N}(\mu, \sigma)$, where μ is the mean and σ is the standard deviation. In comparison, we assume that the extrinsic noise follows

$$\begin{aligned}v_x^N &= v_x + \epsilon_{e,x}, & \epsilon_{e,x} &\sim \mathcal{N}(0, \sigma_e), \\ v_y^N &= v_y + \epsilon_{e,y}, & \epsilon_{e,y} &\sim \mathcal{N}(0, \sigma_e),\end{aligned}\quad (14)$$

where v_x^N and v_y^N represent the horizontal and vertical velocity components upon adding the noise, v_x and v_y are the original velocity components, and $\epsilon_{e,x}$ and $\epsilon_{e,y}$ are drawn from the same Gaussian noise $\mathcal{N}(0, \sigma_e)$.

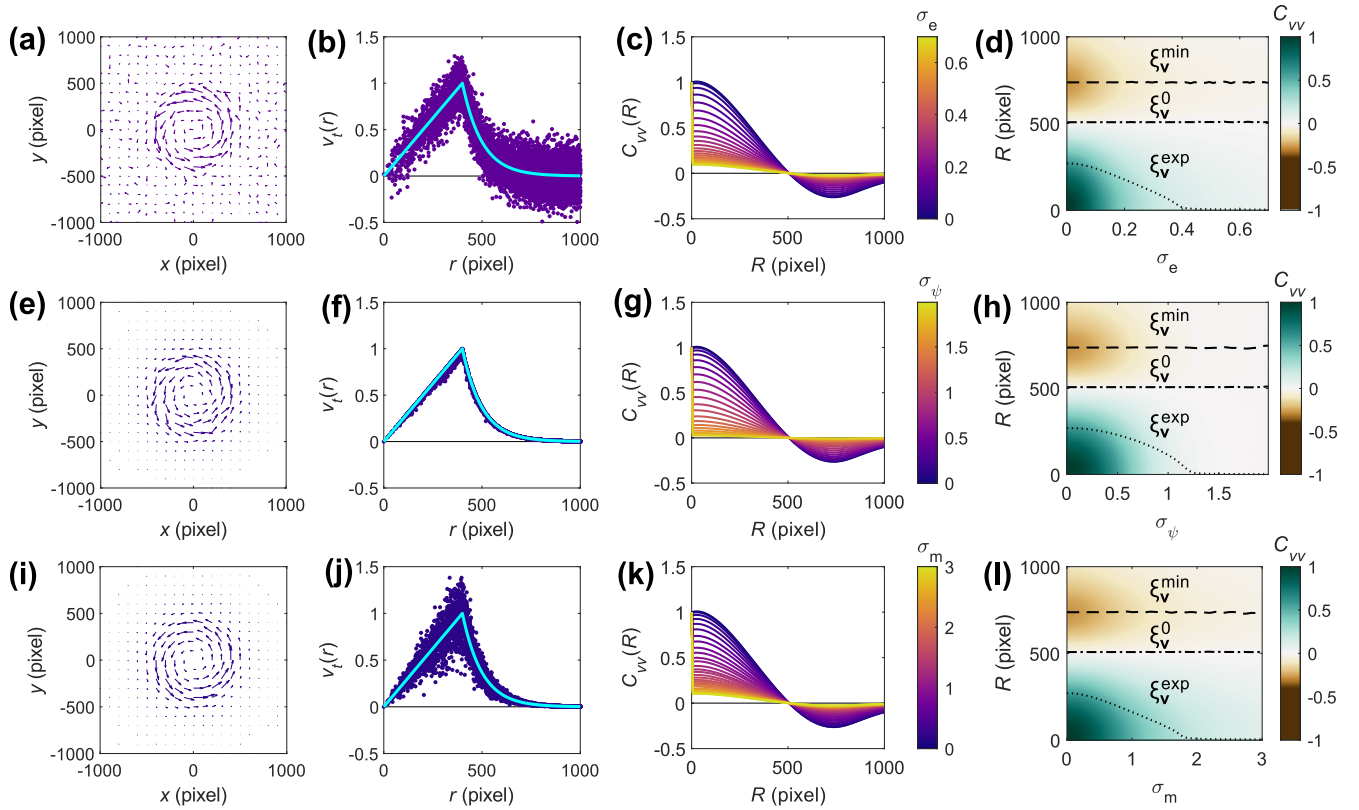


FIG. 4. Impact of noises in velocity on the correlation functions $C_{vv}(R)$. Velocity fields of isolated vortices. The frame size is 2001×2001 pixels, $r_0 = 400$ pixels, and $r_c = 100$ pixels. Three velocity fields with different noises are shown: (a) extrinsic noise ϵ_e ; (e) intrinsic angular noise ϵ_ψ ; and (i) intrinsic magnitude noise ϵ_m . (b), (f), (j) Tangential velocity profiles $v_t(r)$ for each noise condition. The cyan curves show the tangential velocity without noise, and the dots are the tangential velocities of noisy vortices. (c), (g), (k) Corresponding correlation functions $C_{vv}(R)$. The colors represent various noise levels. The colors in the first two columns are chosen from these color bars as well, according to the corresponding noise levels. (d), (h), (l) 2D color maps showing $C_{vv}(R)(\sigma, R)$. The dotted curves show ξ_v^{exp} , the dashed curves show ξ_v^{min} , and the dot-dashed curves show ξ_v^0 .

Now we examine how these three different noises— ϵ_ψ , ϵ_m , and ϵ_e —affect the spatial correlation functions using isolated vortex flows, which have well-defined correlation lengths. As shown in Fig. 4, the extrinsic noise has the same effect everywhere [Figs. 4(a) and 4(b)], while the intrinsic noise is velocity dependent—It increases with the cell speed and vanishes where there is no cell motion [Figs. 4(e), 4(f), 4(i), and 4(j)]. The corresponding correlation functions $C_{vv}(R)$ are shown in Figs. 4(c), 4(g), and 4(k). The curves start at $C_{vv}(R=0) = 1$, and all three types of noise lead to a sudden drop in $C_{vv}(R)$ immediately after $R > 0$, indicating a disturbance of short-range spatial order. At the same time, increasing noise makes the troughs shallower. Consequently, by fitting $C_{vv}(R)$ with an exponential function in Eq. (12), we obtain an increasingly smaller characteristic length ξ_v^{exp} as the noise increases, which is indicated by the dotted curves in Figs. 4(d), 4(h), and 4(l). In contrast, the locations of the minima ξ_v^{min} in the troughs are largely unaffected, so are the places where $C_{vv}(R) = 0$, ξ_v^0 , as indicated by the dashed lines and dot-dashed lines, respectively. This suggests that ξ_v^{min} and ξ_v^0 (if they exist) are more robust measurements when the noise level is high. However, on the other hand, they are not sensitive to changes in fluctuations; therefore, exponential fitting is a better approach when focusing on fluctuations. All three methods work well when the noise level is low.

Note that it is not trivial to directly compare the noise introduced by ϵ_ψ , ϵ_m , and ϵ_e . We calculated the velocity fluctuation $\sigma(\mathbf{v}^N - \mathbf{v}) \equiv \sqrt{\langle (\mathbf{v}^N - \mathbf{v})^2 \rangle}$ in the whole field of view. For the intrinsic noise [Eq. (13)], $\sigma(\mathbf{v}^N - \mathbf{v})$ decreases as the system enlarges, because the noise is proportional to local speed. Furthermore, for the extrinsic noises [Eq. (14)], $C_{vv}(R)$ drops faster from 1 at small R when the system is larger. In Figs. 4(b), 4(f), and 4(j), we set their velocity fluctuations to be identical at the peak, where $r = r_0$, and $\sigma[\mathbf{v}^N(r_0) - \mathbf{v}(r_0)] = 0.2$, as a reference.

VI. EXTRACT CHARACTERISTIC LENGTH SCALES FROM CORRELATION CURVES

In Figs. 3 and 4, we have introduced three ways to extract the characteristic length scale of the vortex flow ξ_v from $C_{vv}(R)$: (1) Exponential fit using Eq. (12) [10,24,38,40,48–50]. (2) Taking the intersection between the correlation curve and a constant, $C_{vv}(\xi_v) = \text{const}$. We selected $C_{vv}(\xi_v^0) = 0$ in Fig. 4, but a more commonly used constant is $1/e$, or $C_{vv}(\xi_v^{\text{cst}}) = 1/e$ [43,44,51–58]. (3) Position of the minimum when there is a trough [45,59–61]. One more method is (4) linear fit [41]: The data in the approximately linear regime between $C_{vv}(R=0) = 1$ and a cutoff value (in this case, $C_{vv} = 0.2$) are fitted to the function $C_{vv}(R) = 1 - R/\xi_v^{\text{lin}}$.

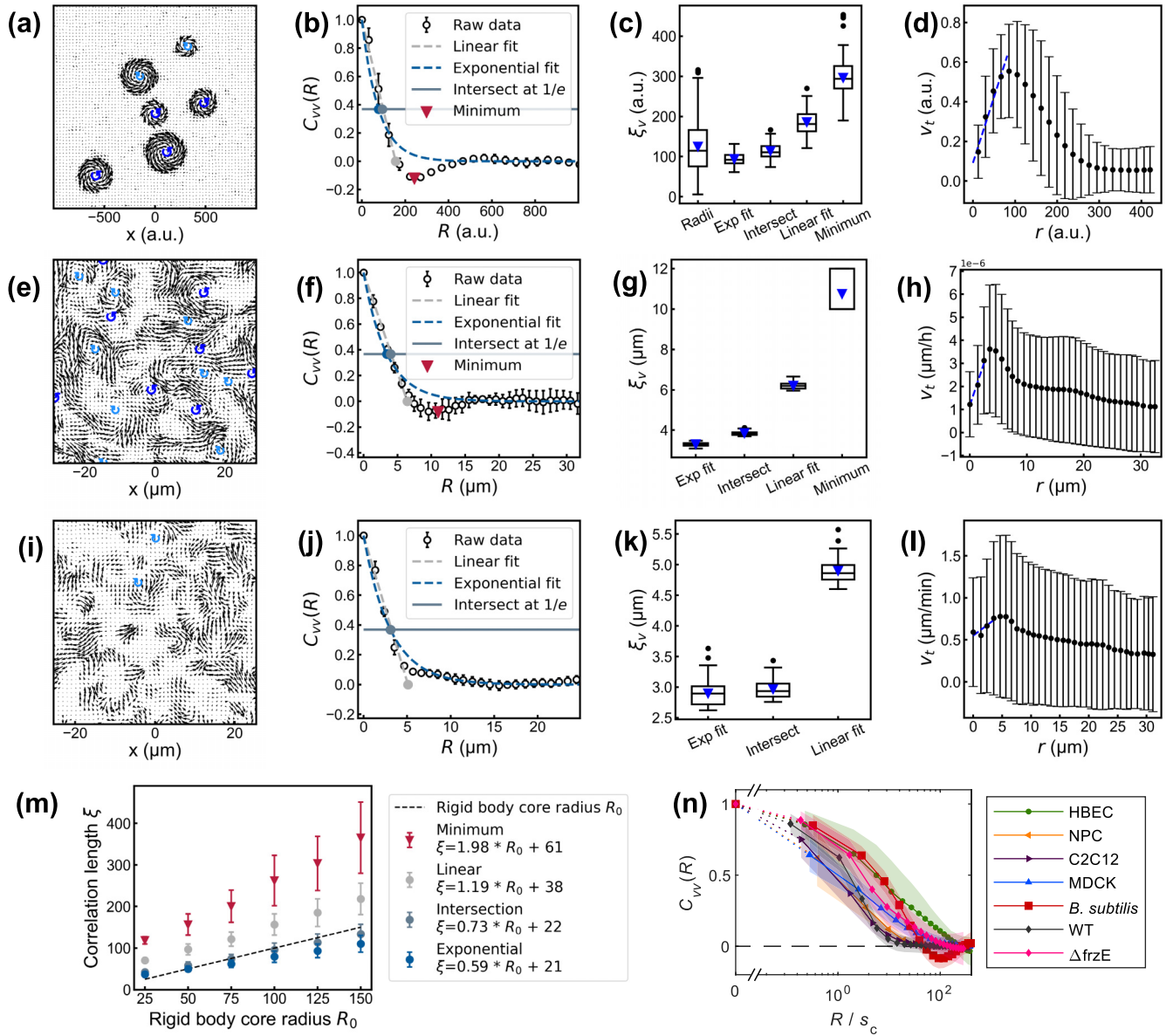


FIG. 5. Correlation lengths of artificial and experimental velocity fields with vortex flows. (a) The velocity field with multiple vortices with random rotational directions. The core radii follow a Rayleigh distribution with the mean of $\langle r_0 \rangle = 125$. The light blue symbols $\odot$ represent the centers of the clockwise vortices, and the blue symbols $\ominus$ represent the centers of counterclockwise vortices. All vortices have the same decay length $r_c = 29.4$. (b) Correlation function $C_{vv}(R)$ corresponding to panel (a). (c) Distribution of various correlation lengths ξ_v extracted from 99 independent vector fields. (d) Averaged tangential velocity $v_t(r)$ of 594 vortices in 99 velocity fields. (e) Velocity field of a monolayer of swarming *B. subtilis* [47], demonstrating turbulent collective swirling motion. (f) Correlation function $C_{vv}(R)$ corresponding to panel (e). (g) Distribution of the correlation lengths ξ_v extracted from 50 frames of velocity fields. (h) Averaged tangential velocity $v_t(r)$ of 889 vortices identified in 50 velocity fields. (i) Velocity field of the ΔfrzE mutant of *M. xanthus*. (j) Correlation function $C_{vv}(R)$ corresponding to panel (i). (k) Distribution of the correlation lengths from 50 consecutive vector fields. In this case, $C_{vv}(R)$ rarely has troughs, so finding the minimum is not an option. (l) Averaged tangential velocity $v_t(r)$ of 82 vortices identified in 50 velocity fields. (m) Summary of the correlation lengths of the artificial multivortices velocity fields with different rigid body core radii R_0 . (n) Summary of the correlation functions from experimental living systems. The detailed correlation lengths and the tangential velocity information are summarized in Table II.

Now we compare these methods using numerically generated [Figs. 5(a)–5(d)] and experimentally measured [Figs. 5(e)–5(l)] velocity fields and test how well they indicate the characteristic length scales of collective cell motion.

Figure 5(a) shows a numerically generated velocity field with randomly distributed vortices. Each vortex has the flow distribution described by Eq. (11), and its direction of rotation

was selected randomly (see Appendix C for more details about how these velocity fields were created). Figure 5(b) shows the resulting correlation function, which is similar to those obtained for an isolated vortex. It decays from 1 as R increases, reaches a minimum with negative correlation, and then gradually approaches 0. We measured the correlation length ξ_v using the four methods described above, as Fig. 5(c) shows.

The exponential fit and the intersection with $1/e$ yield similar results. The linear fit provides a larger ξ_v , but when multiplied by $1 - 1/e$, the rescaled correlation length matches well with ξ_v^{exp} and ξ_v^{est} . The position of the minimum, ξ_v^{min} , is the largest, which is approximately the diameter of the rigid body core $2r_0$, because when two vortices are off by $2r_0$, the peak and negative peak in the tangential velocity overlay, which leads to negative correlations. Finally, we identified the vortices in the velocity fields and calculated their average tangential velocity profiles $\langle v_t(r) \rangle$ [Fig. 5(d)]. The peak in $\langle v_t(r) \rangle$ matches with the previously extracted correlation lengths.

Next, we analyzed the data from several representative experimental systems, including monolayers of swarming bacteria *Bacillus subtilis* [47] [Figs. 5(e)–5(h)], human bronchial epithelial cells (HBECs) [62], and another bacterium *Serratia marcescens* [63] (see Fig. 11 for the HBEC and *S. marcescens* results). *Bacillus subtilis* is a genetically tractable bacterium that exhibits swimming (in bulk liquid) and swarming (as groups on surfaces), both powered by rotating flagella. HBECs are epithelial cells from human airways and are widely used to study respiratory epithelial function and pathology. They migrate by actin-polymerization-driven protrusions coupled to adhesion to the substrate and contractile forces. *Serratia marcescens* is a flagellated pathogen known for strong biofilm formation and frequent multidrug resistance. These systems exhibit velocity fields with rich vortices, and their corresponding correlation curves $C_{vv}(R)$ display distinct troughs with well-defined minima, consistent with our observations in the simplified velocity fields of isolated vortices above.

In contrast, thin *M. xanthus* colonies [Figs. 5(i)–5(l)], murine neural progenitor cells (NPCs) [25], and C2C12 cells (see Fig. 12 for the NPCs and C2C12 results) lack long-lasting vortices. NPCs are developing central nervous system precursor cells, whose migration is driven by cytoskeletal dynamics. Such systems show correlation functions that decay monotonically to zero without a noticeable trough. Furthermore, the peak in the averaged tangential velocity profile of the vortices is less pronounced than in the previous examples.

In summary, the linear fit, intersection with $1/e$, and exponential fit can be utilized to extract ξ_v from all the experimental systems that we have checked. As Fig. 5(n) shows, $C_{vv}(R)$ always provides an apparent characteristic length scale when plotted in log-linear scales. In contrast, the valley position method works well only in vortex-rich systems. As the characteristic length of the vortices in the flow changes, such as their average core size r_0 , the correlation lengths ξ_v obtained with all four methods increase linearly with r_0 [Fig. 5(m)], although their slopes may vary. The minima yield the steepest slope, indicating better sensitivity to varying r_0 . It is also more robust against noise as discussed in Fig. 4, while the other three approaches are more sensitive to the slope change in $C_{vv}(R)$ at short length scales.

VII. RELATIONSHIP BETWEEN DIRECTOR AND VELOCITY CORRELATION LENGTHS

Upon identifying the appropriate methods to calculate spatial correlation functions and extract correlation lengths, we now propose a diagram showing the relationship between

the velocity correlation length ξ_v and director correlation length ξ_d to help understand the features of living systems, as shown in Fig. 6(a). Here, both ξ_v and ξ_d are obtained through exponential fits and normalized by the typical cell size s_c in the corresponding system. This diagram separates systems with different nematic and velocity orders and enables easy tracking of these systems' temporal evolution.

On the left of the diagram are systems with relatively low director order, whereas on the right are those exhibiting long-range nematic alignment. Likewise, a point in the upper region indicates that the cells form large, coherently moving swarms or streams, signifying collective motion over extended distances. Conversely, in the lower region of the diagram, cells move in smaller clusters or more individually. As a result, the four corners of the diagram represent four extreme cases in collective cell motility. In the upper-left region (e.g., point P1), cells form large swarms despite random orientations. In the lower-left region (e.g., point P2), both velocity and orientation correlations are weak, so cells behave more like independent individuals. In the lower-right region (e.g., point P3), cells exhibit long-range nematic alignment but move in opposite directions—whether head to head or side by side. Finally, in the upper-right region (e.g., point P4), cells remain well aligned and move cohesively in the same direction, forming streams. The corresponding experimental examples are demonstrated in Figs. 6(b)–6(e), respectively.

Figure 6(a) shows that cells exhibit distinct collective motility characteristics, displaying a rich variety of behaviors. A rodlike bacterium *B. subtilis*—the only swimming species in our dataset—is in an active turbulence regime with many vortices in the flow (Fig. 3). Its ξ_v and ξ_d are approximately equal and show no significant change over time. The ξ_v – ξ_d relationships for C2C12 cells, NPCs, and WT *M. xanthus* follow similar trends: ξ_v remains smaller than ξ_d , and as time progresses, ξ_d increases significantly while ξ_v changes only mildly or remains unchanged. All three cell types have relatively large aspect ratios. At low density [P2, Fig. 6(c)], C2C12 cells move relatively independently, so both ξ_v and ξ_d are in the range of one to two cell sizes. As the density of the cells increases over time, the orientation of the cells realigns in increasingly larger areas. Similarly, NPCs and WT *M. xanthus* both develop increasingly larger domains with well-aligned cells, as shown in Fig. 6(d) (P3).

The diffusivity of cells also affects the trajectory of the system in the ξ_v – ξ_d diagram. Controlled by an internal clock, WT *M. xanthus* reverses its direction of motion about every 10 min in a nutrient-rich environment. In a $\Delta frzE$ mutant of *M. xanthus*, this clock is disabled, which significantly prolongs their reversal period, and thus the cells move more unidirectionally. This nonreversing strain forms a more turbulentlike flow, with a shorter ξ_d than the WT. Moreover, as time progresses, the cells form collectively moving swarms, and ξ_v increases until it approaches ξ_d .

Finally, we examined the Madin-Darby canine kidney (MDCK) cells, which are an epithelial cell line originally derived from dog kidney tubule tissue and widely used as a model for epithelial polarity, cell-cell adhesion, and collective

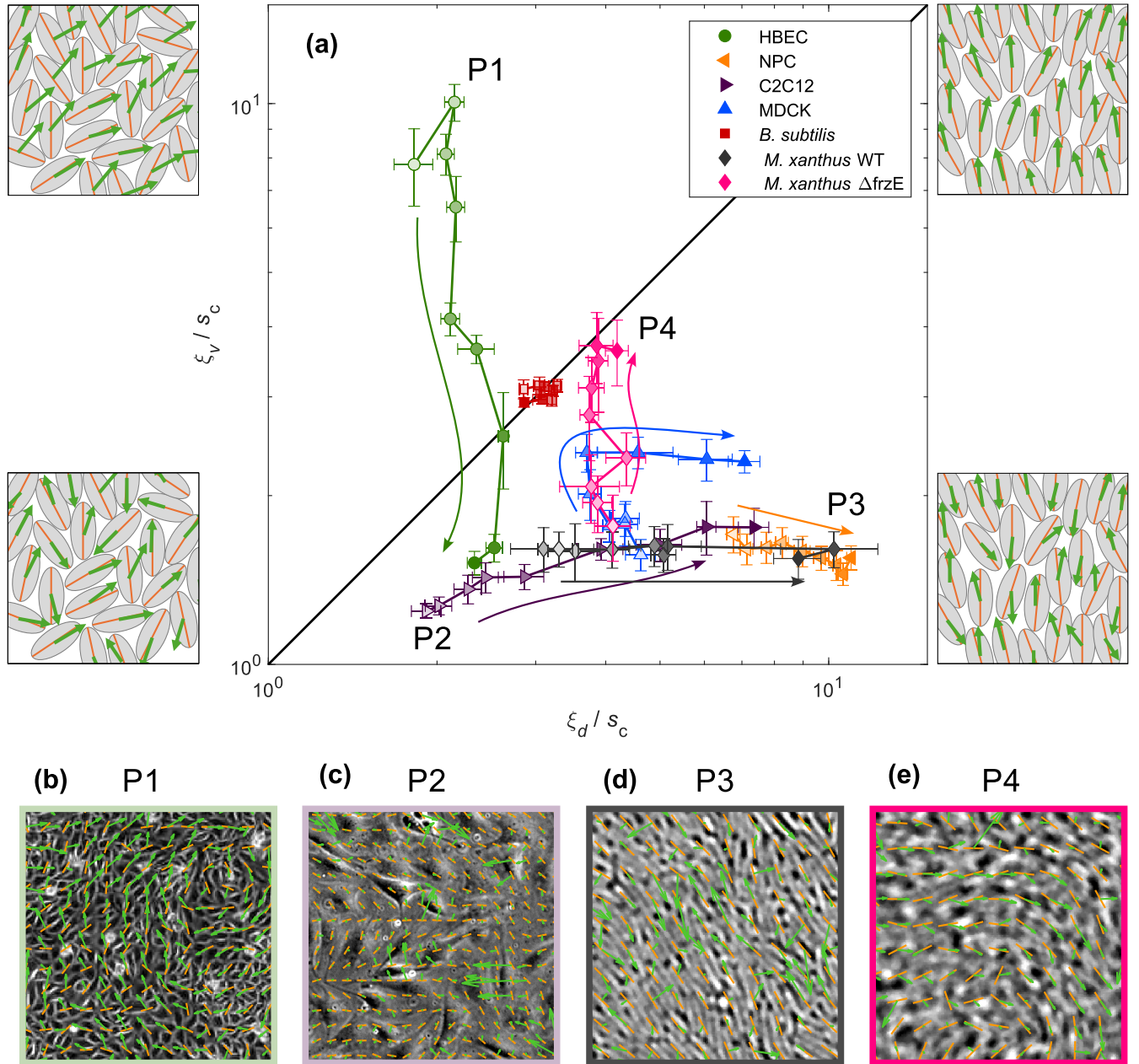


FIG. 6. Categorize various living systems using correlation lengths. (a) ξ_d is the director correlation length, ξ_v is the velocity correlation length, and they are both normalized by the characteristic cell size s_c . The solid black line indicates $\xi_v = \xi_d$. HBEC, NPC, C2C12, and MDCK are mammalian cells, while *B. subtilis* and *M. xanthus* are bacteria. The opacity of the symbols represents temporal evolution—A more opaque color means later in time. The curves with arrowheads point in the direction of time. At the four corners of the diagram, the corresponding “states” of the cells are illustrated schematically. The orange bars represent the orientation of the gray cells, and the green vectors represent their velocities. From the upper-left corner, in the counterclockwise direction, they represent disordered cells with collective flow, disordered cells without collective flow, aligned cells without collective flow, and aligned cells with collective flow, respectively. (b) P1: HBECs. (c) P2: C2C12 cells at low density. (d) P3: wild-type (WT) *M. xanthus* with cell reversal. (e) P4: $\Delta frzE$ *M. xanthus* without cell reversal.

motility [1,28]. In monolayers, MDCK cells migrate primarily via actin-driven lamellipodial crawling. In Fig. 6(a), MDCK cells behave similarly to the $\Delta frzE$ mutant of *M. xanthus* at earlier times— ξ_v increases while ξ_d decreases slightly. However, as the cells become more active over time, ξ_v saturates, and ξ_d begins to increase, reflecting their stronger nematic order.

In the diagram, the solid black line marks an important boundary indicating $\xi_d = \xi_v$. Systems in which cells tend to move along their long axes without forming large swarms are expected to fall below this line, because a director is head-tail symmetric, while a velocity vector is not. Adjacent cells moving in opposite directions can only decrease ξ_v relative to ξ_d . Indeed, of the seven living systems analyzed, most lie

below $\xi_d = \xi_v$, except for HBECs. Initially, HBECs have a ξ_v as long as 10 cell sizes, far exceeding ξ_d , indicating strong collective cell motion despite weak nematic alignment. Over time, as cells pack more densely, large-scale cell flows vanish, ξ_v decreases, and the system transitions downward across $\xi_d = \xi_v$. Figure 6(b) [point P1 in Fig. 6(a)] shows HBECs exhibiting flows spanning ten or more cells when orientation remains disordered.

Figures 6(b)–6(e) imply that the position of the point in the diagram is related to the preferred direction of motion of the cells. At point P2 [Fig. 6(c)], the low-density C2C12 cells tend to move perpendicular to their long axes. In contrast, at P3 and P4 [Figs. 6(d) and 6(e)], *M. xanthus*' velocity aligns well with the local directors. At point P1 [Fig. 6(b)], the direction of cell motility is not strongly affected by the cell orientation—In some regions they are aligned, while in some other regions they are perpendicular. So next, we will investigate the relationship between the direction of cell motility and the correlation length scales.

VIII. CELL MOTILITY ANISOTROPY VERSUS CORRELATION-LENGTH RATIO

The direction of cell motility is indicated by the angle ϕ between its velocity and director. Because of the symmetry, ϕ ranges from $-\pi/2$ to $\pi/2$. In each frame, we can calculate the distribution of ϕ , as shown in Fig. 7. MDCK cells and *B. subtilis* showed little anisotropy, while NPC and *M. xanthus* preferred to move along their long axis ($\phi = 0^\circ$). Furthermore, HBECs and C2C12 cells changed how they move over time—HBECs transitioned from isotropic to anisotropic motion and then back to isotropic; C2C12 cells moved perpendicular to their long axes at low densities and shifted toward moving along their long axes at high densities. We use the ratio between the probability densities at $\phi = 0$ rad and $\phi = \pi/2$ rad, $P(\phi = 0)/P(\phi = \pi/2)$, to represent the anisotropy in cell motion. This enables us to map out the relationship between how cells move and their large-scale collective behaviors.

In Fig. 7(a), we compare the ratio between the velocity and director correlation lengths, ξ_v/ξ_d , and the directional anisotropy $P(\phi = 0)/P(\phi = \pi/2)$. For rod-shaped cells moving on a solid surface, such as NPC and *M. xanthus*, the cell-substrate friction is anisotropic—It is easier for the cell to move along its long axis than its short axis, meaning the friction coefficient is smaller in the longitudinal direction than in the transverse direction [10,25]. The data corresponding to these cells fall in the lower-right quadrant of the diagram. In this regime, cells tend to move along their directors. When the velocity direction is random (e.g., NPC and *M. xanthus* WT), $\xi_v/\xi_d < 1$, while as cells form swarms and align their motion (*M. xanthus* $\Delta frzE$), ξ_v/ξ_d approaches 1. In contrast, for cells with weak motility anisotropy, such as HBEC, MDCK, and *B. subtilis*, ξ_v/ξ_d can be greater than, less than, or equal to 1. HBECs undergo a nonmonotonic transition in motility anisotropy. At early times, cell motility is isotropic, which enables the formation of large-scale collective flows despite low directional order. As time progresses, motility becomes increasingly anisotropic (cells tend to move along their long axis), and ξ_v decreases monotonically. At

later times, the motility becomes isotropic again, while ξ_v continues to decrease. In comparison, MDCK exhibits no preferred direction of motion throughout, but ξ_v/ξ_d varies nonmonotonically—first increasing toward 1, then decreasing again. *Bacillus subtilis* is the only swimming cell shown in the figure and remains in the active turbulence regime over the entire observation period. Its motility stays isotropic, and ξ_v/ξ_d remains approximately 1 throughout. Finally, as discussed above, C2C12 cells transition from moving perpendicular to their long axis at low density to moving along their long axis at high density. Over this process, ξ_v/ξ_d decreases monotonically. When the anisotropy is high, their behavior resembles that of NPC and *M. xanthus*.

IX. DISCUSSION

In summary, we have examined commonly used definitions of correlation functions in the literature, focusing on two primary types of the velocity correlation: the average-first approach $C_{vv}(R)$ and the normalize-first approach $C_{\hat{v}\hat{v}}(R)$. Significant discrepancies arise between $C_{vv}(R)$ and $C_{\hat{v}\hat{v}}(R)$ while processing numerically generated flow fields, such as shear flows, isolated vortices, and multiple vortices. We demonstrate that $C_{vv}(R)$ effectively captures features in the flow field, but $C_{\hat{v}\hat{v}}(R)$ fails to do so. Moreover, $C_{vv}(R)$ exhibits features robust to intrinsic and extrinsic noise, although its decay slope at small R reflects the noise level. In vortex-deficient flow fields, $C_{vv}(R)$ is monotonic. In vortex-rich flow fields, $C_{vv}(R)$ exhibits a trough with negative correlation, and the location of this local minimum is controlled by the vortex diameter. We compared several methods for extracting correlation lengths from $C_{vv}(R)$ —including exponential fits, linear fits, intersections with certain constants, and trough positions—and found that they yield consistent results unless the noise is excessively high.

Next, we used velocity and director correlation lengths, ξ_v and ξ_d , to categorize various systems and reflect their collective behaviors. In the ξ_d – ξ_v plane, *B. subtilis* remains on the diagonal $\xi_d = \xi_v$ in the active-turbulence regime, signaling a single dominant length scale. This agrees with the theoretical prediction that in active turbulence, there is one characteristic length scale determined by the balance of active and elastic stresses [55]. In contrast, the other systems are in various regimes, and the relationship between ξ_v and ξ_d becomes more complex. These transitions from one regime to another cause trajectories to diverge from the simple proportionality $\xi_d \propto \xi_v$. For example, HBECs begin in the upper-left region—with high motility and large ξ_v despite small ξ_d —before transitioning into a jammed state where orientation correlations grow and velocity correlations diminish. Similarly, the two strains of *M. xanthus* (WT and $\Delta frzE$) both occupy the lower-right quadrant but follow distinct paths: The WT remains well below the $\xi_d = \xi_v$ line, while the $\Delta frzE$ mutant evolves upward toward it as the system transitions into a more turbulent regime with swarms.

Finally, by mapping each system's velocity-to-director correlation length ratio (ξ_v/ξ_d) against its motility anisotropy $P(\phi = 0)/P(\phi = \pi/2)$, we can immediately see how alignment bias controls collective flows. Similar to Fig. 6, in Fig. 7, in general, cells with different properties occupy different

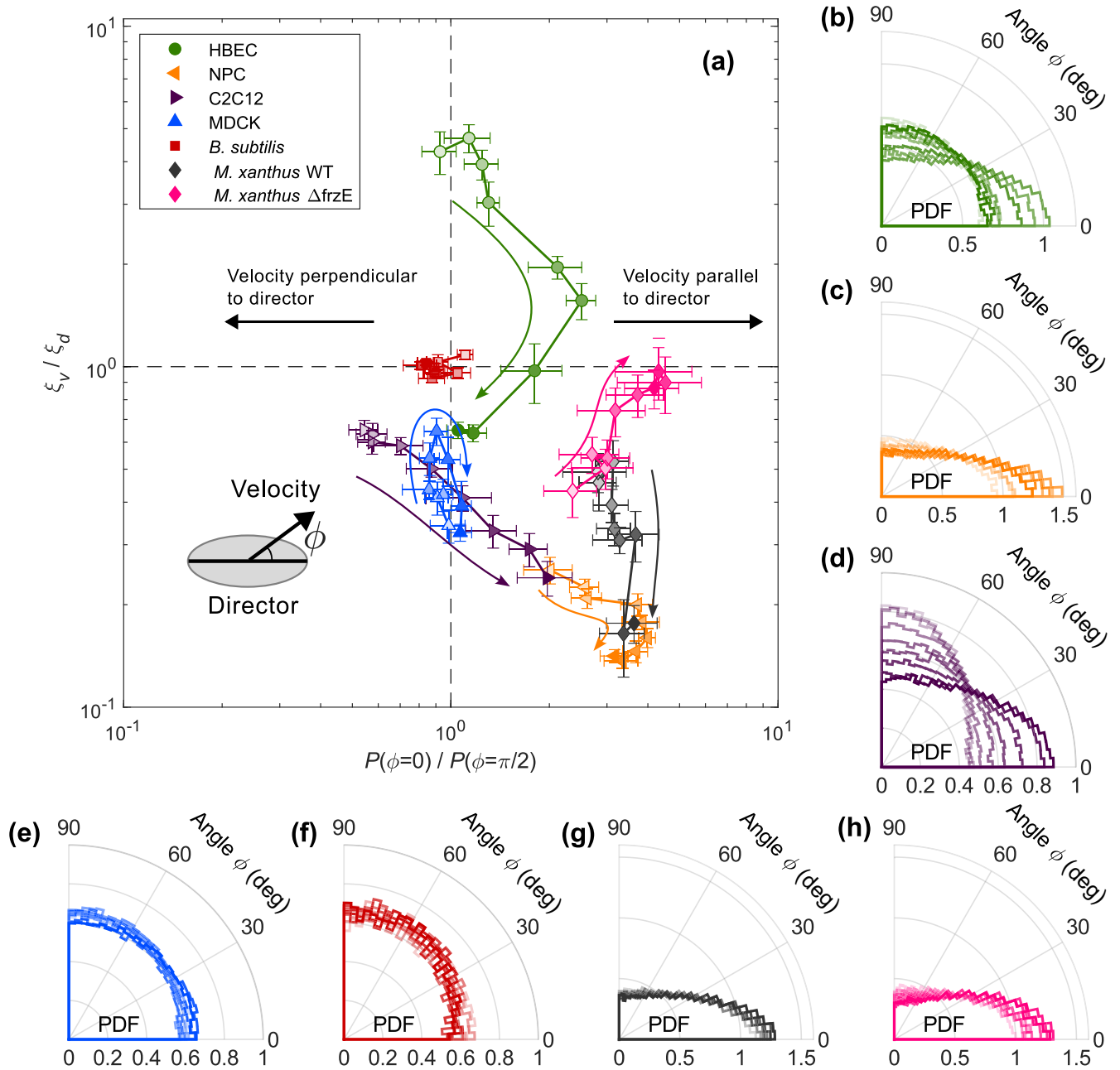


FIG. 7. Correlation lengths are controlled by anisotropy in cell movement. (a) Relationship between the ratio of velocity to director correlation lengths, ξ_v/ξ_d , and anisotropy in cell movement. The motility anisotropy is characterized by the angular distribution ϕ between the local velocity and the director, as illustrated in the diagram. The symbols are identical to those in Fig. 6. (b)–(h) Probability distribution functions of ϕ for different living systems: (b) HBED; (c) NPC; (d) C2C12; (e) MDCK; (f) *B. subtilis*; (g) *M. xanthus* WT; and (h) *M. xanthus* $\Delta frzE$. In each panel, the lighter color represents an earlier time in the video.

areas. As the cells transition from one state to another, such as the HBECs and C2C12 cells, they each leave a unique trajectory in the diagram, showing how the global properties (ξ_v/ξ_d) and local properties [$P(\phi=0)/P(\phi=\pi/2)$] evolve together. The diagram also links seemingly different systems when they enter the same regime. For example, active turbulence, where $\xi_v = \xi_d$, occurs both in cells with strong alignment bias (*M. xanthus* WT) and in those with no apparent bias (*B. subtilis*). This two-dimensional diagram allows for direct comparison between cell types, pinpoints how genetic or mechanical changes shift the flow regime of cells, and

potentially guides the design of perturbations to achieve desired collective behaviors.

Overall, this work provides guidelines for consistently analyzing correlation functions and extracting characteristic lengths in living systems, offering tools and insights to quantify and understand their collective behaviors. By comparing the positions and temporal evolution of these systems on the two proposed diagrams, we can infer the properties of new biological systems and link them to known ones. However, the relationship between cell alignment and direction of motion is governed by multiple factors—including active force

generation, anisotropic cell-substrate friction, cell-cell adhesion, and cell density—and the parameter or combination that controls ξ_v/ξ_d remains unclear. Addressing this will require more comprehensive data on intrinsic polarity, local cell-substrate interactions, and cell-cell mechanical stresses. At the same time, a rigorous data-driven approach is needed to derive robust laws linking global properties such as correlation lengths to cellular motility and mechanics.

ACKNOWLEDGMENTS

We thank Changjin Huang at Nanyang Technological University (NTU) for insightful discussions on extrinsic and intrinsic noise in living systems. The authors thank their labmates for their invaluable suggestions and inspiring discussions. K.J.H., H.X., and E.H. acknowledge financial support from the Ministry of Education (MOE), Singapore, via its Tier 1 Award (Grant No. RG154/23). Y.H. acknowledges the Research Scholarship awarded by NTU. E.H. acknowledges financial support from NTU (Start-Up Grant) and MOE Tier 1 Award (Grant No. RS04/23).

DATA AVAILABILITY

The data that support the findings of this article are openly available [64].

APPENDIX A: EXPERIMENTS WITH C2C12 CELLS

Culture and imaging of C2C12 cells. C2C12 cells were grown on a polydimethylsiloxane (PDMS) substrate in a humidified incubator at 37 °C and 5% CO₂. The phase contrast images were recorded using the Nikon Ti2 microscope with 10×/0.30 objective lens.

Image analysis. The open-source MATLAB toolbox PIVlab [65] was utilized to perform particle image velocimetry (PIV) analysis on phase-contrast image sequences obtained from both literature sources and experimental recordings. The PIV analysis employed the FFT window deformation algorithm. In the first pass, the interrogation area was set to be slightly larger than the size of the cell diameters with a 50% overlap, and the interrogation area for the second pass was reduced to half the first pass. Subpixel displacement estimation was performed using the Gaussian 2 × 3-point estimator, and the correlation robustness was set to the standard level. The vector data were subsequently extracted for further analysis and visualization using Python.

APPENDIX B: EXPERIMENTS WITH *M. XANTHUS*

The experimental and data processing methods to obtain the results with *M. xanthus* are described in detail in the main text and Supplementary Information of Ref. [11]. Here, we just provide a brief description.

Bacterial strains and cultures. We used three *M. xanthus* strains—WT (DK1622) [10], $\Delta pilA$ [66], and $\Delta frzE$ [67]. To grow the bacteria, we took samples from frozen stocks and plated them on 1.5% agar pads in CTTYE (1% casitone, 10 mM Tris-HCl, 1 mM KH₂PO₄, 8 mM MgSO₄, and adjusted its pH to 7.6). After growing on these plates at 32 °C temperature for at least three days, the bacteria were transferred into a

CTTYE solution, incubated at 32 °C overnight with continuous shaking, and harvested for sample preparation when they were in the exponential phase.

Sample preparation. The bacteria were placed on a soft hydrogel substrate made of polyacrylamide (PAA) with a shear modulus of 430 Pa. The samples were prepared in 35 mm diameter Petri dishes with a 0.17-mm-thick glass bottom (Thermo Fisher). The related PAA surfaces were coated with chitosan, washed with DI water, and soaked in sufficient CTTYE twice for about 20 min in total. After soaking, we removed CTTYE and dried the excess liquid. To make the cell monolayer, we took the previously prepared liquid culture, separated the cells by centrifuging, and resuspended the bacteria in DI water so that the cell density was 8× the concentration of OD = 1 (550 nm wavelength). Then, we put a 1.5 μ l droplet of the concentrated cell suspension onto the PAA surface and let it settle for several minutes. To prevent gel dehydration, we created a sealed chamber using a 2-mm-thick acrylic spacer and a 22 × 22 mm cover glass, ensuring the glass did not touch the gel. The chamber was sealed with Valap (1/3 vaseline, 1/3 lanolin, and 1/3 paraffin by mass) and incubated at 32 °C for 2 h before imaging at room temperature.

Microscopic imaging. The images were captured using a Nikon Ti-E microscope with Yokogawa CSU-21 spinning disk confocal module. We used a 60× Plan Apochromat Water Immersion objective (NA 1.20, WD 0.27 mm) and an Andor Zyla camera. Samples were placed in a humidified chamber at 25 °C. While acquiring the time series of bright-field images, the imaging plane was set at the substrate surface, where the bacteria were, and the time between adjacent frames was 15 s.

Data analysis. We used bright-field images to extract the director field and cell flow velocity, applying basic pre-processing steps such as background removal and contrast adjustment. To characterize the cell orientation angle θ_{ij} , we calculated the Hessian matrix

$$H_{ij} = \begin{pmatrix} \left(\frac{\partial I_{ij}}{\partial x}\right)^2 & \frac{\partial I_{ij}}{\partial x} \frac{\partial I_{ij}}{\partial y} \\ \frac{\partial I_{ij}}{\partial x} \frac{\partial I_{ij}}{\partial y} & \left(\frac{\partial I_{ij}}{\partial y}\right)^2 \end{pmatrix} \quad (B1)$$

using the preprocessed bright-field image, where the brightness of a pixel in row i and column j is I_{ij} . Then, the matrix was smoothed by a Gaussian filter with standard deviation $\sigma = 10$ pixels. The eigenvector associated with the smallest eigenvalue of H_{ij} gives the local direction of the smallest brightness gradient, which was the cell orientation. We used the optical flow function *opticalFlowFarneback* in MATLAB (R2019b) to measure cell velocity.

APPENDIX C: NUMERICALLY GENERATED VELOCITY FIELDS WITH VORTICES

The numerically generated velocity field with multiple vortices shown in Fig. 5(a) is created according to the following specifications. The frame size is set to 2000 × 2000 pixels. Each frame contains six vortices, with their rotational directions randomly assigned as either +1 or −1. The radii of the rigid body core follow a Rayleigh distribution, with the scale parameter chosen to achieve an average radius of 125 pixels.

The velocity field is defined such that the maximum tangential velocity $v_t(r_0)$ is 1. Each vortex has a characteristic decay length of $r_c = 29.4$ pixels. Vortex centers were randomly assigned within the simulation frame, subject to boundary and separation constraints. Specifically, candidate centers were sampled from a uniform distribution over an interior region that excluded a 150-pixel margin from each frame boundary. A minimum center-to-center separation, determined by the corresponding vortex radii plus an additional 150 pixels, was then imposed to avoid excessive overlap and clustering. Candidate positions were generated iteratively until all vortices satisfied these constraints, yielding spatially distributed vortices that remained fully inside the computational domain. Additionally, the extrinsic noise is applied to individual vortices with a noise level of $\epsilon_e = 0.02$.

APPENDIX D: DETECTION OF VORTICES IN VELOCITY FIELDS

The centers of vortices are identified by locating regions with low order parameter, defined as $s = \frac{|\langle \vec{v}_i \rangle|}{\langle |\vec{v}_i| \rangle}$, in the velocity field. The extracted regions with s smaller than a certain threshold are subsequently screened again based on their shapes. In this step, we filtered out those with an anisotropy above a certain threshold. The remaining regions with a small s and a relatively round shape are recognized as vortex cores. Finally, we located their centers where the order parameter s reaches the minimum, and we identified the direction of circulation surrounding each vortex.

APPENDIX E: CORRELATION OF DIRECTORS, POLARITIES, AND VELOCITIES

Figure 8 illustrates how the directors, polarities, and velocities are obtained with a pair of cells positioned at $\mathbf{r}_j$ and

$\mathbf{r}_j + \mathbf{R}$, respectively, where $\mathbf{R}$ indicates their separation. In a nematic system, each elongated cell has two identical ends, so its orientation is represented by a director $\hat{\mathbf{n}} = (\cos \theta, \sin \theta)$, where θ is the orientation angle [Fig. 8(a)]. Because each director has a head-tail symmetry, $\theta(\mathbf{r}_j)$ and $\theta(\mathbf{r}_j + \mathbf{R})$ both range from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$, and the relative angle between them ranges from 0 to $\frac{\pi}{2}$. In contrast, when the two ends of a cell are different, its direction is described by a vector with head-tail asymmetry, which we call cell polarity $\mathbf{p}$ [Fig. 8(b)]. For polarity vectors, the orientation angles $\psi(\mathbf{r}_j)$ and $\psi(\mathbf{r}_j + \mathbf{R})$ range from $-\pi$ to π , and the relative angle $|\psi(\mathbf{r}_j) - \psi(\mathbf{r}_j + \mathbf{R})|$ spans from 0 to π .

Nematic order can arise in systems composed either of apolar particles or of polar particles whose head-tail orientations are randomly distributed. In contrast, polar order requires polar entities that, on average, align in the same direction. In experiments, directly measuring the polarity of living matter is often challenging. However, the velocity field can be obtained relatively easily and used to approximate the polarity field. Like polar vectors, velocity vectors break head-tail symmetry, but they also vary in magnitude, providing additional information about the dynamics.

A schematic of the correlations between pairs of directors or vectors is shown in Fig. 8(d). For directors, a correlation value of +1 occurs when they are parallel, while a value of -1 corresponds to perpendicular orientations. For vectors, +1 correlation occurs only when the two are parallel and point in the same direction, whereas parallel but oppositely oriented vectors yield -1 . When two vectors are perpendicular, the correlation is zero.

APPENDIX F: CALCULATING CORRELATION FUNCTIONS NUMERICALLY

For director fields, the spatial correlation function is defined consistently across the literature. In contrast, velocity correlation functions are implemented in different ways, mainly through either a normalize-first approach, in which velocity vectors are converted to unit vectors before averaging, or an average-first approach, in which the original vector magnitudes are retained during averaging. In the average-first approach, the final normalization can be performed in several ways, including the global variance, local variance, geometric mean of local variances, or arithmetic mean of local variances. The equations for calculating director and velocity correlation functions are summarized in Table I.

We use the velocity field $\mathbf{v}$ of *M. xanthus*, shown in Fig. 2, as an example to demonstrate the step-by-step calculation process of the correlation function $C_{vv}(\mathbf{R})$. An illustrative flow chart is presented in Fig. 9. First, the global drift is removed, giving $\delta \mathbf{v} = \mathbf{v} - \langle \mathbf{v} \rangle$. Second, we compute a matrix $N(\mathbf{R}) = \sum_j \mathbf{I}(\mathbf{r}_j) \cdot \mathbf{I}(\mathbf{r}_j + \mathbf{R})$, which represents the number of vector pairs at a given separation $\mathbf{R}$ [Fig. 9(a)]. Here, $\mathbf{I}$ is a matrix with the same dimensions as the velocity field, with all elements equal to 1. For $C_{vv}(\mathbf{R})$, we first normalize all velocities to unit vectors $\delta \mathbf{v} / |\delta \mathbf{v}|$ [Fig. 9(b)], calculate the autocorrelation of the unit vector field, divide it by $N(\mathbf{R})$, and finally perform an azimuthal average to obtain the 1D

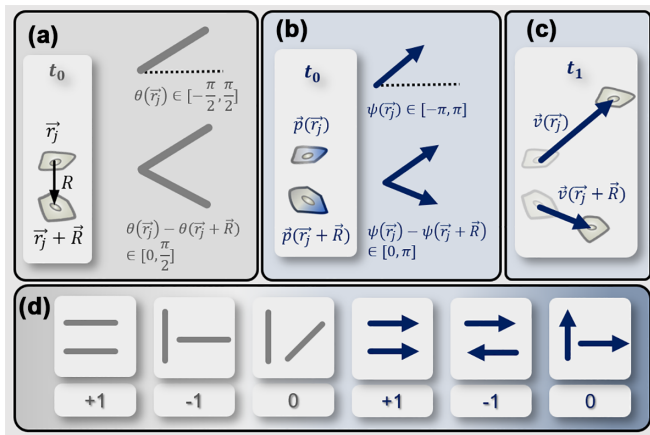


FIG. 8. Schematics of the correlation of a pair of directors or vectors. (a) The directors of two cells at time t_0 and positions $\mathbf{r}_j$ and $\mathbf{r}_j + \mathbf{R}$. The angle $\theta(\mathbf{r}_j) \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ is the local director angle. The relative angle between any two directors ranges from 0 to $\frac{\pi}{2}$. (b) The polarities of two cells. $\psi(\mathbf{r}_j) \in [-\pi, \pi]$ is the angle of $\mathbf{p}(\mathbf{r}_j)$. For polarities, the relative angle $|\psi(\mathbf{r}_j) - \psi(\mathbf{r}_j + \mathbf{R})|$ ranges from 0 to π . (c) The velocity vectors of two cells, $\mathbf{v}(\mathbf{r}_j)$ and $\mathbf{v}(\mathbf{r}_j + \mathbf{R})$, from time t_0 to t_1 . (d) Correlations of two directors and two vectors.

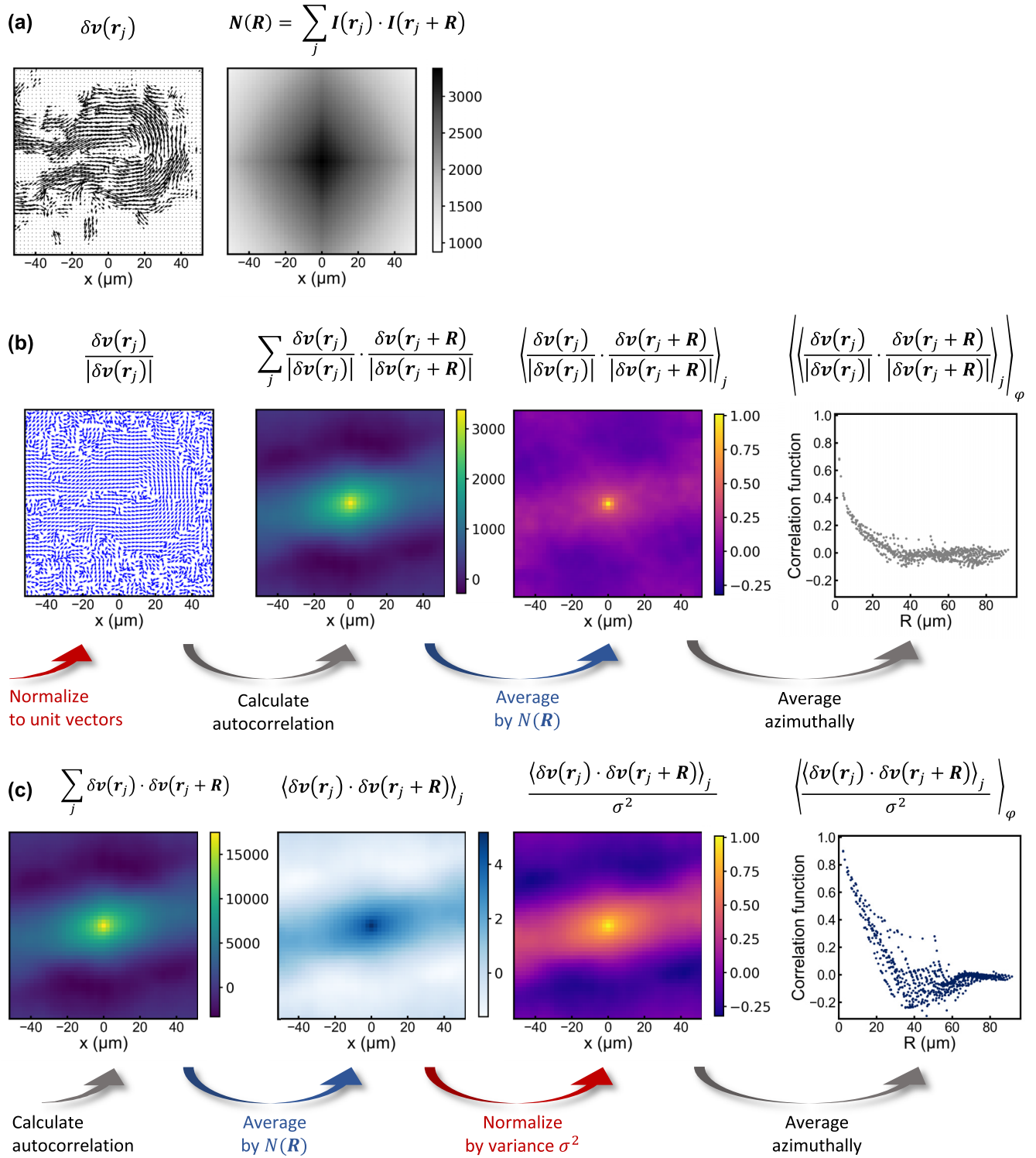


FIG. 9. Calculation process of the correlation functions $C_{\hat{v}\hat{v}}(R)$ and $C_{vv}(R)$, using the velocity field from Fig. 2(a) as an example. (a) Velocity field of the rod-shaped bacterium *M. xanthus* after removing the background drift $\delta \mathbf{v}(\mathbf{r}_j)$. The autocorrelation of two all-ones matrices with the same dimensions as the vector field yields a matrix $N(\mathbf{R})$, representing the number of elements for later averaging. (b) Step-by-step calculation of $C_{\hat{v}\hat{v}}(R)$ according to Eq. (2). The velocity vectors are first normalized to unit vectors. The autocorrelation of the unit vectors is then calculated and averaged using $N(\mathbf{R})$, yielding $C_{\hat{v}\hat{v}}(R)$. Averaging over all azimuthal directions produces the correlation curve $C_{\hat{v}\hat{v}}(R)$. (c) Step-by-step calculation of $C_{vv}(R)$ according to Eq. (G1). The autocorrelation of the velocity deviations $\delta \mathbf{v}(\mathbf{r}_j)$ is calculated and averaged using $N(\mathbf{R})$. The result is then scaled by the variance to normalize the autocorrelation to the range $[-1, 1]$, yielding $C_{vv}(R)$. Finally, averaging over all azimuthal directions gives the correlation function curve $C_{vv}(R)$.

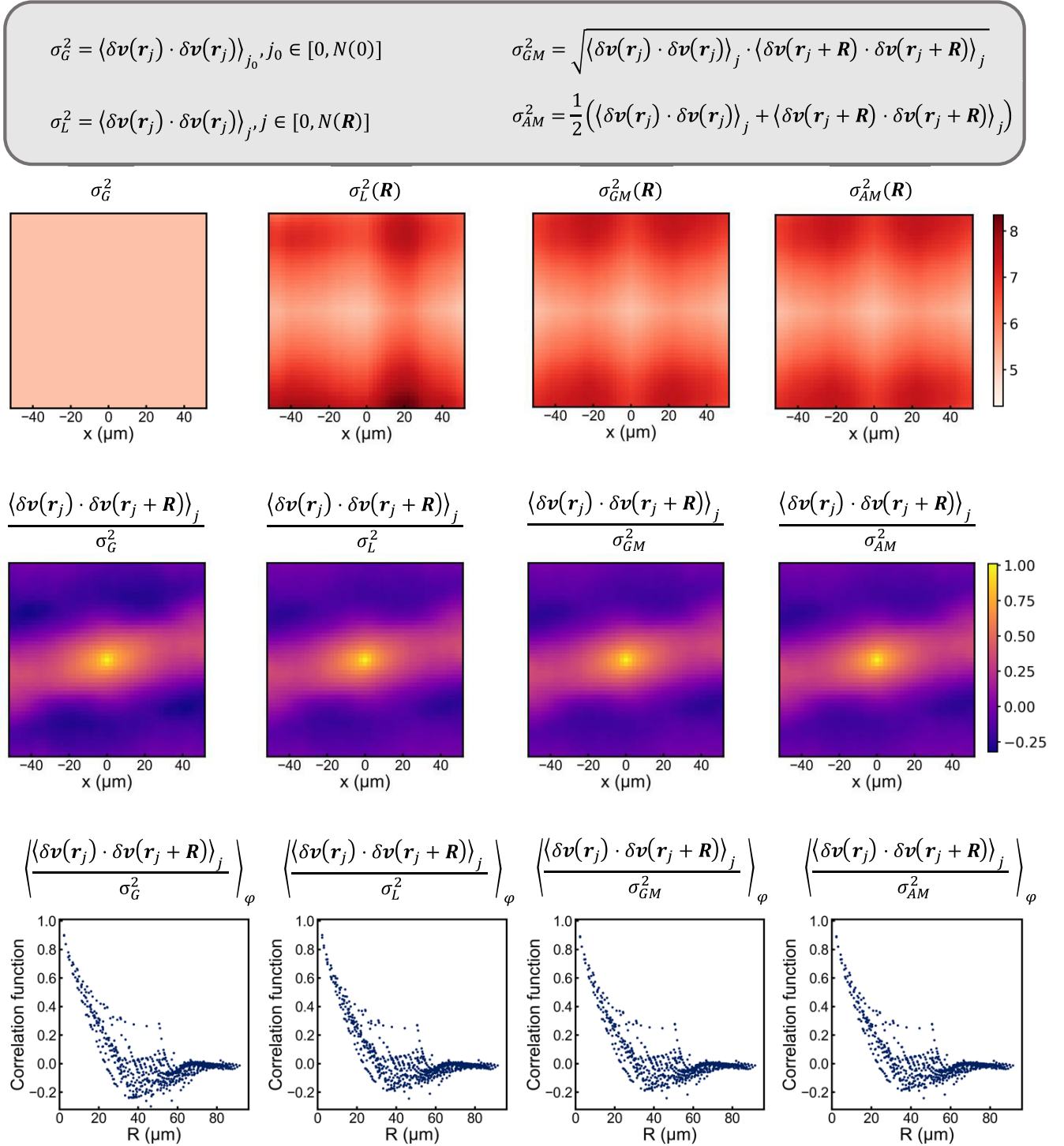


FIG. 10. Comparison of the correlation functions $C_v(R)$ normalized using different variances σ^2 . The global variance σ_G^2 is a constant calculated over the entire frame and depends only on the vector field itself, irrespective of the location $\mathbf{R}$. In contrast, the local variance σ_L^2 considers the variance of the velocity deviations $\delta \mathbf{v}(\mathbf{r}_j)$ associated with vector pairs separated by a distance $\mathbf{R}$. By flipping the local variance both horizontally and vertically, the variance of their pairs, $\langle \delta \mathbf{v}(\mathbf{r}_j + \mathbf{R}) \cdot \delta \mathbf{v}(\mathbf{r}_j + \mathbf{R}) \rangle_j$, can be computed. To minimize the effects of a nonisotropic field, the geometric mean and arithmetic mean of the local variance and its pairwise counterpart are used. The differences among the four normalization methods are subtle. For relatively uniform vector fields, normalizing with the global variance is computationally efficient while maintaining accurate results. However, for systems with strong anisotropy, e.g., when large-magnitude vectors are concentrated in one region while others are extremely small, using the geometric mean or arithmetic mean of the local variances helps reduce these effects while keeping the correlation function within the range $[-1, 1]$.

TABLE I. A selection of spatial correlation functions in the literature.

Field	Calculation	Correlation function formula
Director field	Average only	$C_{\hat{n}\hat{n}}(R) = \langle 2(\cos^2(\theta(\mathbf{r}_j) - \theta(\mathbf{r}_j + \mathbf{R})))_j - 1 \rangle_\varphi$ (F1)
	Normalize-first	$C_{\hat{v}\hat{v}}(R) = \left\langle \left\langle \frac{\delta \mathbf{v}(\mathbf{r}_j)}{ \delta \mathbf{v}(\mathbf{r}_j) } \cdot \frac{\delta \mathbf{v}(\mathbf{r}_j + \mathbf{R})}{ \delta \mathbf{v}(\mathbf{r}_j + \mathbf{R}) } \right\rangle_j \right\rangle_\varphi$ (F2)
	Average-first Normalize by global variance	$C_{vv}^G(R) = \left\langle \frac{\langle \delta \mathbf{v}(\mathbf{r}_j) \cdot \delta \mathbf{v}(\mathbf{r}_j + \mathbf{R}) \rangle_j}{\langle \delta \mathbf{v}(\mathbf{r}_j) \cdot \delta \mathbf{v}(\mathbf{r}_j) \rangle_{j_0}} \right\rangle_\varphi$ (F3)
Velocity field	Average-first Normalize by local variance	$C_{vv}^L(R) = \left\langle \frac{\langle \delta \mathbf{v}(\mathbf{r}_j) \cdot \delta \mathbf{v}(\mathbf{r}_j + \mathbf{R}) \rangle_j}{\langle \delta \mathbf{v}(\mathbf{r}_j) \cdot \delta \mathbf{v}(\mathbf{r}_j) \rangle_j} \right\rangle_\varphi$ (F4)
	Average-first Normalize by geometric mean	$C_{vv}^{GM}(R) = \left\langle \frac{\langle \delta \mathbf{v}(\mathbf{r}_j) \cdot \delta \mathbf{v}(\mathbf{r}_j + \mathbf{R}) \rangle_j}{\sqrt{\langle \delta \mathbf{v}(\mathbf{r}_j) \cdot \delta \mathbf{v}(\mathbf{r}_j) \rangle_j \cdot \langle \delta \mathbf{v}(\mathbf{r}_j + \mathbf{R}) \cdot \delta \mathbf{v}(\mathbf{r}_j + \mathbf{R}) \rangle_j}} \right\rangle_\varphi$ (F5)
	Average-first Normalize by arithmetic mean	$C_{vv}^{AM}(R) = \left\langle \frac{\langle \delta \mathbf{v}(\mathbf{r}_j) \cdot \delta \mathbf{v}(\mathbf{r}_j + \mathbf{R}) \rangle_j}{\frac{1}{2}(\langle \delta \mathbf{v}(\mathbf{r}_j) \cdot \delta \mathbf{v}(\mathbf{r}_j) \rangle_j + \langle \delta \mathbf{v}(\mathbf{r}_j + \mathbf{R}) \cdot \delta \mathbf{v}(\mathbf{r}_j + \mathbf{R}) \rangle_j)} \right\rangle_\varphi$ (F6)

correlation curve

$$C_{\hat{v}\hat{v}}(R) = \left\langle \frac{\text{corr}\left(\frac{\delta \mathbf{v}}{|\delta \mathbf{v}|}, \frac{\delta \mathbf{v}}{|\delta \mathbf{v}|}\right)}{N} \right\rangle_\varphi. \quad (\text{F7})$$

For $C_{vv}(R)$, we compute the autocorrelation of the original vector field $\text{corr}(\delta \mathbf{v}, \delta \mathbf{v})$, take the average, calculate the variance $\langle |\delta \mathbf{v}|^2 \rangle$, and perform normalization. After taking the

azimuthal average, we obtain the correlation curve

$$C_{vv}(R) = \left\langle \frac{\text{corr}(\delta \mathbf{v}, \delta \mathbf{v})}{N \langle |\delta \mathbf{v}|^2 \rangle} \right\rangle_\varphi, \quad (\text{F8})$$

as shown in Fig. 9(c). There are various ways to calculate the variance, which will be discussed next. Overall, their effect on the resulting correlation function is negligible.

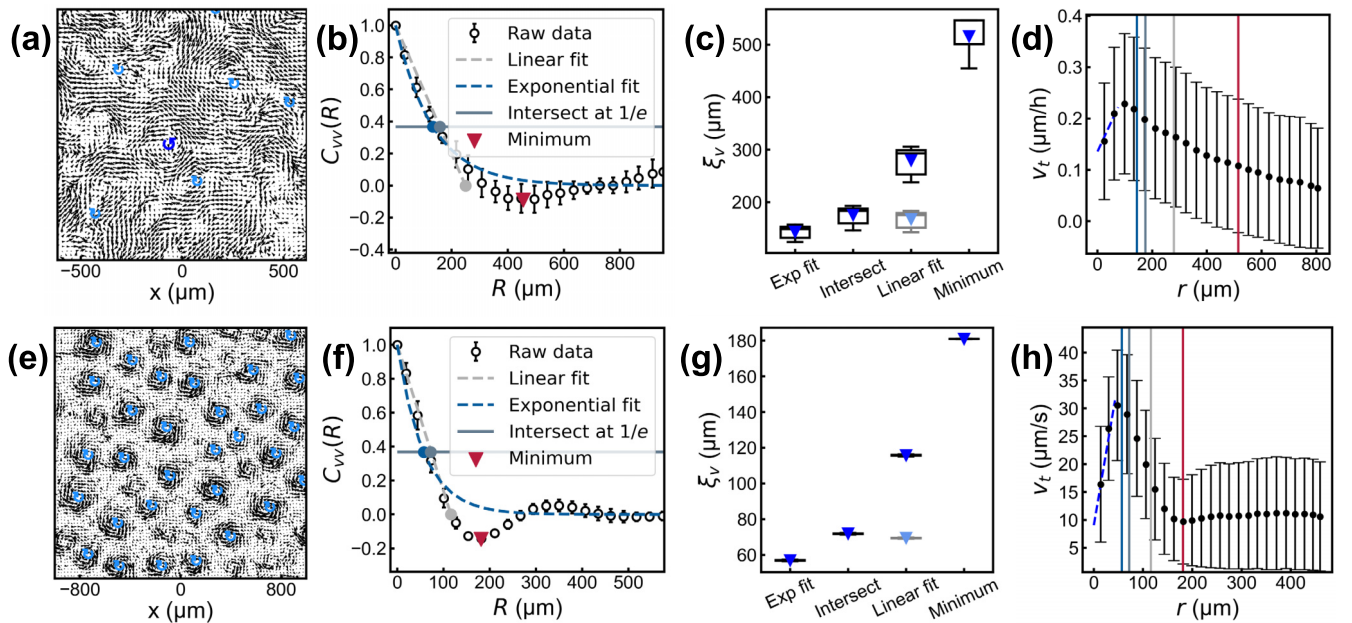


FIG. 11. Correlation lengths of vortices extracted from experimental velocity fields. (a) Velocity field of the human bronchial epithelial cell (HBEC) monolayer [62]. (b) Correlation function curve $C_{vv}(R)$ corresponding to panel (a). (c) Distribution of the correlation lengths. (d) Averaged $v_t(r)$ of the 330 identified vortices from 50 velocity fields. (e) Velocity field of the large-scale vortex lattice in a dense suspension of the bacterium *S. marcescens* [63]. (f) Correlation function curve $C_{vv}(R)$ corresponding to panel (e). (g) Distribution of the correlation lengths. (h) Averaged tangential velocity profile $v_t(r)$ of the 4241 vortices identified from 50 consecutive velocity fields.

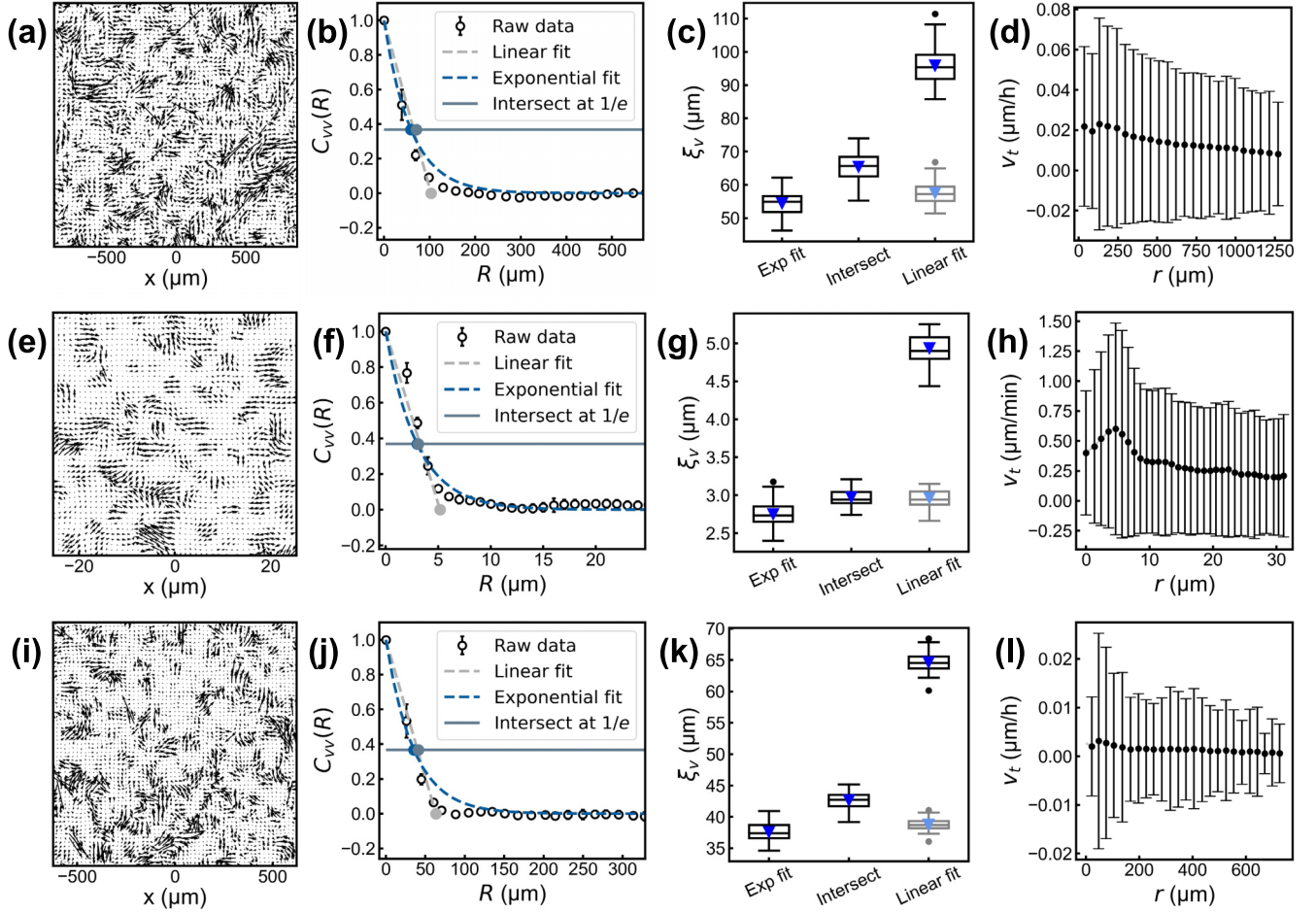


FIG. 12. Correlation lengths of experimental data without apparent vortices in velocity fields. (a), (e), (i) Velocity fields of murine neural progenitor cells (NPCs) [25], *M. xanthus* $\Delta pilA$, and C2C12 cell monolayers. (b), (f), (j) Correlation function curves $C_{vv}(R)$ corresponding to panels (a), (e), and (i). (c), (g), (k) Correlation length distributions from 50 consecutive velocity fields for NPCs, *M. xanthus* $\Delta pilA$, and C2C12 cells, with no valleys detected in any frame. (d), (h), (l) Averaged tangential velocity profiles $v_t(r)$ of identified vortices, showing no rigid body cores. NPCs, *M. xanthus* $\Delta pilA$, and C2C12 cells have 64, 65, and 51 identified vortices, respectively, fewer than those in Fig. 11.

Notably, $C_{\hat{v}\hat{v}}(R)$ and $C_{vv}(R)$ differ in both their two-dimensional heatmaps and their one-dimensional curves.

APPENDIX G: VELOCITY VARIANCE IN CORRELATION FUNCTIONS

Figure 9 shows an illustrative flow chart for calculating the velocity correlation functions Eq. (2) using the “normalize-first” approach [Fig. 9(b)] and Eq. (3) using the “average-first” approach [Fig. 9(c)]. In the “average-first” approach, the variance of velocity $\langle |\delta \mathbf{v}|^2 \rangle$ can be calculated in various ways, as listed in Table I:

(1) The global variance of the entire frame. $C_{vv}^G(R)$ [Eq. (F3)] [10,38,49–51,53,56] uses the global variance, a constant for a given vector field, calculated across the entire frame:

$$\sigma_G^2 = \langle \delta \mathbf{v}(\mathbf{r}_j) \cdot \delta \mathbf{v}(\mathbf{r}_j) \rangle_{j_0}.$$

(2) The local variance of vectors at each separation distance. $C_{vv}^L(R)$ [Eq. (F4)] [31,48,55,59–61,63,68,69] uses the local variance computed only for vectors $\mathbf{v}(\mathbf{r}_j)$ used in the

averaging process:

$$\sigma_L^2 = \langle \delta \mathbf{v}(\mathbf{r}_j) \cdot \delta \mathbf{v}(\mathbf{r}_j) \rangle_j.$$

(3) The geometric mean of the two local variances. $C_{vv}^{GM}(R)$ [Eq. (F5)] [38,52,57,58] uses the geometric mean of the local variances of $\mathbf{v}(\mathbf{r}_j)$ and $\mathbf{v}(\mathbf{r}_j + \mathbf{R})$:

$$\sigma_{GM}^2 = \sqrt{\langle \delta \mathbf{v}(\mathbf{r}_j) \cdot \delta \mathbf{v}(\mathbf{r}_j) \rangle_j \cdot \langle \delta \mathbf{v}(\mathbf{r}_j + \mathbf{R}) \cdot \delta \mathbf{v}(\mathbf{r}_j + \mathbf{R}) \rangle_j}.$$

(4) The arithmetic mean of the two local variances. $C_{vv}^{AM}(R)$ [Eq. (F6)] uses the arithmetic mean of the local variances of $\mathbf{v}(\mathbf{r}_j)$ and $\mathbf{v}(\mathbf{r}_j + \mathbf{R})$:

$$\sigma_{AM}^2 = \frac{1}{2} [\langle \delta \mathbf{v}(\mathbf{r}_j) \cdot \delta \mathbf{v}(\mathbf{r}_j) \rangle_j + \langle \delta \mathbf{v}(\mathbf{r}_j + \mathbf{R}) \cdot \delta \mathbf{v}(\mathbf{r}_j + \mathbf{R}) \rangle_j].$$

Global variance is the most computationally efficient since it is computed once for the whole frame. However, it could fail in highly biased systems. For example, if central vectors are much larger than peripheral ones, the variance can drop sharply at small separations, potentially causing $C_{vv}^G(R)$ to exceed 1. In most experimental datasets, though, $C_{vv}^G(R)$ remains within bounds and is generally reliable. Local variance requires recomputation for each separation $\mathbf{R}$. It is centrosymmetric—flipping it horizontally and vertically

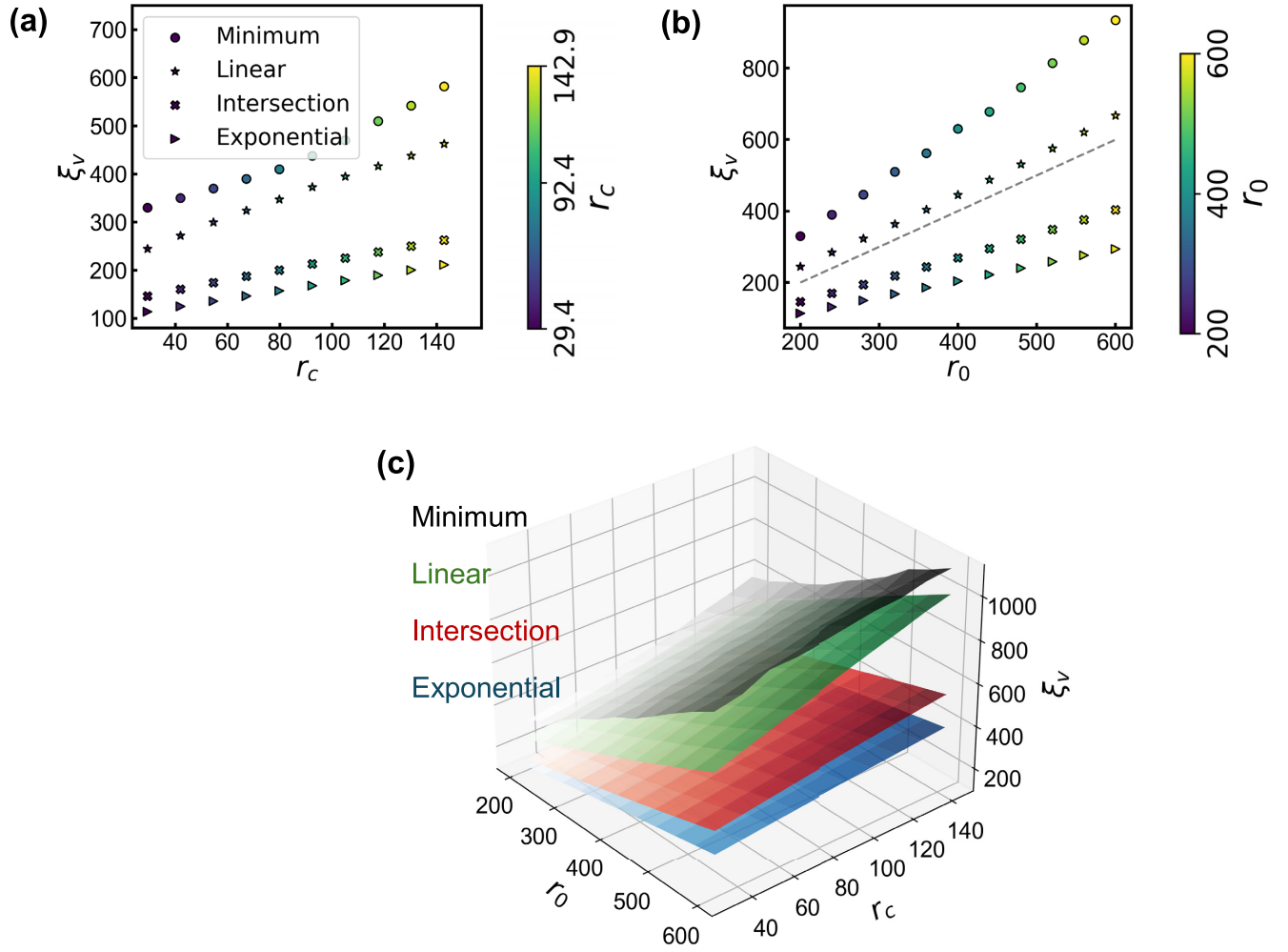


FIG. 13. Plot of the correlation lengths vs the rigid body core radius r_0 and decay parameter r_c of the isolated vortex. (a) For a given rigid body core $r_0 = 200$, the correlation lengths ξ_v change along the decay parameter r_c . (b) For a given decay parameter $r_c = 29.4$, the correlation lengths ξ_v linearly increase along the rigid body core r_0 . The gray dashed line indicates $\xi_v = r_0$. (c) The correlation length ξ_v varies with both the rigid body core radius r_0 and the decay parameter r_c . Among the methods used to calculate ξ_v , the value obtained from the minimum of the valley is consistently the largest, followed by the linear fitting method, which gives the second-largest ξ_v , and the intersection method. The exponential fitting method consistently yields the smallest ξ_v .

yields the same values as for its counterpart at $\mathbf{r}_j + \mathbf{R}$. Biases can arise if the variance distribution is asymmetric; geometric or arithmetic mean normalization mitigates this effect, with minimal difference between the two.

We compared σ^2 , the 2D correlation matrix $C_{vv}(\mathbf{R})$, and the azimuthally averaged curve $C_{vv}(R)$ for all four normalization methods in Fig. 10. The differences are negligible. While the geometric mean is often preferred in the literature for its precision, the global variance offers a good approximation for

TABLE II. Summary of the correlation lengths and the averaged tangential velocities in Figs. 5, 11, and 12.

Figures	Cell line	Correlation lengths ξ_v				Average tangential velocity		
		Exp. fit	Intersect	Linear	Minimum	Vortex number	Rigid body slope	r_0
5(a)–5(d)	Artificial data	95.62 a.u.	116.09 a.u.	189.53 a.u.	303.94 a.u.	5.94	8.45×10^{-3} a.u.	89.69 a.u.
5(e)–5(h)	<i>B. subtilis</i>	3.29 μm	3.84 μm	6.20 μm	10.74 μm	17.78	7.39×10^{-7} h $^{-1}$	3.80 μm
11(a)–11(d)	HBEC	143.57 μm	175.50 μm	279.39 μm	515.18 μm	6.60	1.15×10^{-3} h $^{-1}$	96.58 μm
11(e)–11(h)	<i>S. marcescens</i>	56.87 μm	71.84 μm	115.74 μm	181.00 μm	84.82	1.88×10^3 h $^{-1}$	55.10 μm
5(i)–5(l)	<i>M. xanthus</i> $\Delta frzE$	2.91 μm	3.12 μm	5.17 μm	NA	1.64	3.13 h $^{-1}$	4.66 μm
12(e)–12(h)	<i>M. xanthus</i> $\Delta pilA$	2.75 μm	2.97 μm	4.94 μm	NA	1.30	3.10 h $^{-1}$	4.66 μm
12(a)–12(d)	NPCs	44.39 μm	50.94 μm	74.64 μm	NA	1.28	NA	NA
12(i)–12(l)	C2C12	34.46 μm	36.97 μm	57.03 μm	NA	1.02	NA	NA

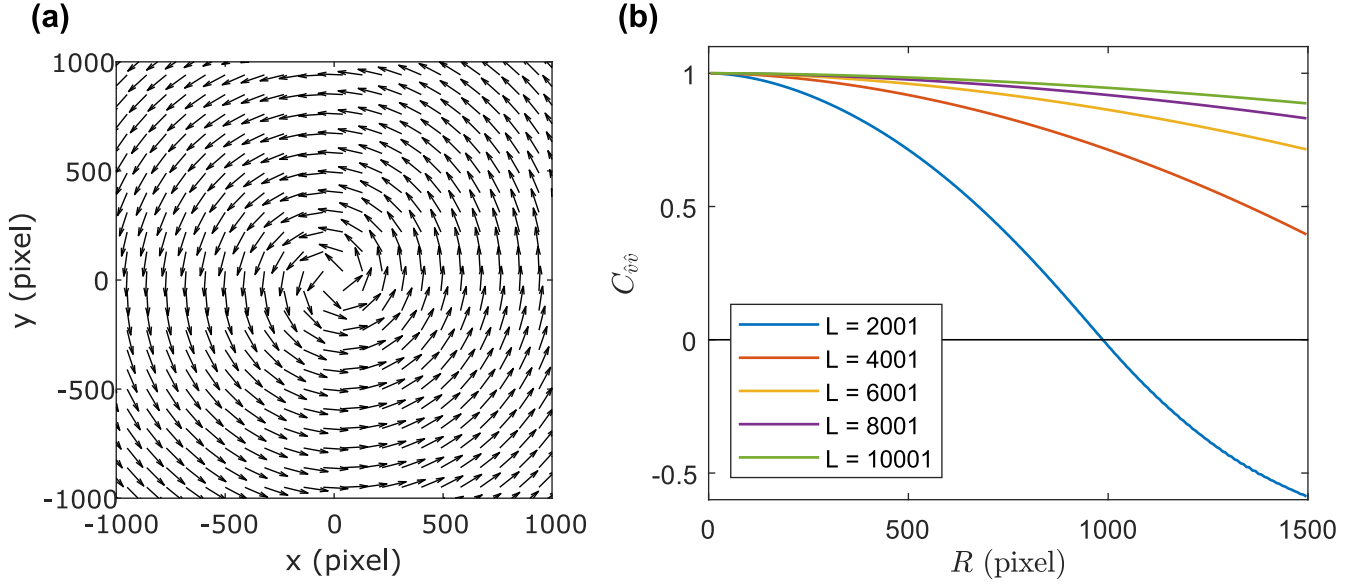


FIG. 14. Velocity correlation function $C_{\hat{v}\hat{v}}$ of an idealized isolated vortex obtained with the normalize-first approach. (a) Flow field of a vortex after normalizing the velocities by their magnitude $\hat{v} = \mathbf{v}/|\mathbf{v}|$. After normalization, the flow field loses its characteristic scales. (b) Correlation functions of a single normalized vortex with different system sizes, where L represents the dimension of the whole flow field used in the calculation. For example, the flow field in panel (a) has the dimension $L = 2001$.

most experimental data. For computational efficiency, we used $C_{vv}^G(R)$ to obtain the results in the main text:

$$C_{vv}(R) = C_{vv}^G(R) = \left\langle \frac{\langle \delta \mathbf{v}(\mathbf{r}_j) \cdot \delta \mathbf{v}(\mathbf{r}_j + \mathbf{R}) \rangle_j}{\langle \delta \mathbf{v}(\mathbf{r}_j) \cdot \delta \mathbf{v}(\mathbf{r}_j) \rangle_{j_0}} \right\rangle_\varphi. \quad (\text{G1})$$

APPENDIX H: CORRELATION LENGTHS

Figure 11 presents the velocity fields of HBEC cells and *S. marcescens*, both of which are abundant in vortices. Consequently, the correlation function curves exhibit a distinct trough with negative values. While plotting the correlation lengths obtained using different methods on top of the averaged tangential velocity profiles, we find that the peak aligns well with the ξ_v calculated via exponential fitting or intersection with $1/e$. Furthermore, the ξ_v determined by the minimum position approximately coincides with the point where the tangential velocity decays close to zero, which is particularly evident in Fig. 11(h). However, this alignment is less apparent in Fig. 11(d), as the vortices are less dominant compared to the other examples.

TABLE III. Summary of the fitting methods to calculate the correlation length from the correlation function curve.

Fitting methods	Formula
Exponential fit	$C_{vv}(R) = e^{-\frac{R}{\xi}}$
Scaled exp. fit	$C_{vv}(R) = A_0 e^{-\frac{R}{\xi}}$
Linear fit	$C_{vv}(R) = 1 - \frac{R}{\xi}$
Intersection	ξ is where $C_{vv}(\xi) = \frac{1}{e}$
Minimum position	ξ is where $C_{vv}(\xi) = \min(C_{vv}(R))$

The correlation lengths ξ_v and the tangential velocity profile information are summarized in Table II. The first four rows correspond to examples with abundant vortices, all of which exhibit a trough in their correlation function curves. Consequently, ξ_v can be determined by identifying the minimum position. Interestingly, fitting the initial increasing portion of the averaged tangential velocity profile reveals that the tangential velocity increases linearly with radial distance until reaching its peak. This observation confirms that, for the examples studied in this paper, the vortices have rigid-body-like cores.

Figure 12 shows examples of correlation function curves that decrease monotonically. These velocity fields exhibit very few identified vortices, and consequently, no troughs are observed in the correlation function curves. In contrast to systems rich in vortices, the examples in Fig. 12 each contain no more than two vortices per frame. The corresponding averaged tangential velocity profiles show large fluctuations, suggesting that these vortices are transient and not persistent over time.

The correlation lengths obtained with different fitting methods are summarized in Table III. Overall, if the correlation function has an evident trough, the minimum position is a good way to obtain the characteristic length scale of the vortex flows; otherwise, if the correlation function decays monotonically, the exponential fit or finding the intersection with a constant are both good methods.

APPENDIX I: IDEALIZED VELOCITY FIELD OF AN ISOLATED VORTEX

Correlation length of isolated vortices. Figure 13 presents the correlation lengths extracted from the velocity fields of

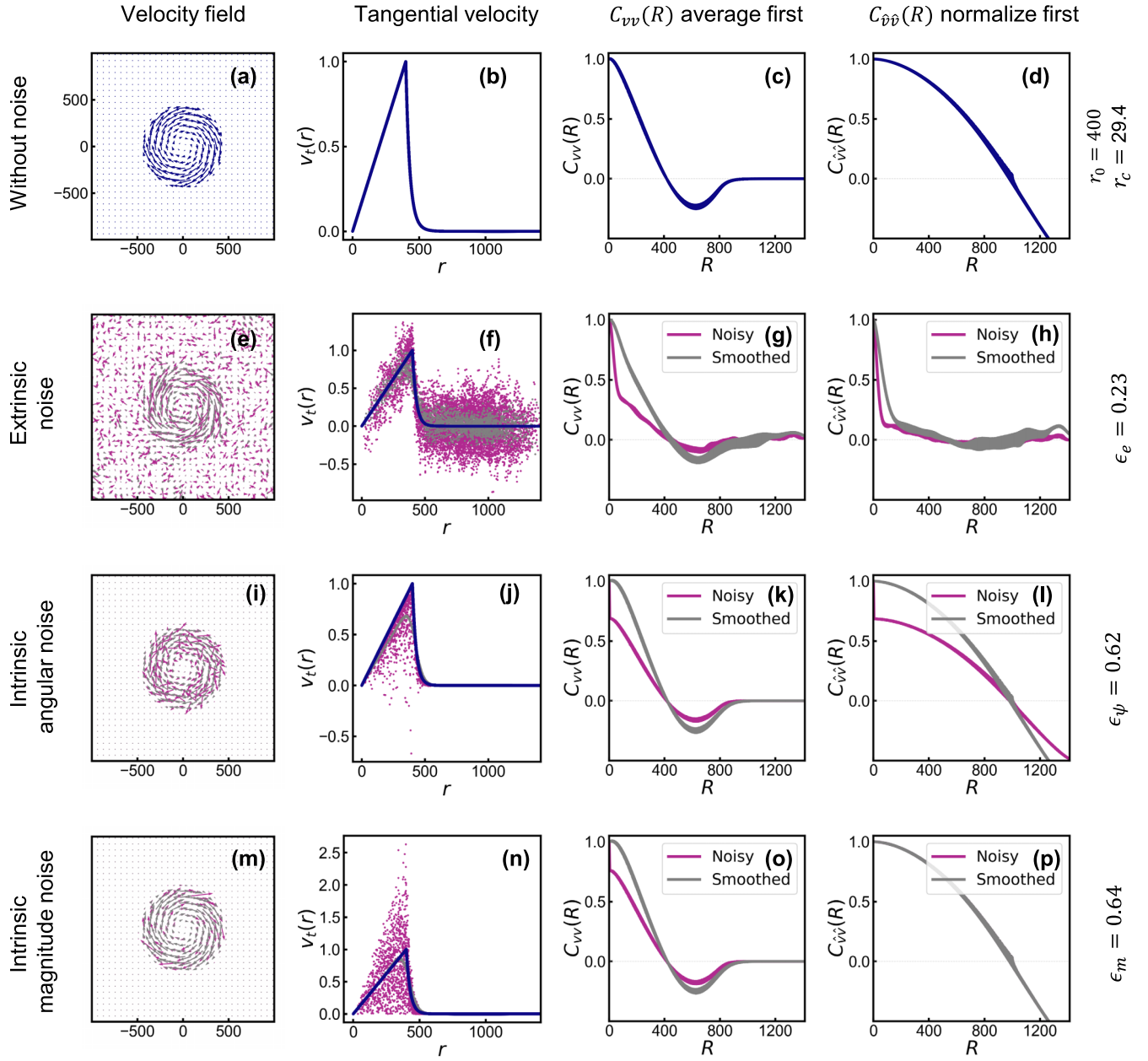


FIG. 15. Effect of smoothing the velocity field on the correlation functions. (a)–(d) The numerically generated noise-free velocity field of an isolated vortex, similar to those in Fig. 4. (e)–(p) After introducing noise, we apply a Gaussian filter to smooth the noisy velocity field and calculate the corresponding correlation function curves. For $C_{vv}(R)$, the smoothed correlation function curve closely matches the noise-free curve. In contrast, for $C_{\hat{v}\hat{v}}(R)$, introducing intrinsic magnitude noise results in identical correlation curves for the noise-free, noisy, and smoothed fields, as normalizing vectors to unit length inherently removes magnitude noise.

isolated vortices in Fig. 3. The correlation lengths ξ_v measured with various methods all increase linearly with the core radius r_0 and also increase with the decay parameter r_c . These results are consistent with those of the multivortex vector fields shown in Fig. 5(m).

Furthermore, the length scale determined by the minimum position of the trough $\xi_v^{\min}$ always yields the largest values, followed by those obtained through linear fitting, ξ_v^{lin} , intersection at $1/e$, ξ_v^{cst} , and exponential fitting, ξ_v^{exp} , which consistently gives the smallest values. This is due to the geometric factors, which can be removed by rescaling these

correlation lengths. After rescaling, $\xi_v^{\min}/2$, $\frac{e-1}{e}\xi_v^{\text{lin}}$, ξ_v^{cst} , and ξ_v^{exp} are approximately the same. This trend holds for both numerically generated velocity fields and experimental data in Fig. 5.

As shown in Fig. 13(c), the correlation length derived from the minimum position is the most sensitive to changes in the core radius r_0 or decay parameter r_c , making it an ideal indicator for the characteristic length of a vortex-rich velocity field.

The normalize-first approach. The normalize-first approach removes any characteristic length scale the original isolated

vortex flow had, as shown in Fig. 14. As a result, the corresponding correlation function C_{vv} decreases monotonically as R increases, but how fast it decreases is only controlled by the system size L , or the dimension of the whole velocity field. The larger the system size, the slower C_{vv} decays. This means that the normalize-first approach does not provide us with any information about the characteristic sizes of the original vortex flow, especially when the flow has apparent symmetries, such as a single vortex.

Effects of smoothing on the correlation function of noisy velocity fields. Figure 15 illustrates the effect of smoothing the velocity field on the correlation function curves. For velocity fields with extrinsic noise, smoothing restores the correlation curves [Fig. 15(g)] to a shape similar to the original [Fig. 15(c)], with only minor distortion. For velocity fields with intrinsic noise [Figs. 15(k) and 15(o)], after smoothing, the correlation function curves closely resemble those of the noise-free case [Fig. 15(c)].

-
- [1] H. Xu, Y. Huo, Q. Zhou, L. A. Wang, P. Cai, B. Doss, C. Huang, and K. J. Hsia, Geometry-mediated bridging drives nonadhesive stripe wound healing, *Proc. Natl. Acad. Sci. USA* **120**, e2221040120 (2023).
 - [2] M. Poujade, E. Grasland-Mongrain, A. Hertzog, J. Jouanneau, P. Chavrier, B. Ladoux, A. Buguin, and P. Silberzan, Collective migration of an epithelial monolayer in response to a model wound, *Proc. Natl. Acad. Sci. USA* **104**, 15988 (2007).
 - [3] T. Chen, A. Callan-Jones, E. Fedorov, A. Ravasio, A. Brugués, H. T. Ong, Y. Toyama, B. C. Low, X. Trepát, T. Shemesh, *et al.*, Large-scale curvature sensing by directional actin flow drives cellular migration mode switching, *Nat. Phys.* **15**, 393 (2019).
 - [4] Q. Wei, X. Shi, T. Zhao, P. Cai, T. Chen, Y. Zhang, C. Huang, J. Yang, X. Chen, and S. Zhang, Actin-ring segment switching drives nonadhesive gap closure, *Proc. Natl. Acad. Sci. USA* **117**, 33263 (2020).
 - [5] Y. Yang and H. Levine, Role of the supracellular actomyosin cable during epithelial wound healing, *Soft Matter* **14**, 4866 (2018).
 - [6] A. Ravasio, I. Cheddadi, T. Chen, T. Pereira, H. T. Ong, C. Bertocchi, A. Brugués, A. Jacinto, A. J. Kabla, Y. Toyama, *et al.*, Gap geometry dictates epithelial closure efficiency, *Nat. Commun.* **6**, 7683 (2015).
 - [7] S. R. K. Vedula, G. Peyret, I. Cheddadi, T. Chen, A. Brugués, H. Hirata, H. Lopez-Menendez, Y. Toyama, L. N. de Almeida, X. Trepát, *et al.*, Mechanics of epithelial closure over non-adherent environments, *Nat. Commun.* **6**, 6111 (2015).
 - [8] V. Nier, M. Deforet, G. Duclos, H. G. Yevick, O. Cochet-Escartin, P. Marcq, and P. Silberzan, Tissue fusion over nonadhering surfaces, *Proc. Natl. Acad. Sci. USA* **112**, 9546 (2015).
 - [9] P. Martin, Wound healing—Aiming for perfect skin regeneration, *Science* **276**, 75 (1997).
 - [10] K. Copenhagen, R. Alert, N. S. Wingreen, and J. W. Shaevitz, Topological defects promote layer formation in *Myxococcus xanthus* colonies, *Nat. Phys.* **17**, 211 (2021).
 - [11] E. Han, C. Fei, R. Alert, K. Copenhagen, M. D. Koch, N. S. Wingreen, and J. W. Shaevitz, Local polar order controls mechanical stress and triggers layer formation in *Myxococcus xanthus* colonies, *Nat. Commun.* **16**, 952 (2025).
 - [12] Y. I. Yaman, E. Demir, R. Vetter, and A. Kocabas, Emergence of active nematics in chaining bacterial biofilms, *Nat. Commun.* **10**, 2285 (2019).
 - [13] Y. Chai, T. Norman, R. Kolter, and R. Losick, An epigenetic switch governing daughter cell separation in *Bacillus subtilis*, *Genes Dev.* **24**, 754 (2010).
 - [14] S. S. Branda, J. E. González-Pastor, S. Ben-Yehuda, R. Losick, and R. Kolter, Fruiting body formation by *Bacillus subtilis*, *Proc. Natl. Acad. Sci. USA* **98**, 11621 (2001).
 - [15] H. Vlamakis, Y. Chai, P. Beauregard, R. Losick, and R. Kolter, Sticking together: Building a biofilm the *Bacillus subtilis* way, *Nat. Rev. Microbiol.* **11**, 157 (2013).
 - [16] B. Qin, C. Fei, A. A. Bridges, A. A. Mashruwala, H. A. Stone, N. S. Wingreen, and B. L. Bassler, Cell position fates and collective fountain flow in bacterial biofilms revealed by light-sheet microscopy, *Science* **369**, 71 (2020).
 - [17] P. J. Keller, A. D. Schmidt, J. Wittbrodt, and E. H. Stelzer, Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy, *Science* **322**, 1065 (2008).
 - [18] D. Ellison, A. Mugler, M. D. Brennan, S. H. Lee, R. J. Huebner, E. R. Shamir, L. A. Woo, J. Kim, P. Amar, I. Nemenman, *et al.*, Cell-cell communication enhances the capacity of cell ensembles to sense shallow gradients during morphogenesis, *Proc. Natl. Acad. Sci. USA* **113**, E679 (2016).
 - [19] M. Behrndt, G. Salbreux, P. Campinho, R. Hauschild, F. Oswald, J. Roensch, S. W. Grill, and C.-P. Heisenberg, Forces driving epithelial spreading in zebrafish gastrulation, *Science* **338**, 257 (2012).
 - [20] S. Nair and T. F. Schilling, Chemokine signaling controls endodermal migration during zebrafish gastrulation, *Science* **322**, 89 (2008).
 - [21] F. Peruani, F. Ginelli, M. Bär, and H. Chaté, Polar vs apolar alignment in systems of polar self-propelled particles, *J. Phys.: Conf. Ser.* **297**, 012014 (2011).
 - [22] A. Doostmohammadi and B. Ladoux, Physics of liquid crystals in cell biology, *Trends Cell Biol.* **32**, 140 (2022).
 - [23] T. B. Saw, W. Xi, B. Ladoux, and C. T. Lim, Biological tissues as active nematic liquid crystals, *Adv. Mater.* **30**, 1802579 (2018).
 - [24] Y. Luo, M. Gu, M. Park, X. Fang, Y. Kwon, J. M. Urueña, J. R. de Alaniz, M. E. Helgeson, M. C. Marchetti, and M. T. Valentine, Molecular-scale substrate anisotropy and crowding drive long-range nematic order of cell monolayers, *J. R. Soc. Interface* **20**, 20230160 (2023).
 - [25] K. Kawaguchi, R. Kageyama, and M. Sano, Topological defects control collective dynamics in neural progenitor cell cultures, *Nature (London)* **545**, 327 (2017).
 - [26] D. Nishiguchi, K. H. Nagai, H. Chaté, and M. Sano, Long-range nematic order and anomalous fluctuations in suspensions of swimming filamentous bacteria, *Phys. Rev. E* **95**, 020601 (2017).
 - [27] N. S. Rossen, J. M. Tarp, J. Mathiesen, M. H. Jensen, and L. B. Oddershede, Long-range ordered vorticity patterns in

- living tissue induced by cell division, *Nat. Commun.* **5**, 5720 (2014).
- [28] Y. Huo, H. Xu, Y. Zhao, E. Han, M. Lin, and K. J. Hsia, Myoblast choreographic alignment, *ACS Nano* **19**, 36245 (2025).
- [29] F. C. Keber, E. Loiseau, T. Sanchez, S. J. DeCamp, L. Giomi, M. J. Bowick, M. C. Marchetti, Z. Dogic, and A. R. Bausch, Topology and dynamics of active nematic vesicles, *Science* **345**, 1135 (2014).
- [30] A. Doostmohammadi, S. P. Thampi, and J. M. Yeomans, Defect-mediated morphologies in growing cell colonies, *Phys. Rev. Lett.* **117**, 048102 (2016).
- [31] G. Duclos, C. Erlenkämper, J.-F. Joanny, and P. Silberzan, Topological defects in confined populations of spindle-shaped cells, *Nat. Phys.* **13**, 58 (2017).
- [32] T. Sanchez, D. T. Chen, S. J. DeCamp, M. Heymann, and Z. Dogic, Spontaneous motion in hierarchically assembled active matter, *Nature (London)* **491**, 431 (2012).
- [33] L. Balasubramaniam, A. Doostmohammadi, T. B. Saw, G. H. N. S. Narayana, R. Mueller, T. Dang, M. Thomas, S. Gupta, S. Sonam, A. S. Yap, *et al.*, Investigating the nature of active forces in tissues reveals how contractile cells can form extensile monolayers, *Nat. Mater.* **20**, 1156 (2021).
- [34] L. Giomi, M. J. Bowick, X. Ma, and M. C. Marchetti, Defect annihilation and proliferation in active nematics, *Phys. Rev. Lett.* **110**, 228101 (2013).
- [35] T. Turiv, J. Krieger, G. Babakhanova, H. Yu, S. V. Shivanovskii, Q.-H. Wei, M.-H. Kim, and O. D. Lavrentovich, Topology control of human fibroblast cells monolayer by liquid crystal elastomer, *Sci. Adv.* **6**, eaaz6485 (2020).
- [36] T. B. Saw, A. Doostmohammadi, V. Nier, L. Kocgozlu, S. Thampi, Y. Toyama, P. Marcq, C. T. Lim, J. M. Yeomans, and B. Ladoux, Topological defects in epithelia govern cell death and extrusion, *Nature (London)* **544**, 212 (2017).
- [37] Y. Huo, K. Guo, M. Paradiso, and K. J. Hsia, Stress distribution in contractile cell monolayers, *Extreme Mech. Lett.* **82**, 102426 (2026).
- [38] T. S. Grigera, Correlation functions as a tool to study collective behaviour phenomena in biological systems, *J. Phys.: Complexity* **2**, 045016 (2021).
- [39] R. Kemkemer, S. Jungbauer, D. Kaufmann, and H. Gruler, Cell orientation by a microgrooved substrate can be predicted by automatic control theory, *Biophys. J.* **90**, 4701 (2006).
- [40] G. Duclos, S. Garcia, H. Yevick, and P. Silberzan, Perfect nematic order in confined monolayers of spindle-shaped cells, *Soft Matter* **10**, 2346 (2014).
- [41] P. Guillamat, C. Blanch-Mercader, G. Pernollet, K. Kruse, and A. Roux, Integer topological defects organize stresses driving tissue morphogenesis, *Nat. Mater.* **21**, 588 (2022).
- [42] T. H. Tan, M. Malik-Garbi, E. Abu-Shah, J. Li, A. Sharma, F. C. MacKintosh, K. Keren, C. F. Schmidt, and N. Fakhri, Self-organized stress patterns drive state transitions in actin cortices, *Sci. Adv.* **4**, eaar2847 (2018).
- [43] B.-C. Wang and G.-K. Xu, Collective migration of cells in geometric spaces: Intrinsic correlation length racing against extrinsic confinement size, *J. Mech. Phys. Solids* **177**, 105327 (2023).
- [44] H. Haga, C. Irahara, R. Kobayashi, T. Nakagaki, and K. Kawabata, Collective movement of epithelial cells on a collagen gel substrate, *Biophys. J.* **88**, 2250 (2005).
- [45] W. Tang, A. Das, A. F. Pegoraro, Y. L. Han, J. Huang, D. A. Roberts, H. Yang, J. J. Fredberg, D. N. Kotton, D. Bi, *et al.*, Collective curvature sensing and fluidity in three-dimensional multicellular systems, *Nat. Phys.* **18**, 1371 (2022).
- [46] J. Muñoz-Dorado, F. J. Marcos-Torres, E. García-Bravo, A. Moraleda-Muñoz, and J. Pérez, Myxobacteria: Moving, killing, feeding, and surviving together, *Front. Microbiol.* **7**, 781 (2016).
- [47] V. Yashunsky, D. J. Pearce, G. Ariel, and A. Be'er, Topological defects in multi-layered swarming bacteria, *Soft Matter* **20**, 4237 (2024).
- [48] S. Garcia, E. Hannezo, J. Elgeti, J.-F. Joanny, P. Silberzan, and N. S. Gov, Physics of active jamming during collective cellular motion in a monolayer, *Proc. Natl. Acad. Sci. USA* **112**, 15314 (2015).
- [49] C. Malinverno, S. Corallino, F. Giavazzi, M. Bergert, Q. Li, M. Leoni, A. Disanza, E. Frittoli, A. Oldani, E. Martini, *et al.*, Endocytic reawakening of motility in jammed epithelia, *Nat. Mater.* **16**, 587 (2017).
- [50] G. Szamel and E. Flenner, Long-ranged velocity correlations in dense systems of self-propelled particles, *Europhys. Lett.* **133**, 60002 (2021).
- [51] D. T. Tambe, C. Corey Hardin, T. E. Angelini, K. Rajendran, C. Y. Park, X. Serra-Picamal, E. H. Zhou, M. H. Zaman, J. P. Butler, D. A. Weitz, *et al.*, Collective cell guidance by cooperative intercellular forces, *Nat. Mater.* **10**, 469 (2011).
- [52] K. Doxzen, S. R. K. Vedula, M. C. Leong, H. Hirata, N. S. Gov, A. J. Kabla, B. Ladoux, and C. T. Lim, Guidance of collective cell migration by substrate geometry, *Integr. Biol.* **5**, 1026 (2013).
- [53] M. Vishwakarma, J. Di Russo, D. Probst, U. S. Schwarz, T. Das, and J. P. Spatz, Mechanical interactions among followers determine the emergence of leaders in migrating epithelial cell collectives, *Nat. Commun.* **9**, 3469 (2018).
- [54] S. P. Thampi, R. Golestanian, and J. M. Yeomans, Velocity correlations in an active nematic, *Phys. Rev. Lett.* **111**, 118101 (2013).
- [55] E. J. Hemingway, P. Mishra, M. C. Marchetti, and S. M. Fielding, Correlation lengths in hydrodynamic models of active nematics, *Soft Matter* **12**, 7943 (2016).
- [56] A. J. Kabla, Collective cell migration: Leadership, invasion and segregation, *J. R. Soc. Interface* **9**, 3268 (2012).
- [57] L. Petitjean, M. Reffay, E. Grasland-Mongrain, M. Poujade, B. Ladoux, A. Buguin, and P. Silberzan, Velocity fields in a collectively migrating epithelium, *Biophys. J.* **98**, 1790 (2010).
- [58] S. R. K. Vedula, M. C. Leong, T. L. Lai, P. Hersen, A. J. Kabla, C. T. Lim, and B. Ladoux, Emerging modes of collective cell migration induced by geometrical constraints, *Proc. Natl. Acad. Sci. USA* **109**, 12974 (2012).
- [59] S.-Z. Lin, S. Ye, G.-K. Xu, B. Li, and X.-Q. Feng, Dynamic migration modes of collective cells, *Biophys. J.* **115**, 1826 (2018).
- [60] T. E. Angelini, E. Hannezo, X. Treppe, J. J. Fredberg, and D. A. Weitz, Cell migration driven by cooperative substrate deformation patterns, *Phys. Rev. Lett.* **104**, 168104 (2010).
- [61] R. Mueller, J. M. Yeomans, and A. Doostmohammadi, Emergence of active nematic behavior in monolayers of isotropic cells, *Phys. Rev. Lett.* **122**, 048004 (2019).

- [62] C. Blanch-Mercader, V. Yashunsky, S. Garcia, G. Duclos, L. Giomi, and P. Silberzan, Turbulent dynamics of epithelial cell cultures, *Phys. Rev. Lett.* **120**, 208101 (2018).
- [63] H. Xu and Y. Wu, Self-enhanced mobility enables vortex pattern formation in living matter, *Nature (London)* **627**, 553 (2024).
- [64] H. Xu, Y. Huo, K. J. Hsia, and E. Han, Repository for: “A robust framework for quantifying collective motion in living matter with spatial correlation functions” DR-NTU, 2026, <https://doi.org/10.21979/N9/POFIDG>.
- [65] E. Stamhuis and W. Thielicke, PIVlab—Towards user-friendly, affordable and accurate digital particle image velocimetry in MATLAB, *J. Open Res. Softw.* **2**, 30 (2014).
- [66] B. Sabass, M. D. Koch, G. Liu, H. A. Stone, and J. W. Shaevitz, Force generation by groups of migrating bacteria, *Proc. Natl. Acad. Sci. USA* **114**, 7266 (2017).
- [67] C. Kaimer and D. R. Zusman, Regulation of cell reversal frequency in *Myxococcus xanthus* requires the balanced activity of CheY-like domains in FrzE and FrzZ, *Mol. Microbiol.* **100**, 379 (2016).
- [68] H. Xu, Y. Huang, R. Zhang, and Y. Wu, Autonomous waves and global motion modes in living active solids, *Nat. Phys.* **19**, 46 (2023).
- [69] S. P. Thampi, R. Golestanian, and J. M. Yeomans, Vorticity, defects and correlations in active turbulence, *Philos. Trans. R. Soc. A* **372**, 20130366 (2014).