



OPEN

Comparison of particle image velocimetry and the underlying agents dynamics in collectively moving self propelled particles

Uday S. Basak¹, Sulimon Sattari²✉, Md. Motaleb Hossain^{2,3}, Kazuki Horikawa⁴, Mikito Toda^{5,6,7} & Tamiki Komatsuzaki^{1,8,9,10,11,12}✉

Collective migration of cells is a fundamental behavior in biology. For the quantitative understanding of collective cell migration, live-cell imaging techniques have been used using e.g., phase contrast or fluorescence images. Particle tracking velocimetry (PTV) is a common recipe to quantify cell motility with those image data. However, the precise tracking of cells is not always feasible. Particle image velocimetry (PIV) is an alternative to PTV, corresponding to Eulerian picture of fluid dynamics, which derives the average velocity vector of an aggregate of cells. However, the accuracy of PIV in capturing the underlying cell motility and what values of the parameters should be chosen is not necessarily well characterized, especially for cells that do not adhere to a viscous flow. Here, we investigate the accuracy of PIV by generating images of simulated cells by the Vicsek model using trajectory data of agents at different noise levels. It was found, using an alignment score, that the direction of the PIV vectors coincides with the direction of nearby agents with appropriate choices of PIV parameters. PIV is found to accurately measure the underlying motion of individual agents for a wide range of noise level, and its condition is addressed.

Cells perform organized, ‘orchestrated’ movement known as collective cell migration^{1,2}. Their migration is crucial during embryo and organ development^{3,4}, wound healing^{1,5}, cancer growth⁶, and immune defense⁷. The failure of cells to cooperate may result in life-threatening concerns^{1,8}. How do cells harmonize with each other during migration, given that individual cells have limited sensory, communication, and decision-making abilities? In some cases, collective behavior is coordinated by leader cells, or cells whose behavior is more influential than others, however, the existence and role of leader cells are still widely unknown for different cell lines^{9–12}. Quantitative measurement of cell velocities can shed light on how they coordinate. Recently, improvements in image acquisition have resulted in high-resolution, high-frame rate, images with unprecedented viewing range^{13–19}, which provide rich information on collective cell motion.

Particle tracking velocimetry (PTV) is a well-known technique where individual cells are tracked, resulting in a so-called “Lagrangian description” of the velocity dynamics. PTV usually contains two steps—a segmentation step, where the locations of cells and their boundaries are identified^{20–22}, and a tracking step, where the location

¹Pabna University of Science and Technology, Pabna 6600, Bangladesh. ²Research Center of Mathematics for Social Creativity, Research Institute for Electronic Science, Hokkaido University, Kita 20, Nishi 10, Kita-ku, Sapporo 001-0020, Japan. ³University of Dhaka, Dhaka 1000, Bangladesh. ⁴Department of Optical Imaging, Advanced Research Promotion Center, Tokushima University, Kuramoto-cho 3-18-15, Tokushima, Tokushima 770-8503, Japan. ⁵Faculty Division of Natural Sciences, Research Group of Physics, Nara Women’s University, Kita-Uoya-Nishimachi, Nara 630-8506, Japan. ⁶Graduate School of Information Science, University of Hyogo, 7-1-28 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan. ⁷Research Institute for Electronic Science, Hokkaido University, Kita 20 Nishi 10, Kita-Ku, Sapporo 001-0020, Japan. ⁸Graduate School of Life Science, Transdisciplinary Life Science Course, Hokkaido University, Kita 12, Nishi 6, Kita-ku, Sapporo 060-0812, Japan. ⁹Institute for Chemical Reaction Design and Discovery (WPI-ICReDD), Hokkaido University, Kita 21 Nishi 10, Kita-ku, Sapporo, Hokkaido 001-0021, Japan. ¹⁰Graduate School of Chemical Sciences and Engineering Materials Chemistry and Engineering Course, Hokkaido University, Kita 13, Nishi 8, Kita-ku, Sapporo 060-0812, Japan. ¹¹Institute for Open and Transdisciplinary Research Initiatives, Osaka University, Yamadaoka, Suita, Osaka 565-0871, Japan. ¹²The Institute of Scientific and Industrial Research, Osaka University, Mihogaoka, 8-1, Osaka, Ibaraki 567-0047, Japan. ✉email: sulimon510@gmail.com; tamiki@es.hokudai.ac.jp

of each cell in one frame is associated with the location of the same cell in the next frame. PTV is onerous when done manually for many cells, and, when automated using supervised and unsupervised techniques, can lead to erroneous results when cell density is high or cells move rapidly compared to the image acquisition rate^{23–25}. In segmentation, a cell may be missed, or something which is not a cell can be marked as a cell. In tracking, a cell can be associated with another (incorrect) cell in the next frame. Furthermore, for the supervised methods, a fair amount of training data set is needed, whose labels must be assigned manually²⁰.

Alternatively, one can obtain an Eulerian description of the velocity field, where, instead of identifying individual cells, flow properties are expressed as a field. Particle image velocimetry (PIV) is a technique that has been used widely in cell migration analysis^{26–30} to obtain Eulerian velocity field dynamics from sequences of images. In PIV, each image is first divided into interrogation regions (grids), and each grid area in a given frame is associated with the corresponding grid area in the next frame with the highest cross-correlation³¹. Previously, the efficacy of the PIV technique has been assessed to a sequence of images generated by applying a known flow mapping to images of cells, to see how the vectors outputted by PIV can accurately track the deformations in the image by the flow field over time. It was shown that PIV works well in these cases²³. However, it is yet to be scrutinized the performance of PIV when cells move in an incoherent fashion, or when the system changes from disordered motions to some coherent state. In many cell lines, such as *Dictyostelium discoideum*, cells do not move by any prescribed viscous flow, and instead, travel like stochastic particles with some couplings. In order to verify the effectiveness of PIV in reconstructing the trajectories of cells which do not travel in a viscous flow, one may take the help of simulation models where the actual trajectories are known.

In this paper, we investigate the accuracy of PIV in capturing the motion of particles, not in a prescribed viscous flow, but the particles spontaneously move with their mutual interference under the presence of noise by the Vicsek Model (VM)³². Here noise emerges from the particles' "failure" to accurately follow the motion of their neighbors. Using the trajectory data of particles at different noise levels, we produce images of the agents by plotting filled ovals at each particle location which are oriented following the interaction rules of the VM. The PIV technique is applied to the images alone without having access to the original trajectory data. We then compare the velocities computed by PIV to the velocities known from the trajectories using an alignment score, A^R , we introduce in this paper. Here, R denotes a given radius of a circle whose center coincides with each PIV position so that all particles dynamics passing through inside the circle are taken into account. The alignment score has revealed that the PIV vectors' direction matches the direction of the adjacent particles, especially at low noise levels where the system exhibits global coherence. For higher noise cases, the selection of a precise PIV grid size and R values have proven to be crucial. It has been seen that the alignment score increases significantly upon incorporating such R values such that agents are relatively coherent with each other within that radius. As the noise level gets higher, the agents get more randomized so that the larger the PIV grid size, the more the PIV vectors quickly deviate from the underlying agents' dynamics. In other words, PIV can still capture the dynamics inherent to the entities of the agents if we use the right PIV grid size in which just a single agent exists on average.

Averaged normalized velocity and alignment score

We now describe the PIV scheme and also several measures that have been investigated as ways to quantify the efficiency of PIV in scrutinizing collective migration.

Particle image velocimetry (PIV). PIV computes Eulerian velocity fields, that is, it measures a flow field as the fluid passes an observation point that is fixed in space. PIV does not use image segmentation, instead, it computes velocity vectors by searching for a zone in a future frame that matches best with a given grid in the current frame. Figure 1 shows two consecutive images of moving particles at time t (Fig. 1a) and $t + 1$ (Fig. 1b). At the beginning of the process, each image is divided into grids as shown in Fig. 1, and the center of each grid remains the same in all images. For each grid at time t , say the red grid in Fig. 1a, our goal is to find the zone at a subsequent image at time $t + 1$ where the image in the grid at time t matches best with the image in the zone at time $t + 1$. Instead of searching the whole image to find the best match of the red grid, an interrogation zone which is typically a square of size $N_g \times N_g$ pixels is defined in the succeeding image (shaded region in Fig. 1b) where N is chosen to be roughly equal to the largest possible displacement of agents within two successive frames. Then PIV creates a temporary grid (we refer to it as a window) within the search zone at time $(t + 1)$ (blue box in Fig. 1b) of the same size as the red grid. We then search the window to match the orientation of particles within the red grid. The location of the window is then varied within the search zone to find the best match. The process of finding the best match for the displacements of grids between two subsequent images is performed by cross-correlation having the following form²³:

$$C(k, l) = \frac{\sum_{i=1}^Z \sum_{j=1}^Z [I^t(i, j) - \mu_{I^t}] [I^{t+1}(i + k, j + l) - \mu_{I^{t+1}}]}{\sqrt{\sum_{i=1}^Z \sum_{j=1}^Z [I^t(i, j) - \mu_{I^t}]^2} \sqrt{\sum_{i=1}^Z \sum_{j=1}^Z [I^{t+1}(i + k, j + l) - \mu_{I^{t+1}}]^2}} \quad (1)$$

where $I^t(i, j)$ and $I^{t+1}(i + k, j + l)$ denote the intensity of images at the positions (i, j) and $(i + k, j + l)$ at time t and $t + 1$, respectively. Here Z represents the number of pixels along one side of the (square) grid. Here μ_{I^t} represents the average intensity of a grid at time t (red grid, say), and $\mu_{I^{t+1}}$ is the average intensity of a grid located within the interrogation window (e.g., blue grid) at time $t + 1$. (k, l) is the displacement of the (red) grid from its original location. The correlation function $C(k, l)$ defined in Eq. (1) is maximized when the overlapping part between the red grid (Fig. 1a) and its best matched location at two consecutive images is maximum, and the location of this peak in the correlation then defines how far the (red) grid has moved between the two consecutive images. Finally, PIV draws a vector starting from the center of the (red) grid to the center of its best-matched

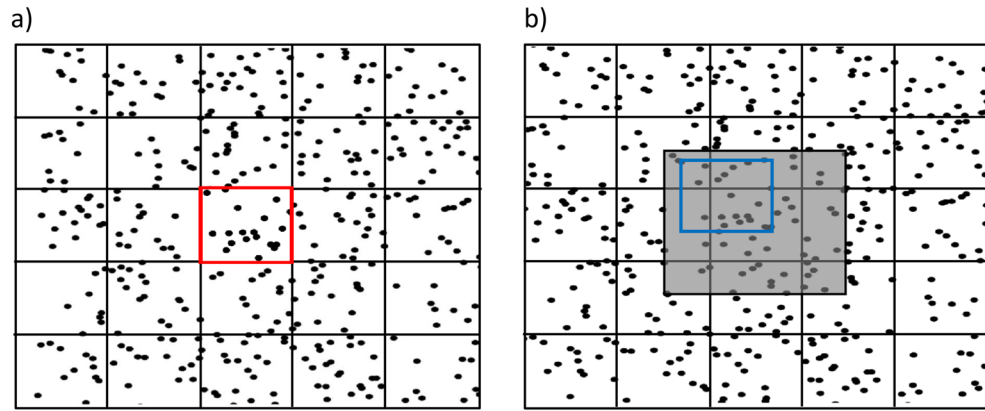


Figure 1. Schematic diagram of moving particles at time (a) t and (b) $t + 1$. Small black boxes in both figures represent a division of the original image into grids, and for each grid, a PIV vector is defined that estimates the mean displacement vector of particles within that grid. The shaded region in (b) represents the interrogation zone of the red grid in (a) within which the matched area may exist. The size of the interrogation window largely depends on the distance traveled by the particles in the time interval between two successive images. The blue box inside the search zone represents the window whose center is varied within the search zone. The orientation of agents within the red grid and windows is computed to find the best-matched window.

window. The length of the vector represents the distance that the particles within the (red) grid travel during these consecutive time frames. Hence, a PIV vector characterizes the aggregate motion of particles located within a grid, and by studying all the PIV vectors in the system one can surmise the way particles move.

Averaged normalized velocity. The VM is based on a simple rule where every agent in the swarm aligns its velocity with the average velocity of its neighbors within a certain radius^{32,33} (see “Methods” section “Vicsek model” for a mathematical description of the VM). The alignment of an agent is not perfect, however, there is some noise present in the system. When the intensity of the noise present in the system is low, the system quickly gets aligned in the same direction and moves coherently. But the coherent motion disappears as the noise increases. Particles in the system move randomly at very high noise. This means a phase transition, from ordered motion to disordered states, occurs. This transition can be studied by an order parameter equal to the average normalized velocity having the following form³²:

$$v_a(t) = \frac{1}{N} \left| \sum_{i=1}^N \frac{\vec{v}_i(t)}{|\vec{v}_i(t)|} \right|, \tag{2}$$

where N is the number of particles present in the system, and $\vec{v}_i(t)$ represents the velocity of the particle i at time t . If the particles in the system move randomly, the order parameter v_a tends to be zero, and it tends to be 1 if all the particles move in the same direction³². This order parameter $v_a(t)$ is defined in such a way that it characterizes the global behavior of the system at time t by incorporating the action of all particles present in the system. In general, coherence may have a length scale, that is, all the particles in the system may not move coherently, but locally some coherence can be observed. To characterize this local behavior, we introduce a variable R that portrays the radius of the circle within which the local coherent movement of particles $v_a^{\vec{x}_j, R}$ is evaluated:

$$v_a^{\vec{x}_j, R}(t) = \frac{1}{N_j^R(t)} \left| \sum_{i \text{ s.t. } |\vec{x}_j - \vec{x}_i| \leq R}^{N_j^R(t)} \frac{\vec{v}_i(t)}{|\vec{v}_i(t)|} \right|, \tag{3}$$

where \vec{x}_j is the center of the circle which is chosen at the initial point \vec{x}_j of a PIV vector j , and $N_j^R(t)$ is the number of particles lying within a distance R from the PIV vector j at time t . Point to be noted that the PIV vectors are equally spaced all over the field. We finally take the mean of all local $v_a^{\vec{x}_j, R}$ ’s as follows:

$$v_a^R(t) = \frac{1}{M} \sum_j \left(v_a^{\vec{x}_j, R}(t) \right), \tag{4}$$

where M represents the total number of PIV vectors defined for each image data.

Alignment score. As stated in “Particle image velocimetry (PIV)”, PIV uses a statistical measure in the form of cross-correlation to find the best match between two consecutive images, which requires a sufficient amount of data to determine the displacement accurately. Hence, the grid size is chosen so that each grid contains multi-

ple particles²³, and PIV estimates the mean displacement of a group of particles located within the grid. How well can PIV characterize the underlying particles' motion? To investigate this issue we introduce a measure named 'alignment score' which is the average cosine of the angle between a PIV vector and its neighboring moving particles. The alignment score $A^R(t)$ has the following form:

$$A^R(t) = \frac{1}{M} \sum_{j=1}^M \frac{1}{N_j^R(t)} \sum_{i \text{ s.t. } |\vec{x}_j - \vec{x}_i| \leq R} \frac{\vec{v}_j^t \cdot \vec{w}_i^t}{|\vec{v}_j^t| |\vec{w}_i^t|}, \quad (5)$$

where \vec{v}_j^t and \vec{w}_i^t denote the j th PIV vector and the velocity vector of particle i , respectively, at time t , and ' \cdot ' represents the dot product between the two vectors. The second sum in Eq. (5) is taken over all particles i located within a distance R from the initial point of the PIV vector j at time t . The value of A^R lies between -1 and 1 , where 1 means that all the particles within the circle are perfectly aligned to the PIV vectors, 0 means that the particles have no tendency to be aligned with the PIV vectors, and -1 means that the particles move in the exact opposite direction of PIV vectors (see Fig. 2 for a visual illustration). Whenever the value of A^R is greater than zero, it means that PIV vectors have some alignment on average with the particle motion.

Results

Since, in simulating the VM, we know the positions and velocities of individual particles at each time instance, we can use these trajectories to test the performance of PIV in capturing the single particle's motion. We mimic images of the particles by plotting markers representing agents at each particle location which are oriented in the direction of particle motion. The PIV technique is applied to the images, and we can compare the velocities computed by PIV to the velocities known from the trajectories using the alignment score (defined in Eq. 5).

Figure 3 represents some snapshots of moving particles (black arrows with filled ovals) along with PIV vectors (red arrows) for noise $\eta_0 = \frac{\pi}{6}$ at time (a) $t = 2$, and (b) $t = 150$. In the VM, the initial direction and position of particles are chosen randomly. Consequently, $t = 2$ is too early to observe a global feature of moving particles. Since the particles within two nearby PIV grids are not aligned with one another, the PIV vectors associated with nearby grids are also not aligned with one another. Hence, at time $t = 2$, PIV vectors do not exhibit global coherence (Fig. 3a). When looking closer at the PIV vectors and their nearby particles (insets of Fig. 3a), some coherence in the orientations of particles is observed at such smaller length-scales, and the PIV vectors appear to be aligned with their nearby particles. As time passes, the coherent movement of particles begins to be visible (Fig. 3b). At this time, small groups of particles are observed which are directed in the same direction. Hence, the direction of PIV vectors placed within a cluster of particles will be the same as the direction in which the particles are heading. As more time passes, all the particles in the system start to move in the same direction, and the whole system acts as a single cluster that produces the PIV vectors directed in the same direction at this noise level. Note that in Fig. 3b, even though there are the same amount of particles in each image, there can be some vacant regions. The PIV vector is not computed and not drawn at a given time t if there is no any particle in a grid at either the time t or the consecutive time $(t + 1)$.

How precisely can the PIV vectors dictate the movement of neighboring particles? Figure 4 shows the alignment score A^R , defined in "Alignment score", as a function of time t at noise level (a) $\eta_0 = \frac{\pi}{6}$, (b) $\eta_0 = \frac{3\pi}{6}$, (c) $\eta_0 = \pi$, and (d) $\eta_0 = \frac{11\pi}{6}$. It appears that at low noise ($\eta_0 = \frac{\pi}{6}$), A^R quickly converges to its maximum value 1 for small R values, and the A^R value decreases as the radius R rises at early times (Fig. 4a). The global coherence of particles takes some time to form since, according to the VM's interaction rule, particles cannot interact globally straight away as shown in the v_a vs time t figure (see Supplementary Figure 1). This delay in the coordinated movement of agents is a very common phenomenon in nature as information cascade among group members

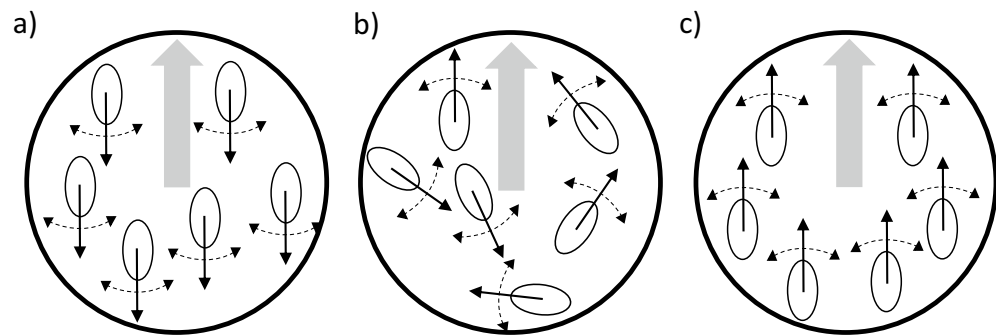


Figure 2. Schematic representation of groups of particles with varying alignment score A^R where the thick-shaded arrow represents a PIV vector defined over a grid, and the ovals with arrows stand for the moving particles along with their directions of movement. (a) The direction of the PIV vector is the opposite of the direction in which the particles move. Therefore, A^R will be close to -1 where R is the radius of the circle. (b) Particles move randomly hence A^R will be close to 0 for such cases. (c) The direction of the PIV vector is aligned to the direction in which the particles move, hence A^R will be close to unity.

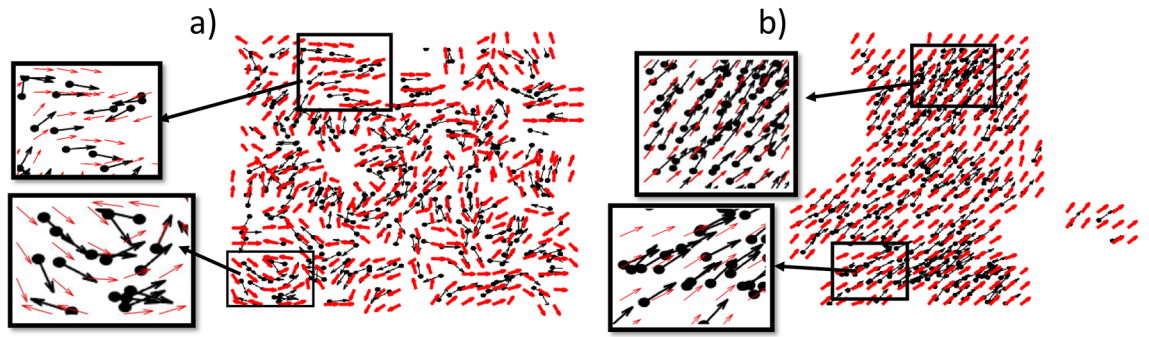


Figure 3. Moving particles (filled ovals) along with PIV vectors (red arrows) at time (a) $t = 2$, and (b) $t = 150$ for noise $\eta_0 = \frac{\pi}{6}$. Black arrows indicate the headings of respective particles. (a) Initially, particles move randomly, producing chaotically directed PIV vectors. (b) Particles (as well as PIV vectors) are getting aligned in a specific direction. Note that, if there is no any particle in a grid at either time t or the subsequent time $(t + 1)$, the PIV vector is not computed at time t , and hence, no arrow is drawn.

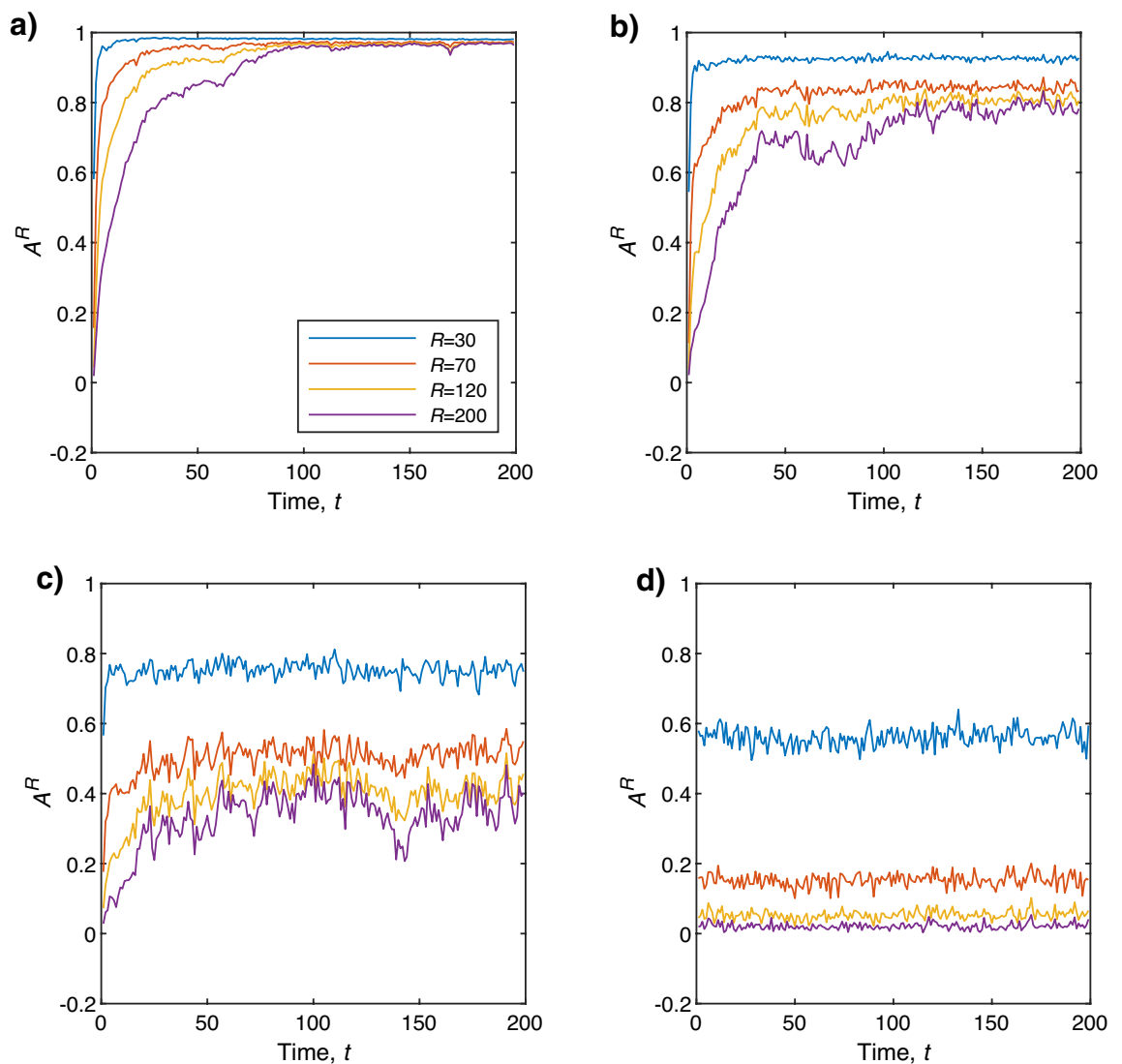


Figure 4. Alignment score A^R for different radius R values as a function of time t at noise (a) $\eta_0 = \frac{\pi}{6}$, (b) $\eta_0 = \frac{3\pi}{6}$, (c) $\eta_0 = \pi$, and (d) $\eta_0 = \frac{11\pi}{6}$. The value of N for this calculation is set to 300 and the image size is 615×615 pixels. Here the blue, red, yellow, and purple colored lines correspond to $R = 30$, $R = 70$, $R = 120$, and $R = 200$ pixels, respectively.

requires some time. However, the local coherent motion of the particles can be seen as illustrated in Fig. 5a. If the values of R are set to less than the size of the local collectives (black circles in Fig. 5a), then A^R will be very high (~ 1). Conversely, if the values of R are set to higher than the size of the local collectives (red circle in Fig. 5a), the value of A^R decreases as the particles oriented in different directions are considered. Since PIV measures the average displacement of particles located within a grid, hence average cosine of the angle between a PIV vector and its neighboring particles decreases as more randomly moving particles are considered. On the contrary, if all the particles in the system move in the same direction ($v_a \sim 1$ (Supplementary Figure 1)), then the value of A^R will always be ~ 1 regardless of the values of R (Fig. 5b) as the direction of all PIV vectors will also be the same as the particles. Therefore, once the system reaches global coherence, A^R does not depend on R to such a considerable level. As the noise η_0 increases the value of A^R decreases (Fig. 4b–d). The transition from order to disorder is found to be at a noise level of around π (see Supplementary Figure 2). In the original paper by Vicsek³², it was reported that $v_a \approx 0.5$ at noise $\eta_0 = 3$ for $N = 100$, and η_0 is slightly less than 3 for $N = 400$. Our model has slightly different parameter values such as the box size L , interaction domain R , N , and we consider our finding to be consistent with Ref.³². At this noise level, it was found that the alignment score A^R fluctuates over time (see Supplementary Figure 3c), which indicates that transient coherent and incoherent motions are intermingled. At very high noise ($\eta_0 = \frac{11\pi}{6}$) A^R is close to zero as the noise dominates the system and particles move randomly (Fig. 4d). It has been found that even for very high noise ($\eta_0 = \frac{11\pi}{6}$), A^R is relatively high (~ 0.5) for $R = 30$. Thus at high levels of noise, it is important to choose R such that the number of particles within the radius R is not too large as well as the size of grids to evaluate PIV vectors in order to capture and the underlying particle's motility and properly characterize the accuracy of PIV using the quantity A^R .

How to determine the optimal R and the grid size in the PIV estimation at high noise cases? PIV measures the average displacement of particles located within a grid of size $\gamma \times \gamma$ ($\gamma \in \mathbb{R}$). Hence, PIV cannot precisely measure the velocities when there are several particles in a PIV grid because, in the case of very high noise, particles move randomly. In principle, if a PIV grid contains only a single particle, then PIV should be able to characterize the motion of that particle even for very high noise cases. Let $\gamma_0(N)$ represents the typical size of the PIV grid for a given number of particles N , which contains only a single particle on average, which has the following form:

$$\gamma_0(N) = \sqrt{\frac{\Gamma \times \Gamma}{N}},$$

where Γ represents the length of the image data. Likewise, under the assumption of uniform distribution, $R_0(N) = \gamma_0(N)/\sqrt{\pi}$ is the radius size of the circle centered at each pixel for a given number of particles N , having only a single particle on average. Figure 6 represents the A^R landscape as a function of normalized PIV grid size, $\frac{\gamma}{\gamma_0(N)}$, and normalized radius, $\frac{R}{R_0(N)}$ for (a) $N = 100$ and (b) $N = 300$ with $\eta_0 = \frac{11\pi}{6}$ (high noise case). The A^R values get closer to unity when $\frac{\gamma}{\gamma_0(N)}$ and $\frac{R}{R_0(N)}$ are proximate to unity, which results from the fact that, on average, the PIV grid of size γ contains a single particle, and A^R estimation is also carried out with a small circle within which only a single particle exists on average. Consequently, the PIV vector captures the motion of the particle lying inside that grid appropriately. On the other hand, a $\frac{\gamma}{\gamma_0(N)}$ value greater than 1 means that,

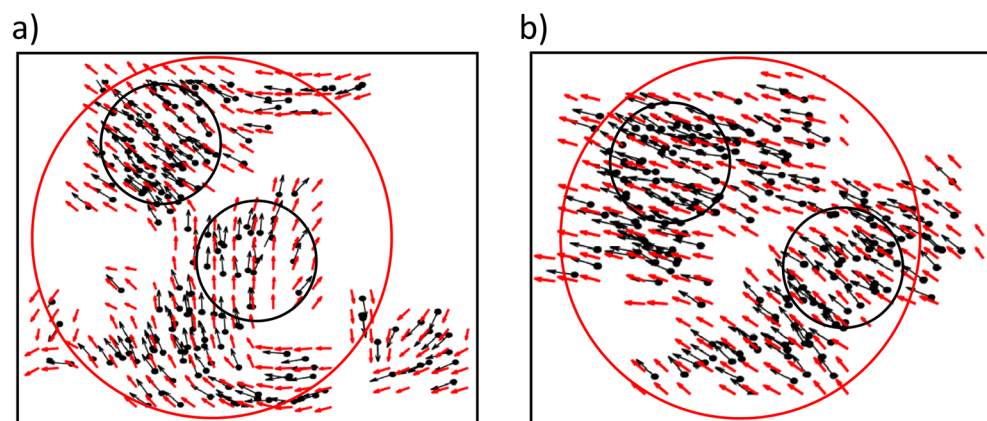


Figure 5. Schematic diagram of R dependency of A^R . Two clusters of particles (inside black circles) are shown, and within each cluster, all the particles are heading in the same direction. (a) The directions of particles within clusters are different. (b) The directions of particles within both clusters are the same. Red-colored arrows represent the PIV vectors associated with each cluster. If A^R is computed for the black circles, then in both cases A^R would be very high as within each black circle the direction of the PIV vector and particles are the same. But if A^R is computed for the red circle then for the 1st case A^R will be much lower as the particles are directed in different directions even though for the 2nd case A^R would be close to 1. Note that, if there is no any particle in a grid at either time t or the subsequent time $(t + 1)$, the PIV vector is not computed at time t , resulting in no arrow.

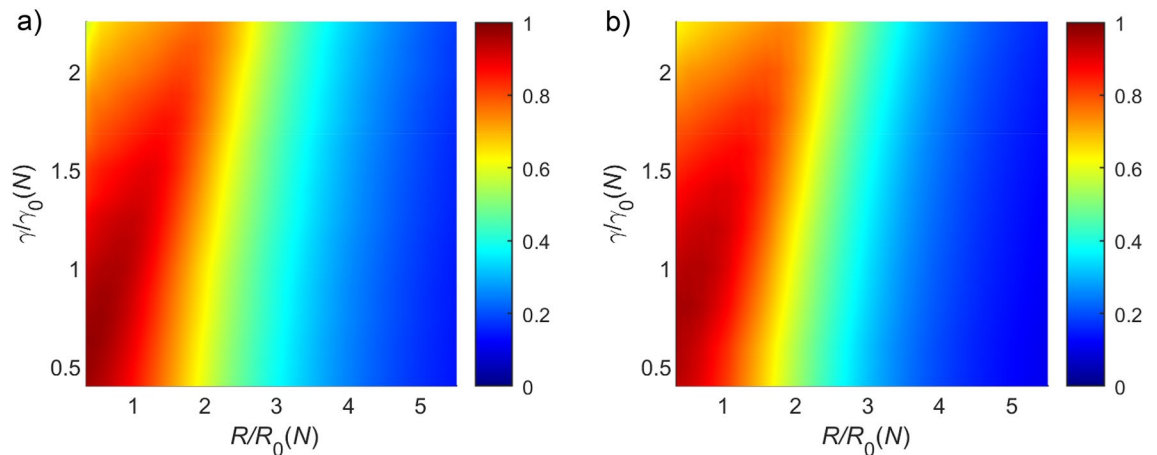


Figure 6. A^R landscape at $\eta_0 = \frac{11\pi}{6}$ (high noise) as a function of normalized PIV grid size, $\frac{\gamma}{\gamma_0(N)}$ and normalized radius, $\frac{R}{R_0(N)}$ for (a) $N = 100$ (corresponding to $\gamma_0 = 61.5$ pixels and $R_0 = 34.7$ pixels), (b) $N = 300$ ($\gamma_0 = 35.5$, $R_0 = 20$). It is apparent that A^R values of close to 1 are obtained when both $\frac{\gamma}{\gamma_0(N)}$ and $\frac{R}{R_0(N)}$ are close to 1. When both $\frac{\gamma}{\gamma_0(N)}$ and $\frac{R}{R_0(N)}$ exceed 1, the value of A^R begins to decrease. A similar trend in A^R values has been found for different N values (see also Supplementary Figure 4 for $N = 200, 500, 700$, and 1000). The value of A^R decreases along the direction normal to the slope $\frac{d\gamma}{dR} \sim 2$ approximately in this normalized plane, which implies that the radius R becoming two times larger is equivalent to an increase in γ four times. This is due to the geometrical reason that one circle of radius R is involved in the grid of $\gamma \times \gamma$. The intervals along the horizontal and vertical axes are not the same for every N since the values of γ_0 and R_0 vary with N with a series of different, discrete γ values. The figures were plotted using the MATLAB ‘shading interp’ function to interpolate the data points or grids for different number of grids.

on average, a PIV grid of size γ may contain multiple, randomly-moving particles, hence the A^R value decreases (for all R values) as the value of $\frac{\gamma}{\gamma_0(N)}$ increases over 1, at such high noise levels where the particles have no collectives. Likewise, an $\frac{R}{R_0(N)}$ value of greater than 1 may signify that several randomly-moving particles are being taken into account, which results in low A^R values. A universal trend in A^R values has been found for different N values (see Supplementary Figure 4).

To characterize the local coherent movement of particles we defined $v_a^R(t)$ in Eq. (4), where the headings of particles located within a distance of R from the initial point of the PIV vector j were considered. The same quantity can also be used to check the alignment of PIV vectors located within a circle of radius R . Now we have two different $v_a^R(t) \in [0, 1]$, one is for the moving particles (self-propelled particles (SPPs)), we denote it by $v_a^{R,SPP}(t)$, and the other one is for the PIV vectors which is denoted by $v_a^{R,PIV}(t)$. Hence the difference defined by

$$D^R(t) = |v_a^{R,SPP}(t) - v_a^{R,PIV}(t)|$$

quantifies the difference in coherent movement of SPPs and alignment of PIV vectors located within a distance of R . D^R can take a value between 0 and 1 and, for example, $D^R \sim 0$ means that both the PIV vectors and nearby particles are directed in the same direction. Figure 7 shows D^R as a function of time t for different R/R_0 values at noise (a) $\eta_0 = \frac{\pi}{6}$, (b) $\eta_0 = \frac{3\pi}{6}$, (c) $\eta_0 = \pi$, and (d) $\eta_0 = \frac{11\pi}{6}$. Here the value of N is set to 300, and the grid size γ corresponds to $\gamma = 0.986\gamma_0$. Regardless of normalized radius R/R_0 values, D^R rises as the noise level increases (Fig. 7). Here, $R/R_0 \simeq 1$ denotes that there is typically just one particle present within the circle (on average under uniform distribution), and more noticeable is that, the value of D^R increases for $R/R_0 = 1$ faster than for other R/R_0 values as the noise level η_0 gets larger. Here, the PIV grid size γ was set to 64 pixels, which means the distance between two neighboring PIV vectors is 64 pixels and $R/R_0 = 1$ corresponds to 35.5 pixels in the current settings. Hence the circle of radius $R \leq R_0$ pixels centered at the initial point of a PIV vector contains only one PIV vector that produces $v_a^{R,PIV} = 1$ by definition. However, that circle may contain more than one SPP. Here $v_a^{R,SPP} \approx 1$ for low noise and relatively longer time as the particles within the circle quickly get aligned, resulting in a small D^R over the long duration (Fig. 7a). As the noise increases, $v_a^{R,PIV}$ remains the same (i.e., 1) for $R = R_0$, but the particles inside the circle begin to disintegrate which lowers the values of $v_a^{R,SPP}$. As a result, D^R increases as the noise increases for $R = R_0$ (Fig. 7b–d). But if the value of R is set larger than R_0 (larger than the PIV grid size γ), the circle contains multiple PIV vectors. Consequently, the value of $v_a^{R,PIV}$ decreases as the noise increases. Hence, D^R also decreases for $R/R_0 > 1$, because of the difference between the small numbers. Similar results have been found for different grid sizes and normalized radii.

The landscape of A^R as a function of noise η_0 and normalized radius R/R_0 is shown in Fig. 8a. The alignment score A^R was computed as the average taken over 30 trials (initial conditions) and the time t was set to 300, which

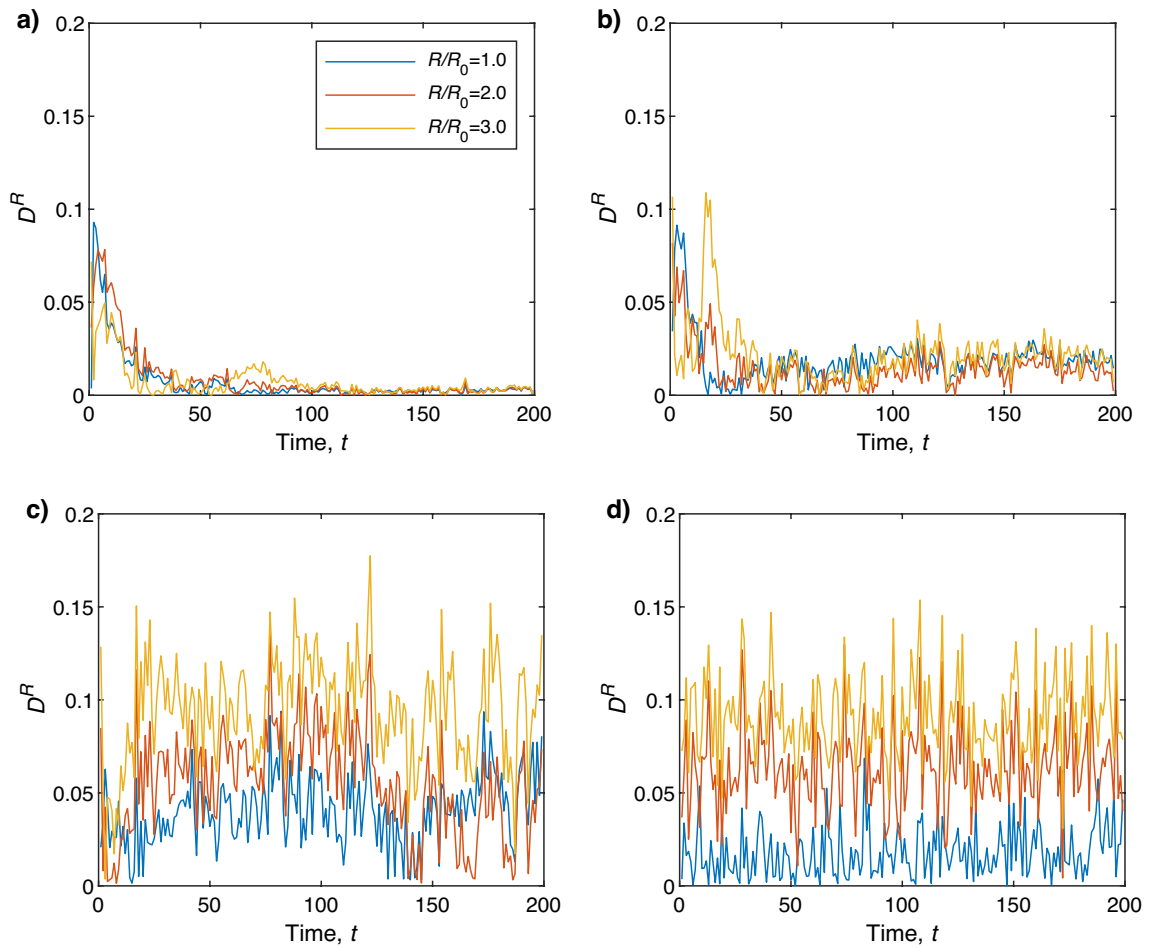


Figure 7. D^R as a function of time t for different R values at noise (a) $\eta_0 = \frac{\pi}{6}$, (b) $\eta_0 = \frac{3\pi}{6}$, (c) $\eta_0 = \pi$, and (d) $\eta_0 = \frac{11\pi}{6}$. For this calculation, the grid size γ and the value of N are set to 64 pixels and 300, respectively. Here the blue, red, yellow, and purple colored lines correspond to $R/R_0 = 1$, $R/R_0 = 2$, and $R/R_0 = 3$, respectively.

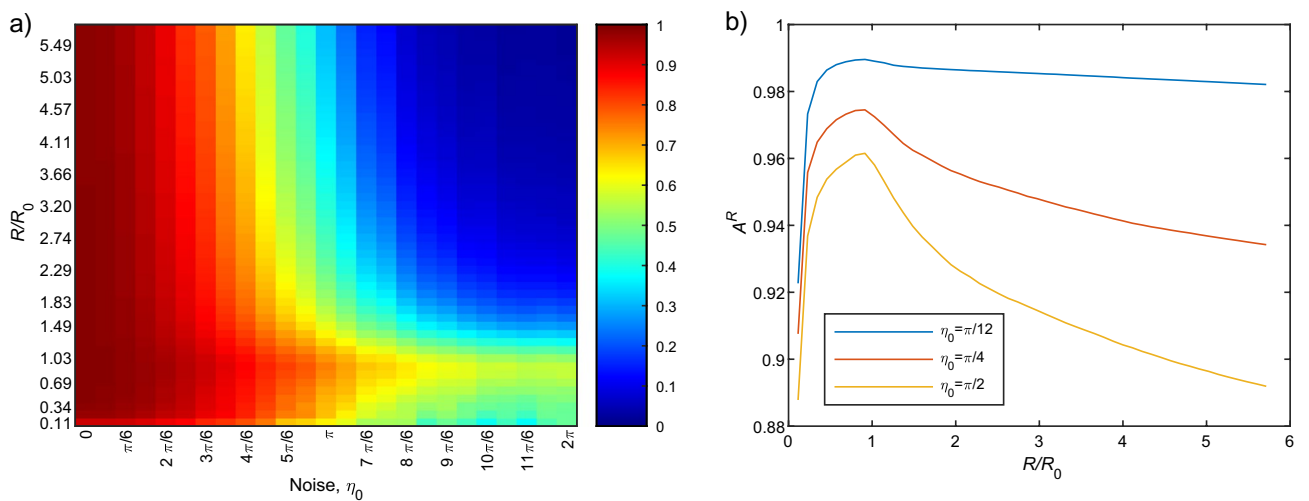


Figure 8. (a) Alignment score A^R landscape as a function of normalized radius R/R_0 and noise η_0 for $N = 300$ and $\gamma = 64 (= 0.986\gamma_0)$ pixels. The A^R can take a value between -1 and 1 in principle, but the smallest observed value was 0.0205 so that we plot A^R landscape for $A^R \in [0 : 1]$. (b) A^R as a function of normalized radius R/R_0 for low and moderate noise levels.

means PIV (for A^R computation) was applied on images at time $t = 299$ and $t = 300$ only. For all noise levels, it can be seen that v_a stabilizes within this time frame (see Supplementary Figure 5). It was found that at the low level of noise ($\eta_0 \leq \pi/2$) A^R is very high (≈ 1) (see Fig. 8a). At low noise levels, particles tune very quickly in the same direction. Consequently, the value of A^R does not rely greatly on radius R . As the noise increases, particles move more randomly instead of coherently. Consequently, the value of A^R drops as the noise increases. On the other hand, a smaller radius means only the close particles, close to the PIV vector, are taken into consideration. Hence, even at relatively higher noise, a relatively high A^R was found for $R/R_0 \lesssim 1.0$. As the radius increases, distant particles begin to be incorporated that resulting in low A^R values (≈ 0). It may be noted worthy that a peak in A^R values near $R/R_0 \approx 1$ can be observed, which is pronounced for the noise interval $\pi \leq \eta_0 \leq 2\pi$. The reason of this peak near $R/R_0 \approx 1$ lies in the PIV computation. The PIV grid size γ was set to 64 pixels, which means a PIV vector quantifies the average motion of the particles located within a box of size 64×64 . The maximum radius of the circle that fits into the box is ≈ 36 pixels. Some particles that the PIV vector depends on are ignored if the value of R is set lower than 36 pixels. In contrast, if we increase R over 36 pixels, we start to take into account those particles from which the PIV vector is independent. A^R would therefore reach its maximum at $R \approx 36$ pixels. It should be noted that the value of R_0 , the typical radius size of a circle that contains only a single particle on average, is 35.5 pixels. Hence, A^R reaches to its maximum at $R/R_0 \approx 1$. Figure 8b demonstrates that even in low and moderate noise cases, this characteristic of A^R remains unchanged as one can find a peak near $R/R_0 \approx 1$.

In principle, our comparison of PIV performance at different noise levels showed that PIV has appropriate length scales to detect the underlying agents' motility dependent on the noise levels and works best whenever single particles are isolated from one another on average in the alignment grid. This should also be relevant to what timescale the series of the images are acquired. For example, an overdamped Langevin equation is represented by its difference form as $x(t + \Delta t) = x(t) + \frac{1}{\gamma} \{F(x) + \xi(t)\} \Delta t$ where x , $F(x)$, γ , and $\xi(t)$ denote the coordinate of the system, the mean force, friction coefficient, random force, respectively. That is, the difference form of the overdamped Langevin equation tells us that random force exerted on the coordinate $x(t)$ is regarded as being proportional to time increment Δt . The time increment of Vicsek model Δt and the recording time step were both unity in our simulation, and high noise in Vicsek model (e.g. $\eta_0 = 11\pi/6$) corresponded to "high temperature" at which the movements of particles are subject to significant changes in their directions within the time unit Δt . What we showed is that, especially in such case, the desired grid size corresponds to the situation that only one agent exists on average within each grid. In turn, low or moderate noises in Vicsek model corresponds to "low temperatures". In such cases, within the time unit Δt the movements of particles are less subject to large changes in their directions. Low or moderate noises in Vicsek model might mimic "a shorter timescale" than the timescale when a system finds the substantial changes in their configuration. Supplementary Figure 6 shows the effect of sampling rate of the images on the time averaged alignment score $\langle A^R \rangle$ at intermediate noise $\eta_0 = \pi/2$ in which the recording time step τ is varied while the time increment of the Vicsek model Δt is unity. The figure demonstrates that, as the recording time step τ increases, the PIV analysis misses the details about the flow field, leading to low $\langle A^R \rangle$ values. It may be noted worthy that the decreasing rate of the average A^R with respect to τ tends to get more pronounced when the radius R is deviated from R_0 .

Conclusion

In this study, we have scrutinized the use of PIV in assessing the motion of agents at a variety of noise levels. As is known for the VM, the system undergoes a transition from ordered to disordered phase as the increase of the noise value. As expected, the PIV vector aligns well with the motion of agents when the noise level is low, resulting in global alignment. When the noise level becomes higher, different behavior in PIV vectors is observed. PIV can generally pick up the local agent motion, and alignment scores are high at low radius even when the noise level is high. At higher radii, however, the alignment score drops because a PIV vector cannot characterize the motion of agents that are far away at high noise levels. We have also computed the alignment score using the optical flow (OF) constraint scheme²³, which measures the motion of individual pixels, at different noise levels. It has been observed that the alignment scores using OF data are almost identical to that using PIV data (see Supplementary Figure 7). Theoretically, PIV can be considered as the coarse-grained scheme of OF. Hence one may use OF to get deep-rooted information of the flow field. Nevertheless, one should also keep in mind that computationally OF, which computes displacement vector pixel-wise, is not as efficient as PIV. Hence depending on individual research perspectives, one may choose a suitable scheme.

An interesting future area of study would be to characterize cell motion and its relation to PIV at a more coarse-grained level in these parameter ranges. This is one of the forthcoming subjects to be addressed along this line. Another important subject is concerning how to design and choose the appropriate grid size to compute the underlying agents dynamics in real applications under the constraint one cannot access the underlying ground truth motility of the agents. As Fig. 6 shows, as the grid size γ decreases and approaches to γ_0 , it was shown that the PIV vectors can capture the underlying agents dynamics irrespective of the noise level. In practice, while one may not be able to access the value of γ_0 , we expect that one can choose the appropriate grid size by checking the convergence of resultant PIV vectors in real applications. In addition, the forthcoming subjects to be solved for real applications such as fluorescence cell images are, for example, cell division, mortality, and also optical artifacts like out-of-focus effects, noise during image acquisition, which affect the estimation of optimal grid size to retrieve cell motility. These are some of the imminent issues to be addressed in the future.

Methods

Vicsek model. A flocking model was proposed by Vicsek et al.³² to study the phase transition of a group of collectively moving particles. The model consists of a two-dimensional square box of length L , and N self-propelled particles move within the box with periodic boundary conditions³⁴. For simplicity, it was assumed that all the particles move with the same constant speed v_0 and can move in any direction. The initial (at time $t = 0$) positions of particles are distributed randomly in the square box, and the orientations are also chosen randomly within $[0, 2\pi]$. The position of the i -th particle at time $t + 1$ is denoted by $\vec{r}_i^{t+1} = (x_i(t + 1), y_i(t + 1))$ and is given by the following discrete-time stochastic system:

$$\vec{r}_i^{t+1} = \vec{r}_i^t + \vec{v}_i^t \Delta t,$$

where $\vec{v}_i^t = (v_0 \cos \theta_i(t), v_0 \sin \theta_i(t))$ is the velocity of the particle i at time t , and $\theta_i(t)$ be its orientation. At each time step the orientation angle of particle i is computed as the average of the orientations of the particles which are located within a circle of radius r centered at \vec{r}_i^t including the particle i itself with some random perturbation added. Hence the direction update rule of the particle i is as follows:

$$\theta_i(t + 1) = \langle \theta_i(t) \rangle_r + \Delta\theta.$$

Here $\Delta\theta$ is the random perturbation of the system distributed uniformly in the range $[-\frac{\eta_0}{2}, \frac{\eta_0}{2}]$, where η_0 is considered as a temperature-like parameter. Suppose $\mathcal{N}_i(t)$ be the set of all neighboring particles (within radius r) of i at time t including itself. Then the average direction is given by³⁵:

$$\langle \theta_i(t) \rangle_r = \arctan \left[\frac{\sum_{j \in \mathcal{N}_i(t)} \sin \theta_j(t)}{\sum_{j \in \mathcal{N}_i(t)} \cos \theta_j(t)} \right]$$

In this paper, the value of the box size L is set to be 5 arb. units (615 pixels), and the constant speed of particles v_0 being 0.03 arb. units (3.7 pixels per frame). The value of the interaction radius r is set to be 0.5 arb. units (61.5 pixels). Results shown in “Results” are for $N = 300$ besides Fig. 6. But it was found that in the range of $100 \leq N \leq 2000$, the alignment score A^R was found to be in similar order (see Supplementary Figure 8).

PIV parameters. PIV technique relies on the intensity variation of image data. In this study, gray-scale images are generated by plotting black-filled ovals at each particle location. Hence, a pixel with an intensity of 0 represents a part of a particle, whereas an intensity value of 255 represents the background. The size of a particle is set to be 8×6 pixels, where 8 indicates the length of the major axis and 6 the length of the minor axis. Since PIV estimates the mean displacement of particles, hence it is directly related to the size of the interrogation zone (grid size, γ). It was found that the alignment score A^R remains very consistent as long as the particle size is considerably less than γ (see Supplementary Figure 9). It is easily understandable that PIV is unable to detect intensity variation when a particle covers the entire PIV grid. Hence, A^R should decrease if the particle size gets much closer to the grid size, γ . For the statistical purpose, the grid size is chosen large enough so that it includes a significant number of particles^{36,37}. Hence, the grid size was set to be 64×64 pixels, unless stated otherwise. The step size which defines the interrogation/search zone was set to be 32×32 pixels (50% of the grid size). That means the size of the search zone was set to be 96×96 pixels. It was found that the grid and step size do not have much effect on the alignment score A^R for a low level of noise (not shown). Finally, the images were analyzed using the particle image velocimetry (PIV) technique with PIVlab³⁸, and the A^R computation is carried out by excluding the outlier PIV vectors using the ‘vector validation’ function.

Data availability

The data that support the findings of this study are available from the corresponding author upon request.

Code availability

The code employed to create all simulations is accessible at <https://zenodo.org/record/8174506>.

Received: 27 March 2023; Accepted: 28 July 2023

Published online: 02 August 2023

References

1. Heinrich, M. A. et al. Size-dependent patterns of cell proliferation and migration in freely-expanding epithelia. *Elife* **9**, e58945 (2020).
2. Chen, Y., Dodd, S. J., Tangrea, M. A., Emmert-Buck, M. R. & Koretsky, A. P. Measuring collective cell movement and extracellular matrix interactions using magnetic resonance imaging. *Sci. Rep.* **3**, 1–9 (2013).
3. Dormann, D. & Weijer, C. J. Imaging of cell migration. *EMBO J.* **25**, 3480–3493 (2006).
4. Chevalier, N. et al. How tissue mechanical properties affect enteric neural crest cell migration. *Sci. Rep.* **6**, 1–18 (2016).
5. Li, L., He, Y., Zhao, M. & Jiang, J. Collective cell migration: Implications for wound healing and cancer invasion. *Burns Trauma* **1**, 2321–3868 (2013).
6. Miyazaki, K. et al. Cancer cell migration on elongate protrusions of fibroblasts in collagen matrix. *Sci. Rep.* **9**, 1–15 (2019).
7. Pijuan, J. et al. In vitro cell migration, invasion, and adhesion assays: From cell imaging to data analysis. *Front. Cell Dev. Biol.* **7**, 107 (2019).
8. Horwitz, R. & Webb, D. Cell migration. *Curr. Biol.* **13**, R756–R759 (2003).
9. Yang, Y. & Levine, H. Leader-cell-driven epithelial sheet fingering. *Phys. Biol.* **17**, 046003 (2020).

10. Riahi, R. *et al.* Notch1-dll4 signalling and mechanical force regulate leader cell formation during collective cell migration. *Nat. Commun.* **6**, 1–11 (2015).
11. Sattari, S. *et al.* Modes of information flow in collective cohesion. *Sci. Adv.* **8**, eabj1720 (2022).
12. Roy, S., Howes, K., Müller, R., Butail, S. & Abaid, N. Extracting interactions between flying bat pairs using model-free methods. *Entropy* **21**, 42 (2019).
13. Ohta, Y., Furuta, T., Nagai, T. & Horikawa, K. Red fluorescent camp indicator with increased affinity and expanded dynamic range. *Sci. Rep.* **8**, 1–9 (2018).
14. Ichimura, T. *et al.* Exploring rare cellular activity in more than one million cells by a transscale scope. *Sci. Rep.* **11**, 1–16 (2021).
15. Van Valen, D. A. *et al.* Deep learning automates the quantitative analysis of individual cells in live-cell imaging experiments. *PLoS Comput. Biol.* **12**, e1005177 (2016).
16. Nguyen, T. H. *et al.* Halo-free phase contrast microscopy. *Sci. Rep.* **7**, 1–10 (2017).
17. Yücelen, E., Lazić, I. & Bosch, E. G. Phase contrast scanning transmission electron microscopy imaging of light and heavy atoms at the limit of contrast and resolution. *Sci. Rep.* **8**, 1–10 (2018).
18. Wei, Q. *et al.* Plasmonics enhanced smartphone fluorescence microscopy. *Sci. Rep.* **7**, 1–10 (2017).
19. Llopis, P. M. *et al.* Best practices and tools for reporting reproducible fluorescence microscopy methods. *Nat. Methods* **20**, 1–14 (2021).
20. Sakane, A., Yoshizawa, S., Yokota, H. & Sasaki, T. Dancing styles of collective cell migration: Image-based computational analysis of JRB/MICAL-L2. *Front. Cell Dev. Biol.* **6**, 4 (2018).
21. Nasser, L. & Boudier, T. A novel generic dictionary-based denoising method for improving noisy and densely packed nuclei segmentation in 3d time-lapse fluorescence microscopy images. *Sci. Rep.* **9**, 1–13 (2019).
22. Kobitski, A. Y. *et al.* An ensemble-averaged, cell density-based digital model of zebrafish embryo development derived from light-sheet microscopy data with single-cell resolution. *Sci. Rep.* **5**, 1–10 (2015).
23. Vig, D. K., Hamby, A. E. & Wolgemuth, C. W. On the quantification of cellular velocity fields. *Biophys. J.* **110**, 1469–1475 (2016).
24. Liang, J., Cai, S., Xu, C., Chen, T. & Chu, J. Deepptv: Particle tracking velocimetry for complex flow motion via deep neural networks. *IEEE Trans. Instrum. Meas.* **71**, 1–16 (2021).
25. Pereira, F., Stüer, H., Graff, E. C. & Gharib, M. Two-frame 3d particle tracking. *Meas. Sci. Technol.* **17**, 1680 (2006).
26. Atkins, M. Velocity field measurement using particle image velocimetry (PIV). In *Application of Thermo-Fluidic Measurement Techniques* 125–166 (Elsevier, 2016).
27. Petitjean, L. *et al.* Velocity fields in a collectively migrating epithelium. *Biophys. J.* **98**, 1790–1800 (2010).
28. Nematollahi, O., Samsam-Khayani, H., Nili-Ahmadabadi, M., Yoon, S. Y. & Kim, K. C. A novel self-seeding method for particle image velocimetry measurements of subsonic and supersonic flows. *Sci. Rep.* **10**, 1–9 (2020).
29. Reinken, H. *et al.* Organizing bacterial vortex lattices by periodic obstacle arrays. *Commun. Phys.* **3**, 1–9 (2020).
30. Lin, S.-Z., Zhang, W.-Y., Bi, D., Li, B. & Feng, X.-Q. Energetics of mesoscale cell turbulence in two-dimensional monolayers. *Commun. Phys.* **4**, 1–9 (2021).
31. Park, H., Yeom, E. & Lee, S. J. X-ray piv measurement of blood flow in deep vessels of a rat: An in vivo feasibility study. *Sci. Rep.* **6**, 1–8 (2016).
32. Vicsek, T., Czirók, A., Ben-Jacob, E., Cohen, I. & Shochet, O. Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Lett.* **75**, 1226 (1995).
33. Basak, U. S., Sattari, S., Hossain, M. M., Horikawa, K. & Komatsuzaki, T. An information-theoretic approach to infer the underlying interaction domain among elements from finite length trajectories in a noisy environment. *J. Chem. Phys.* **154**, 034901 (2021).
34. De Lellis, P. & Porfiri, M. Inferring the size of a collective of self-propelled vicsek particles from the random motion of a single unit. *Commun. Phys.* **5**, 1–6 (2022).
35. Lu, X., Zhang, C., Huang, C. & Qin, B. Research on swarm consistent performance of improved vicsek model with neighbors' degree. *Phys. A* **588**, 126567 (2022).
36. Tian, J.-D. & Qiu, H.-H. Eliminating background noise effect in micro-resolution particle image velocimetry. *Appl. Opt.* **41**, 6849–6857 (2002).
37. Lacassagne, T., Vatteville, J., Degouet, C., El Hajem, M. & Simoëns, S. Ptv measurements of oscillating grid turbulence in water and polymer solutions. *Exp. Fluids* **61**, 1–15 (2020).
38. Thielicke, W. & Stamhuis, E. Pivlab-towards user-friendly, affordable and accurate digital particle image velocimetry in matlab. *J. Open Res. Softw.* **2**, 25 (2014).

Acknowledgements

This work was supported by a Grant-in-Aid for Scientific Research on Innovative Areas ‘Singularity Biology (No. 8007)’ (Grant No. 18H05413), MEXT, the research program of ‘Five star Alliance’ in ‘NJRC Matter and Dev’ (Grant No. 20191062-01), the JSPS (Grant Nos. 25287105 and 25650044, T.K.), and the JST/CREST (Grant No. JPMJCR1662, T.K.). U.S.B is supported by the University Grants Commission (UGC) of Bangladesh research Grant (Grant no. Physical Science-72-2021). M.T. is supported by the Research Program of “Dynamic Alliance for Open Innovation Bridging Human, Environment and Materials” in “Network Joint Research Center for Materials and Devices” and a Grant-in-Aid for Scientific Research (C) (No. 22654047, No. 25610105, and No. 19K03653) from JSPS.

Author contributions

U.S.B.: data curation, formal analysis, software, visualization, writing-original draft, S.S.: conceptualization, supervision, methodology, software, validation, writing-original draft, M.M.H.: software, writing review and editing, K.H.: investigation, supervision, writing review and editing, M.T.:supervision, writing review and editing, T.K.: conceptualization, supervision, project administration, resources, funding acquisition, writing-original draft.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-39635-z>.

Correspondence and requests for materials should be addressed to S.S. or T.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023