A dynamical analysis of the alignment mechanism between two interacting cells

Vivienne Leech[1], Mohit P. Dalwadi[1,2*†], Angelika Manhart[1,3*†]

[1*]Department of Mathematics, University College London, London, UK.
[2]Mathematical Institute, University of Oxford, Oxford, UK.
[3]Department of Mathematics, University of Vienna, Vienna, Austria.

*Corresponding author(s). E-mail(s): m.dalwadi@ucl.ac.uk; angelika.manhart@univie.ac.at;
Contributing authors: vivienne.leech.16@ucl.ac.uk;
†Co-corresponding authors

Abstract

We introduce an analytical framework that can be used to understand the alignment mechanism of ellipse-shaped cells interacting via self-propulsion and overlap avoidance. To do this, we consider two interacting particles with certain symmetries imposed to make the problem analytically tractable. This results in a dynamical system we can mathematically analyse in detail. We find that there is a half-stable steady state corresponding to a cell configuration that depicts perfect alignment, and that the presence of a separatrix splits the domain into two regions. These two regions characterise the outcome of a trajectory as moving towards this state or not. Understanding the asymptotic case corresponding to a small amount of self-propulsion offers an insight into the timescales at play when a trajectory is moving towards the point of perfect alignment. We find that the two cells initially move apart to avoid overlap over a fast timescale, and then the presence of self-propulsion causes them to move towards a configuration of perfect alignment over a much slower timescale. Overall, our analysis highlights how the interaction between self-propulsion and overlap avoidance is sufficient to generate alignment.

Keywords: dynamical systems, cell alignment, self-propulsion, overlap avoidance

1 Introduction

Alignment of particles in a system is a phenomenon that can be observed in many contexts. In biology, this ranges from the large scale e.g. schools of fish (Herbert-Read et al., 2011), down to the microscopic scale e.g. cells and bacteria (Balagam and Igoshin, 2015; Dartsch and Betz, 1989; Kenny et al., 2023). This alignment of particles, especially when it occurs collectively, can play key roles in their migration e.g. hydrodynamic benefits as a result of collective alignment can aid migration in schools of fish (Lopez et al., 2012) and alignment of fibroblasts can affect key mechanical properties of the tissue in which they are found (Erdogan et al., 2017).

Many alignment models (Peruani et al., 2006; Baskaran and Marchetti, 2008; Kraikivski et al., 2006) include self-propulsion and some kind of overlap avoidance or volume exclusion as model ingredients, suggesting that these are key components needed for alignment between particles to occur. This can be justified intuitively, since cell interactions provide a way for cells to change their orientation and self-propulsion provides a way for cells to continually interact with each other; hence their interaction allows for alignment to propagate through the population.

Motivated by experiments on the alignment of fibroblasts in Kenny et al. (2023), an agent-based model was recently developed in Leech et al. (2024) to investigate the mechanism behind the collective alignment of self-propelled interacting particles. The model comprises N ellipse-shaped cells with fixed area, which move in the direction of their orientation and change their position, orientation and shape in order to avoid overlap. Through computational analysis of the model, it is found that these model components lead to some amount of collective alignment, dependent on model parameters.

The most common approach for the analysis of agent-based models tends to be computational (as was the case in Leech et al. (2024)), since there are fewer analytic tools for analysing agent-based models. For an overview of different approaches to modelling and analysing pattern formation, see e.g. Deutsch and Dormann (2005). One method is to take the mean-field limit to obtain an equivalent continuum model, where the positions and orientations of cells are translated into cell density and mean orientation across continuous space (Albi and Pareschi, 2013; Degond and Motsch, 2008; Großmann et al., 2016). This has been done for the Vicsek model (Bolley et al., 2012), but is mathematically challenging and poses challenges for discontinuous coefficients, such as those that arise in Leech et al. (2024).

Instead of considering the large-cell-number limit, in this work we look in the other direction and consider two interacting cells, with the goal of making analytic progress. Since the agent-based model in Leech et al. (2024) requires computing the points of intersection between the boundaries of overlapping ellipses, it is helpful to impose some symmetry in relative ellipse orientation and position to facilitate analysis. Specifically, this allows us to analytically determine the overlap points of the two overlapping ellipses, which allows us to write down explicit governing equations. We are able to make significant analytic progress in understanding the non-trivial dynamic interaction between two interacting cells, and how the different aspects of the model contribute to alignment between two cells.

In this work, we will mathematically analyse the model derived in Leech et al. (2024) by considering two interacting cells with some symmetry imposed. We begin by introducing the full model, and then the analytical framework which leads to a 3 dimensional dynamical system (Sec. 2). We then analyse the full 3D dynamical system which allows for changes to the aspect ratio, as well as changes in position and orientation as a result of overlap avoidance (Sec. 3). We then reduce the system to 2D by taking the limit of no shape change, which allows us to fully understand the effect of the self-propulsion speed on alignment, as well as visualise the dynamical system more easily (Sec. 4).

2 Full Model Derivation

2.1 Model summary

In Leech et al. (2024), an energy minimisation approach is used to derive a system of governing equations that describe the behaviour of self-propelled ellipse-shaped cells that want to avoid cell overlap upon interacting with one other. While full details of the derivation are provided in Leech et al. (2024), we summarize the important aspects below for convenience.

Below, we introduce governing equations that describe how cells can change their position $\mathbf{X}$, orientation α, and aspect ratio r (while keeping their area constant), in order to avoid overlap. They are non-dimensionalised with respect to a reference length $\sqrt{A/\pi}$ and reference time $A\eta/\sigma$, where A is the cell area, η is the strength of friction that a cell experiences with the substrate, and σ is the strength of overlap avoidance. This reference length and time is chosen since the area A remains constant over time (in contrast to e.g. axis lengths), and hence both of these quantities are constant over time. The governing equations are

$$\frac{\mathrm{d}\mathbf{X}_i}{\mathrm{d}t} = -\sum_{j\in\mathcal{N}_i}\sum_{k=1}^{K_{ij}}(\mathbf{P}^{ij}_{2k-1} - \mathbf{P}^{ij}_{2k})^{\perp} + \nu\mathbf{e}(\alpha_i), \tag{1a}$$

$$\frac{\mathrm{d}\alpha_i}{\mathrm{d}t} = \frac{2r_i}{r_i^2+1}\sum_{j\in\mathcal{N}_i}\sum_{k=1}^{K_{ij}}(|\mathbf{X_i} - \mathbf{P}^{ij}_{2k}|^2 - |\mathbf{X_i} - \mathbf{P}^{ij}_{2k-1}|^2), \tag{1b}$$

$$\frac{\mathrm{d}r_i}{\mathrm{d}t} = \frac{4r_i^2}{r_i^2+1}\sum_{j\in\mathcal{N}_i}\sum_{k=1}^{K_{ij}}[\sin(2\theta^{ij}_{2k-1}) - \sin(2\theta^{ij}_{2k})] + \frac{16}{\gamma}\frac{1+\bar{r}r_i^3}{1+r_i^2}\left(1-\frac{r_i}{\bar{r}}\right), \tag{1c}$$

where $\perp$ denotes the left-turned normal vector and

$$\nu = \frac{v\eta}{\sigma}\sqrt{\pi A}, \qquad \gamma = \frac{A\sigma}{\pi g}. \tag{2}$$

These non-dimensional quantities ν and γ defined in (2) depend on the self-propulsion speed v, the cell area A, the strength of overlap avoidance σ, the strength of friction with the substrate η, and the strength of relaxation to the preferred aspect ratio g. The

quantity ν can be interpreted as the ratio of the self-propulsion speed to the strength of repulsion in the presence of friction. A larger value of ν means a faster self-propulsion, or a smaller overlap avoidance strength. The quantity γ can be interpreted as the ratio between the strength of overlap avoidance and the strength of shape restoration. A larger value of γ means that the strength of relaxation to the preferred aspect ratio g is smaller and hence cell aspect ratios away from $\bar{r}$ will be punished less. Equation (1a) describes how the position of each cell changes in time, which depends on the points of overlap between cell i and cell j, $\mathbf{P}_k^{ij}$, with $k \in \{1, 2, 3, 4\}$ points of overlap. Some care is required with defining notation appropriately. In equation (1a), $K_{ij} = 1$ or $K_{ij} = 2$ denotes the number of overlap point pairs between cell i and cell j (having one or three points of overlap can be reduced to having zero or two points of overlap by neglecting the overlap point corresponding to a repeated root). $\mathcal{N}_i$ is the set of indices of all cells overlapping with cell i. The overlap points are ordered anticlockwise around the boundary of the cell. In addition, we choose $\mathbf{P}_1^{ij}$ and $\mathbf{P}_2^{ij}$ such that the boundary segment of cell i between these points is contained within the domain of cell j, see Fig. 1A, and similarly for $\mathbf{P}_3^{ij}$ and $\mathbf{P}_4^{ij}$. The direction vector of cell i is given by $\mathbf{e}(\alpha_i) = (\cos\alpha_i, \sin\alpha_i)^T$. The right-hand side of equation (1a) represents cell movement to reduce overlap, and a self-generated motion of the cell in the direction of its orientation, α_i. Equation (1b) denotes how the orientation of the cell changes over time and depends on the same points of overlap $\mathbf{P}_k^{ij}$ as in (1a). The right-hand side of (1b) represents cell rotation to avoid overlap. This is the only mechanism that affects the orientation of the cell. Equation (1c) depends on θ_k^{ij} which corresponds to the θ value that parameterises the overlap point $\mathbf{P}_k^{ij} = \mathbf{X}_i + \mathbf{R}(\alpha_i)\mathbf{k}(\theta_k^{ij})$, where

$$\mathbf{R}(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}, \quad \mathbf{k}(\theta) = \begin{pmatrix} \sqrt{r}\cos\theta \\ \frac{1}{\sqrt{r}}\sin\theta \end{pmatrix}. \tag{3}$$

The first term on the right-hand side of (1c) represents the change in aspect ratio to avoid overlap, where the cell will either shorten or lengthen to minimise overlap, depending on the overlap configuration. The term involving γ on the right-hand side of (1c) represents the shape-restoring force, restoring the cell to its preferred aspect ratio $\bar{r}$ while keeping the cell area constant. We provide a schematic of the model components in Fig. 1.

2.2 Reduction to two cells

We can reduce (1) to consider two interacting cells, which simplifies the system significantly. We initially consider two points of overlap in order to gain some intuition for the system, then we introduce the full reduction to two interacting cells, considering zero, two and four points of overlap. For two points of overlap, (1) becomes

$$\frac{\mathrm{d}\mathbf{X}}{\mathrm{d}t} = -(\mathbf{P}_1 - \mathbf{P}_2)^{\perp} + \nu\mathbf{e}(\alpha), \tag{4a}$$

$$\frac{\mathrm{d}\alpha}{\mathrm{d}t} = \frac{2r}{r^2+1}(|\mathbf{X}-\mathbf{P}_2|^2 - |\mathbf{X}-\mathbf{P}_1|^2), \tag{4b}$$

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \frac{4r^2}{r^2+1}[\sin(2\theta_1) - \sin(2\theta_2)] + \frac{16}{\gamma}\frac{1+\bar{r}r^3}{1+r^2}\left(1-\frac{r}{\bar{r}}\right). \tag{4c}$$

From equation (4a), we see that the cell with centre $\mathbf{X}$ is being pushed perpendicular to the vector connecting the points of overlap $\mathbf{P}_1$ and $\mathbf{P}_2$. We can see from (4b) that the cell is rotated by an amount proportional to the difference in the square of the lengths of the lines connecting the cell centre and the overlap points, thus turning the cell in the direction from the shorter to the longer line. We can also observe that the larger the value of $r > 1$, the slower the rate of change of orientation, due to a larger moment of inertia. The first term in (4c) leads to cells shortening if the cell overlap is near the ends of the cells and lengthening if the overlap is along the sides. This will happen at a faster rate for larger values of r. The final term on the right-hand side of (4c) acts to restore the cell to its preferred aspect ratio $\bar{r}$.

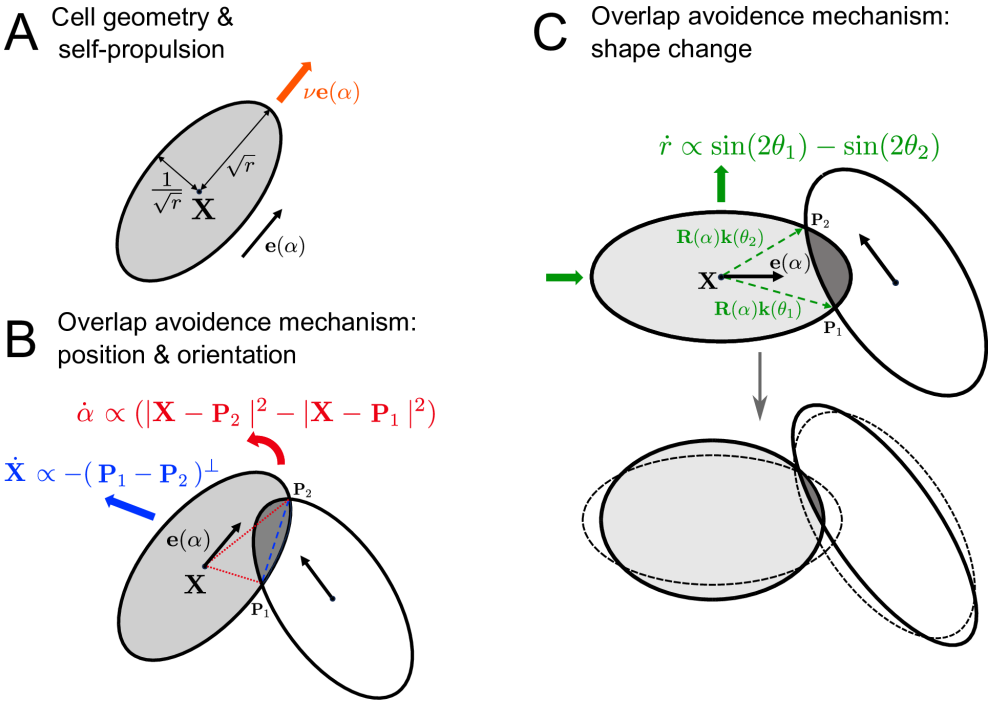

Fig. 1 A: Diagram to show non-dimensional cell geometry. B,C: Diagrams to show self-propulsion and overlap avoidance mechanisms.

2.3 Analytical framework

To further analyse and understand system (1), we consider two symmetric interacting cells. This allows us to obtain explicit expressions for the points of overlap, and therefore analytically tractable equations. We consider one cell with centre $\mathbf{X} = (x(t), y(t))$, orientation $\alpha(t)$ and aspect ratio $r(t)$ interacting with another cell with centre $(x(t), -y(t))$, orientation $-\alpha(t)$, and aspect ratio $r(t)$ as shown in Fig. 3A. Since this is a configuration that leads to perfect alignment, we note that we show cell 1 as having negative y position and $\alpha \in [0, \pi/2)$. Alternatively we could have cell 1 with positive y position, but then we would need $\alpha \in [-\pi/2, 0)$ to lead to perfect alignment. Both cells have equal self-propulsion speeds ν. Depending on the values of y, α and r (neglecting borderline cases), the two cells can have zero, two or four points of intersection. Cells 1 and 2 can be parameterised by

$$\mathbf{r}_1(t, s, \theta) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} + s\mathbf{R}(\alpha(t))\mathbf{k}(\theta), \quad s \in [0, 1], \theta \in [0, 2\pi), \tag{5a}$$

$$\mathbf{r}_2(t, s, \theta) = \begin{pmatrix} x(t) \\ -y(t) \end{pmatrix} + s\mathbf{R}(-\alpha(t))\mathbf{k}(\theta), \quad s \in [0, 1], \theta \in [0, 2\pi), \tag{5b}$$

where θ can be thought of as a modified arclength and s quantifies the normalized distance from the centre to the boundary of the ellipse. We note that there are other parameterization options; we use this parameterization for analytic convenience. We emphasize that θ is not necessarily equivalent to the angle around the ellipse β, but it relates to this via the relation $\tan\theta = r\tan\beta$.

By considering where the cell boundaries (5a) and (5b) intersect, we obtain the following points of intersection (in Cartesian coordinates)

$$\mathbf{P}_B^{\pm} = \left(-\frac{(r^2-1)y\sin\alpha\cos\alpha}{\delta_1^2} \pm \sqrt{r}\frac{\sqrt{\delta_1^2 - ry^2}}{\delta_1^2}, 0 \right), \tag{6a}$$

$$\mathbf{P}_C^{\pm} = \left(-\frac{\delta_2^2 y}{(r^2-1)\sin\alpha\cos\alpha}, \pm\sqrt{r}\frac{\sqrt{(r^2-1)^2\sin^2\alpha\cos^2\alpha - r\delta_2^2 y^2}}{\delta_2(r^2-1)\sin\alpha\cos\alpha} \right), \tag{6b}$$

where we have defined

$$\delta_1^2 = r^2\sin^2\alpha + \cos^2\alpha, \quad \delta_2^2 = \sin^2\alpha + r^2\cos^2\alpha. \tag{7}$$

Given the square roots contained within (6a) and (6b), these points of intersection only exist in certain regions of (α, y, r)-space. We denote these regions by $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ and illustrate them in Fig. 3B; in region $\mathcal{A}$ there are no points of overlap, in region $\mathcal{B}$ there are two points of overlap, $\mathbf{P}_B^{\pm}$, and in region $\mathcal{C}$ there are four points of overlap, $\mathbf{P}_B^{\pm}$ and $\mathbf{P}_C^{\pm}$. The boundaries between these regions occur when the square root terms

in (6a) and (6b) vanish, which allows us to write these boundaries using the curves

$$\Gamma_{\mathcal{AB}}(\alpha) = \frac{\delta_1(\alpha)}{\sqrt{r}}, \quad \Gamma_{\mathcal{BC}}(\alpha) = \frac{(r^2-1)\sin\alpha\cos\alpha}{\sqrt{r}\delta_2(\alpha)}. \tag{8}$$

Finally, this means that we can formally define the regions via

$$\mathcal{A} = \left\{(\alpha, y) \mid \Gamma^2_{\mathcal{AB}}(\alpha) < y^2\right\}, \tag{9a}$$
$$\mathcal{B} = \left\{(\alpha, y) \mid \Gamma^2_{\mathcal{BC}}(\alpha) < y^2 < \Gamma_{\mathcal{AB}}(\alpha)^2\right\}, \tag{9b}$$
$$\mathcal{C} = \left\{(\alpha, y) \mid y^2 < \Gamma^2_{\mathcal{BC}}(\alpha)\right\}. \tag{9c}$$

With (9) we can substitute the points of overlap in (6a) and (6b) into the two-cell versions of the full governing equations (1) to formulate the explicit two-cell governing equations for $(\alpha, y, r) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+$ with given initial conditions $(\alpha(0), y(0), r(0)) = (\alpha_0, y_0, r_0)$. Further details regarding the ordering of the points in regions $\mathcal{B}$ and $\mathcal{C}$ are provided in Appendix A.

$$\dot{x} = \nu\cos\alpha, \qquad \text{for } (\alpha, y, r) \in \mathcal{A}, \mathcal{B}, \mathcal{C} \tag{10a}$$

$$\dot{y} = \nu\sin\alpha + \begin{cases} 0, & \text{for } (\alpha, y, r) \in \mathcal{A} \\ 2\sqrt{r}\frac{\sqrt{\delta_1^2 - ry^2}}{\delta_1^2}\operatorname{sign}(y), & \text{for } (\alpha, y, r) \in \mathcal{B} \\ 2\frac{r^2}{\delta_1^2(r^2-1)|\sin\alpha\cos\alpha|}y, & \text{for } (\alpha, y, r) \in \mathcal{C} \end{cases} \tag{10b}$$

$$\dot{\alpha} = \begin{cases} 0, & \text{for } (\alpha, y, r) \in \mathcal{A} \\ -8\frac{r^2-1}{r^2+1}|y|r\sqrt{r}\sin\alpha\cos\alpha\frac{\sqrt{\delta_1^2 - ry^2}}{\delta_1^4}, & \text{for } (\alpha, y, r) \in \mathcal{B} \\ \frac{4r\operatorname{sign}(\sin\alpha\cos\alpha)}{r^2+1}\Big[\frac{(r^2-1)^2y^2\sin^2\alpha\cos^2\alpha}{\delta_1^4} & \\ +r\frac{\delta_1^2 - ry^2}{\delta_1^4} - \frac{(\delta_2^4 - r^2)y^2}{(r^2-1)^2\sin^2\alpha\cos^2\alpha} - \frac{r}{\delta_2^2}\Big], & \text{for } (\alpha, y, r) \in \mathcal{C} \end{cases} \tag{10c}$$

$$\dot{r} = -\frac{1}{\gamma}\frac{16(1+\bar{r}r^3)}{\bar{r}(r^2+1)}(r-\bar{r}) + \begin{cases} 0, & \text{for } (\alpha, y, r) \in \mathcal{A} \\ \dfrac{16r^2}{r^2+1}|y|\sqrt{r}(\cos^2\alpha \\ -r^2\sin^2\alpha)\dfrac{\sqrt{\delta_1^2 - ry^2}}{\delta_1^4}, & \text{for } (\alpha, y, r) \in \mathcal{B} \\ 4r\sin\alpha\cos\alpha\dot{\alpha} + \dfrac{8r^2y}{1+r^2}(\dot{y} \\ -\nu\sin\alpha)(\sin^2\alpha - \cos^2\alpha), & \text{for } (\alpha, y, r) \in \mathcal{C} \end{cases} \tag{10d}$$

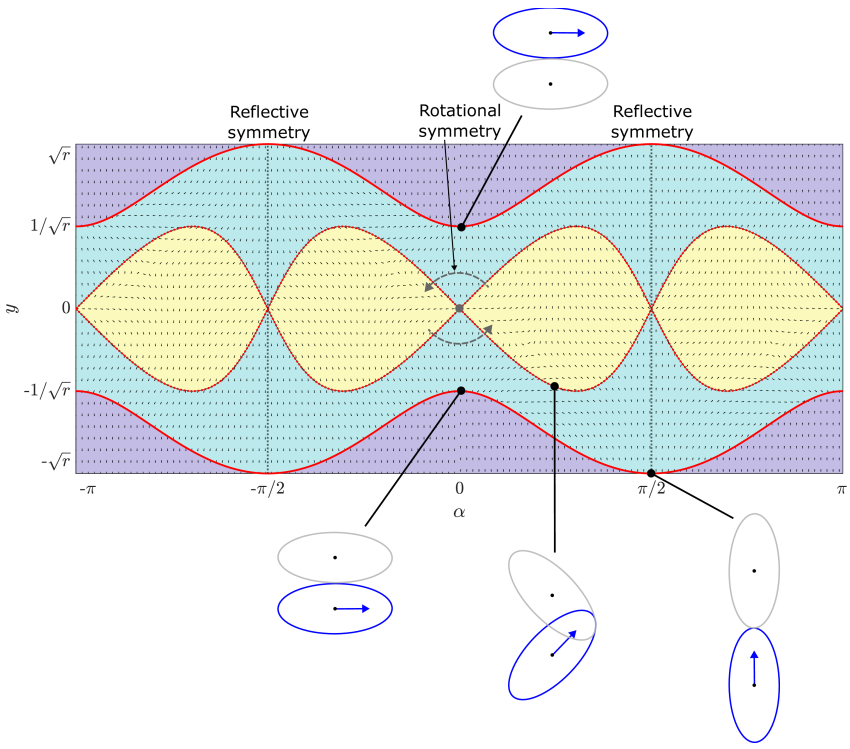

Fig. 2 Diagram to show symmetries in the (α, y) plane for $r = 2, \nu = 2$. We also see cell diagrams corresponding to points labelled on the phase portrait.

We start by discussing the general behaviour of the system (10) to gain initial insight, before presenting several analytical results in Sections 3 and 4. We note that the system is invariant to the transformation $(\alpha, y) \to (-\alpha, -y)$, which means that there is rotational symmetry about the line $\alpha = y = 0$. We also note that the system has reflectional symmetry about $\alpha = \pi/2$ and $\alpha = -\pi/2$. This can be easily visualised

in the $\alpha - y$ plane as seen in Fig. 2. We can consequently restrict our subsequent analysis to the region $\alpha \in [0, \pi/2]$, noting that all the results can be applied analogously to the ranges $\alpha \in [-\pi, 0], \alpha \in [\pi/2, \pi]$ via symmetry arguments. It is also helpful for interpretation to relate the dynamical system to the physical configuration of the cells. We indicate in Fig. 2 points on the phase portrait with corresponding cell configurations, including those at $\alpha = 0$ that represent perfect alignment.

With a better understanding of the phase space, and how this corresponds to cell configuration, we now discuss the equations in (10). The first terms in (10a) and (10b) represent self-propulsion that leads to movement in the direction of $(\cos\alpha, \sin\alpha)$. This is proportional to the non-dimensional parameter ν that has been defined in (2). Inspecting (10b) further, we note that if $r > 1$ and cell 1 is lower than cell 2 ($y < 0$), then the second term in (10b) is negative and cell 1 will move downwards, and vice versa for $y > 0$. This is a result of overlap avoidance pushing the cells apart. Equation (10c) describes how the cell orientation changes over time. If we consider $r > 1, \alpha \in (0, \pi/2)$, we see that $\dot{\alpha} < 0$ in region $\mathcal{B}$. This means that overlap avoidance in region $\mathcal{B}$ causes a clockwise rotation and drives the system towards $\alpha = 0$, i.e. towards velocity alignment. In region $\mathcal{C}$, both signs of $\dot{\alpha}$ are possible. The first term of (10d) describes how the cells will relax back to their natural aspect ratio $\bar{r}$, with $\dot{r} > 0$ if $r < \bar{r}$ and $\dot{r} < 0$ if $r > \bar{r}$. We see that in regions $\mathcal{B}$ and $\mathcal{C}$, overlap avoidance causes the aspect ratio r to change. If α is near 0, or if $\cos^2\alpha - r^2\sin^2\alpha > 0$, then $\dot{r} > 0$ and the cells will elongate to avoid overlap. This is due to the cells being side by side in this configuration and hence wanting to elongate to avoid overlap. If α is near $\pi/2$ and the cells are head to head, or if $\cos^2\alpha - r^2\sin^2\alpha < 0$, then $\dot{r} < 0$ and the cells will shorten to avoid overlap. The behaviour in Region $\mathcal{C}$ is more complex in general, aside from noting that both signs of $\dot{r}$ are possible depending on the configuration of the two cells. Finally, we note that equation (10a) is decoupled from (10b), (10c) and (10d), hence it suffices to analyse the (α, y, r)-system for $\alpha \in [0, 2\pi], y \in \mathbb{R}, r > 0$.

3 Full model with shape change: $\gamma \neq 0$

We first explore the full shape-change model (10) where $\gamma \neq 0$ and $\nu = O(1)$. This involves explicitly accounting for the restoration time of the aspect ratio r to its preferred value $\bar{r}$, where γ can be thought of as the timescale of shape restoration. The full model is a 3D dynamical system where y, α and r vary in time in response to cell overlap, self-propulsion and shape restoration.

The system (10) has steady states at the points $(\alpha, y, r) = (0, \pm 1/\sqrt{\bar{r}}, \bar{r})$. This corresponds to the cells having their preferred aspect ratio $\bar{r}$ and being positioned side-by-side, with their orientations parallel and their boundaries just touching, illustrated in Fig. 2. We note that there are additional steady states, but only consider the two listed above in the subsequent analysis since these correspond to cell alignment, which is the focus of this work. To determine the stability of these steady states, we perturb the system around the points $(\alpha, y, r) = (0, \pm 1/\sqrt{\bar{r}}, \bar{r})$. Importantly, a standard linear stability analysis would not be sufficient to determine their stability. In fact, such an analysis would be degenerate, and determining the stability of these states is non-trivial, as we will see below.

3.1 Stability Analysis

We consider the steady point $(\alpha, y, r) = (0, -1/\sqrt{\bar{r}}, \bar{r})$, noting that the analysis of the steady point $(\alpha, y, r) = (0, 1/\sqrt{\bar{r}}, \bar{r})$ will follow via the rotational symmetry of the system. In order to perform the stability analysis, we perturb around the steady point $(\alpha, y, r) = (0, -1/\sqrt{\bar{r}}, \bar{r})$ by a small amount (to be defined below) and calculate the subsequent dynamics of the system. This steady point is degenerate, so the scalings of our perturbation and its subsequent dynamics are non-standard.

To capture the richest dynamics, we consider the distinguished asymptotic limit in which as many mechanisms as possible balance at the same time. If we define the (small) perturbation in α to be of $O(\varepsilon)$, where $\varepsilon \ll 1$, then with the benefit of hindsight and justified *a posteriori*, distinguished asymptotic limits occur when $\gamma = O(\varepsilon)$ and when $\gamma = O(1/\varepsilon)$. The former is the physically relevant case, since $\gamma = O(1/\varepsilon)$ would allow for larger deformations to the aspect ratio r, which are not observed biologically. We henceforth focus on the distinguished limit $\gamma = O(\varepsilon)$ and therefore scale $\gamma = \varepsilon\Gamma$, where $\Gamma = O(1)$. In this case, the appropriate asymptotic scalings for the perturbations

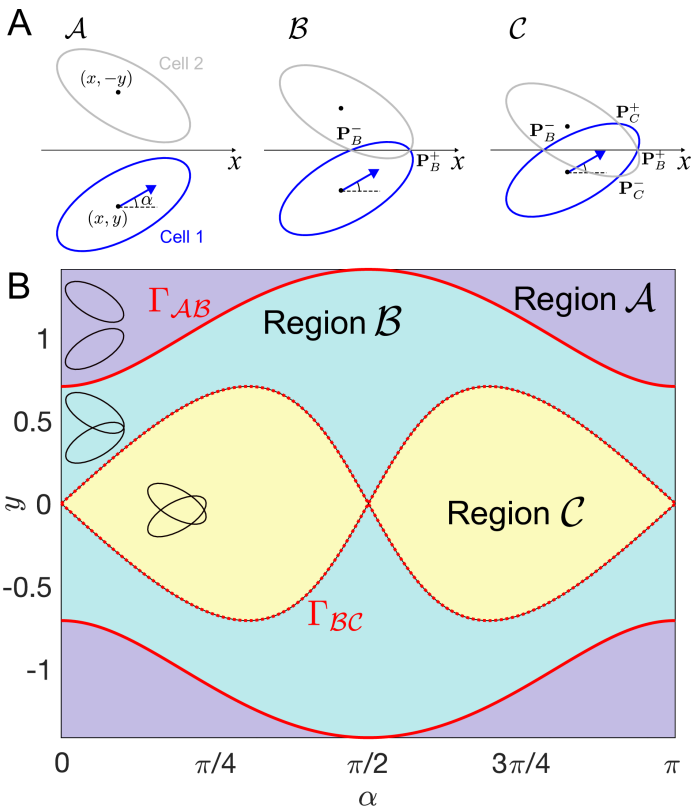

Fig. 3 A: Diagram to show analytical set up and how cells could have 0, 2 or 4 points of overlap. B: α, y space split into regions $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ corresponding to 0, 2 and 4 points of overlap for $r = 2$.

(justified *a posteriori*) are

$$\alpha(t) = \varepsilon A(t), \quad y(t) = -\frac{1}{\sqrt{\bar{r}}} + \varepsilon^2 B(t), \quad r(t) = \bar{r} + \varepsilon^2 C(t), \tag{11}$$

where A, B, and C are perturbations in their respective variables, and are functions of time that we will calculate. Understanding their dynamical behaviours will determine the stability of the steady point. To get an idea of the asymptotic structure of the solution before we go into the details, it is helpful to note that there are two distinguished timescales of interest in the system: the 'early time' where $t = O(\varepsilon)$ and the 'late time' where $t = O(1/\varepsilon)$. Over the early timescale, we will show below that $A(t)$ remains unchanged, $B(t)$ is driven by overlap avoidance and $C(t)$ is driven by overlap avoidance and restoration to aspect ratio $\bar{r}$. Using the early time results, we will then show that over the late timescale, $A(t)$ is affected by overlap avoidance and, along with $B(t)$ and $C(t)$, decays to zero algebraically, demonstrating that the system is stable.

We can substitute (11) into (10), noting that perturbing (α, y, r) by a small amount as defined in (11) pushes the system into region $\mathcal{B}$, to obtain

$$\frac{\mathrm{d}A}{\mathrm{d}t} = -\varepsilon 8 \frac{\bar{r}^2 - 1}{\bar{r}^2 + 1} \bar{r} A \sqrt{(\bar{r}^2 - 1)A^2 + 2\sqrt{\bar{r}}B - \frac{C}{\bar{r}}} + O(\varepsilon^3), \qquad A(0) = a, \tag{12a}$$

$$\varepsilon \frac{\mathrm{d}B}{\mathrm{d}t} = \nu A - 2\sqrt{\bar{r}} \sqrt{(\bar{r}^2 - 1)A^2 + 2\sqrt{\bar{r}}B - \frac{C}{\bar{r}}} + O(\varepsilon^2), \qquad B(0) = b, \tag{12b}$$

$$\varepsilon \frac{\mathrm{d}C}{\mathrm{d}t} = -\frac{1}{\Gamma} \frac{16(1 + \bar{r}^4)}{\bar{r}(1 + \bar{r}^2)} C + \frac{16\bar{r}^2}{1 + \bar{r}^2} \sqrt{(\bar{r}^2 - 1)A^2 + 2\sqrt{\bar{r}}B - \frac{C}{\bar{r}}} + O(\varepsilon^2), \quad C(0) = c. \tag{12c}$$

We now analyse the system (12), starting with the early time.

3.1.1 Early time.

We start our analysis under the early timescale $\tau = O(1)$, defined via $t = \varepsilon\tau$. We indicate the early timescale variables with overhats, and therefore write

$$A(t) = \hat{A}(\tau), \quad B(t) = \hat{B}(\tau), \quad C(t) = \hat{C}(\tau), \quad t = \varepsilon\tau. \tag{13}$$

On substituting (13) into our governing equations (10) and taking the limit $\varepsilon \to 0$, we obtain the following leading-order equations.

$$\frac{\mathrm{d}\hat{A}}{\mathrm{d}\tau} = 0, \qquad \hat{A}(0) = a, \tag{14a}$$

$$\frac{\mathrm{d}\hat{B}}{\mathrm{d}\tau} = \nu \hat{A} - 2\sqrt{\bar{r}} \sqrt{(\bar{r}^2 - 1)\hat{A}^2 + 2\sqrt{\bar{r}}\hat{B} - \frac{\hat{C}}{\bar{r}}}, \qquad \hat{B}(0) = b, \tag{14b}$$

$$\frac{\mathrm{d}\hat{C}}{\mathrm{d}\tau} = -\frac{1}{\Gamma}\frac{16(1+\bar{r}^4)}{\bar{r}(1+\bar{r}^2)}\hat{C} + \frac{16\bar{r}^2}{1+\bar{r}^2}\sqrt{(\bar{r}^2-1)\hat{A}^2 + 2\sqrt{\bar{r}}\hat{B} - \frac{\hat{C}}{\bar{r}}}, \qquad \hat{C}(0) = c. \tag{14c}$$

It is straightforward to use (14a) to determine that $\hat{A}(\tau) = a$ over the early time, and hence to deduce that the orientation is not affected over this timescale. The remaining 2D system (14b)–(14c) governs the dynamics of $\hat{B}(\tau)$ and $\hat{C}(\tau)$, and can be solved computationally. The first term on the right-hand side of (14b) indicates that self-propulsion causes a change in $\hat{B}$ over the early time depending on the sign of a. This will be an increase if $a > 0$ since this means that the cell is inclined slightly upwards. The second term on the right-hand side of (14b) represents overlap avoidance, suggesting that cells will move apart to avoid overlap, consequently causing a decrease in $\hat{B}$. The relative size of these two terms will determine whether $\hat{B}_\tau$ is initially positive or negative. The first term on the right-hand side of (14c) leads to a decrease in magnitude of $\hat{C}$, essentially restoring r to $\bar{r}$. The second term on the right-hand side of (14c) is positive, corresponding to cell elongation to avoid overlap. This elongation forcing occurs because the cells are close to side-by-side near the steady state, and therefore elongation reduces overlap. We also note that the non-trivial dynamics over this early timescale justify the ε scalings we initially imposed in (11).

Finally, we note that the system (14) becomes independent of early time τ in the large-τ limit. We can calculate the specific constants to which the solutions tend by setting the left-hand sides of (14b)–(14c) to zero and solving the resulting algebraic equations. This procedure yields the following $\tau \to \infty$ results

$$\hat{A}(\tau) = a, \quad \hat{B}(\tau) \to \frac{\nu^2 a^2}{8\bar{r}^{3/2}} - \frac{(\bar{r}^2-1)a^2}{2\sqrt{\bar{r}}} + \frac{\Gamma\nu a\bar{r}}{4(1+\bar{r}^4)}, \quad \hat{C}(\tau) \to \frac{\Gamma\nu a\bar{r}^{5/2}}{2(1+\bar{r}^4)}. \tag{15}$$

The early time 'far-field' conditions (15) will be required to asymptotically match into the late-time dynamics we consider next.

3.1.2 Late time.

The system (12) behaves as the early time far-field (15) until $t = O(1/\varepsilon)$, when we have a new distinguished timescale that we refer to as the 'late time'. By inspecting (12), we note that this is the timescale over which the dynamics of A become relevant. Formally, the late timescale is defined by the new variable $T = O(1)$, where $t = T/\varepsilon$. We retain the perturbation scalings (11), but now use tildes to denote late-time variables, defining

$$A(t) = \tilde{A}(T), \quad B(t) = \tilde{B}(T), \quad C(t) = \tilde{C}(T), \quad t = \frac{T}{\varepsilon}. \tag{16}$$

Over this timescale, in the limit $\varepsilon \to 0$, our governing equations (12) become

$$\frac{\mathrm{d}\tilde{A}}{\mathrm{d}T} = -8\frac{\bar{r}^2-1}{\bar{r}^2+1}\bar{r}\tilde{A}\sqrt{(\bar{r}^2-1)\tilde{A}^2 + 2\sqrt{\bar{r}}\tilde{B} - \frac{\tilde{C}}{\bar{r}}}, \tag{17a}$$

$$0 = \nu\tilde{A} - 2\sqrt{\bar{r}}\sqrt{(\bar{r}^2-1)\tilde{A}^2 + 2\sqrt{\bar{r}}\tilde{B} - \frac{\tilde{C}}{\bar{r}}}, \tag{17b}$$

$$0 = -\frac{16(1+\bar{r}^4)}{\Gamma\bar{r}(1+\bar{r}^2)}\tilde{C} + \frac{16\bar{r}^2}{1+\bar{r}^2}\sqrt{(\bar{r}^2-1)\tilde{A}^2 + 2\sqrt{\bar{r}}\tilde{B} - \frac{\tilde{C}}{\bar{r}}}, \tag{17c}$$

with the following 'initial' conditions, obtained by matching with the early-time far-field conditions (15),

$$\lim_{T\to 0^+} \tilde{A}(T) = a \tag{18a}$$

$$\lim_{T\to 0^+} \tilde{B}(T) = \frac{\nu^2 a^2}{8\bar{r}^{3/2}} - \frac{(\bar{r}^2-1)a^2}{2\sqrt{\bar{r}}} + \frac{\Gamma\nu a\bar{r}}{4(1+\bar{r}^4)} \tag{18b}$$

$$\lim_{T\to 0^+} \tilde{C}(T) = \frac{\Gamma\nu a\bar{r}^{5/2}}{2(1+\bar{r}^4)} \tag{18c}$$

Although the differential-algebraic system (17)–(18) is nonlinear, we can simplify the nonlinearity in the differential equation (17a) using the algebraic equation (17b). This procedure reduces (17a) to the following nonlinear but separable differential equation in $\tilde{A}$

$$\frac{\mathrm{d}\tilde{A}}{\mathrm{d}T} = -4\nu\sqrt{\bar{r}}\frac{\bar{r}^2-1}{\bar{r}^2+1}\tilde{A}^2. \tag{19}$$

Solving (19), and substituting into the remaining algebraic equations (17b)–(17c), we obtain the late-time solutions

$$\tilde{A}(T) = \frac{a(\bar{r}^2+1)}{4a\nu\sqrt{\bar{r}}(\bar{r}^2-1)T + \bar{r}^2 + 1}, \tag{20a}$$

$$\tilde{B}(T) = \left(\frac{\nu^2 + 4a\bar{r} - 4\bar{r}^3}{8\bar{r}^{3/2}}\right)\tilde{A}^2(T) + \frac{\Gamma\nu\bar{r}}{4(\bar{r}^4+1)}\tilde{A}(T), \tag{20b}$$

$$\tilde{C}(T) = \frac{\Gamma\nu\bar{r}^{5/2}}{2(\bar{r}^4+1)}\tilde{A}(T). \tag{20c}$$

The late-time solutions (20) show that $\tilde{A}(T), \tilde{B}(T)$ and $\tilde{C}(T)$ all decay to zero algebraically as $T \to \infty$, for $a > 0$, but diverges for $a < 0$. This demonstrates that the point $(0, -1/\sqrt{\bar{r}}, \bar{r})$ is half-stable: if the point is perturbed in the positive α direction, then the point is stable, but if it is perturbed in the negative α direction then it is unstable. We also emphasize that the solutions (20) yield an algebraic decay for $a > 0$, rather than an exponential decay that might arise from a standard linear stability analysis. This *a posteriori* justifies our claim that a non-standard stability analysis is required to determine the system stability. We note that the stability of the other steady state $(0, 1/\sqrt{\bar{r}}, \bar{r})$ can be analysed analogously. The only difference is a change

of sign in (12b) such that it becomes

$$\varepsilon\frac{\mathrm{d}B}{\mathrm{d}t} = \nu A + 2\sqrt{\bar{r}}\sqrt{(\bar{r}^2-1)A^2 + 2\sqrt{\bar{r}}B - \frac{C}{\bar{r}}} + O(\varepsilon^2), \qquad B(0) = b. \tag{21}$$

This change of sign is reflected in the analysis that follows by yielding

$$\tilde{A}(T) = -\frac{a(\bar{r}^2+1)}{4a\nu\sqrt{\bar{r}}(\bar{r}^2-1)T - \bar{r}^2 - 1}, \tag{22a}$$

$$\tilde{B}(T) = \left(\frac{\nu^2 + 4a\bar{r} - 4\bar{r}^3}{8\bar{r}^{3/2}}\right)\tilde{A}^2(T) - \frac{\Gamma\nu\bar{r}}{4(\bar{r}^4+1)}\tilde{A}(T), \tag{22b}$$

$$\tilde{C}(T) = -\frac{\Gamma\nu\bar{r}^{5/2}}{2(\bar{r}^4+1)}\tilde{A}(T). \tag{22c}$$

This late time solution (22) shows that $\tilde{A}(T)$, and consequently $\tilde{B}(T)$ and $\tilde{C}(T)$, diverge as T increases for $a > 0$, and decay to zero algebraically for $a < 0$. This demonstrates that the point $(0, 1/\sqrt{\bar{r}}, \bar{r})$ is also half-stable. Given this, we focus the rest of our analysis on understanding more details surrounding the half-stable point $(0, -1/\sqrt{\bar{r}}, \bar{r})$, again noting that all results can be applied analogously to the point $(0, 1/\sqrt{\bar{r}}, \bar{r})$ via appropriate symmetry arguments. The characterisation of these two points as being half-stable shows that the stability of these points would be sensitive to noise fluctuations (which we do not consider in this work) and suggests that alignment is fleeting in practice. In the many-cell system which has been computationally analysed in Leech et al. (2024), there is additional noise due to random initial conditions and the potential for many cell interactions to happen at once. This means that while two cells may initially be in a configuration which would lead to perfect alignment ($a > 0$), they could be pushed out of this configuration into a configuration where $a < 0$ by a neighbouring cell, meaning that perfect alignment would not be achieved in practice. This could explain why we see pockets of alignment, but no global alignment in the computational simulations in Leech et al. (2024). The points of perfect alignment between two neighbouring cells are half-stable, suggesting that alignment is temporary in practice and the noise in the system means that cells can easily be pushed into the unstable region and hence off the path to alignment.

3.1.3 Composite solution.

Combining the solutions (14) and (20) in both timescales, we can obtain a uniformly regular (additive) composite asymptotic solution (Van Dyke, 1975) for y, α and r in terms of t.

$$\alpha(t) \sim \varepsilon\tilde{A}(\varepsilon t), \tag{23a}$$

$$y(t) + \frac{1}{\sqrt{\bar{r}}} \sim \varepsilon^2\hat{B}\left(\frac{t}{\varepsilon}\right) + \varepsilon^2\tilde{B}(\varepsilon t) - \varepsilon^2\left(\frac{\nu^2 + 4\bar{r} - 4\bar{r}^3}{8\bar{r}^{3/2}}\right)a^2 - \varepsilon\frac{\gamma\nu\bar{r}a}{4(\bar{r}^4+1)}, \tag{23b}$$

$$r(t) - \bar{r} \sim \varepsilon^2 \hat{C}\left(\frac{t}{\varepsilon}\right) + \varepsilon^2 \tilde{C}(\varepsilon t) - \varepsilon \frac{\gamma \nu \bar{r}^{5/2} a}{2(\bar{r}^4 + 1)}, \tag{23c}$$

where the hatted (early-time) variables are defined in (14) and the tilded (late-time) variables are defined in (20). In Fig. 4 we verify that the quantities A, B and C all decay to zero as $t \to \infty$, showing that the point $(\alpha, y, r) = (0, -\frac{1}{\sqrt{\bar{r}}}, \bar{r})$ is stable for $a > 0$. On comparing our analytical solution for the stability analysis with the computational full solution, solved using `ode15s` in MATLAB, we see that we have a good agreement between the two for $\varepsilon = 0.1$ and for $\gamma = 0.01, 0.1, 1$. This is expected from our analysis since $\gamma = O(\varepsilon)$ is a distinguished limit of the system, and hence we expect our analysis to hold in its sublimits until a new distinguished limit is reached. Specifically for this problem, our above analysis allows us to straightforwardly understand the subcases $\gamma \ll \varepsilon$ and $\varepsilon \ll \gamma \ll 1/\varepsilon$ as regular sublimits of our analysis. We also note that the solution for $A(t)$ is unaffected by γ, suggesting that having shape change in the model does not affect the change in orientation that two cells will experience in response to overlap avoidance. Since a change in orientation is needed for cells to align with each other, this suggests that varying the non-dimensional shape change parameter γ, at least for $\gamma < \nu$, has little effect on cell alignment. This is contrary to what is found in the computational analysis in Leech et al. (2024), suggesting that these computational results are a result of the collective behaviour of the system, and not the underlying alignment mechanism.

3.1.4 Asymptotic sublimits.

From the above stability analysis, we can conclude that the position y and aspect ratio r initially respond to overlap over the early timescale $t = O(\varepsilon)$, but that the orientation responds over the late timescale $t = O(1/\varepsilon)$. More precisely, the overlap avoidance timescale is $t = O(\varepsilon)$, and the aspect ratio restoration timescale is $t = O(\gamma)$. These latter two timescales overlap for the distinguished limit $\gamma = O(\varepsilon)$ we consider above. This overlap is why $\gamma = O(\varepsilon)$ is a distinguished limit, and is also why there is another distinguished limit when $\gamma = O(1/\varepsilon)$. The new distinguished limit corresponds to extremely "squishy" cells where shape change is punished very little, generating cells with aspect ratios far away from the natural shape $\bar{r}$ as a result of overlap avoidance. This distinguished limit is less biologically relevant, so we do not consider this case further.

It is instructive to briefly summarize what happens in the sublimits of the distinguished limit we analysed above. In the sublimit of $\varepsilon \ll \gamma \ll 1/\varepsilon$, equivalently $1 \ll \Gamma \ll 1/\varepsilon^2$ (the upper limits of which correspond to the different distinguished limit noted above), the different balancing mechanisms over the early timescale diverge into two separate timescales: $t = O(\varepsilon)$ and $t = O(\gamma)$. Over the standard early timescale $t = O(\varepsilon)$, only overlap avoidance affects C. Over this timescale, B and C grow without a restoring force. The timescale over which this restoring force becomes important is then a new intermediate timescale $t = O(\gamma)$, which kicks in when C gets large enough for the overlap avoidance term to balance the aspect-ratio-restoring term. Over this intermediate timescale, B and C reach the equivalent steady states to those in the full behaviour we determined above. Finally, the late timescale $t = O(1/\varepsilon)$ is essentially

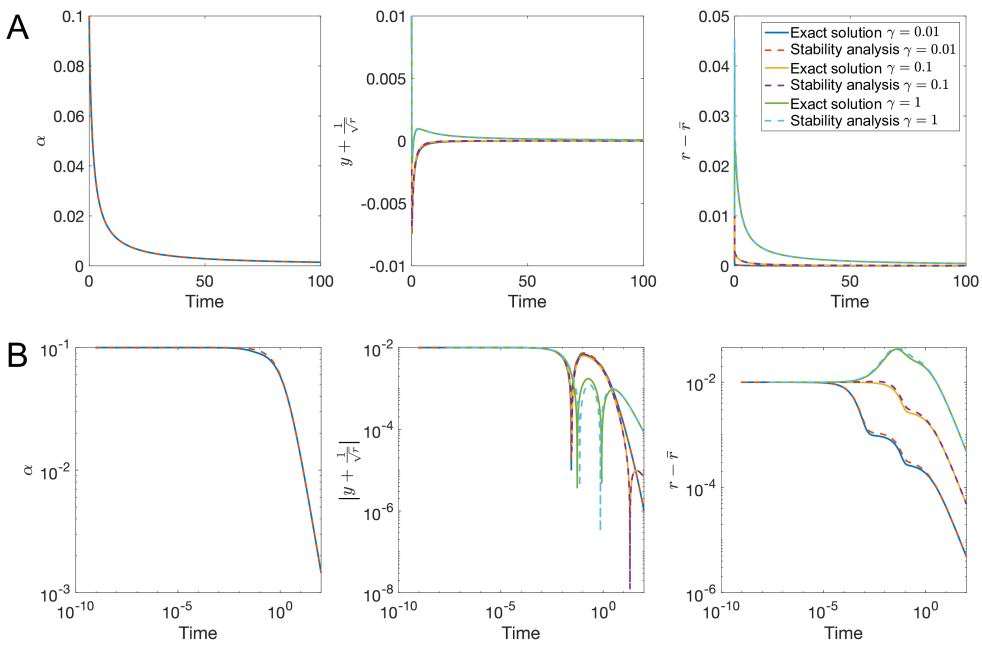

Fig. 4 A, B: Plots of the exact solution of system (10) and the stability analysis approximation solution given in (23) against time for $\gamma = 0.01, 0.1, 1$. Legend applies to all plots. Other parameters are $\bar{r} = 2, \nu = 2, a = b = c = 1$.

unchanged from the full analysis in this sublimit; it is this timescale over which $A(t)$ varies. The full analysis for $\gamma = O(1)$ is outlined in Appendix B.

The opposite sublimit of $\gamma \ll \varepsilon$ (equivalently $\Gamma \ll 1$), corresponds to cells that restore their natural aspect ratio very quickly (over a timescale of $t = O(\gamma)$), and then behave as non-deformable cells with distinct early ($t = O(\varepsilon)$) and late ($t = O(1/\varepsilon)$) time behaviours. This sublimit is tractable for a more general dynamic analysis, not just near the steady points, and is of biological interest. Therefore, we examine this case in more detail in the next section.

4 No shape change, $\gamma = 0$

In this section, we consider the no-shape-change limit, taking $\gamma \to 0$. As noted in the previous section, there is a very fast timescale of $t = O(\gamma)$ over which the cells restore and then maintain their natural aspect ratio of $r = \bar{r}$. We will proceed by taking $\gamma = 0$ and $r \equiv \bar{r}$, essentially analysing (10) after the very fast $t = O(\gamma)$ timescale. This reduces (10) to a two-dimensional dynamical system in α and y. This reduction is both biologically relevant and more straightforward to analyse and interpret. In the rest of this section, we will denote $\bar{r}$ by r for notational convenience, where r no longer has a time dependence. The (α, y) state space is split into regions $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$. We show an

example of the state space for a fixed value of $r = 2$ in Fig. 3B. We start our investigation by confirming that we obtain the same stability result from Section 3 before exploring the resulting 2D dynamical system computationally. Then, to gain deeper insight into the precise effect of self-propulsion, which has dimensionless strength ν, we consider the case of $\nu = 0$ before using asymptotic methods to understand the case where $\nu \ll 1$, noting that the limit $\nu \to 0$ is singular.

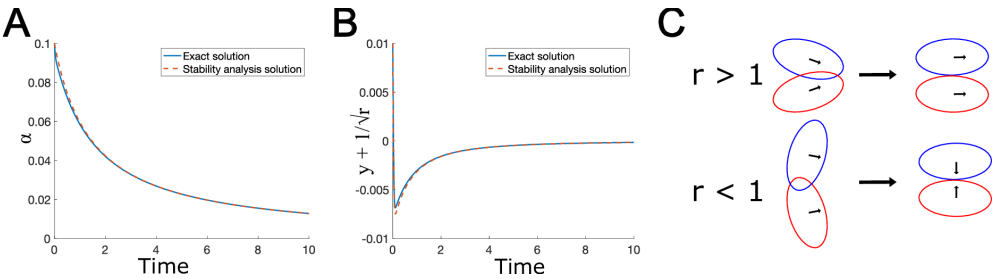

Fig. 5 A,B: plots comparing the exact solution to system (10) with $\gamma = 0$ to the solution obtained doing stability analysis with $\varepsilon = 0.1, \nu = 2, r = 2, a = b = 1$. C: diagram depicting the effect of r on final alignment.

4.1 Stability analysis.

For $\nu \neq 0, \nu = O(1)$, the stability of $(\alpha, y) = (0, -\frac{1}{\sqrt{r}})$ is inherited directly as a sublimit $\Gamma \to 0$ of the full shape change stability analysis we conducted in Section 3 above. Specifically, if we take $\gamma \to 0$ in (10d), or equivalently $\Gamma \to 0$ in (14c), we find that $C(t) \to 0$ over $t = O(\gamma)$, which is much quicker compared to the other two timescales of $O(\varepsilon)$ and $O(1/\varepsilon)$. Then, over the fast timescale $t = O(\varepsilon)$, (14) reduces to

$$\hat{A}(\tau) = a, \quad \frac{\mathrm{d}\hat{B}}{\mathrm{d}\tau} = \nu a - 2\sqrt{r}\sqrt{(r^2-1)a^2 + 2\sqrt{r}B}, \tag{24}$$

and over the slow timescale $t = O(1/\varepsilon)$, (17) becomes

$$\tilde{A}(T) = \frac{a(r^2+1)}{4a\nu\sqrt{r}(r^2-1)T + r^2 + 1}, \quad \tilde{B}(T) = \left(\frac{\nu^2 + 4r - 4r^3}{8\bar{r}^{3/2}}\right)\tilde{A}^2(T). \tag{25}$$

We can see that if $r > 1, a > 0$ then $\bar{A}, \bar{B} \to 0$ algebraically as $T \to \infty$. For $a < 0$, $\bar{A}$, and consequently $\bar{B}$ diverge. The point $(0, -\frac{1}{\sqrt{r}})$ is therefore half-stable. If $r < 1$ then $A(T)$ and $B(T)$ both diverge for $a > 0$ and decay to zero for $a < 0$, and we have the opposite behaviour. To understand physically why this is the case, consider the diagrams in Fig. 5C. If $r > 1, a > 0$ then the combination of overlap avoidance and self-propulsion with the geometry of the cells causes the cells to more towards an orientation of $\alpha = 0$. If, on the other hand, $r < 1, a > 0$, we can see that the opposite happens and though they move to the same configuration geometrically, this

corresponds to an orientation of $\alpha = \frac{\pi}{2}$ and hence the orientation moves away from $\alpha = 0$. We can see in Fig. 5A,B that we have a good agreement between the exact solution and the approximation that we obtain from the stability analysis.

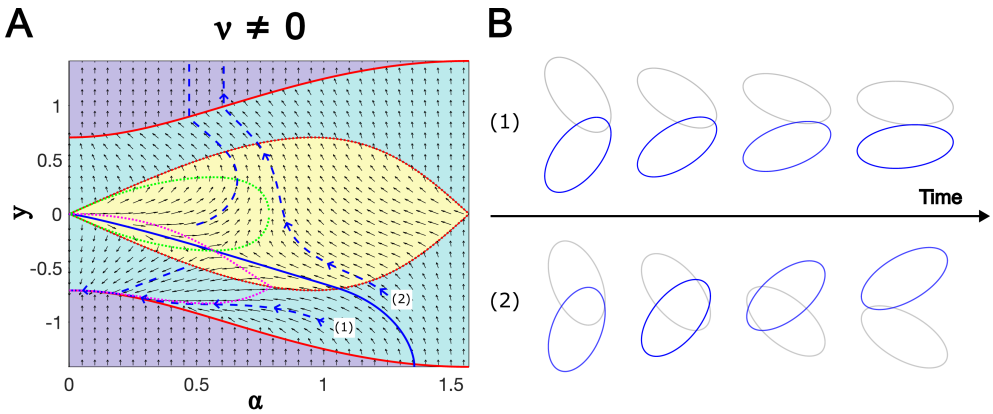

Fig. 6 A: Phase portrait for $\nu = 2, r = 2$ showing nullclines (dotted pink and green), boundaries between regions (red) and example trajectories (blue). B: Schematics showing snapshots of the system over time for two example trajectories labelled (1) and (2) in A. The cell in blue corresponds to the trajectory in blue.

We can then explore the system more generally by computationally generating a phase portrait with some example trajectories in Fig. 6B. Nullclines corresponding to $\dot{y} = 0$ (pink) and $\dot{\alpha} = 0$ (green) are shown as dotted lines. We see that in region $\mathcal{A}$ (purple) the arrows all point vertically upwards. This is because this is the region of no overlap and hence the cell will only change its position (as a result of self-propulsion) and not its orientation. The arrows are pointed upwards since $\alpha \in [0, \pi/2]$ and hence this leads to an increase in y since the cell is inclined upwards. If we focus our attention around the point $(\alpha, y) = (0, -1/\sqrt{r})$, we see that all arrows in the surrounding area are directed towards this point, illustrating that this point is an attractor from the side where $\alpha > 0$. Example trajectories are shown in blue and from these we see that there are two outcomes depending on the initial conditions: (1) the trajectory moves towards the half-stable point $(\alpha, y) = (0, -\frac{1}{\sqrt{r}})$, or (2) the trajectory crosses region $\mathcal{C}$ and ends up leaving the overlap region with $y > 0$. The line separating these two outcomes is a non-trivial separatrix, indicated by the solid blue line in Fig. 6B. We can calculate this separatrix computationally, by starting near the separatrix and solving the system (10) backwards in time (Fay and Joubert, 2010). Computing the area under this separatrix quantifies the proportion of initial conditions that lead to perfect alignment. In Fig. 7A, we see that this area decreases nonlinearly as ν increases, showing that we get less alignment for large values of ν. This matches with the computational results of the full model in Leech et al. (2024), where it is found that alignment decreases as ν increases, though only for $\nu > 0.2$. While the simplified two-cell model we consider here does not account for the numerous collective mechanisms in the many-cell model, these

findings provide a partial explanation for the collective results. That is, an increase in cell velocity means that there are fewer initial configurations that result in perfect alignment.

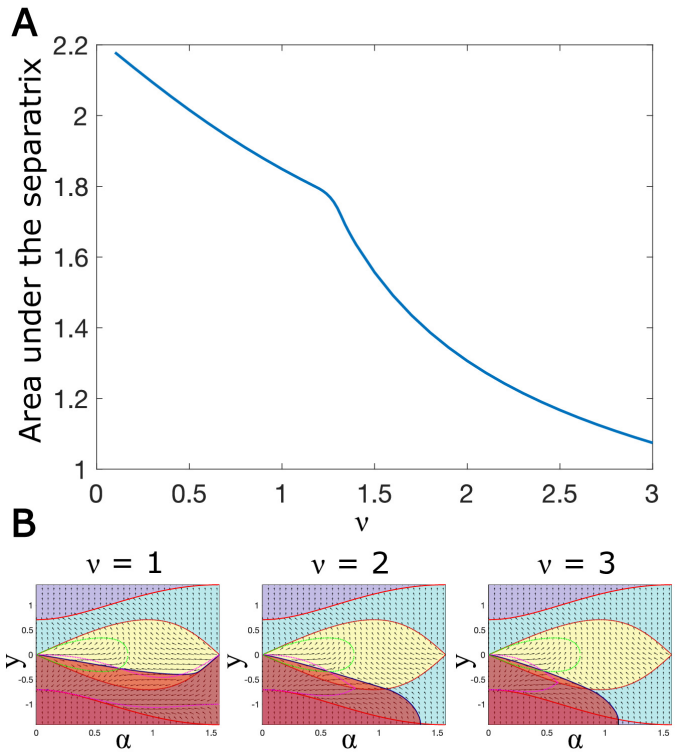

Fig. 7 A: plot showing the dependence of ν on the area under the separatrix (a way to quantify the proportion of initial conditions that move to perfect alignment). B: Phase portraits for different values of ν with the area under the separatrix shaded. $r = 2$.

To summarise, we have shown that, similarly to the 3D case, the steady point $(\alpha, y) = (0, -\frac{1}{\sqrt{r}})$ is half-stable for $r > 1$, with stability occurring for $\alpha > 0$. For $r < 1$ we have the opposite behaviour where the point is stable for $\alpha < 0$. We can plot a phase portrait in 2D for $r > 1$ and find that there is a separatrix which splits the phase space into two sections. Any trajectory that starts under the separatrix will end up at the half-stable steady state given by $(\alpha, y) = (0, -\frac{1}{\sqrt{r}})$, which corresponds to the cells being perfectly aligned. Trajectories which start above the separatrix do not end up in a configuration of perfect alignment. We note that this is only true for $r > 1$ since the point $(\alpha, y) = (0, -\frac{1}{\sqrt{r}})$ is only a stable steady state for $r > 1, \alpha > 0$. For $r < 1$, further analysis would need to be done to analyse steady states and investigate the separatrix. Since this case is less biologically relevant, it is not the focus of this work. The area under the separatrix can be used as a way of quantifying the proportion of initial conditions which lead to perfect alignment. We find that as ν increases, this

proportion decreases, suggesting that smaller values of ν will lead to more alignment. This corresponds to a smaller self-propulsion speed or a larger strength of overlap avoidance. Smaller self-propulsion speeds will mean that the cells are interacting for a longer duration, and a larger overlap avoidance strength means that cells will move and reorient to avoid overlap more quickly.

4.2 No self-propulsion, $\nu = 0$

To investigate the effect of propulsion on (10) with $\gamma = 0$, we now investigate the singular change in the system in the limit of small ν. Specifically, we are interested in the strong difference in system behaviour for zero self-propulsion ($\nu = 0$) and small self-propulsion ($0 < \nu \ll 1$). We start by considering zero self-propulsion, $\nu = 0$. The governing equations (10) with $\nu = 0$ are

$$\dot{y} = \begin{cases} 0, & \text{for } (\alpha, y) \in \mathcal{A} \\ 2\sqrt{r}\frac{\sqrt{\delta_1^2 - ry^2}}{\delta_1^2}\text{sign}(y), & \text{for } (\alpha, y) \in \mathcal{B} \\ 2\frac{r^2}{\delta_1^2(r^2-1)|\sin\alpha\cos\alpha|}y, & \text{for } (\alpha, y) \in \mathcal{C} \end{cases} \tag{26a}$$

$$\dot{\alpha} = \begin{cases} 0, & \text{for } (\alpha, y) \in \mathcal{A} \\ -8\frac{r^2-1}{r^2+1}|y|r\sqrt{r}\sin\alpha\cos\alpha\frac{\sqrt{\delta_1^2 - ry^2}}{\delta_1^4}, & \text{for } (\alpha, y) \in \mathcal{B} \\ \frac{4r\text{sign}(\sin\alpha\cos\alpha)}{r^2+1}\left[\frac{(r^2-1)^2y^2\sin^2\alpha\cos^2\alpha}{\delta_1^4} + r\frac{\delta_1^2 - ry^2}{\delta_1^4} - \frac{(\delta_2^4 - r^2)y^2}{(r^2-1)^2\sin^2\alpha\cos^2\alpha} - \frac{r}{\delta_2^2}\right], & \text{for } (\alpha, y) \in \mathcal{C} \end{cases} \tag{26b}$$

The system (26) is symmetric around $y = 0$. Together with the symmetry considerations for α, it therefore suffices to consider $y \leq 0$, $\alpha \in [0, \pi/2]$.

Below, we outline the behaviours of the system (26) descriptively before going through and deriving each result. We first note that y and α remain unchanged in region $\mathcal{A}$. This is because region $\mathcal{A}$ corresponds to no cell overlap; if there is no self-propulsion, there is no mechanism for the cells to move. In region $\mathcal{B}$, there is a line of stationary points (where $\dot{y} = \dot{\alpha} = 0$) when $\delta_1^2 = ry^2$. This line exactly corresponds to the boundary between regions $\mathcal{A}$ and $\mathcal{B}$. In fact, since $\dot{y}$ takes a finite negative value in regions $\mathcal{B}$ and $\mathcal{C}$, the trajectory will always move in the direction of decreasing y and terminate at the region $\mathcal{A}-\mathcal{B}$ boundary in finite time. This is true for trajectories that start in region $\mathcal{C}$; these always move into region $\mathcal{B}$ in finite time and then cannot move back into region $\mathcal{C}$. This can all be seen in the phase portrait in Fig.8A.

We let $\nu = 0$, $r > 1$ and consider system (10) for $(\alpha, y) \in \mathcal{S} := \{(\alpha, y) \mid y < 0, \alpha \in [0, \pi/2]\}$. For given initial data $(\alpha(0), y(0)) = (\alpha_0, y_0) \in \mathcal{S}$ we find that:

i If $(\alpha_0, y_0) \in \mathcal{S} \cap \mathcal{A}$ then $(\alpha(t), y(t)) = (\alpha_0, y_0)$ for all time.

ii If $(\alpha_0, y_0) \in \mathcal{S} \cap \mathcal{C}$, then there exists a finite time $\bar{t} > 0$ such that $(\alpha(t), y(t)) \in \mathcal{S} \cap \mathcal{B}$ for $t > \bar{t}$.

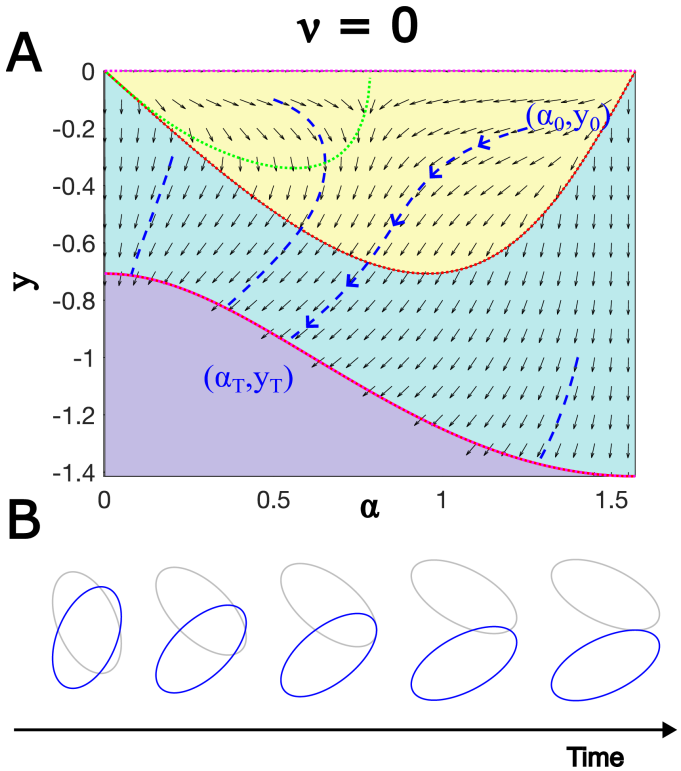

Fig. 8 A: Phase portrait in (α, y) space for $\nu = 0$. Example trajectories shown in blue. Boundaries between regions shown in red. Nullclines shown in pink and green. B: snapshots of the cell configurations over time corresponding to the trajectory in A with initial conditions (α_0, y_0).

iii If $(\alpha_0, y_0) \in \mathcal{S} \cap B$, then there exists a finite time $T > 0$ such that $(\alpha(t), y(t)) = (\alpha(T), y(T))$ for $t > T$. Further $y(T) = -\Gamma_{\mathcal{AB}}(\alpha(T))$, as defined in (8).

Part (i): From equations (26), we see that $\dot{\alpha} = \dot{y} = 0$ in Region A, and therefore α and y remain unchanged in region $\mathcal{A}$.

Part (ii): From (26a) we see that $\dot{y} < 0$ when $y < 0$ within region $\mathcal{C}$ and on the boundary of region $\mathcal{C}$. We also have that $|\dot{y}| \geq 4|y|/(r^2 - 1)$ which shows that $|\dot{y}|$ is bounded away from zero (we note that $r > 1$). Hence, we can conclude that if $(\alpha_0, y_0) \in \mathcal{S} \cap \mathcal{C}$ then there exists a finite time $\bar{t} > 0$ where the trajectory $(\alpha(t), y(t))$ will satisfy $y(\bar{t})^2 = \Gamma^2_{\mathcal{BC}}(\alpha(\bar{t}))$ and hence move into region $\mathcal{B}$. Further, region $\mathcal{B}$ is invariant, in the sense that once the trajectory is in region $\mathcal{B}$, it remains in region $\mathcal{B}$ for the rest of time. To show that this is the case, we begin by characterising the solution curves in region $\mathcal{B}$. We consider the case where (α_0, y_0) and (α, y) are in $\mathcal{S} \cap \mathcal{B}$ and derive an equation for the resultant trajectory as $y(\alpha)$. Dividing (26b) by (26a) we obtain

$$\frac{\mathrm{d}y}{\mathrm{d}\alpha} = -\frac{1}{4r}\frac{r^2+1}{r^2-1}\frac{r^2\sin^2\alpha + \cos^2\alpha}{r\sin\alpha\cos\alpha}\frac{1}{y}. \tag{27}$$

Equation (27) can be solved explicitly to yield

$$y(\alpha) = -\sqrt{|y_0|^2 + \frac{1}{2r}\frac{r^2+1}{r^2-1}\log\left(\frac{(\cos\alpha)^{r^2}}{\sin\alpha}\frac{\sin\alpha_0}{(\cos\alpha_0)^{r^2}}\right)}. \tag{28}$$

To show that region $\mathcal{B}$ is invariant, we want to show that once the trajectory moves from region $\mathcal{C}$ into region $\mathcal{B}$, it is not possible for the trajectory to move back into region $\mathcal{C}$ again. To do this, we will show that if we start on the boundary of region $\mathcal{B}$ and C, given by $y^2 = \Gamma^2_{\mathcal{BC}}$ as defined in (8), we will enter region $\mathcal{B}$ and will not intersect the line $y^2 = \Gamma_{\mathcal{BC}}$ again, showing that it is not possible for the trajectory to cross this line a second time and enter back into region $\mathcal{C}$. We substitute our initial conditions $y_0^2 = \Gamma^2_{\mathcal{BC}}(\alpha_0)$ into (28) and want to solve $y^2 = \Gamma^2_{\mathcal{BC}}$ to see whether the trajectory y could intersect the boundary $\Gamma_{\mathcal{BC}}$ again. If there are no additional solutions to this equation, then the trajectory does not cross the boundary between regions $\mathcal{B}$ and $\mathcal{C}$ again and hence does not re-enter region $\mathcal{C}$. From this procedure we obtain

$$\Gamma^2_{\mathcal{BC}}(\alpha) = \Gamma^2_{\mathcal{BC}}(\alpha_0) + \frac{1}{2r}\frac{r^2+1}{r^2-1}\log\left(\frac{(\cos\alpha)^{r^2}}{\sin\alpha}\frac{\sin\alpha_0}{(\cos\alpha_0)^{r^2}}\right), \tag{29}$$

where $\Gamma_{\mathcal{BC}}$ is defined in (8). Rearranging (29) and defining $s = \sin\alpha$, we obtain

$$\begin{aligned}\frac{s}{(1-s^2)^{r^2/2}}\exp\left[\frac{2(r^2-1)^3 s^2(1-s^2)}{(r^2+1)(s^2+r^2(1-s^2))}\right] \\ = \frac{\sin\alpha_0}{(\cos\alpha_0)^{r^2}}\exp\left[\frac{2(r^2-1)^3\sin^2\alpha_0\cos^2\alpha_0}{(r^2+1)(\sin^2\alpha_0+r^2\cos^2\alpha_0)}\right].\end{aligned} \tag{30}$$

We define the left-hand side of (30) as $s \mapsto f(s)$ with domain $[0,1)$. Differentiating f with respect to s shows that f is a strictly monotonically increasing function of s with $f(0) = 0$ and $f(s) \to \infty$ for $s \to 1^-$. Since the right-hand side is always positive, there exists a unique solution $s_{sol} \in [0,1)$ for each (α_0, y_0) and consequently a solution $\alpha_{sol} = \arcsin s_{sol} \in [0, \pi/2]$ such that $y(\alpha_{sol}) = -\Gamma_{\mathcal{BC}}(\alpha_{sol})$. However, since we know that the initial condition satisfies these requirements, the unique solution is the initial condition i.e. $\alpha_{sol} = \alpha_0$. Therefore, after intersecting the curve $y^2 = \Gamma^2_{\mathcal{BC}}$ and moving into region $\mathcal{B}$, the trajectory will not intersect this boundary curve again. Hence, once the trajectory is in region $\mathcal{B}$ it will remain in region $\mathcal{B}$.

Part (iii): We start by noting that the curve $y = -\Gamma_{AB}$ describes a line of stationary points (where $\dot{y} = \dot{\alpha} = 0$). Showing that $y(\alpha)$ always intersects $-\Gamma_{\mathcal{AB}}(\alpha)$ is equivalent to showing that $y^2(\alpha) = \Gamma^2_{\mathcal{AB}}(\alpha)$ has a solution. Using $y(\alpha)$ as defined in (28) and $\Gamma_{\mathcal{AB}}$ as defined in (8) this can be written as

$$y_0^2 + \frac{1}{2}\frac{r^2+1}{r^2-1}\log\left(\frac{(\cos\alpha)^{r^2}}{\sin\alpha}\frac{\sin\alpha_0}{(\cos\alpha_0)^{r^2}}\right) = r^2\sin\alpha^2 + \cos\alpha^2. \tag{31}$$

We need to show that (31) has a solution $\alpha \in [0, \pi/2]$. We define $s = \sin\alpha$ and use this to rewrite (31) as

$$\frac{s}{(1-s^2)^{r^2/2}} \exp\left[2\frac{(r^2-1)^2}{r^2+1}s^2\right] = \frac{\sin\alpha_0}{\cos\alpha_0{}^{r^2}} \exp\left[2\frac{r^2-1}{r^2+1}(y_0^2-1)\right]. \tag{32}$$

We define the left-hand side as $s \mapsto h(s)$ with domain $[0, 1)$. Differentiating $h(s)$ we find that $h'(s) > 0$ and can therefore conclude that h is a strictly monotonically increasing function of s with $h(0) = 0$ and $h(s) \to \infty$ for $s \to 1^-$, i.e. its range is $[0, \infty)$. Since the right-hand side is always positive, there exists a unique solution $s_T \in [0, 1)$ for each (α_0, y_0) and consequently a solution $\alpha_T = \arcsin s_T \in [0, \pi/2]$ such that $y(\alpha_T) = -\Gamma_{\mathcal{AB}}(\alpha_T)$. Finally, we need to show that (α_T, y_T) is reached in finite time, where we define $y_T := y(\alpha_T)$. To show this, we substitute (28) into (26b), the governing equation for $\dot{\alpha}$, separate variables and integrate over time from 0 to t, obtaining

$$8tr\sqrt{r}\frac{r^2-1}{r^2+1} = \int_{\alpha_0}^{\alpha(t)} \frac{\delta_1^4(\alpha)}{y(\alpha)\sin\alpha\cos\alpha\sqrt{\delta_1^2(\alpha) - ry^2(\alpha)}}\,\mathrm{d}\alpha =: I(\alpha(t)). \tag{33}$$

Our final task is to show that $\lim_{\alpha\to\alpha_T^+} I(\alpha) =: I(\alpha_T) < \infty$, which will imply that (α_T, y_T) is reached in finite time $T = I(\alpha_T)\frac{r^2+1}{r^2-1}\frac{1}{8r\sqrt{r}}$. We use the substitution $s = \sin\alpha$ in (33) and obtain

$$I(\alpha_T) = \lim_{s_t\to s_T^+} \int_{s_0}^{s_t} \frac{\delta_1^4(s)}{y(s)s(1-s^2)\sqrt{\delta_1^2(s) - ry^2(s)}}\,\mathrm{d}s. \tag{34}$$

To understand whether (34) converges, we need to understand the nature of the singularity at $s = s_T$ where $\delta^2(s_T) - ry^2(s_T) = 0$. Taylor expanding this term around $s = s_T$ gives

$$\delta_1^2(s) - ry^2(s) = (s - s_T)\left[2(r^2-1)s_T + \frac{(r^2+1)(s_T^2(r^2-1)+1)}{2(r^2-1)s_T(1-s_T^2)}\right] + \mathcal{O}(s-s_T)^2. \tag{35}$$

Since the coefficient of the linear $s - s_T$ term in (35) does not vanish, the singularity in the integrand is an inverse square root, and therefore is integrable. That is, $I(\alpha_T) < \infty$ and hence $(\alpha(t), y(t)) = (\alpha_T, y_T)$ for $t = T < \infty$.

4.3 Understanding a small amount of self-propulsion, $0 < \nu \ll 1$

In the case of a small but finite ν, the system is singular. We can investigate its behaviour using asymptotic methods in the small-ν limit. To briefly summarise before going into details, the main differences between the small- and the zero-ν cases is that small-ν dynamics do not vanish in region $\mathcal{A}$, nor do they terminate at the point

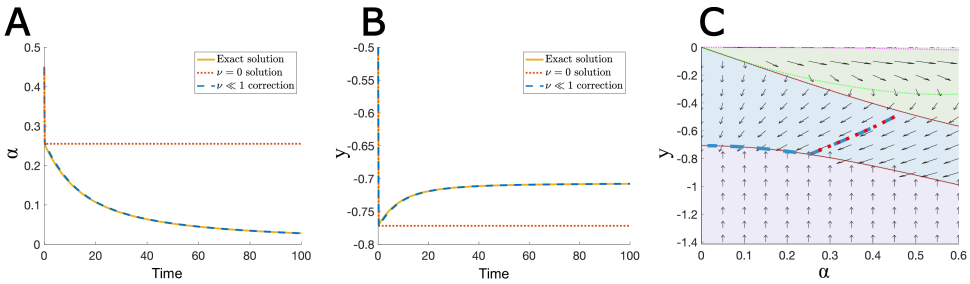

Fig. 9 A, B: results of asymptotic analysis for $\nu \ll 1$. C: zoomed in region of phase portrait showing corresponding trajectory for A and B.

at which they first hit the region $\mathcal{A} - \mathcal{B}$ boundary. Instead, in the small-ν case the trajectories travel within region $\mathcal{A}$ and along the region $\mathcal{A} - \mathcal{B}$ boundary over slow timescales of $O(1/\nu)$. So, in the small-ν case a trajectory starting in region $\mathcal{C}$ will follow the zero-ν trajectory determined in section 4.2 with an $O(\nu)$ correction, and move into region $\mathcal{B}$ and then to the region $\mathcal{A} - \mathcal{B}$ boundary in finite time. A trajectory starting in region $\mathcal{A}$ will also move into region $\mathcal{B}$ over a slower time of $O(1/\nu)$. Given this, we focus our analysis on understanding what happens in region $\mathcal{B}$ in further detail.

To understand the system behaviour in the limit of small ν, we first write both y and α as asymptotic series in powers of ν. That is $y(t) \sim y_0(t) + \nu y_1(t)$ and $\alpha(t) \sim \alpha_0(t) + \nu\alpha_1(t)$. Substituting these into the governing equations (10) and considering the leading-order equations, we obtain exactly the case $\nu = 0$ that we analysed in section 4.2. In this case, we know that $y_0(t) \to \bar{y}$ and $\alpha_0(t) \to \bar{\alpha}$ in some finite time T, where $\delta_1^2(\bar{\alpha}) = r\bar{y}^2$. In the case where $\nu = 0$, the analysis ends here. However, in the singular case of small but finite ν, there is an additional slow timescale after this finite time.

To analyse this slow timescale behaviour, we transform into the slow timescale regime defined by $t = T + \tau/\nu$, where $\tau = O(1)$, and now expand y and α as $Y(\tau) \sim Y_0(\tau) + \nu Y_1(\tau)$ and $A(\tau) \sim A_0(\tau) + \nu A_1(\tau)$. On substituting the slow timescale and these new expansions into the governing equations (10), we obtain

$$\nu \frac{\mathrm{d}Y}{\mathrm{d}\tau} = \nu \sin A - f(Y, A), \tag{36a}$$

$$\nu \frac{\mathrm{d}A}{\mathrm{d}\tau} = g(Y, A) f(Y, A), \tag{36b}$$

where we introduce the following shorthand functions

$$f(Y, A) = 2\sqrt{r}\frac{\sqrt{\delta_1^2(A) - rY^2}}{\delta_1^2(A)}, \tag{37a}$$

$$g(Y, A) = 4\frac{r^2 - 1}{r^2 + 1}\frac{Yr \sin A \cos A}{\delta_1^2(A)}. \tag{37b}$$

The $O(1)$ terms in equations (36) generate a duplication of information, with both giving $f(Y_0, A_0) = 0$, or equivalently

$$\delta_1^2(A_0(\tau)) = rY_0^2(\tau). \tag{38}$$

This tells us that the leading-order slow-time dynamics are confined to the line $\delta_1^2(\alpha) = ry^2$, the boundary between regions $\mathcal{A}$ and $\mathcal{B}$. The remaining goal of our analysis is to determine the precise dynamics of this slow motion. There are several ways to proceed at this point. One method involves continuing to the next asymptotic order and deriving an appropriate solvability condition. Another way involves combining the full equations (36) in a way that removes the duplication of information, then taking the limit $\nu \to 0$ to obtain an independent evolution equation. We proceed via the latter, since this involves significantly less algebra. By substituting (36a) into (36b) and dividing through by ν, we obtain

$$\frac{\mathrm{d}A}{\mathrm{d}\tau} = g(Y, A)\left(\sin A - \frac{\mathrm{d}Y}{\mathrm{d}\tau}\right). \tag{39}$$

The equation (39) has leading-order form

$$\frac{\mathrm{d}A_0}{\mathrm{d}\tau} = 4\frac{r^2-1}{r^2+1}\frac{Y_0 r \sin A_0 \cos A_0}{\delta_1^2(A_0)}\left(\sin A_0 - \frac{\mathrm{d}Y_0}{\mathrm{d}\tau}\right). \tag{40}$$

The equations (38) and (40) represent enough information to determine the slow evolution along the region $\mathcal{A} - \mathcal{B}$ boundary. It is more straightforward to see this if we combine (40) with the direct time derivative of $\delta_1^2(A_0) = rY_0^2$, to obtain

$$\sin A_0 \cos A_0 (r^2-1)\frac{\mathrm{d}A_0}{\mathrm{d}\tau} = rY_0\frac{\mathrm{d}Y_0}{\mathrm{d}\tau}, \tag{41}$$

from which we can deduce that

$$\left(4\frac{(r^2-1)^2}{r^2+1}\sin^2 A_0 \cos^2 A_0 + rY_0^2\right)\frac{\mathrm{d}A_0}{\mathrm{d}\tau} = 4r\frac{(r^2-1)}{r^2+1}Y_0 \sin^2 A_0 \cos A_0, \tag{42a}$$

$$\left(4\frac{(r^2-1)^2}{r^2+1}\sin^2 A_0 \cos^2 A_0 + rY_0^2\right)\frac{\mathrm{d}Y_0}{\mathrm{d}\tau} = 4\frac{(r^2-1)^2}{r^2+1}\sin^3 A_0 \cos^2 A_0. \tag{42b}$$

Then, since we are considering $\alpha \in [0, \pi/2]$ and therefore $A_0 \in [0, \pi/2]$, we let $s_0(\tau) = \sin(A_0(\tau))$, where $s_0 \in [0, 1]$. Under this substitution, (42) can be rewritten as

$$\frac{\mathrm{d}s_0}{\mathrm{d}\tau} = \frac{4r(r^2-1)Y_0 s_0^2(1-s_0^2)}{4(r^2-1)^2 s_0^2(1-s_0^2) + r(r^2+1)Y_0^2}, \tag{43a}$$

$$\frac{\mathrm{d}Y_0}{\mathrm{d}\tau} = \frac{4(r^2-1)^2 s_0^3(1-s_0^2)}{4(r^2-1)^2 s_0^2(1-s_0^2) + r(r^2+1)Y_0^2}. \tag{43b}$$

Solving (43) numerically, we can compare our numerical solution of the reduced problem to a numerical solution of the full problem. In Fig. 9, we see that the asymptotic

solution gives good agreement with the full numerical solution for $\nu = 0.1$, solved using `ode15s` in MATLAB. We also see from this that for all initial conditions, the trajectory will move to the point $(\alpha, y) = (0, -\frac{1}{\sqrt{r}})$, corresponding to perfect alignment.

If we substitute the solution of the algebraic constraint (38), $Y_0 = -\frac{1}{\sqrt{r}}\sqrt{(r^2-1)s_0^2+1}$ into (43a) we can reduce the dynamics to a single ODE:

$$\frac{\mathrm{d}s_0}{\mathrm{d}\tau} = -\frac{4\sqrt{r}(r^2-1)\sqrt{(r^2-1)s_0^2+1}s_0^2(1-s_0^2)}{4(r^2-1)^2s_0^2(1-s_0^2)+(r^2+1)(r^2-1)s_0^2+r^2+1}. \tag{44}$$

On separating variables and integrating, we find that

$$\begin{aligned} &4\sqrt{r}(r^2-1)\tau \\ &= \int_{s_0(0)}^{s_0(\tau)} \frac{4(r^2-1)^2s_0^2(1-s_0^2)+(r^2+1)(r^2-1)s_0^2+r^2+1}{s_0^2(1-s_0^2)\sqrt{(r^2-1)s_0^2+1}}\,\mathrm{d}s_0 =: I(s_0(\tau)). \end{aligned} \tag{45}$$

We note that we have shown in Section 4 that the point $(\alpha, y) = (0, -\frac{1}{\sqrt{r}})$ is a stable steady state for $\alpha > 0$. We also know that all trajectories solving (43) move towards this point. To understand this long-term behaviour in more detail, we must understand what happens as $\alpha \to 0^+$ or equivalently $\lim_{s_0\to 0^+} I(s_0(\tau))$. If this integral converges, then perfect alignment in reached in finite time. However, if this integral diverges, then it takes infinite time to reach perfect alignment where $(s_0(\tau), Y_0(\tau)) = (0, -\frac{1}{\sqrt{r}})$. As $s_0 \to 0$, the integrand in (45) behaves like $\frac{r^2+1}{s_0^2}$ and hence the integral $I(s_0(\tau))$ diverges. Therefore, the point $(\alpha, y) = (0, -\frac{1}{\sqrt{r}})$ of perfect alignment is only reached in infinite time.

To summarise, the analysis of this section allows us to obtain a full understanding of the effect of ν on the alignment mechanism between two cells. We find that for $\nu = 0$ trajectories do not move towards the point $(\alpha, y) = (0, -1/\sqrt{r})$, and hence perfect alignment is not achieved when there is no self-propulsion. This highlights the importance of self-propulsion in the alignment mechanism and shows that both overlap avoidance and self-propulsion are needed for alignment. For $0 < \nu \ll 1$, we have a singular problem, meaning that there will be some region of initial conditions that lead to perfect alignment no matter how small ν, though the alignment does take infinite time. Finally, computational analysis of the phase portrait and how this varies with ν shows us that alignment decreases with increasing ν. Overall, this demonstrates very similar behaviour to what is seen in the computational results for the collective behaviour of the same system in Leech et al. (2024): little alignment is observed for $\nu = 0$ and then alignment decreases as ν increases, albeit with a peak of maximum alignment around $\nu = 0.2$. Though we do not obtain this maximum value in our analysis, suggesting that this is a consequence of the collective dynamics of the system, the same qualitative dependence of the alignment on ν can be explained by our analytical analysis of the alignment mechanism.

5 Discussion

In this work, we have presented a number of results to help us understand the mechanism between two interacting cells in the model presented in Leech et al. (2024). These results have lead to a more in depth understanding of the mechanisms involved, specifically the interplay between self-propulsion and overlap avoidance and how and when this can lead to the alignment of two self-propelled particles interacting according to an overlap avoidance mechanism.

The main components of the cell-cell interactions are self-propulsion and overlap avoidence in which cells will change their position, orientation and aspect ratio in order to avoid overlap. We derive an analytical framework which leads to a dynamical system, for which there are lots of mathematical tools available to help us to analyse the system. In the full 3D system, which includes shape change, a classic linear stability analysis is degenerate. Using a dynamical asymptotic analysis, we find that two cells being perfectly aligned side-by-side with their natural aspect ratio is a half-stable steady state of the system. This stability analysis also gives us insight into the timescales at play during the cell interactions. We find that the relative sizes of the two non-dimensional parameters determines the ordering of the change in aspect ratio and position, and that the change in orientation always occurs last, over a much slower timescale. We find that this change in orientation is unaffected by the shape change parameter, suggesting that shape change is not essential for alignment to occur. In the computational analysis done in Leech et al. (2024), it is found that allowing for changes in aspect ratio leads to greater alignment. Since we do not see this in the two-cell analysis, we can conclude that this result is due to the collective behaviour of the system in the computational analysis, instead of being due to the underlying alignment mechanism. Lastly, we find that the smaller the non-dimensional shape change parameter is, the faster the cells restore to their natural aspect ratio, after which the system behaves as though there is no shape change. This is reflected in the computational results in Leech et al. (2024) by noting that a strong shape-restoring force leads to little change in behaviour from the rigid cell case. We then consider a 2D limit of the 3D system in more detail, specifically, the limit in which the cells are rigid and keep their natural aspect ratio. This 2D system is more straightforward to analyse in general, and we computationally analyse the effect of the non-dimensional self-propulsion speed on alignment. We find that as this increases, there is a smaller proportion of initial conditions in the domain that lead to perfect alignment, allowing us to conclude that as self-propulsion speed increases, alignment decreases. Together with the fact that no self-propulsion leads to no perfect alignment, this suggests a plausible mechanism for the numerically determined collective behaviour of the many-cell system in Leech et al. (2024) (acknowledging that many of the complex details of the collective behaviour in Leech et al. (2024) are not captured in the analysis for two cells here).

We are able to make a significant amount of analytic progress in the limit of rigid cells. In particular, we are able to understand the singular nature of the small self-propulsion limit, which allows us to better understand how overlap avoidance combined with self-propulsion leads to alignment. We find that the cells are quickly pushed to a single point of overlap over a fast timescale, before the small self-propulsion causes cells to align while maintaining a single point of overlap over a much slower timescale.

The modelling framework presented in Leech et al. (2024) is not specific to ellipse-shaped cells and could be applied to a number of different cell shapes. Provided the points of overlap could be found analytically, a similar approach to analyse and understand the system in greater depth could be taken. It would be interesting to apply the same analytical framework to different cell shapes to see how the results differ. This would be of particular interest if we wanted to apply the modelling framework to bacteria, for example, which tend to be modelled as spherocylinders instead of ellipses (Volfson et al., 2008; You et al., 2021). It would be interesting to see how the geometry of the cell would affect the analysis. A similar approach could also be taken to understand the dynamics of more than two cells with some symmetry imposed in such a way that the overlap points could be analytically found, e.g. three cells placed in such a way that their centres of mass form an equilateral triangle. Finally, looking to other methods of mathematically analysing agent-based models, it would be beneficial to see whether any progress could be made on coarse-graining the model to obtain an equivalent continuum model such that further results on the collective behaviour of the system for many cells could be obtained. We hope the results from this paper will be used to offer a deeper insight into the mechanisms of the model and better understanding of the interplay between self-propulsion and overlap avoidance.

Acknowledgements. This work was supported by the Engineering and Physical Sciences Research Council (grant numbers EP/N509577/1 to VL and EP/T517793/1 to VL) which paid the salary of VL.

Appendix A Full model details

Here we give full details of how the governing analytical equations are derived. For two points of overlap, we have the governing equations

$$\frac{\mathrm{d}\mathbf{X}}{\mathrm{d}t} = -(\mathbf{Y}_1 - \mathbf{Y}_2)^{\perp} + \nu\mathbf{e}(\alpha), \tag{A1a}$$

$$\frac{\mathrm{d}\alpha}{\mathrm{d}t} = \frac{2r}{r^2+1}(|\mathbf{X} - \mathbf{Y}_2|^2 - |\mathbf{X} - \mathbf{Y}_1|^2), \tag{A1b}$$

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \frac{4r^2}{r^2+1}[\sin(2\theta_1) - \sin(2\theta_2)] + 16\gamma\frac{1+\bar{r}r^3}{1+r^2}\left(1 - \frac{r}{\bar{r}}\right), \tag{A1c}$$

where $\mathbf{Y_1} = \mathbf{Y}_B^+$, $\mathbf{Y_2} = \mathbf{Y}_B^-$ and $\begin{pmatrix} \sqrt{r}\cos\theta \\ \frac{1}{\sqrt{r}}\sin\theta \end{pmatrix} = \mathbf{R}(\alpha)^{-1}(\mathbf{Y} - \mathbf{X})$. On substituting these points in, we obtain the governing analytical equations for region B. For four points of overlap, we have the governing equations

$$\frac{\mathrm{d}\mathbf{X}}{\mathrm{d}t} = -(\mathbf{Y}_1 - \mathbf{Y}_2 + \mathbf{Y}_3 - \mathbf{Y}_4)^{\perp} + \nu\mathbf{e}(\alpha), \tag{A2a}$$

$$\frac{\mathrm{d}\alpha}{\mathrm{d}t} = \frac{2r}{r^2+1}(|\mathbf{X} - \mathbf{Y}_2|^2 - |\mathbf{X} - \mathbf{Y}_1|^2 + |\mathbf{X} - \mathbf{Y}_4|^2 - |\mathbf{X} - \mathbf{Y}_3|^2), \tag{A2b}$$

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \frac{4r^2}{r^2+1}[\sin(2\theta_1) - \sin(2\theta_2) + \sin(2\theta_3) - \sin(2\theta_4)] + 16\gamma\frac{1+\bar{r}r^3}{1+r^2}\left(1 - \frac{r}{\bar{r}}\right), \tag{A2c}$$

where $\mathbf{Y}_1 = \mathbf{Y}_B^+$, $\mathbf{Y}_2 = \mathbf{Y}_C^-$, $\mathbf{Y}_3 = \mathbf{Y}_B^-$, and $\mathbf{Y}_4 = \mathbf{Y}_C^+$. Substituting in the analytical form of the points $\mathbf{Y}_B^+, \mathbf{Y}_B^-, \mathbf{Y}_C^+, \mathbf{Y}_C^-$ into (A2) we obtain the governing analytical equations for region C.

Appendix B Full stability analysis for $\gamma = O(1)$

Here we go through the full stability analysis of the steady state $(\alpha, y, r) = (0{-}1/\sqrt{\bar{r}}, \bar{r})$ for $\gamma = O(1)$. This is outlined briefly in the main text. We begin by perturbing the point $(\alpha, y, r) = (0, -1/\sqrt{\bar{r}}, \bar{r})$ by a small amount, defined by

$$\alpha(t) = \varepsilon A(t), \quad y(t) = -\frac{1}{\sqrt{\bar{r}}} + \varepsilon^2 B(t), \quad r(t) = \bar{r} + \varepsilon^2 C(t) \tag{B3}$$

Understanding how $A(t), B(t)$ and $C(t)$ change over time will determine whether the point is stable or unstable. Under the conditions $\varepsilon \ll \gamma \ll 1/\varepsilon$ with $\nu = O(1)$, that are three distinguished timescales of interest in the system: early time ($t = O(\varepsilon)$); intermediate time ($t = O(\gamma)$); and late time ($t = O(1/\varepsilon)$). Over the early timescale, $A(t)$ remains unchanged, and $B(t)$ and $C(t)$ are driven by overlap avoidance. Over the intermediate timescale, $A(t)$ still remains unchanged, $B(t)$ and $C(t)$ are still driven by overlap avoidance, and a restorative force restoring the aspect ratio to $\bar{r}$ comes into the equation for C. Over the late timescale, the dynamics of $A(t)$ become important and are driven by overlap avoidance. We see that the perturbations decay to 0 over the late timescale, proving that the steady state is stable. On substituting (B3) into (10) we obtain

$$\frac{\mathrm{d}A}{\mathrm{d}t} = -\varepsilon 8\frac{\bar{r}^2-1}{\bar{r}^2+1}\bar{r}A\sqrt{(\bar{r}^2-1)A^2 + 2\sqrt{\bar{r}}B - \frac{C}{\bar{r}}} + O(\varepsilon^3), \qquad A(0) = a, \tag{B4a}$$

$$\varepsilon\frac{\mathrm{d}B}{\mathrm{d}t} = \nu A - 2\sqrt{\bar{r}}\sqrt{(\bar{r}^2-1)A^2 + 2\sqrt{\bar{r}}B - \frac{C}{\bar{r}}} + O(\varepsilon^2), \qquad B(0) = b, \tag{B4b}$$

$$\varepsilon\frac{\mathrm{d}C}{\mathrm{d}t} = -\varepsilon\frac{1}{\gamma}\frac{16(1+\bar{r}^4)}{\bar{r}(1+\bar{r}^2)}C + \frac{16\bar{r}^2}{1+\bar{r}^2}\sqrt{(\bar{r}^2-1)A^2 + 2\sqrt{\bar{r}}B - \frac{C}{\bar{r}}} + O(\varepsilon^2), \quad C(0) = c. \tag{B4c}$$

B.1 Early time

We start our analysis over the early time, defined via $t = \varepsilon\tau$ for $\tau = O(1)$. We indicate the early timescale variables with hats and write

$$A(t) = \hat{A}(\tau), \quad B(t) = \hat{B}(\tau), \quad C(t) = \hat{C}(\tau), \quad t = \varepsilon\tau. \tag{B5}$$

We obtain the following leading order equations by substituting (B5) into (10).

$$\frac{\mathrm{d}\hat{A}}{\mathrm{d}\tau} = 0, \qquad \hat{A}(0) = a \tag{B6a}$$

$$\frac{\mathrm{d}\hat{B}}{\mathrm{d}\tau} = \nu\hat{A} - 2\sqrt{\bar{r}}\sqrt{(\bar{r}^2-1)\hat{A}^2 + 2\sqrt{\bar{r}}\hat{B} - \frac{\hat{C}}{\bar{r}}}, \qquad \hat{B}(0) = b \tag{B6b}$$

$$\frac{\mathrm{d}\hat{C}}{\mathrm{d}\tau} = \frac{16\bar{r}^2}{1+\bar{r}^2}\sqrt{(\bar{r}^2-1)\hat{A}^2 + 2\sqrt{\bar{r}}\hat{B} - \frac{\hat{C}}{\bar{r}}}, \qquad \hat{C}(0) = c. \tag{B6c}$$

The first term in equation (B6b) shows that self-propulsion will lead $\hat{B}$ to increase or decrease over time, depending on the sign of a. If $a > 0$ this means that the cell is inclined slightly upwards and hence self-propulsion will cause $\hat{B}$ to increase over time. The second term in (B6b) represents overlap avoidance, suggesting that cells will move apart to avoid overlap and consequently $\hat{B}$ will decrease. The size of these two terms will determine whether $\hat{B}_\tau$ is positive or negative. $\hat{C}_\tau > 0$ in (B6c) means that the cell elongates over this timescale. This is because the orientation is near zero, so the cells are side by side and consequently elongate to avoid overlap. From (B6) we can solve (B6a) directly, and then combine (B6b) and (B6c) and integrate to obtain

$$\hat{A}(\tau) = a, \tag{B7a}$$

$$\hat{B}(\tau) + \frac{1+\bar{r}^2}{8\bar{r}^{3/2}}\hat{C}(\tau) = \nu a\tau + b + \frac{1+\bar{r}^2}{8\bar{r}^{3/2}}c. \tag{B7b}$$

Equation (B7a) shows that $\hat{A}$ remains constant over the early timescale, meaning that the orientation is not affected over this timescale. Equation (B7b) shows that $\hat{B}$ and $\hat{C}$ both grow linearly in time. We also find that, while $\hat{B}$ and $\hat{C}$ grow in time, the quantity

$$2\sqrt{\bar{r}}B - \frac{C}{\bar{r}} \sim \frac{\nu^2(1+\bar{r}^2)^2}{4\bar{r}(5+\bar{r}^2)^2}a^2 - (\bar{r}^2-1)a^2 \text{ as } \tau \to \infty. \tag{B8}$$

Specifically, this means that $2\sqrt{\bar{r}}\hat{B} - \frac{\hat{C}}{\bar{r}}$ tends to a constant as τ grows. Combining (B8), with the algebraic constraint for $\hat{B}$ and $\hat{C}$ in (B7b) we are able to solve for $\hat{B}$ and $\hat{C}$ (as $\tau \to \infty$) and deduce that

$$\hat{A}(\tau) = a, \tag{B9a}$$

$$\hat{B}(\tau) \sim \frac{4\nu a}{5+\bar{r}^2}\tau \text{ as } \tau \to \infty, \tag{B9b}$$

$$\hat{C}(\tau) \sim \frac{8\bar{r}^{3/2}\nu a}{5+\bar{r}^2}\tau \text{ as } \tau \to \infty, \tag{B9c}$$

The fact that $\hat{B}$ and $\hat{C}$ grow in time indicates the scaling for the next time region

B.2 Intermediate time.

We now consider the intermediate timescale defined via $t = \gamma\bar{t}$, where $\bar{t} = O(1)$, recalling that $\varepsilon \ll \gamma \ll 1/\varepsilon$. We also redefine our scalings for B and C as

$$\alpha(t) = \varepsilon\bar{A}(\bar{t}), \quad y(t) = -\frac{1}{\sqrt{\bar{r}}} + \gamma\varepsilon\bar{B}(\bar{t}), \quad r(t) = \bar{r} + \gamma\varepsilon\bar{C}(\bar{t}), \quad t = \gamma\bar{t}, \tag{B10}$$

motivated by the large-τ results (B9) in the early time in Sec. B.1. Substituting (B10) into the governing equations (10) we obtain

$$\frac{\mathrm{d}\bar{A}}{\mathrm{d}\bar{t}} = -\gamma 8\frac{\bar{r}^2-1}{\bar{r}^2+1}\bar{r}\bar{A}\sqrt{\varepsilon^2(\bar{r}^2-1)\bar{A} + \gamma\varepsilon 2\sqrt{\bar{r}}\bar{B} - \gamma\varepsilon\frac{\bar{C}}{\bar{r}}}, \tag{B11a}$$

$$\frac{\mathrm{d}\bar{B}}{\mathrm{d}\bar{t}} = \nu\bar{A} - \frac{2\sqrt{\bar{r}}}{\varepsilon}\sqrt{\varepsilon^2(\bar{r}^2-1)\bar{A} + \gamma\varepsilon 2\sqrt{\bar{r}}\bar{B} - \gamma\varepsilon\frac{\bar{C}}{\bar{r}}}, \tag{B11b}$$

$$\frac{\mathrm{d}\bar{C}}{\mathrm{d}\bar{t}} = -\frac{16(1+\bar{r}^4)}{\bar{r}(1+\bar{r}^2)}\bar{C} + \frac{16\bar{r}^2}{\varepsilon(1+\bar{r}^2)}\sqrt{\varepsilon^2(\bar{r}^2-1)\bar{A} + \gamma\varepsilon 2\sqrt{\bar{r}}\bar{B} - \gamma\varepsilon\frac{\bar{C}}{\bar{r}}}. \tag{B11c}$$

To leading order we obtain

$$0 = -\gamma 8\frac{\bar{r}^2-1}{\bar{r}^2+1}\bar{r}\bar{A}\sqrt{\gamma\varepsilon 2\sqrt{\bar{r}}\bar{B} - \gamma\varepsilon\frac{\bar{C}}{\bar{r}}}, \tag{B12a}$$

$$0 = 2\sqrt{\bar{r}}\sqrt{\gamma\varepsilon 2\sqrt{\bar{r}}\bar{B} - \gamma\varepsilon\frac{\bar{C}}{\bar{r}}}, \tag{B12b}$$

$$0 = \frac{16\bar{r}^2}{1+\bar{r}^2}\sqrt{\gamma\varepsilon 2\sqrt{\bar{r}}\bar{B} - \gamma\varepsilon\frac{\bar{C}}{\bar{r}}}. \tag{B12c}$$

Solving (B12) we see that we only obtain one piece of information:

$$2\sqrt{\bar{r}}\bar{B}(\bar{t}) = \frac{\bar{C}(\bar{t})}{\bar{r}}. \tag{B13}$$

To obtain the additional information required, a formal analysis would involve expanding $\bar{B}$ and $\bar{C}$ in higher powers of ε, and deriving appropriate solvability conditions. However, we can circumvent this laborious calculation by instead combining the equations in (B11) to remove the duplication of information at leading order (essentially the square root term in each equation), and generate a system of distinct ODEs

for $\bar{A}, \bar{B}$ and $\bar{C}$. We can eliminate the square root term in (B11a), (B11b) and (B11c) to obtain

$$\frac{\mathrm{d}\bar{A}}{\mathrm{d}\bar{t}} = \varepsilon\gamma 4\sqrt{\bar{r}}\frac{\bar{r}^2-1}{\bar{r}^2+1}\bar{A}\left(\frac{\mathrm{d}\bar{B}}{\mathrm{d}\bar{t}} - \nu\bar{A}\right), \tag{B14a}$$

$$\frac{\mathrm{d}\bar{B}}{\mathrm{d}\bar{t}} = \nu\bar{A} - \frac{1+\bar{r}^2}{8\bar{r}^{3/2}}\left(\frac{\mathrm{d}\bar{C}}{\mathrm{d}\bar{t}} + \frac{16(1+\bar{r}^4)}{\bar{r}(1+\bar{r}^2)}\bar{C}\right), \tag{B14b}$$

$$\frac{\mathrm{d}\bar{C}}{\mathrm{d}\bar{t}} = -\frac{16(1+\bar{r}^4)}{\bar{r}(1+\bar{r}^2)}\bar{C} - \frac{8\bar{r}^{3/2}}{1+\bar{r}^2}\left(\frac{\mathrm{d}\bar{B}}{\mathrm{d}\bar{t}} - \nu\bar{A}\right). \tag{B14c}$$

We can then decouple (B14b) and (B14c) using (B13). On doing this, to leading order we obtain

$$\frac{\mathrm{d}\bar{A}}{\mathrm{d}\bar{t}} = 0, \qquad \lim_{\hat{t}\to 0^+} \bar{A}(\bar{t}) = a, \tag{B15a}$$

$$\frac{\mathrm{d}\bar{B}}{\mathrm{d}\bar{t}} = \nu\bar{A} - \frac{1+\bar{r}^2}{4}\left(\frac{\mathrm{d}\bar{B}}{\mathrm{d}\bar{t}} + \frac{16(1+\bar{r}^4)}{\bar{r}(1+\bar{r}^2)}\bar{B}\right), \qquad \lim_{\hat{t}\to 0^+} B(\bar{t}) = 0 \tag{B15b}$$

$$\frac{\mathrm{d}\bar{C}}{\mathrm{d}\bar{t}} = 2\bar{r}^{3/2}\nu\bar{A} - \frac{1+\bar{r}^2}{4}\left(\frac{\mathrm{d}\bar{C}}{\mathrm{d}\bar{t}} + \frac{16(1+\bar{r}^4)}{\bar{r}(1+\bar{r}^2)}\bar{C}\right), \qquad \lim_{\hat{t}\to 0^+} C(\bar{t}) = 0. \tag{B15c}$$

The initial conditions are matching conditions from the previously solved early timescale τ. Solving (B15) we obtain

$$\bar{A}(\bar{t}) = a, \tag{B16a}$$

$$\bar{B}(\bar{t}) = \frac{\nu a\bar{r}}{4(1+\bar{r}^4)}\left(1 - \exp\left(-\frac{16(1+\bar{r}^4)}{\bar{r}(5+\bar{r}^2)}\bar{t}\right)\right), \tag{B16b}$$

$$\bar{C}(\bar{t}) = \frac{\nu a\bar{r}^{5/2}}{2(1+\bar{r}^4)}\left(1 - \exp\left(-\frac{16(1+\bar{r}^4)}{\bar{r}(5+\bar{r}^2)}\bar{t}\right)\right). \tag{B16c}$$

From (B16) we see that, again, $\bar{A}$ remains constant over the intermediate time, and that $\bar{B}$ and $\bar{C}$ decay exponentially in time, reaching constant values given by

$$\bar{A}(\bar{t}) = a, \quad \bar{B}(\bar{t}) \to \frac{\nu a\bar{r}}{4(1+\bar{r}^4)}, \quad \bar{C}(\bar{t}) \to \frac{\nu a\bar{r}^{5/2}}{2(1+\bar{r}^4)}, \tag{B17}$$

as $\bar{t} \to \infty$. These will be used as matching conditions for the next timescale.

B.3 Late time.

Finally, we consider the late timescale defined via $t = \tilde{t}/\varepsilon$, where $\tilde{t} = O(1)$. Given the large-$\bar{t}$ results (B17), we retain the scalings from the intermediate time region in (B10).

$$\alpha(t) = \varepsilon\tilde{A}(\tilde{t}), \quad y(t) = -\frac{1}{\sqrt{\bar{r}}} + \gamma\varepsilon\tilde{B}(\tilde{t}), \quad r(t) = \bar{r} + \gamma\varepsilon\tilde{C}(\tilde{t}), \quad t = \tilde{t}/\varepsilon, \tag{B18}$$

This is the timescale over which the dynamics of A become important. Over this timescale equations (10) become

$$\frac{\mathrm{d}\tilde{A}}{\mathrm{d}\tilde{t}} = -\frac{8}{\varepsilon}\frac{\bar{r}^2-1}{\bar{r}^2+1}\bar{r}\tilde{A}\sqrt{\varepsilon^2(\bar{r}^2-1)\tilde{A}+\gamma\varepsilon 2\sqrt{\bar{r}}\tilde{B}-\gamma\varepsilon\frac{\tilde{C}}{\bar{r}}}, \tag{B19a}$$

$$\frac{\mathrm{d}\tilde{B}}{\mathrm{d}\tilde{t}} = \frac{1}{\varepsilon}\nu\tilde{A} - \frac{2}{\varepsilon^2}\sqrt{\bar{r}}\sqrt{\varepsilon^2(\bar{r}^2-1)\tilde{A}+\gamma\varepsilon 2\sqrt{\bar{r}}\tilde{B}-\gamma\varepsilon\frac{\tilde{C}}{\bar{r}}}, \tag{B19b}$$

$$\frac{\mathrm{d}\tilde{C}}{\mathrm{d}\tilde{t}} = -\frac{1}{\varepsilon}\frac{16(1+\bar{r}^4)}{\bar{r}(1+\bar{r}^2)}\tilde{C} + \frac{1}{\varepsilon^2}\frac{16\bar{r}^2}{1+\bar{r}^2}\sqrt{\varepsilon^2(\bar{r}^2-1)\tilde{A}+\gamma\varepsilon 2\sqrt{\bar{r}}\tilde{B}-\gamma\varepsilon\frac{\tilde{C}}{\bar{r}}}. \tag{B19c}$$

To leading order we obtain

$$0 = -8\frac{\bar{r}^2-1}{\bar{r}^2+1}\bar{r}\tilde{A}\sqrt{\gamma 2\sqrt{\bar{r}}\tilde{B}-\gamma\frac{\tilde{C}}{\bar{r}}}, \tag{B20a}$$

$$0 = -2\sqrt{\bar{r}}\sqrt{\gamma 2\sqrt{\bar{r}}\tilde{B}-\gamma\frac{\tilde{C}}{\bar{r}}}, \tag{B20b}$$

$$0 = \frac{16\bar{r}^2}{1+\bar{r}^2}\sqrt{\gamma 2\sqrt{\bar{r}}\tilde{B}-\gamma\frac{\tilde{C}}{\bar{r}}}. \tag{B20c}$$

Again, we obtain only one piece of information:

$$2\sqrt{\bar{r}}\tilde{B}(\tilde{t}) = \frac{\tilde{C}(\tilde{t})}{\bar{r}} \tag{B21}$$

Since this information is duplicated in all three equations, we would have to expand $\tilde{B}$ and $\tilde{C}$ in higher orders of ε to obtain solvability conditions. Alternatively, as in the intermediate time region, we can rearrange the equations to eliminate the square root terms and generate a system of ODEs for $\tilde{A}, \tilde{B}$ and $\tilde{C}$. On doing so we obtain

$$\frac{\mathrm{d}\tilde{A}}{\mathrm{d}\tilde{t}} = -4\nu\sqrt{\bar{r}}\frac{\bar{r}^2-1}{\bar{r}^2+1}\tilde{A}^2 + \varepsilon 4\sqrt{\bar{r}}\frac{\bar{r}^2-1}{\bar{r}^2+1}\tilde{A}\frac{\mathrm{d}\tilde{B}}{\mathrm{d}\tilde{t}}, \tag{B22a}$$

$$\frac{\mathrm{d}\tilde{B}}{\mathrm{d}\tilde{t}} = \frac{1}{\varepsilon}\nu\tilde{A} - \frac{1+\bar{r}^2}{8\bar{r}^{3/2}}\frac{\mathrm{d}\tilde{C}}{\mathrm{d}\tilde{t}} - \frac{1}{\varepsilon}\frac{2(1+\bar{r}^4)}{\bar{r}^{5/2}}\tilde{C}, \tag{B22b}$$

$$\frac{\mathrm{d}\tilde{C}}{\mathrm{d}\tilde{t}} = -\frac{1}{\varepsilon}\frac{16(1+\bar{r}^2)}{\bar{r}(1+\bar{r}^2)}\tilde{C} + \frac{1}{\varepsilon}\frac{8\bar{r}^{3/2}}{1+\bar{r}^2}\nu\tilde{A} - \frac{8\bar{r}^{3/2}}{1+\bar{r}^2}\frac{\mathrm{d}\tilde{B}}{\mathrm{d}\tilde{t}}. \tag{B22c}$$

To leading order we have

$$\frac{\mathrm{d}\tilde{A}}{\mathrm{d}\tilde{t}} = -4\nu\sqrt{\bar{r}}\frac{\bar{r}^2-1}{\bar{r}^2+1}\tilde{A}^2, \qquad \lim_{\tilde{t}\to 0^+}\tilde{A}(\tilde{t}) = a \tag{B23a}$$

$$0 = \nu\tilde{A} - \frac{2(1+\bar{r}^4)}{\bar{r}^{5/2}}\tilde{C}, \qquad \lim_{\tilde{t}\to 0^+} \tilde{B}(\tilde{t}) = \frac{\nu a\bar{r}}{4(1+\bar{r}^4)} \tag{B23b}$$

$$0 = -\frac{16(1+\bar{r}^4)}{\bar{r}(1+\bar{r}^2)}C + \frac{8\bar{r}^{3/2}}{1+\bar{r}^2}, \qquad \lim_{\tilde{t}\to 0^+} \tilde{C}(\tilde{t}) = \frac{\nu a\bar{r}^{5/2}}{2(1+\bar{r}^4)}, \tag{B23c}$$

where the matching conditions come from the long time behaviour (B17) in the previous timescale. Solving (B23) using (B21) we find that

$$\tilde{A}(\tilde{t}) = \frac{a(1+\bar{r}^2)}{4\nu a\sqrt{\bar{r}}(\bar{r}^2-1)\tilde{t} + \bar{r}^2 + 1}, \tag{B24a}$$

$$\tilde{B}(\tilde{t}) = \frac{\nu\bar{r}}{4(1+\bar{r}^4)}\tilde{A}(\tilde{t}), \tag{B24b}$$

$$\tilde{C}(\tilde{t}) = \frac{\nu\bar{r}^{5/2}}{2(1+\bar{r}^4)}\tilde{A}(\tilde{t}). \tag{B24c}$$

We can see that $\tilde{A}(\tilde{t})$, $\tilde{B}(\tilde{t})$ and $\tilde{C}(\tilde{t})$ all decay to zero algebraically as $\tilde{t} \to \infty$ for $a > 0$, but diverge for $a < 0$. This shows that the point we are interested in is half-stable. We note that this analysis for $\gamma = O(1)$ is just a splitting of the early timescale in the analysis done in the main text.

References

Albi, G., Pareschi, L.: Modeling of self-organized systems interacting with a few individuals: from microscopic to macroscopic dynamics. Appl. Math. Lett. **26**(4), 397–401 (2013) https://doi.org/10.1016/j.aml.2012.10.011

Bolley, F., Cañizo, J.A., Carrillo, J.A.: Mean-field limit for the stochastic vicsek model. Appl. Math. Lett. **25**(3), 339–343 (2012) https://doi.org/10.1016/j.aml.2011.09.011

Balagam, R., Igoshin, O.A.: Mechanism for collective cell alignment in myxococcus xanthus bacteria. PLoS Comput. Biol. **11**(8), 1004474 (2015) https://doi.org/10.1371/journal.pcbi.1004474

Baskaran, A., Marchetti, M.C.: Enhanced diffusion and ordering of self-propelled rods. Phys. Rev. Lett. **101**(26), 268101 (2008) https://doi.org/10.1103/PhysRevLett.101.268101

Dartsch, P., Betz, E.: Response of cultured endothelial cells to mechanical stimulation. Basic Res. Cardiol. **84**(3), 268–281 (1989) https://doi.org/10.1007/BF01907974

Deutsch, A., Dormann, S.: Mathematical Modeling of Biological Pattern Formation. Springer, New York (2005). https://doi.org/10.1007/978-1-4613-0133-2

Degond, P., Motsch, S.: Continuum limit of self-driven particles with orientation interaction. Math. Models Methods Appl. Sci. **18**(supp01), 1193–1215 (2008) https://doi.org/10.1142/S0218202508003005

Erdogan, B., Ao, M., White, L.M., Means, A.L., Brewer, B.M., Yang, L., Washington, M.K., Shi, C., Franco, O.E., Weaver, A.M., *et al.*: Cancer-associated fibroblasts promote directional cancer cell migration by aligning fibronectin. J. Cell Bio. **216**(11), 3799–3816 (2017) https://doi.org/10.1083/jcb.201704053

Fay, T.H., Joubert, S.V.: Separatrices. Int. J. Math. Educ. Sci. Technol. **41**(3), 412–418 (2010) https://doi.org/10.1080/00207390903564637

Großmann, R., Peruani, F., Bär, M.: Mesoscale pattern formation of self-propelled rods with velocity reversal. Phys. Rev. E **94**(5), 050602 (2016) https://doi.org/10.1103/PhysRevE.94.050602

Herbert-Read, J.E., Perna, A., Mann, R.P., Schaerf, T.M., Sumpter, D.J., Ward, A.J.: Inferring the rules of interaction of shoaling fish. Proc. Natl. Acad. Sci. **108**(46), 18726–18731 (2011) https://doi.org/10.1073/pnas.1109355108

Kraikivski, P., Lipowsky, R., Kierfeld, J.: Enhanced ordering of interacting filaments by molecular motors. Phys. Rev. Lett. **96**(25), 258103 (2006) https://doi.org/10.1103/PhysRevLett.96.258103

Kenny, F.N., Marcotti, S., De Freitas, D.B., Drudi, E.M., Leech, V., Bell, R.E., Easton, J., Fleck, R., Allison, L., Philippeos, C., *et al.*: Autocrine il-6 drives cell and extracellular matrix anisotropy in scar fibroblasts. Matrix Biol. **123**, 1–16 (2023) https://doi.org/10.1016/j.matbio.2023.08.004

Lopez, U., Gautrais, J., Couzin, I.D., Theraulaz, G.: From behavioural analyses to models of collective motion in fish schools. Interface Focus **2**(6), 693–707 (2012) https://doi.org/10.1098/rsfs.2012.0033

Leech, V., Kenny, F., Marcotti, S., Shaw, T., Stramer, B., Manhart, A.: Derivation and simulation of a computational model of active cell populations: How overlap avoidance, deformability, cell-cell junctions and cytoskeletal forces affect alignment. bioRxiv, 2024–02 (2024)

Peruani, F., Deutsch, A., Bär, M.: Nonequilibrium clustering of self-propelled rods. Phys. Rev. E. Stat. Nonlin. Soft Matter Phys. **74**(3), 030904 (2006) https://doi.org/10.1103/PhysRevE.74.030904

Volfson, D., Cookson, S., Hasty, J., Tsimring, L.S.: Biomechanical ordering of dense cell populations. Proc. Natl. Acad. Sci. **105**(40), 15346–15351 (2008) https://doi.org/10.1073/pnas.0706805105

Van Dyke, M.: Perturbation Methods in Fluid Mechanics. Stanford, CA: Parabolic Press, California (1975)

You, Z., Pearce, D.J., Giomi, L.: Confinement-induced self-organization in growing bacterial colonies. Sci. Adv. **7**(4), 8685 (2021) https://doi.org/10.1126/sciadv.

abc8685