

1 Derivation and simulation of a computational model of active cell
2 populations: How overlap avoidance, deformability, cell-cell
3 junctions and cytoskeletal forces affect alignment

4 Vivienne Leech¹, Fiona N Kenny², Stefania Marcotti², Tanya J Shaw², Brian M Stramer²,
5 and Angelika Manhart^{1,3}

6 ¹Department of Mathematics, University College London, London, UK

7 ²Randall Centre for Cell and Molecular Biophysics, King's College London, London, UK

8 ³Faculty of Mathematics, University of Vienna, Vienna, Austria

9 **Abstract**

10 Collective alignment of cell populations is a commonly observed phenomena in biology. An important
11 example are aligning fibroblasts in healthy or scar tissue. In this work we derive and simulate a mechanistic
12 agent-based model of the collective behaviour of actively moving and interacting cells, with a focus
13 on understanding collective alignment. The derivation strategy is based on energy minimisation. The
14 model ingredients are motivated by data on the behaviour of different populations of aligning fibroblasts
15 and include: Self-propulsion, overlap avoidance, deformability, cell-cell junctions and cytoskeletal forces.
16 We find that there is an optimal ratio of self-propulsion speed and overlap avoidance that maximises
17 collective alignment. Further we find that deformability aids alignment, and that cell-cell junctions by
18 themselves hinder alignment. However, if cytoskeletal forces are transmitted via cell-cell junctions we
19 observe strong collective alignment over large spatial scales.

20 **Author summary.** *Collective dynamics is the study of how the interaction of individual animals, humans,
21 cells, etc. can lead to patterns on a scale much larger than the individuals themselves. Prominent examples
22 are flocking birds, schools of fish or the patterns of bacterial colonies. Since the behaviour of the group is very
23 difficult to intuit from individual interaction rules, mathematical models are invaluable in testing hypotheses,
24 making predictions, and suggesting explanations. In our collaborative work between mathematicians and
25 biologists, we use a mathematical model, whose ingredients are motivated by the experimental observation
26 of fibroblasts. Fibroblasts are cells that are part of the connective tissue in animals and play an important
27 role e.g. in wound healing. Under certain conditions, e.g. in scar tissue or near tumours, these spindle-
28 shaped cells form strongly aligned populations on large spatial scales for unclear reasons. By simulating our*

29 mathematical model, we predict how the cells' self-propulsion, overlap avoidance, deformability, and other
30 interactions influence the alignment dynamics. The results help to understand alignment of cell populations
31 and apply to many other cells or organisms.

32 1 Introduction

33 **The challenge of active particles.** The ability of particles to align with their neighbours is observed
34 in many contexts and in many scales in biology. Famous examples include flocking of birds, schools of fish
35 or even motion of large groups of people [1, 2, 3]. Here we focus on alignment on a cellular scale. Cell
36 alignment has been observed in bacterial swarms [4], e.g. in myxobacteria [5], as well as in amoeboid cells
37 [6], or fibroblasts [7]. Alignment of different types of particles has also been observed and modelled in
38 physics [8, 9, 10], however alignment in cells is more challenging to understand, since cells can exhibit much
39 more complicated behaviour than passive matter, see e.g. reviews [11, 12]. In particular, cells can actively
40 self-propel, can change their shape, can interact biochemically via signalling and mechanically via physical
41 junctions with their neighbours and with the extracellular matrix, with several levels of potential feedback
42 involved. Physically, these systems can be thought of as consisting of agents which are able to convert energy
43 into movement. The ability to do this, combined with interactions between neighbouring particles can give
44 rise to collective behaviour, often in the form of alignment. The novelty of this is that energy input into
45 the system is on a local scale, which pushes the system out of equilibrium leading to changes which can
46 propagate through the whole system and lead to emergent structures and behaviour. Understanding the
47 details of how this works, and looking for theoretical descriptions of living matter has been of interest for a
48 number of years [13].

49 **Causes of alignment.** Larger, more complex organisms, such as birds or fish, are typically able to perceive
50 their neighbours using senses such as sight and can adjust their own direction and speed correspondingly.
51 Some interesting work for causes of alignment for such species include [14, 15]. Cells, the main focus of this
52 work, can perceive their surroundings in several different ways. The main ones include chemotaxis, durotaxis,
53 signalling or mechanics. Several experimental works [16, 17] have shown that e.g. the extracellular matrix
54 (ECM) can transmit forces, aiding cytoskeletal alignment. This has been modelled e.g. in [18, 19]. Further,
55 the shape of the environment itself can influence alignment by introducing constraints to cell orientation
56 on the boundary [7]. Finally, contact-based alignment can be caused by particles avoiding overlap. For
57 movement in e.g. a fluid this would be a hard constraint, while for crawling cells "overlap" could imply cells
58 moving on top of each other.

59 **Basic types of alignment models.** Dynamics and patterns emerging from the interactions of many
60 individuals are difficult to intuit from microscopic interaction rules, hence mathematical modelling and
61 simulations are a powerful tool to shed light on the involved mechanisms. One common model type is

continuum models, where the system is described in terms of continuous space and time dependent macroscopic quantities like cell density, mean direction, etc. The book [20] offers an excellent overview of models and biomedical applications. Continuum models have the advantage of ease of analysis and have been used extensively to investigate alignment and pattern formation in cell populations, for some examples see [21, 22, 23, 24, 25, 26, 27]. Agent-based models, on the other hand, where cells are discrete objects (“agents”) and each cell is equipped with its own set of equations, are particularly suitable for mechanistic hypothesis testing, since biological assumptions can be translated in a relatively straight-forward way (see e.g. reviews [28, 29]). Amongst the classical agent-based models for flocking ([30, 31]), the most famous agent-based alignment model is probably the Vicsek model [32]. This model assumes self-propelling particles align their orientation with their neighbours and produces large scale alignment whenever the alignment force is large compared to the orientational noise. While the Vicsek model and its variants have been applied to many biological problems [33, 34], it isn’t suitable to test mechanisms of alignment, since alignment is already a model ingredient.

Shape deformations. For individual cells or small groups of cells, there exist several modelling frameworks capable of describing cell shapes in a flexible and biologically well-motivated manner. Prominent approaches include 1. Phase-field models [35, 36, 37, 38, 39], where the cell in- and outsides are characterised by a continuous, but steep phase-field variable, 2. Models using the immersed boundary method [40, 41, 42], where the cell boundary is an explicit curve or surface interacting with the surrounding fluid, 3. Cellular Potts models [43, 44, 45, 46], where each cell is a collection of pixels, whose dynamics follow an energy minimization, or 4. Vertex models [47, 48, 49, 50], which describe sheets of cells via cell-cell boundaries. While phase-field and immersed boundary models allow for the description of a large class of cell shapes, they are also computationally costly and hence less suitable to investigate large numbers of cells. Cellular Potts models are not as computationally costly as they used to be (see e.g. Morpheus [51]), however the fact that they are lattice based makes analytical studies as well as the inclusion of mechanical effects more challenging. The effect of self-propulsion and cell-cell adhesions in Cellular Potts models have been investigated e.g. in [52], however it is unclear how the actin cables in Sec 4 could be included in this framework. Vertex models, on the other hand, are mostly used for simulating tissue dynamics and are less suitable for describing individual cell movement. An example that explores cell alignment in vertex models includes the work in [50]. This is a suitable approach in the context of epithelial monolayers, however the model varies from ours in the fact that cell overlap and cell-cell collisions are not included. A similar approach to ours is described in [53, 54], where populations of deformable ellipsoidal cells in 3D are modelled, which experience changes in aspect ratio as a result of cell overlap.

This work and paper overview. In this work, we will mathematically model alignment of collectives of cells moving in two space dimensions (2D). As the main cause of alignment, we assume cells want to avoid

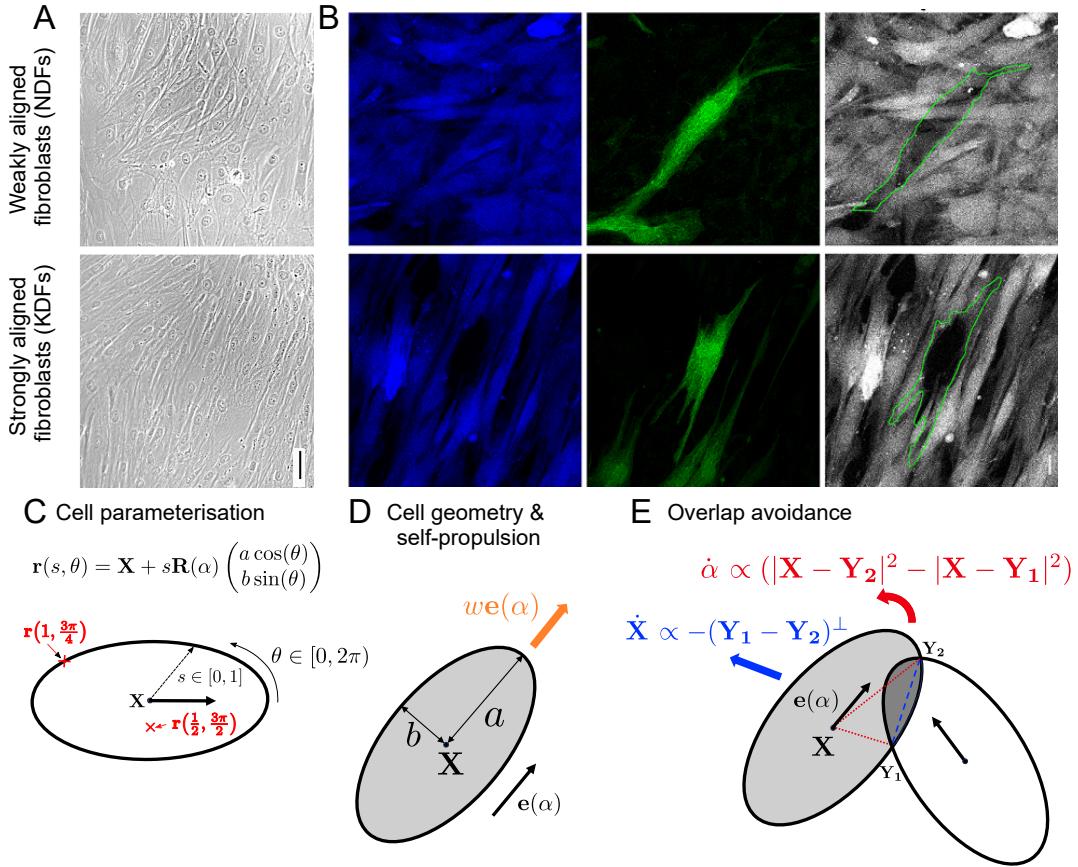


Figure 1: A,B: Experimental figures for weakly aligned normal dermal fibroblasts (NDFs), top row, and strongly aligned keloid derived fibroblasts (KDFs), bottom row. A: Phase microscopy pictures, scale bar $50 \mu\text{m}$. B: Mosaically labelled cells with two different probes in order to investigate cell overlap (left: CellTrace Violet, middle: CellTrace Green, right: overlay), scale bar $20 \mu\text{m}$. C,D,E: Model schematics. C: Ellipse parameterisation. D: Elliptic cell geometry of a single cell with centre \mathbf{X} , dimensions a and b , orientation α and self-propulsion speed w . E: Effect of overlap avoidance upon one cell overlapping with another cell.

overlap in a tuneable (i.e. non-perfect) way (Sec. 2). Motivated by biologically observed, distinct behaviour of different populations of fibroblasts [55], we will assess the influence of the following characteristics of active matter on alignment: Self-propulsion (Sec. 2), deformability (Sec. 3), cell-cell junctions (Sec. 4) and cytoskeletal forces (Sec. 4).

2 The base model: Self-propulsion and overlap avoidance

2.1 Biological background & model ingredients

Experimental motivation. While alignment processes in active particles are relevant in many contexts, we will focus on the particular example of fibroblasts. Fibroblasts are cells in the connective tissue in animals and are responsible for making and remodelling the ECM. Fibroblast and ECM alignment is observed during

105 various scarring pathologies. In [55] we have investigated the difference in alignment behaviour of fibroblasts
106 in healthy tissue (normal dermal fibroblasts, NDFs) as compared to dermofibroblasts in certain scar tissue
107 (keloid derived fibroblasts, KDFs). KDFs were found to show stronger alignment over larger length scales,
108 Fig. 1A,B. Further we found that KDFs show less tendency to crawl on top of each other and form aligned
109 supracellular actin bundles via cell-cell junctions, spanning multiple cells. Using mathematical modelling,
110 we found in [55] that the increase in overlap avoidance can explain the stronger alignment. In this work we
111 will use the differences found in NDFs and KDFs to motivate further model extensions. However, the model
112 ingredients are applicable to other cells and situations and hence the findings are relevant beyond fibroblasts.

113 **Model ingredients.** We build a mechanistic, agent-based model to describe the motion of individual cells
114 interacting with neighbouring cells in 2D, where cells are approximated as ellipses. While typically elongated,
115 real fibroblast shapes are of course more complex and include ruffles, and protruding and contracting
116 lamellipodia [56, 57]. We argue that, over the timescale we are interested in, ellipses are a rough, but appropriate
117 approximation of real cell shapes. We also choose ellipses, because they allow for a straight-forward
118 description of dynamic cell shape changes (see Sec. 3). Further our modelling approach can be extended
119 to more complex cell shapes, which will be the subject of future work. In this work we do not model cell
120 divisions, however, they could be included in a straight-forward manner.

- 121
- 122 The key model ingredients are
- 123 • *Environmental friction:* As usual in cell biology, we assume a friction-dominated regime. As a consequence, velocities (not accelerations) are proportional to forces. The strength of the friction with the substrate is given by η , which effectively sets a time scale.
 - 126 • *Self-propulsion:* In the absence of interactions, cells move with fixed speed w in the direction of their orientation. Orientational noise could be included in a straight forward manner, but is omitted in this work for the sake of simplicity.
 - 129 • *Overlap avoidance:* When placed on a 2D substrate, many cell types will tend to avoid moving on top of other cells. The term *contact inhibition of locomotion* is sometimes used in this context. However, contact inhibition of locomotion more commonly refers to an active change of direction upon contact, as opposed to a more passive reaction, which is what we model here. Another commonly used term in this context is *repulsion*, however, in this work will use the term *overlap avoidance* to emphasise that the effect is short-ranged and driven by cell overlap. Note that “overlap” in 2D can be interpreted either as being positioned partly on top of each other, or allowing for some cell softness. We allow overlap avoidance to be tuneable, its strength given by parameter σ . If $\sigma = 0$, cells have no overlap avoidance, and for $\sigma \rightarrow \infty$, cells would behave as solid objects that never overlap/move on top of each other. To avoid overlap cells can

- 139 – move to change their location,
 140 – turn to change their orientation, or
 141 – change their shape (\rightarrow Sec. 3).

142 All these effects will be a consequence of the minimisation of a common energy term.

- 143 • *Cell-cell junctions:* In Sec. 4 we model and investigate the effect of cell-cell junctions, where cells are
 144 elastically tethered to each other, which can affect their orientation and position.
 145 • *Actin forces:* Also in Sec. 4 we describe the presumed effect of supracellular actin cables that lead to
 146 cytoskeletal forces affecting cell orientation.

147 In this work we explore the effect of different parameters on the model without the inclusion of additional
 148 noise. However, we have added a short study on the effect of orientational noise in S1 Appendix, Sec. 3 and
 149 in [55] we fitted the noise parameter to experimental data. In principle, other sources of stochasticity could
 150 be included in the model to e.g. explore randomly occurring protrusions.

151 2.2 Model derivation

152 We will show the derivation excluding cell deformability and cell-cell junctions. These will be considered
 153 in Sec. 3 and Sec. 4. More derivation details can be found in S1 Appendix, Sec. 1 and a summary of
 154 model parameter names and meaning can be found in S1 Appendix Tab. 1. We consider N cells within
 155 the fixed domain $\Omega \in \mathbb{R}^2$, each with centroid position $\mathbf{X}_i = (X_i, Y_i) \in \mathbb{R}^2, i = 1, \dots, N$ and orientation
 156 $\alpha_i \in [0, 2\pi), i = 1, \dots, N$. Each cell is described by an ellipse with semi-major axis a and semi-minor axis b
 157 as shown in Fig. 1D. The cell's area is given by $A = ab\pi$. In the absence of other cells, each cell self-propels
 158 with constant velocity w in direction $\mathbf{e}(\alpha_i) = (\cos(\alpha_i), \sin(\alpha_i))^T$, where superscript T denotes the transpose.
 159 We assume that this self propulsion encompasses the mechanism of an individual cell moving in a directed
 160 way across the substrate. An alternative approach would be to model cell motion as a persistent random
 161 walk, as has been done in [58, 59].

162
 We derive the governing equations using energy minimisation. An alternative model derivation based on
 force balance, which leads to the same governing equations, can be found in S1 Appendix, Sec. 1. Focusing
 on one cell positioned at $\mathbf{X}(t)$ with orientation $\alpha(t)$ at time t , we parameterise the points inside the cell, \mathbf{x} ,
 by

$$\mathbf{x}(t, s, \theta) = \mathbf{X}(t) + s\mathbf{R}(\alpha(t))\mathbf{k}(\theta), \quad s \in [0, 1], \theta \in [0, 2\pi), \quad (1)$$

where s encodes the distance from the centre of the cell \mathbf{X} to the point \mathbf{x} as a proportion of the distance
 from the centre of the cell to the boundary and θ encodes the angle parameter, see Fig. 1C. The rotation

matrix $\mathbf{R}(\alpha)$ and the shape vector $\mathbf{k}(\theta)$ are defined by

$$\mathbf{R}(\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \quad \text{and} \quad \mathbf{k}(\theta) = \begin{pmatrix} a \cos(\theta) \\ b \sin(\theta) \end{pmatrix}.$$

We assume that at every time step Δt , the system minimizes a total energy E_{tot} , which, for the base model, is the sum of contributions from friction E_{friction} , from overlap avoidance E_{overlap} and from self-propulsion E_{prop} . All terms inside the integrals below represent the effect of each contribution on one point \mathbf{x} inside the ellipse. We then obtain the total energy for one cell by integrating over whole (elliptic) cell area, i.e. with respect to s and θ . The chosen parametrisation given in (1) leads to the appearance of the area element abs (which accounts for the fact that closer to the cell centre one step in s -direction contributes less area) in the integrals in (2), (3) and (4). E_{friction} models friction with the environment by comparing how much points have moved between time t and time $t - \Delta t$:

$$E_{\text{friction}} = \eta \int_0^{2\pi} \int_0^1 abs \frac{|\mathbf{x}(t, s, \theta) - \mathbf{x}(t - \Delta t, s, \theta)|^2}{2\Delta t} ds d\theta. \quad (2)$$

The overlap avoidance term E_{overlap} is modelled by an energy potential V which includes overlap avoidance interactions with all other cells. The choice of V will be discussed below.

$$E_{\text{overlap}} = \int_0^{2\pi} \int_0^1 abs V(\mathbf{x}(t, s, \theta)) ds d\theta. \quad (3)$$

Finally, we want to model self-propulsion. One way to do this is to include it in the energy formulation by prescribing a force \mathbf{F} acting on the cell. In the course of the derivation it is chosen to be $\mathbf{F} = w\eta\mathbf{e}(\alpha)$, i.e. acting in the direction of the orientation and proportional to the experienced friction η . This choice of \mathbf{F} leads to a self-propulsion speed that is independent of friction (choosing \mathbf{F} not proportional to η would only lead to a different definition of the non-dimensional quantities below).

$$E_{\text{prop}} = - \int_0^{2\pi} \int_0^1 abs \mathbf{F} \cdot \mathbf{x}(t, s, \theta) ds d\theta. \quad (4)$$

The total energy is then given by summing (2), (3) and (4)

$$E_{\text{tot}} = E_{\text{friction}} + E_{\text{overlap}} + E_{\text{prop}}. \quad (5)$$

163 We obtain governing equations by minimising this energy in each time step. In other words, in each time step
164 the cell can change its characteristics (position, orientation, shape) to decrease the energy. Calculation details
165 can be found in S1 Appendix Sec. 1. The main derivation steps for the base model are: 1. Differentiation
166 with respect to \mathbf{X} and α respectively (treating all other variables in the energy potential as constants). 2.
167 Setting the derivative to zero and taking the limit $\Delta t \rightarrow 0$. 3. Evaluation of the integrals. We then obtain
168 the following differential equations for the motion of one cell

$$\frac{d\mathbf{X}}{dt} = -\frac{1}{\eta\pi} \int_0^{2\pi} \int_0^1 s \nabla V ds d\theta + w\mathbf{e}(\alpha), \quad (6a)$$

$$\frac{d\alpha}{dt} = -\frac{4}{\eta\pi(a^2 + b^2)} \int_0^{2\pi} \int_0^1 s^2 \nabla V \cdot (\mathbf{R}\mathbf{k}(\theta))^{\perp} ds d\theta. \quad (6b)$$

170 The superscript \perp describes the left-turned normal vector. The two equations in (6) show how the position
 171 and orientations are influenced by the force and torque associated with V respectively. Note that we are
 172 working in a friction-dominated regime, which is why velocity and angular velocity (as opposed to acceleration
 173 and angular acceleration) are proportional to force and torque.

174 **Choice of overlap potential V .** The potential V describes the influence of overlap, where $V > 0$ describes
 175 overlap avoidance and $V < 0$ overlap preference. Many choices of V are possible: e.g. since cells might be
 176 thicker closer to the cell center, overlap closer to the cell center could be punished more than further away.
 177 However, due to the governing equations in (6) being formulated in terms of integrals of the gradient of the
 178 potential V , complicated shapes of V are computationally harder to evaluate, especially in the context of
 179 collective dynamics when this will need to be computed numerous times at each time step. We therefore
 180 choose V to be constant with value σ in regions of overlap and zero elsewhere: For two overlapping ellipses
 181 with domains \mathcal{A} and \mathcal{B} we define $V(\mathbf{x}) = \sigma \mathbb{1}_{\mathcal{A} \cap \mathcal{B}}(\mathbf{x})$, where $\mathbb{1}_{\mathcal{A} \cap \mathcal{B}}(\mathbf{x})$ is the indicator function which equals 1
 182 if $\mathbf{x} \in \mathcal{A} \cap \mathcal{B}$ and 0 otherwise. The strength of this potential is $\sigma \in \mathbb{R}$. If $\sigma > 0$, the cells experience repulsion
 183 in response to overlap, and if $\sigma < 0$, the cells experience attraction. In this work $\sigma > 0$. As a result of this
 184 choice of potential V , two cells only experience overlap avoidance upon overlapping with each other, hence
 185 we define \mathcal{N}_i as the set of indices of cells that overlap with the i -th cell.

186 **Final base model.** We non-dimensionalise the model using as reference time $\frac{A\eta}{\sigma}$, as reference length $\sqrt{\frac{A}{\pi}}$
 187 and define $r = a/b$ as the cell's aspect ratio. The above choice of V allows to evaluate the integrals in (6)
 188 explicitly (for calculation details see S1 Appendix, Sec. 1). The resulting equations can be formulated such
 189 that they depend only on the points of overlap between cells i and j , denoted by \mathbf{Y}_k^{ij} , where up to $k = 4$
 190 points of overlap are possible. This is computationally advantageous since only the points of overlap need to
 191 be found, instead of areas of overlap which would be more computationally costly. In the following $K_{ij} = 1$
 192 or $K_{ij} = 2$ denotes the number of overlap point pairs between cell i and cell j (having one or three points of
 193 overlap can be reduced to having zero or two points of overlap). The overlap points are ordered such that
 194 they traverse the boundary of the cell in an anti-clockwise direction. \mathbf{Y}_1^{ij} and \mathbf{Y}_2^{ij} (and \mathbf{Y}_3^{ij} and \mathbf{Y}_4^{ij}) are
 195 chosen in such a way that the boundary segment of cell i between these pairs of points is contained in the
 196 domain of cell j , see Fig. 1D.

$$\frac{d\mathbf{X}_i}{dt} = - \sum_{j \in \mathcal{N}_i} \sum_{k=1}^{K_{ij}} (\mathbf{Y}_{2k-1}^{ij} - \mathbf{Y}_{2k}^{ij})^\perp + \nu \mathbf{e}(\alpha_i), \quad (7a)$$

$$\frac{d\alpha_i}{dt} = \frac{2r}{r^2 + 1} \sum_{j \in \mathcal{N}_i} \sum_{k=1}^{K_{ij}} (|\mathbf{X}_i - \mathbf{Y}_{2k}^{ij}|^2 - |\mathbf{X}_i - \mathbf{Y}_{2k-1}^{ij}|^2), \quad (7b)$$

198 where the non-dimensional quantity ν is given by $\nu = \frac{w\eta}{\sigma} \sqrt{\pi A}$ and can be interpreted as comparing the
 199 strength of repulsion in the presence of friction to the self-propulsion speed (see more interpretation in the
 200 results section below). These two governing equations are supplemented with initial conditions and boundary

201 conditions. Throughout this work the domain is a square box with side length L and cells are initially placed
 202 randomly inside the box with a random orientation. Further, we use periodic boundary conditions.

203 **Interpretation for two cells.** To understand the equations better, we consider a situation where there
 204 is only interaction between one cell with center \mathbf{X} and orientation α , and one other cell. If there is only one
 205 pair of overlap points, \mathbf{Y}_1 and \mathbf{Y}_2 , then (7) reduces to

$$\frac{d\mathbf{X}}{dt} = -(\mathbf{Y}_1 - \mathbf{Y}_2)^\perp + \nu \mathbf{e}(\alpha), \quad (8a)$$

$$\frac{d\alpha}{dt} = \frac{2r}{r^2 + 1} (|\mathbf{X} - \mathbf{Y}_2|^2 - |\mathbf{X} - \mathbf{Y}_1|^2). \quad (8b)$$

206 We see in (8a) that the cell's center is being pushed in the direction normal to the vector connecting the
 207 points of overlap. Further, we see in (8b) that the change in orientation depends on the difference in lengths
 208 of the segments connecting the cell center with the intersection points, turning the cell in the direction from
 209 the shorter to the longer one, see Fig. 1D. This shows that cells will both move away from each other, and
 210 reorient themselves in order to minimise cell overlap. These are both behaviours that can be observed in
 211 experimental videos, see [55]. Compared to a more *ad hoc* model of simple repulsion between centroids there
 212 are two main differences: Firstly the direction of movement of the centroids would be different: movement
 213 would occur along the line connecting the centroids as opposed to $-(\mathbf{Y}_1 - \mathbf{Y}_2)^\perp$. Likely this difference
 214 would still lead to a similar behaviour in terms of cell populations spreading out and forming a monolayer.
 215 However, secondly (and more importantly), our model also provides an equation for how the cell orientation
 216 changes without the need of extra assumptions.

218 2.3 Results 1: The base model

219 Computational details can be found in S1 Appendix, Sec. 2. We have made some qualitative comparisons of
 220 the model with experimental data in [55]. In this paper, we focus on the computational results of the model.

221 **Alignment increases over time and for larger aspect ratios.** We start by demonstrating basic model
 222 behaviour. In Fig. 2A,B we see the typical behaviour of the base model with overlap avoidance. Over time
 223 cell overlap decreases and alignment increases until a dynamic equilibrium is reached. At this point cells
 224 still move, but the alignment parameter stays relatively constant. Further, we observe that cells tend to be
 225 aligned with their direct neighbours, but this alignment is local and doesn't typically go beyond one or two
 226 cell lengths. This is related to packing problems, where one studies how and how densely objects of a certain
 227 shape can be placed in space without overlapping. Such problems are highly non-trivial, but well-studied
 228 for non-moving, completely solid particles, and for symmetric and elongated particles in 2D and 3D (see
 229 e.g. [8, 9, 10]). Next we inspect how the aspect ratio impacts alignment. We find that increasing the aspect
 230 ratio r of the cells leads to increased alignment. This can be seen in Fig. 2C,D. This is a result that can also
 231 be found in [18]. It is similarly found experimentally in [60] and computationally in [61] that the ability of

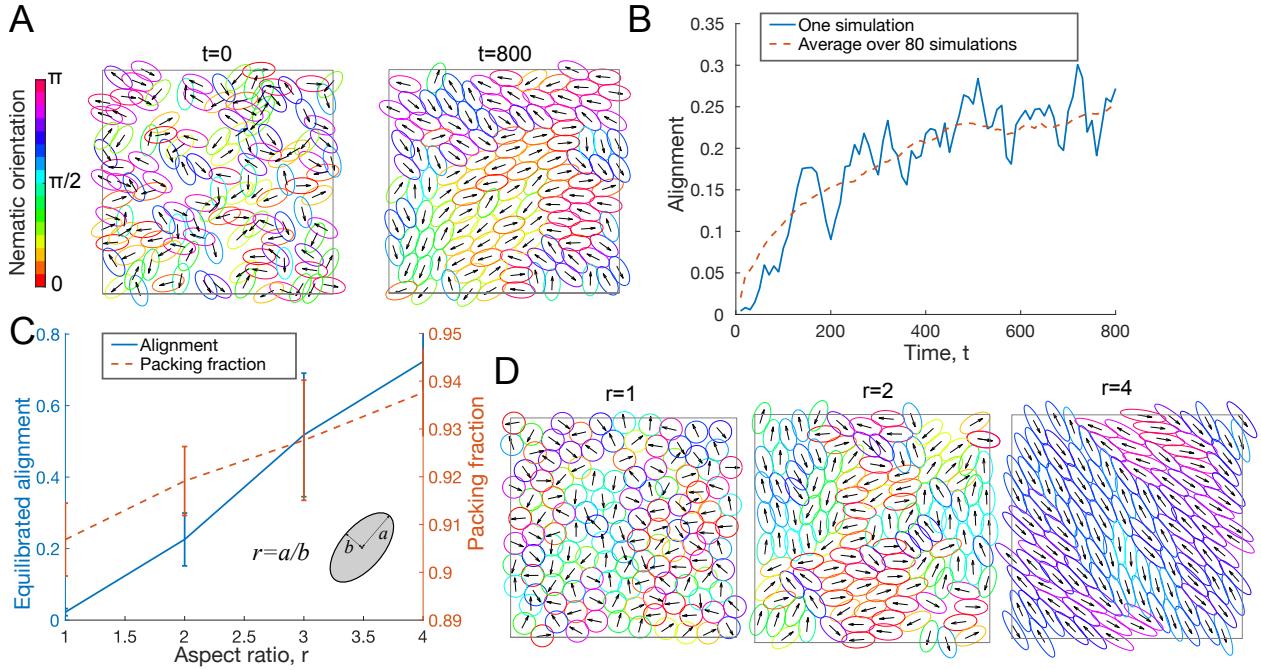


Figure 2: A: Simulation snapshots at time $t = 0$ and $t = 800$ showing cells for an example simulation, color indicates nematic orientation, arrows indicate orientation. See also S2 Video. B: Alignment parameter over time as defined in S1 Appendix, Eq. (1) shown for one individual simulation (blue, solid) and averaged over 80 simulations (orange, dashed). Parameters for A,B: $\nu = 0.2$, $N = 125$, $L = 20$, $r = 2$. C: Alignment parameter and packing fraction (see S1 Appendix, Sec. 2), measured at the dynamic equilibrium, plotted against the cell aspect ratio r , averaged over 60 simulations, error bars represent standard deviation. D: Simulation snapshot at the (final) time points for three different aspect ratios. Parameters for C,D: $\nu = 0.5$, $N = 125$, $L = 20$.

bacteria to swarm effectively is modulated by the cell aspect ratio. Another model that shows that aspect ratio increases alignment can be found in [62]. Investigating this further, we see in Fig. 2C that increasing the aspect ratio of the cells also leads to a higher packing fraction, meaning that there is less cell overlap in the population. Interestingly, for non-moving, solid ellipses [10] found a similar dependence of the alignment parameter on the aspect ratio, albeit with a peak near $r \approx 1.3$.

Optimal ratio of cell speed to overlap avoidance for alignment. The non-dimensional parameter ν is proportional to the self-propulsion speed divided by the strength of overlap avoidance, $\nu \propto \frac{w}{\sigma}$, and can be interpreted as the ratio between two time scales t_1/t_2 , where $t_1 = \frac{A\eta}{\sigma}$ is the time scale of the movement caused by overlap avoidance acting against friction. The time scale $t_2 = w/\sqrt{A/\pi}$ is the time it takes a self-propelling cell of speed w to move one reference length $\sqrt{A/\pi}$. Varying ν and measuring the resulting alignment parameter at the dynamic equilibrium, we find a non-monotone dependence on ν , with a maximal alignment at $\nu = 0.2$, see Fig. 3A,E. We hypothesised that if there is too little self-propulsion, overlap avoidance pushes cells into a little or no overlap configuration, after which cells do not move much and the alignment doesn't increase further. In that situation, pairs of cells might therefore interact with each other

246 for a long duration, but each cell doesn't interact with a large number of cells. Self-propulsion, on the other
 247 hand, might lead to a re-shuffling of cell contacts, leading to shorter-lived, but more numerous interactions.
 248 To test this, we quantified the number of cell contacts over 100 time points for a duration of $t=10$, and the
 249 typical interaction duration (see S1 Appendix, Sec. 2 for details on the quantification). Indeed, we found that
 250 the number of interaction partners increases with ν (Fig. 3B) and that the interaction duration decreases
 251 with ν (Fig. 3C). This leads us to suggest the following explanation for the optimal value ν found in Fig. 3A:
 252 Effective alignment requires cells: 1. To be in contact with their neighbours over a sufficiently long duration
 253 (for overlap avoidance to take effect and cause cells to re-orient locally, time scale given by $t_1 = \frac{A\eta}{\sigma}$) and 2.
 254 To be in contact with sufficiently numerous different cells (time scale given by $t_2 = w/\sqrt{A/\pi}$) in order for the
 255 local order to be propagated beyond immediate neighbours. This finding underlines an important difference
 256 in alignment dynamics between active, self-propelling matter, and passive, non self-propelling matter: Self-
 257 propulsion will generally increase the number of different neighbours a given cell will interact with, while
 258 without self-propulsion there will be fewer interaction partners, but potentially longer interaction times.
 259

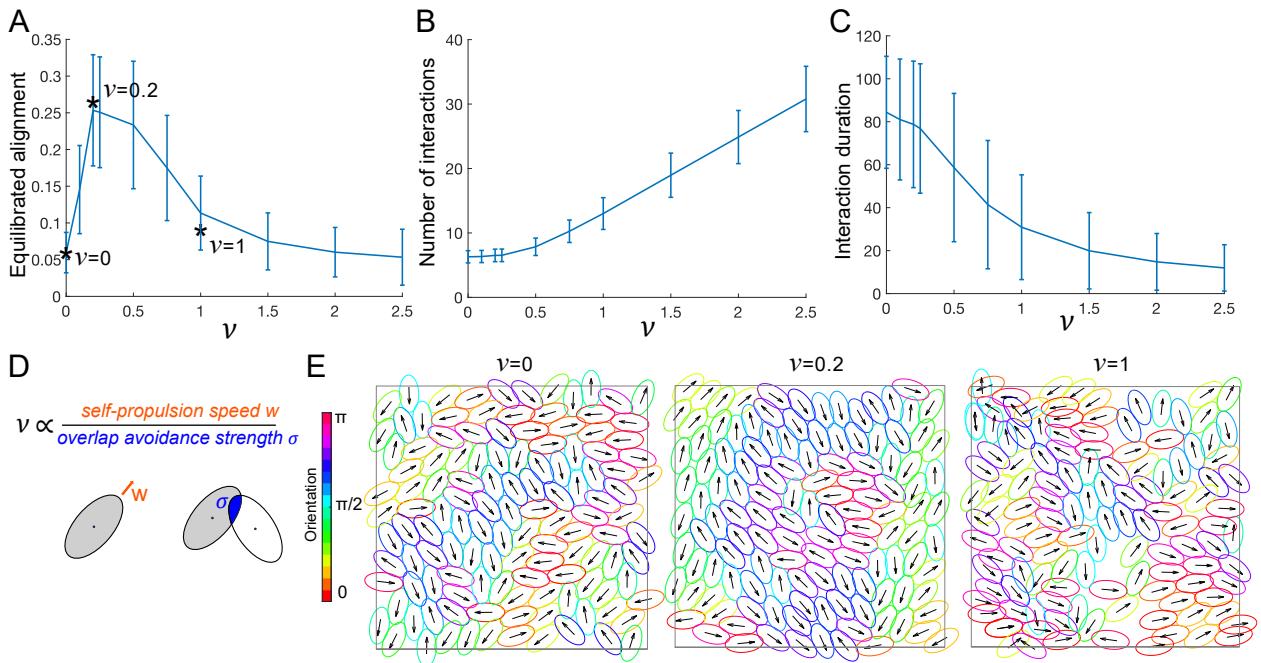


Figure 3: A: Alignment parameter measured at the dynamic equilibrium as a function of ν , averaged over 80 simulations, error bars show standard deviation, stars mark simulations in E. B,C: Number of interaction partners per cell (B) and duration of each pair-wise cell interaction (C) between $t = 0$ and $t = 10$ plotted against ν , averaged over all 125 cells and 5 simulations. Error bars show standard deviation. D: Schematic explanation of ν . E: Simulation snapshots at final time for ν -values marked with a star in A. Colors and arrows as in Fig. 2. Fixed parameters: $r = 2$, $N = 125$, $L = 20$.

260 **Results are consistent with similar models.** In relation to the field of polar active matter (see e.g.
261 review paper [13] and [18, 63, 64, 65, 4] for applications to biology) our model can be thought of as a 'dry'
262 nematic system, which means that momentum is not conserved and that the interactions do not have a polar
263 preference. The latter being a consequence of the elongated shape of the particles and that the interactions
264 arises through overlap avoidance. Several of our results are consistent with simulations of self-propelled hard
265 rods [66, 67, 68], such as the appearance of nematically ordered regions. In [66] it was noted that self-
266 propulsion enhances nematic order, which is also what we observe for small propulsion speeds, however in
267 our model self-propulsion can also hinder large scale alignment. This might be a consequence of our particles
268 being "soft", i.e. that overlap is merely punished, not forbidden. In [69], where soft interactions are modelled,
269 it is found that increasing self-propulsion speed causes cells to transition from a solid like state to a liquid
270 like state. This can be compared to our results by noting that cells do indeed become jammed for very low
271 values of ν in which we see little alignment. As ν increases the cells are able to move around more and hence
272 behave more like a liquid. Also other works have analysed phase transitions between ordered and unordered
273 state occurring at a critical density which depends on the particle size, noise and self-propulsion speed (see
274 e.g. [66, 70]). While this was not the focus of this work (we work at high densities and with no noise), we
275 anticipate that our model would show the same kind of behaviour. We also expect similar behaviour with
276 regards to orientational defects. As a future direction, it would be interesting to investigate whether our
277 model replicates general alignment properties of continuum models of active nematic systems. These tend
278 to be developed from hydrodynamic theory, as reviewed in [13], and can be useful to analyse the overarching
279 macroscopic properties of these systems (such as the stability of steady states), without going into details
280 on a cell level.

281 3 Modelling shape changes

282 3.1 Biological motivation.

283 Cells vary in their ability to change shape. Many bacteria, for example, are surrounded by a stiff cell wall,
284 leading to few shape deformations. Most other cell types, including fibroblasts, are soft and deformable, and
285 change their shape dynamically due to internal changes, or in reaction to their surroundings. For example,
286 they can get squished together when confined or become elongated when attached. While cells can also
287 change shape in the absence of other cells, here we only consider shape changes in reaction to interactions
288 with other cells.

289 3.2 Model derivation.

We investigate the effect of allowing cells to dynamically change their shape in response to overlap. A summary of model parameter names and meaning can be found in S1 Appendix, Tab. 1. We restrict allowed

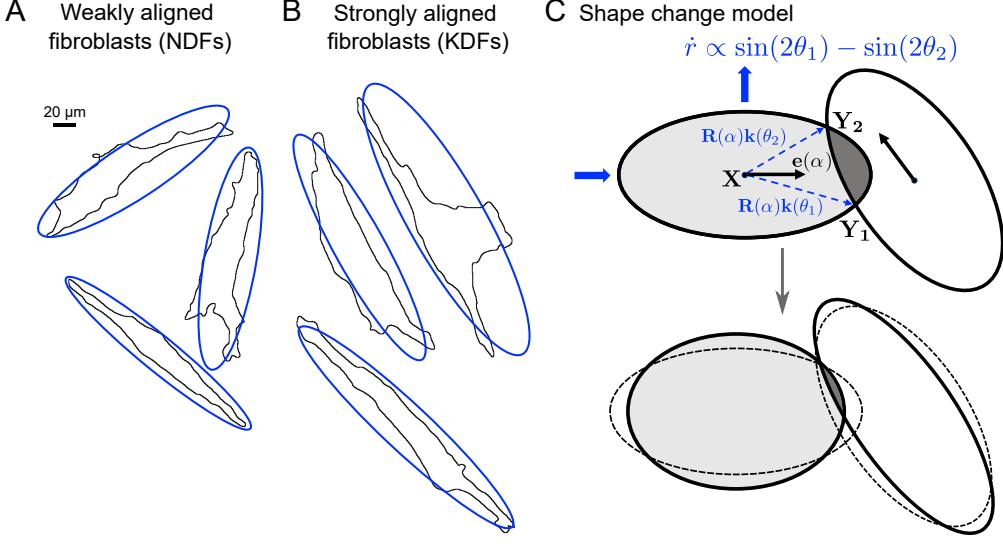


Figure 4: A,B: Single weakly aligned fibroblast (A) and strongly aligned fibroblast (B) cells show a range of different shapes in confluent cultures. No difference can be observed between the two samples [55]. C: Schematic of effect of shape change model.

cell shapes to changing the aspect ratio $r = a/b$, where $r(t)$ is now a function of t , while maintaining a constant cell area A . To avoid unrealistically large (or small) aspect ratios, we add a term to the energy that punishes deviations from some preferred aspect ratio \bar{r} , which we set to $\bar{r} = 2$. In other words, in the absence of interactions, cell shape will relax towards having the preferred aspect ratio. The strength of this relaxation is given by g . This adds an extra term E_{shape} to the E_{tot} of the base model given in (5),

$$E_{\text{shape}} = \frac{g}{2} \left[(r - \bar{r})^2 + \left(\frac{1}{r} - \frac{1}{\bar{r}} \right)^2 \right].$$

Note that E_{shape} is symmetric with respect to $r \rightarrow \frac{1}{r}$, $\bar{r} \rightarrow \frac{1}{\bar{r}}$. This is to ensure that the shape relaxation behaviour is the same along both axes. The model derivation now follows the same steps as before, for some more details see S1 Appendix, Sec. 1. The equations for \mathbf{X} and α remain unchanged by this. For one cell, we obtain the following equation for how the aspect ratio r changes over time as a result of overlap avoidance and shape relaxation:

$$\dot{r} = -\frac{16r^3}{A\eta(r^2+1)} \int_0^{2\pi} \int_0^1 s^2 \nabla V \cdot \left(\mathbf{R}(\alpha) \frac{d\mathbf{k}}{dr} \right) ds d\theta + \frac{16\pi g(1+\bar{r}r^3)}{A^2\eta(1+r^2)} \left(1 - \frac{r}{\bar{r}} \right).$$

Using the same overlap potential as defined above, we obtain (written in non-dimensional form)

$$\dot{r}_i = \frac{4r^2}{r^2+1} \sum_{j \in \mathcal{N}_i} \sum_{k=1}^{K_{ij}} [\sin(2\theta_{2k-1}^{ij}) - \sin(2\theta_{2k}^{ij})] + 16\gamma \frac{1+\bar{r}r^3}{1+r^2} \left(1 - \frac{r}{\bar{r}} \right),$$

where $\gamma = \frac{\pi g}{A\sigma}$ and θ_k^{ij} corresponds to the θ value that parameterises the overlap point $\mathbf{Y}_k^{ij} = \mathbf{X}_i + \mathbf{R}(\alpha_i)\mathbf{k}(\theta_k^{ij})$.

Interpretation for two cells. In the case of the interaction between only two cells with two overlap points (as in Fig. 4C) we have

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

If, as in Fig. 4C, we have that $\theta_1 \in (-\frac{\pi}{2}, 0)$ and $\theta_2 \in (0, \frac{\pi}{2})$, then $\sin(2\theta_1) - \sin(2\theta_2) < 0$ and the first term will cause the aspect ratio r to decrease. This is a result of the cell shortening to avoid overlap. The second term will always act to restore the aspect ratio towards \bar{r} . Note that there is only one additional parameter, γ , quantifying the restoring force, but no extra parameter quantifying the initial deformation. The reason for this is that shape changes are also driven by minimising overlap and hence based on the same energy term, E_{overlap} , as the other dynamics (turning and non-propulsion driven translations) driven

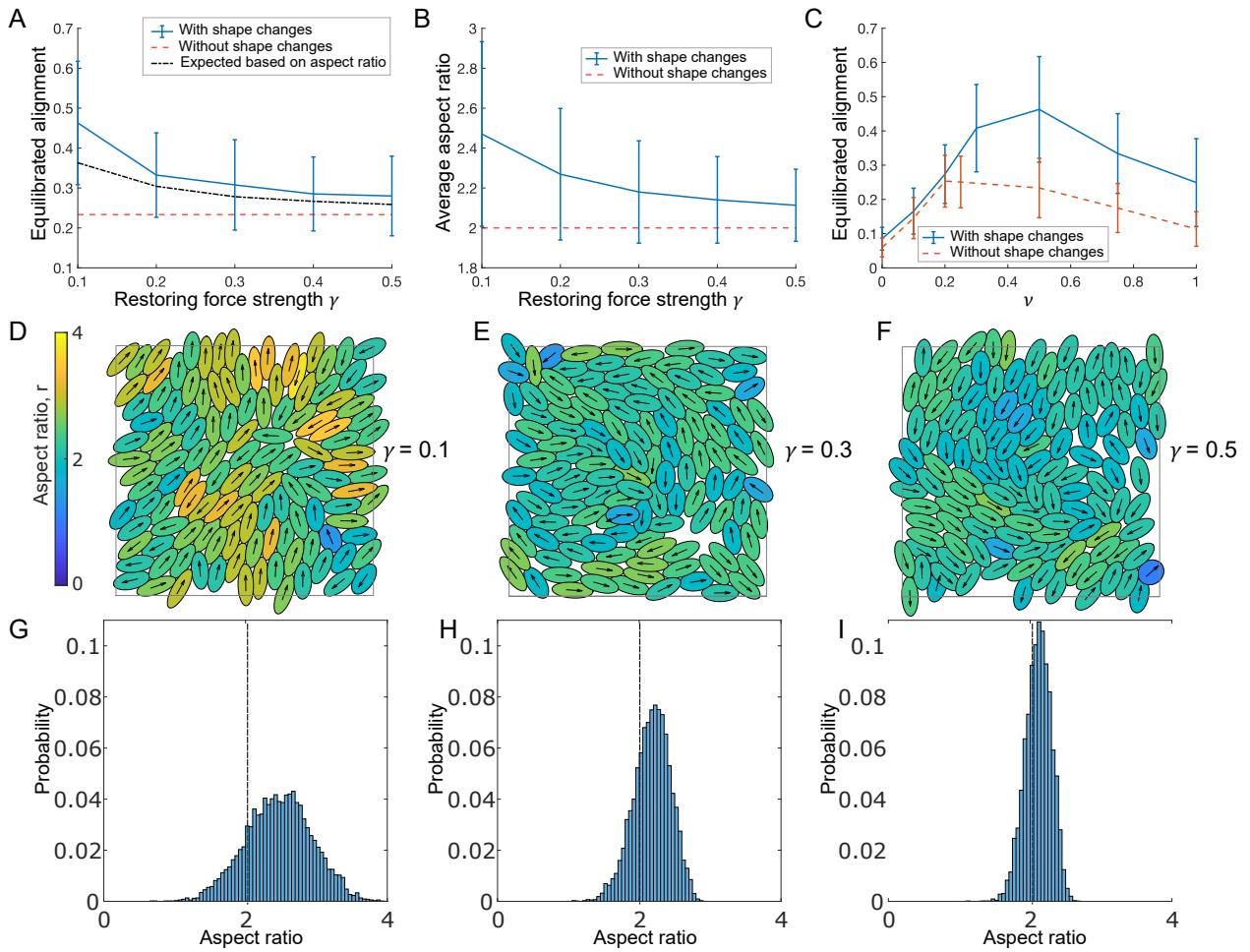


Figure 5: A,B: Equilibrated alignment (A) and average aspect ratio (B) measured for different values of γ and compared to non-deformable cells with $\nu = 0.5$. In A: Black, dash-dotted line shows expected alignment based on the measured average aspect ratio (see text). C: Equilibrated alignment for different values of ν for deformable and non-deformable cells with $\gamma = 0.1$. D-F: Simulation snapshots at equilibrium for three different values of γ . See also S3 Video. Colors indicate aspect ratio. G-I: Distribution of aspect ratio measured for γ -values in D-F using values pooled from 60 simulations at equilibrium. Fixed parameters: $N = 125$, $L = 20$, $\bar{r} = 2$.

302 by overlap avoidance.

303 3.3 Results 2: The base model with shape changes

304 **Deformable cells lead to increased alignment.** We investigate the effect of γ , the strength of the
305 shape restoring force. Small γ means cells are more easily deformable, while the limit $\gamma \rightarrow \infty$ corresponds
306 to non-deformable cells with fixed aspect ratio \bar{r} (corresponding to the base model discussed in Sec. 2).
307 Fig. 5A shows that higher deformability correlates with more alignment, suggesting that changes in aspect
308 ratio aid alignment. Further, we found that allowing cells to deform more increases the average aspect ratio
309 in the population, Fig. 5B. We also assessed whether the optimal value of ν found in Fig. 3A is affected
310 by deformability. Fig. 5C shows that indeed, maximal alignment is now reached for larger values of ν , i.e.
311 bigger self-propulsion speeds or smaller overlap avoidance.

312 4 Modelling cell-cell junctions and actin forces

313 4.1 Biological motivation.

314 In some types of cells, neighbouring cells can form cell-cell junctions. These provide a mechanical coupling
315 as well as a way for cells to exchange signals, i.e. communicate. We model two potential effects of cell-
316 cell junctions: 1. Elastic, reversible connections between two points on the edges of neighbouring cells. 2.
317 Formation of supracellular actin bundles that lead to a bending force. The actin cytoskeleton is the main
318 force generator for moving cells and plays a major role in determining cell shape and polarity. In [55] we
319 observed that the actin network within a cell shares the orientation of the cell itself, with the major bundles
320 typically running along the long axis. Further we found that for keloid fibroblasts (KDFs), neighbouring
321 cells can form supracellular actin bundles, i.e. that the actin bundles of neighbouring cells visually appear to
322 be connected to each other in a smooth manner, without an abrupt change in direction of the actin bundles
323 at the cell-cell junction. This suggests a potential mechanical linkage mediated by cell-cell junctions that
324 causes cells to align their cytoskeleton, and therefore themselves. We will focus on the latter, including
325 elastic connections as a means for these supracellular actin bundles to form. Since the actin bundles are
326 predominantly found to run along the length of each cell, we will only consider cell-cell junctions forming at
327 the front and back of each cell, but see S1 Appendix Sec. 3 for some results on cell-cell junctions forming at
328 the sides.

329 4.2 Model derivation.

330 A summary of model parameter names and meaning can be found in S1 Appendix, Tab. 1.

331 **Elastic connections.** In principle cell-cell junctions could form whenever two points in the domain of the
 332 ellipses are in close contact. However, since we are interested in cell-cell junctions as a way to mediate the
 333 formation of supracellular actin bundles running parallel to the long axis of the cell, we will focus on cell-cell
 334 junctions at the front and the rear of cells. We therefore restrict ourselves to only allowing front-to-back,
 335 front-to-front and back-to-back junctions. We model these connections as Hookean springs with rest length
 336 zero. The effect of cell-cell junctions at different locations around the cell boundary has been explored in S1
 337 Appendix Sec. 3, where we allowed cell-cell junctions to also form at two points at the cell sides. In principle
 338 such connections could be created and broken stochastically, with a distance (or force) dependent breakage
 339 rate. However these processes likely happen on a much faster timescale than the overall alignment dynamics.
 340 Such a stochastic formulation would also introduce additional parameters, the values of which are hard to
 341 infer from experimental data. Further these additional unknown parameters would significantly complicated
 342 the analysis of the model. We therefore assume deterministic springs that form and stay in place whenever

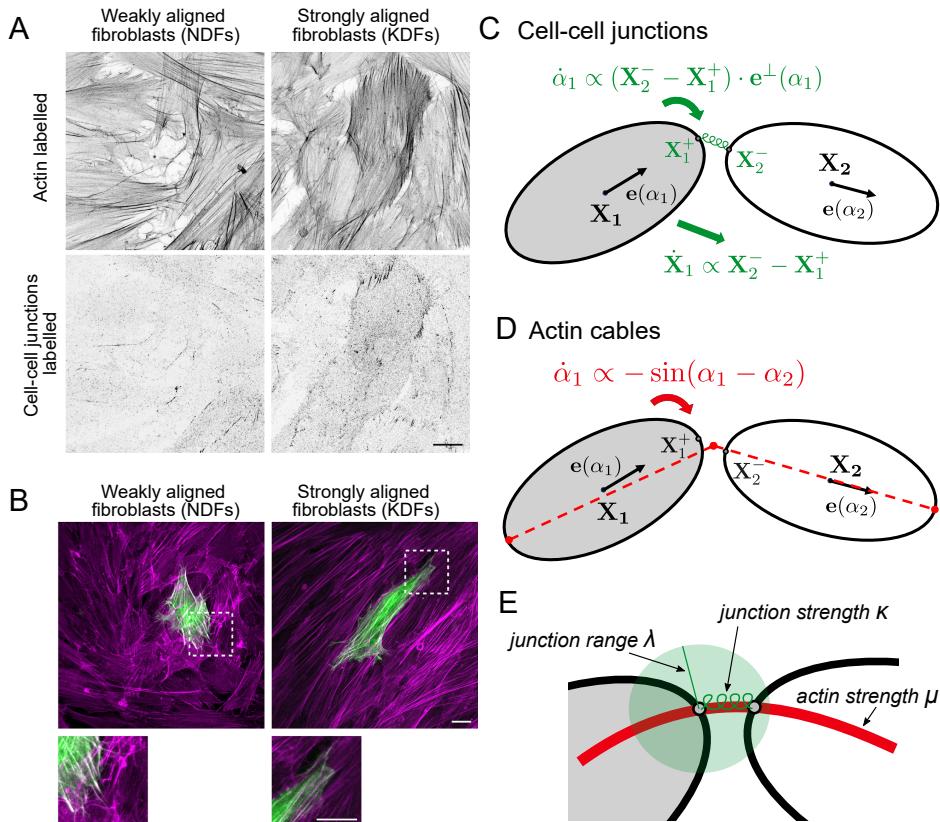


Figure 6: A: NDFs and KDFs stained for F-actin and N-cadherin (a marker for a type of cell-cell junction), revealing an enhanced localisation of N-cadherin at cell-cell junctions in KDF. Scale bar 20 μm . B: Supracellular actin bundles. NDF and KDF cultured in vitro for 48 hours, mosaically transfected with EGFP-LifeAct (green) and stained for F-actin (magenta). KDF display more aligned actin bundles spanning multiple cells. Scale bar 20 μm . C,D: Schematics showing how the cell-cell junctions (C) and actin forces (D) affect changes in cell position and orientation. E: Explanation of new non-dimensional parameters λ , κ and μ .

³⁴³ the distance between connection points are below some critical distance, and break when stretched beyond
³⁴⁴ that distance.

Trans-cellular actin cables. Detailed models of actin bundles within cells have been developed e.g. in [71, 72]. Here we are only interested in the potential effect of trans-cellular actin cables on the orientation of neighbouring cells. We therefore don't provide a more detailed model of the dynamics of the actin bundle *within* each cell, but rather focus on the potential effect of the *connected* actin cable. Inside the cells, we just approximate each bundle as a straight inextensible rod. Motivated by the biological findings in [55], summarised above, we model the trans-cellular actin cable as an inextensible rod with a given bending stiffness. For a general rod discretised with a uniform step length q , resulting in the points $\mathbf{x}_i \in \mathbb{R}^2$, $i = 1, \dots, K$, the bending energy of strength m is given by

$$\frac{m}{2} \sum_{i=2}^{K-1} \left(\frac{|\mathbf{x}_{i-1} - 2\mathbf{x}_i + \mathbf{x}_{i+1}|}{q^2} \right)^2 q.$$

Note that its continuous counterpart would have an integral instead of the sum and the norm of the second derivative instead of the quotient. We discretise the supracellular actin bundle using three points, the two end points not involved in the cell junction, plus the midpoint of the cell-cell junction points. Further we use $q \approx 2a$. This formulation has the advantage that it leads to a very simple bending energy. For example, if the front of a cell positioned at \mathbf{X}_1 with orientation α_1 is connected to the rear of a cell positioned at \mathbf{X}_2 with orientation α_2 , the corresponding bending energy would be

$$\frac{m}{2} \left(\frac{|\mathbf{X}_1 - a\mathbf{e}(\alpha_1) - 2\frac{\mathbf{X}_1 + a\mathbf{e}(\alpha_1) + \mathbf{X}_2 - a\mathbf{e}(\alpha_2)}{2} + \mathbf{X}_2 + a\mathbf{e}(\alpha_2)|}{(2a)^2} \right)^2 2a = \frac{m}{4a} |\mathbf{e}(\alpha_1) - \mathbf{e}(\alpha_2)|^2.$$

³⁴⁵ This shows that the bending will only affect the cells' orientation, not their positions.

Incorporation into the full model. We denote the strength of the Hookean springs describing the cell-cell junctions by k and assume junctions will exist whenever potential connection points are within distance l of each other. To distinguish between front and back connections we define the front and back ends of a cell by $\mathbf{X}^\pm := \mathbf{X} \pm a\mathbf{e}(\alpha)$, and the two relevant index sets by

$$\mathcal{N}^{\text{fb}} = \{j \in \{1, \dots, N\} \mid |\mathbf{X}^\mp - \mathbf{X}_j^\pm| < l\}, \quad \mathcal{N}^{\text{ff,bb}} = \{j \in \{1, \dots, N\} \mid |\mathbf{X}^\pm - \mathbf{X}_j^\pm| < l\}. \quad (9)$$

The set \mathcal{N}^{fb} describes front-to-back junctions and $\mathcal{N}^{\text{ff,bb}}$ describes front-to-front and back-to-back junctions. The new contribution to the total energy for a cell positioned at \mathbf{X} with orientation α now takes the form

$$\begin{aligned} E_{\text{junction}} = & \sum_{j \in \mathcal{N}^{\text{fb}}} \left[\frac{k}{2} |\mathbf{X}^\pm - \mathbf{X}_j^\mp|^2 + \frac{m}{4a} |\mathbf{e}(\alpha) - \mathbf{e}(\alpha_j)|^2 \right] \\ & + \sum_{j \in \mathcal{N}^{\text{ff,bb}}} \left[\frac{k}{2} |\mathbf{X}^\pm - \mathbf{X}_j^\pm|^2 + \frac{m}{4a} |\mathbf{e}(\alpha) + \mathbf{e}(\alpha_j)|^2 \right]. \end{aligned}$$

³⁴⁶ The model derivation now follows the same steps as described in Sec. 2 and yields, in non-dimensional
³⁴⁷ form,

$$\begin{aligned} \frac{d\mathbf{X}_i}{dt} &= - \sum_{j \in \mathcal{N}_i} \sum_{k=1}^{K_{ij}} (\mathbf{Y}_{2k-1}^{ij} - \mathbf{Y}_{2k}^{ij})^\perp + \nu \mathbf{e}(\alpha) + \kappa \left[\sum_{j \in \mathcal{N}_i^{\text{fb}}} (\mathbf{X}_j^\mp - \mathbf{X}_i^\pm) + \sum_{j \in \mathcal{N}_i^{\text{ff,bb}}} (\mathbf{X}_j^\pm - \mathbf{X}_i^\pm) \right], \\ \frac{d\alpha_i}{dt} &= \frac{2r}{r^2 + 1} \sum_{j \in \mathcal{N}_i} \sum_{k=1}^{K_{ij}} (|\mathbf{X}_i - \mathbf{Y}_{2k}^{ij}|^2 - |\mathbf{X}_i - \mathbf{Y}_{2k-1}^{ij}|^2) \\ &\quad + 4\kappa \frac{r^{3/2}}{r^2 + 1} \left[\sum_{j \in \mathcal{N}_i^{\text{fb}}} (\mathbf{X}_j^\mp - \mathbf{X}_i^\pm) \cdot (\pm \mathbf{e}^\perp(\alpha_i)) + \sum_{j \in \mathcal{N}_i^{\text{ff,bb}}} (\mathbf{X}_j^\pm - \mathbf{X}_i^\pm) \cdot (\pm \mathbf{e}^\perp(\alpha_i)) \right] \\ &\quad + \mu \frac{\sqrt{r}}{r^2 + 1} \left[- \sum_{j \in \mathcal{N}_i^{\text{fb}}} \sin(\alpha_i - \alpha_j) + \sum_{j \in \mathcal{N}_i^{\text{ff,bb}}} \sin(\alpha_i - \alpha_j) \right], \end{aligned}$$

348 where the index sets \mathcal{N}^{fb} and $\mathcal{N}^{\text{ff,bb}}$ are as defined in (9) with l replaced by the non-dimensional $\lambda = \frac{l\sqrt{\pi}}{\sqrt{A}}$.
 349 Further we have defined the two non-dimensional quantities $\kappa = \frac{k}{\sigma}$ and $\mu = \frac{2m}{\sigma} \left(\frac{\pi}{A} \right)^{3/2}$, which compare the
 350 junction strength and bending strength to the strength of overlap avoidance. We assume that cell aspect
 351 ratio, r , is constant.

Interpretation for two cells. To understand the effect of the new terms, we can return to the situation from above, where the front of cell 1 has a junction with the back of cell 2 (see Fig. 6C,D). Dropping all other terms, this leads to

$$\begin{aligned} \frac{d\mathbf{X}_1}{dt} &= \kappa(\mathbf{X}_2^- - \mathbf{X}_1^+), \\ \frac{d\alpha_1}{dt} &= 4\kappa \frac{r^{3/2}}{r^2 + 1} (\mathbf{X}_2^- - \mathbf{X}_1^+) \cdot \mathbf{e}^\perp(\alpha_1) - \mu \frac{\sqrt{r}}{r^2 + 1} \sin(\alpha_1 - \alpha_2). \end{aligned}$$

352 Inspecting the sine-term in the equation for α_1 , we see that the cytoskeletal coupling always aids alignment,
 353 however larger aspect ratios lead to slower alignment. For the effect of the cell-cell junctions, we see that
 354 they cause the center of cell 1 to be pulled along the vector connecting the two junction points. Further,
 355 the junction also causes cell turning, however it is not obvious whether this will aid or hinder alignment.
 356 The answer becomes even less clear in a multi-cell context and in the context of overlap avoidance and
 357 self-propulsion. For this we turn to simulations.

358 4.3 Results 3: The base model with cell-cell junctions and actin forces

359 **Cell-cell junctions alone hinder alignment.** For cell-cell junctions forming at cell heads and tails there
 360 are two new (non-dimensional) parameters introduced to the model: The junction range λ (i.e. maximal
 361 length over which junctions can form) and the junction strength κ . We varied λ between 0 and 0.8 (i.e.
 362 between 0-28% of one cell length) and κ between 0 and 2.5 (i.e. between 0-2.5 times the strength of overlap
 363 avoidance), initially without actin forces. First we measured the number of junctions formed and found,
 364 as expected, that more junctions are formed as λ or κ increase, Fig. 7A,C. Next, we found that alignment

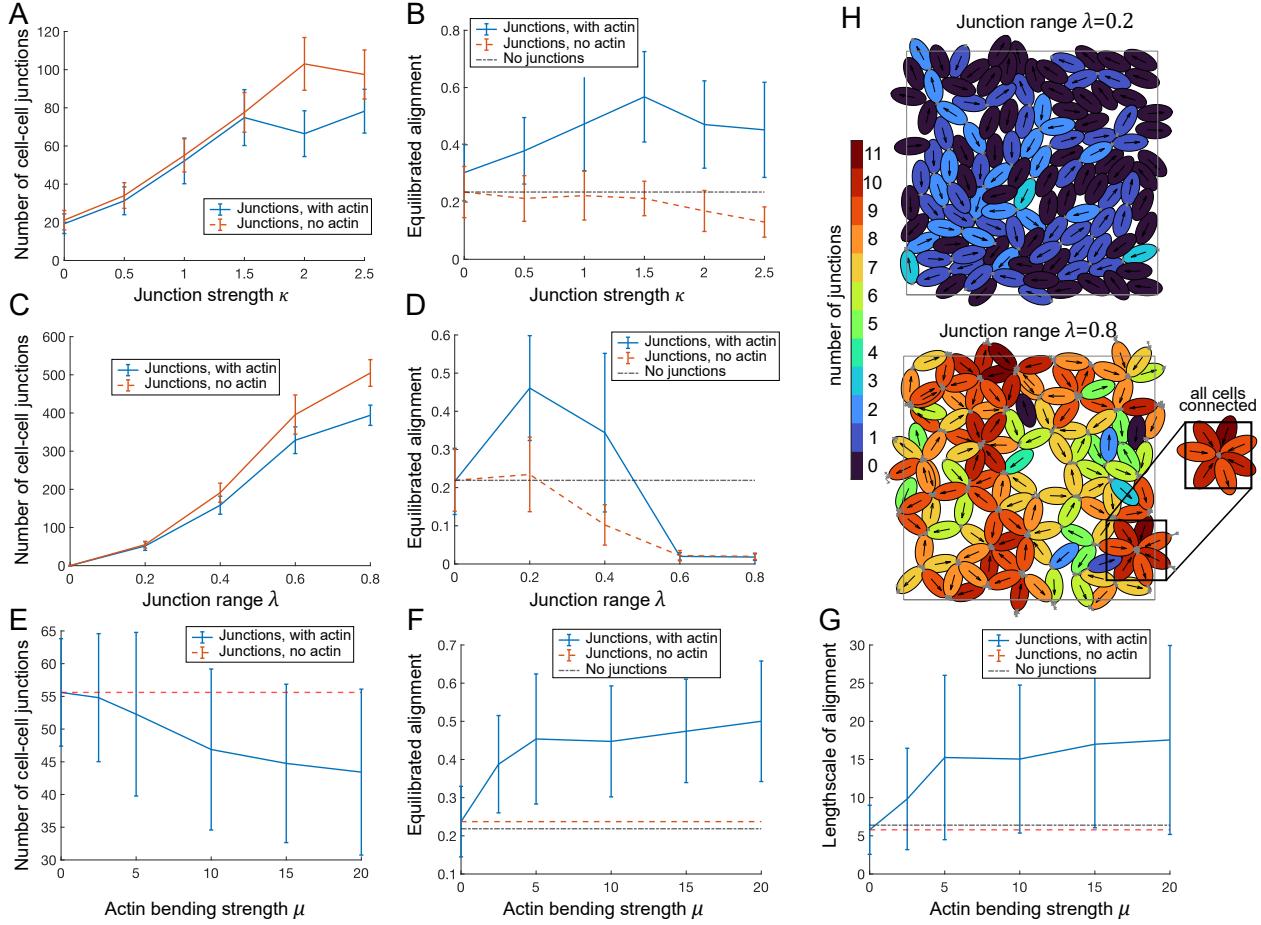


Figure 7: A-F: Number of cell-cell junctions (A,C,E) and value of equilibrated alignment (B,D,F) in dependence of the junction strength κ (A,B), the junction range λ (C,D) and the actin force μ (E,F) with and without an actin force ($\mu = 0$ and $\mu = 5$ respectively). Average over 60 simulations, error bars show standard deviation. G: Length scale of alignment in dependence on actin force μ (see S1 Appendix, Sec. 2). Base parameters are $\lambda = 0.2$, $\nu = 0.5$, $\kappa = 1$, $N = 125$, $L = 20$. H: Simulation snapshots at equilibrium for $\lambda = 0.2$ (top) and $\lambda = 0.8$ (bottom). Cell junctions are marked in grey, color corresponds to number of junctions per cell. Other parameters are $\nu = 0.5$, $\mu = 0$, $\kappa = 1$, $N = 125$, $L = 20$.

365 decreases in both cases, Fig. 7B,D. It appears that the junctions hinder alignment, because they lead to cells
 366 forming clumps where more than two cells are joined at one point, which acts against alignment. Indeed,
 367 we found many more instances of cell clumps for larger junction range than for lower junction range, as
 368 demonstrated in Fig. 7H.

369 **Supracellular actin can greatly aid alignment.** Next we investigated how adding an actin force of
 370 strength μ , representing the effect of supracellular actin bundles, would affect the dynamics. We found that
 371 the actin force has only a small effect on the number of junctions formed, with a slight tendency to reduce
 372 the number of junctions formed, Fig. 7A,C,E. However, we found that the actin force can greatly increase
 373 the equilibrated alignment, Fig. 7B,D,F. We found that as actin forces increase, so does alignment, with

374 alignment values plateauing for large actin forces, Fig. 7F. Further we found that in the presence of actin,
 375 there is an optimal junction strength (at $\kappa \approx 1.5$, i.e the junction strength is $1.5 \times$ the strength of overlap
 376 avoidance), Fig. 7B, and an optimal junction range (at $\lambda \approx 0.2$, corresponding to about 7% of one cell
 377 length), Fig. 7D. Further we found that not only the value of the alignment parameter, but also its length
 378 scale, measured as defined in S1 Appendix, Sec. 2, Eq. (2), increases from around 5 (corresponding to 1-2
 379 cell lengths) up to about 17 (corresponding to about 6 cell lengths). In Fig. 7F,G we see a large variation in
 380 the alignment values and length scales measured for a given parameter set. We speculated that the reason
 381 might be differences in the populations' junction structure. This we explored next.

382 **High alignment at long length scales is driven by linear chains of cells.** To understand the junction
 383 structure of a cell population better, we represented the simulated cell populations as graphs, where cells
 384 and connections between them are represented by the nodes and edges of the graph respectively. This allows
 385 for a visual representation of the population structure as well quantification of graph properties, such as
 386 the number of junctions per cell (the degree of the node). For a cell to be of degree 0 means it has no

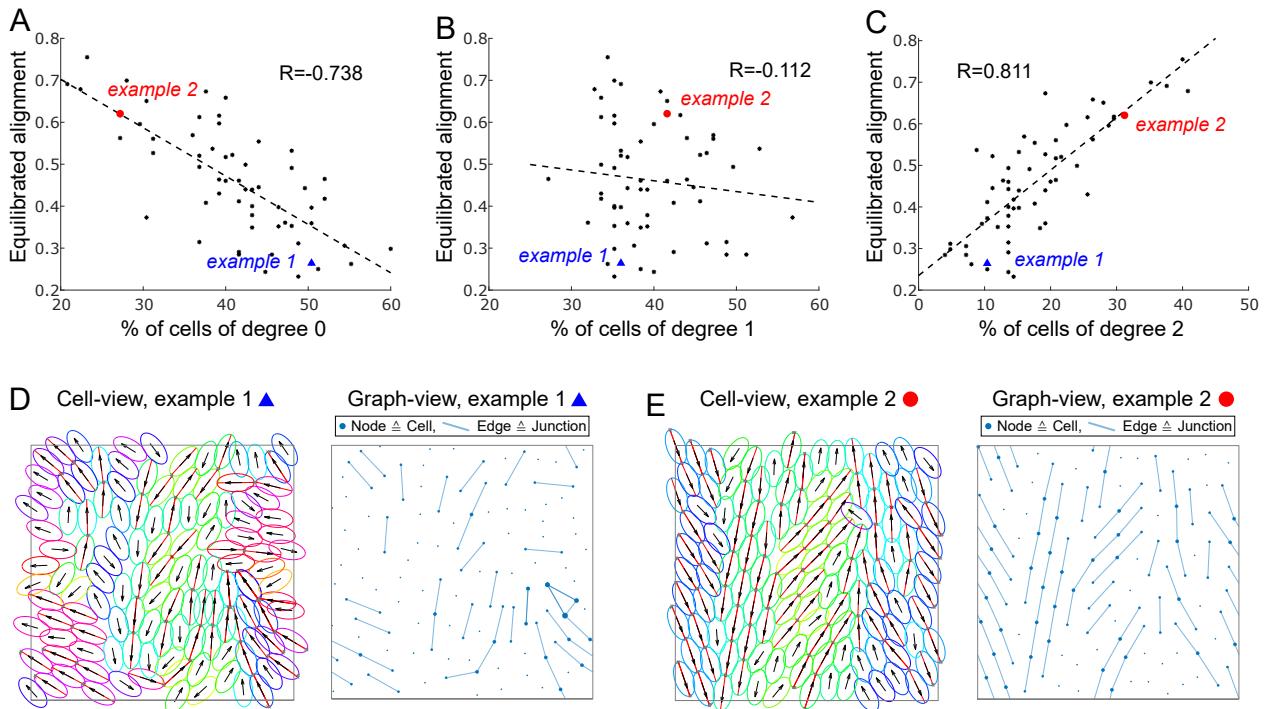


Figure 8: A-C: Scatter plot of equilibrated alignment against % of cells of degree 0 (A), degree 1 (B) and degree 2 (C) at the final time point $T = 400$. Each black dots represents one simulation run. The blue triangle and the red dot mark the examples in D and E respectively. R gives the correlation coefficient and the dotted line gives the linear least squares fit. D,E: Examples marked in A-C in cell-view (left, red lines mark actin, colors and arrows as in Fig. 2) and graph-view (right, dots mark cells/nodes, lines mark edges/junctions). See also S4 Video. Other parameters are $\nu = 0.5$, $\mu = 5$, $\kappa = 1$, $\lambda = 0.2$, $N = 125$, $L = 20$.

387 connections to other cells, a degree 1 cell has a connection to one other cell, etc. For a given parameter
388 set, we then produced 60 repetitions of the same numerical experiments (with random initial conditions)
389 and inspected, at equilibrium, the correlation between the % of cells of a given degree with the alignment,
390 Fig. 8A-C. Strikingly, we found that for alignment, there is a strong negative correlation with % of degree
391 0 cells, a strong positive correlation with the % of degree 2 cells, and almost no correlation with the % of
392 degree 1 cells. We saw exactly the same correlation trends for the alignment length scale (not shown). This
393 means for high alignment and high length scales of alignment one needs few unconnected cells and many
394 cells being connected to exactly 2 other cells. In example simulations, Fig. 8D,E we see that, indeed, degree
395 2 cells form long chains that explain both the high alignment and the long alignment length scales.

396 5 Discussion

397 **A flexible modelling framework.** In this work we have developed a framework to mechanistically model
398 the collective behaviour of active, elliptically shaped particles, that self-propel, avoid overlap, deform and
399 form cell-cell junctions that communicate cytoskeletal forces. The framework is based on energy minimisation
400 and can easily be extended or adapted to include e.g. other cell shapes or different types of cell-cell
401 interactions. A strength of the framework is that the derived equations strike a useful balance between being
402 complex enough to capture the desired phenomena, while being simple enough to be interpretable.

403
404 We simulated the emerging collective dynamics for a large ensemble of cells and analysed the numerical
405 results with an emphasis on alignment dynamics. The main computational findings of this work are:

- 406 • **Cell alignment needs a balance of self-propulsion and overlap avoidance.** We found that,
407 to maximise collective cell alignment, there is an optimal ratio of self-propulsion speed and overlap
408 avoidance. This allows cells to have a sufficiently long contacts with a sufficient number of cells, which
409 aids alignment.
- 410 • **Deformability aids collective alignment.** We found that allowing for flexible cell shapes can aid
411 alignment. We hypothesise that this is because it leads to more elongated cells (which are associated
412 with more alignment) and a more flexibility of the use of space.
- 413 • **Cell-cell junctions alone hinder alignment.** We found that modelling spring-like cell-cell junctions
414 at the cell heads and tails hinders alignment. The reason seems to be the formation of clumps of cells.
- 415 • **Actin forces lead to strong, long-scale alignment.** If actin forces are communicated via the
416 cell-cell junctions, this can significantly increase alignment. In this case alignment will happen on a
417 much larger length scale. The reason seem to be long, linear chains of connected cells

418 **Future work.** The derived equations can be used to study e.g. the interaction of only two cells in more
419 depth: Such a simplified system could then be analysed using analytical methods, such as stability analysis,
420 asymptotics or determining long-term behaviour. This is the subject of current ongoing work. The results
421 will give further insights into the involved time scales of movement, the role of self-propulsion or the effect of
422 deformability. Since the analysis involves a technical treatment of degenerate limits this work is reserved for
423 a separate publication, which is in preparation. In terms of modelling, we are planning on following several
424 directions, such as: 1) We will extend the cell-cell junction model to investigate the effect of cell-cell junctions
425 forming along the whole cell surface. 2) We will investigate how cell-cell junctions affect cell shape. 3) We
426 will derive and analyse a more detailed model of cytoskeletal dynamics within the cell and its interaction
427 with the substrate. Further, we will test our insights in an experimental setting: Our work in [55], where we
428 compared two types of fibroblasts with different overlap avoidance, is a first step in this direction. However,
429 we will also experimentally test several of the other theoretical predictions in this work.

430 6 Material and methods

431 For more detailed biological methods, please refer to [55]. The primary dermal fibroblasts studied in this work
432 were isolated from normal skin and keloid scar tissue from adult patients providing written informed consent.
433 This tissue collection was ethically approved by the National Research Ethics Service (UK) (14/NS/1073).
434 The study was conducted in accordance with the ethical standards as set out in the WMA Declaration of
435 Helsinki and the Department of Health and Human Services Belmont Report. Normal dermal fibroblasts
436 and keloid derived fibroblasts were isolated from the collected tissue and cultured in vitro. To visualise
437 cell overlap, cells were labelled with CellTrace reagents in two colours (Violet and CFSE, ThermoFisher).
438 Subsequently, cells were plated on imaging substrates with a ratio of 9:1 Violet:CFSE and fixed after 24
439 hours. Imaging was obtained using a Zeiss LSM 880 confocal microscope (20x NA 0.8 Plan-Apochromat
440 air objective). To evaluate cell shape, single fibroblasts were observed within confluent monolayer cultures.
441 Mosaic expression to highlight single cells was obtained by transfecting cells with EGFP-LifeAct, followed
442 by fixation after 48 hours, and immunostaining with phalloidin to visualise F-Actin (Life Technologies). To
443 visualise cell-cell adhesions, cells were labelled with N-cadherin (Cell Signaling Technologies) after fixation.
444 Imaging was obtained using a Zeiss LSM 880 confocal microscope (40x NA 1.3 Plan-Apochromat oil objective,
445 40x NA 1.1 LD C-Apochromat water objective, or 63x NA 1.4 Plan-Apochromat oil objective).

446 **Data Availability Statement.** All code written in support of this publication is publicly available at
447 https://github.com/angelikamanhart/Code_Alignment_Ellipses

448 7 Supporting information

449 **S1 Appendix:** Derivation and computational details.

450 **S2 Video:** Video corresponding to snapshots shown in Fig. 2A. Parameters are $\nu = 0.2$, $N = 125$, $L = 20$,
451 $r = 2$, $T = 800$, time step = 0.01.

452 **S3 Video:** Video corresponding to snapshot shown in Fig. 5D. Parameters are $\nu = 0.5$, $\gamma = 0.1$, $N = 125$,
453 $L = 20$, $\bar{r} = 2$, $T = 400$, time step = 0.01.

454 **S4 Video:** Video corresponding to snapshot shown in Fig. 8E. Parameters are $\nu = 0.5$, $\lambda = 0.2$, $\mu = 5$,
455 $\kappa = 1$, $N = 125$, $L = 20$, $\bar{r} = 2$, $T = 400$, time step = 0.01.

456 8 Funding and acknowledgements

457 This work was supported by the Engineering and Physical Sciences Research Council (grant numbers
458 EP/N509577/1 to VL and EP/T517793/1 to VL) which paid the salary of VL, the Wellcome Trust (grant
459 number 107859/Z/15/Z to BS) which paid the salary of FK and SM, the European Research Council (ERC)
460 under the European Union’s Horizon 2020 research and innovation program (grant agreement no. 681808
461 to BS) and BBSRC project grant (BB/V006169/1, to BS), which paid the salary of SM. The funders had
462 no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.
463 The authors wish to thank Antoine Nicolas Diez for his helpful thoughts and comments and Mohit Dalwadi
464 for this supervision support.

465 References

- 466 [1] Ginelli F, Peruani F, Pillot MH, Chaté H, Theraulaz G, Bon R. Intermittent collective dynamics emerge
467 from conflicting imperatives in sheep herds. PNAS. 2015;112(41):12729-34.
- 468 [2] Makris NC, Ratilal P, Jagannathan S, Gong Z, Andrews M, Bertsatos I, et al. Critical population
469 density triggers rapid formation of vast oceanic fish shoals. Science. 2009;323(5922):1734-7.
- 470 [3] Parrish JK, Edelstein-Keshet L. Complexity, pattern, and evolutionary trade-offs in animal aggregation.
471 Science. 1999;284(5411):99-101.
- 472 [4] Zhang HP, Be’er A, Florin EL, Swinney HL. Collective motion and density fluctuations in bacterial
473 colonies. PNAS. 2010;107(31):13626-30.
- 474 [5] Balagam R, Igoshin OA. Mechanism for collective cell alignment in *Myxococcus xanthus* bacteria. PLoS
475 Comput Biol. 2015;11(8):e1004474.
- 476 [6] Gruler H, Dewald U, Eberhardt M. Nematic liquid crystals formed by living amoeboid cells. Eur Phys
477 J B. 1999;11:187-92.

- 478 [7] Duclos G, Garcia S, Yevick H, Silberzan P. Perfect nematic order in confined monolayers of spindle-shaped cells. *Soft Matter*. 2014;10(14):2346-53.
- 479
- 480 [8] Donev A, Cisse I, Sachs D, Variano EA, Stillinger FH, Connelly R, et al. Improving the density of
481 jammed disordered packings using ellipsoids. *Science*. 2004;303(5660):990-3.
- 482 [9] Delaney G, Weaire D, Hutzler S, Murphy S. Random packing of elliptical disks. *Philos Mag Lett*.
483 2005;85(2):89-96.
- 484 [10] Rocks S, Hoy RS. Structure of jammed ellipse packings with a wide range of aspect ratios. *Soft Matter*.
485 2023;19:5701-10.
- 486 [11] Ramaswamy S. The mechanics and statistics of active matter. *Annu Rev Condens Matter Phys*.
487 2010;1(1):323-45.
- 488 [12] Shaebani MR, Wysocki A, Winkler RG, Gompper G, Rieger H. Computational models for active matter.
489 *Nat Rev Phys*. 2020;2(4):181-99.
- 490 [13] Marchetti MC, Joanny JF, Ramaswamy S, Liverpool TB, Prost J, Rao M, et al. Hydrodynamics of soft
491 active matter. *Reviews of modern physics*. 2013;85(3):1143.
- 492 [14] Akoi I. A simulation study on the schooling mechanism in fish. *Nippon Suisan Gakkaishi*.
493 1982;48(8):1081-8.
- 494 [15] Shu R, Tadmor E. Anticipation breeds alignment. *Arch Ration Mech Anal*. 2021;240:203-41.
- 495 [16] Wang WY, Pearson AT, Kutys ML, Choi CK, Wozniak MA, Baker BM, et al. Extracellular matrix
496 alignment dictates the organization of focal adhesions and directs uniaxial cell migration. *APL Bioeng*.
497 2018;2(4).
- 498 [17] Maruthamuthu V, Sabass B, Schwarz US, Gardel ML. Cell-ECM traction force modulates endogenous
499 tension at cell-cell contacts. *PNAS*. 2011;108(12):4708-13.
- 500 [18] Li X, Balagam R, He TF, Lee PP, Igoshin OA, Levine H. On the mechanism of long-range orientational
501 order of fibroblasts. *PNAS*. 2017;114(34):8974-9.
- 502 [19] Wershof E, Park D, Jenkins RP, Barry DJ, Sahai E, Bates PA. Matrix feedback enables diverse higher-
503 order patterning of the extracellular matrix. *PLoS Comput Biol*. 2019;15(10):e1007251.
- 504 [20] Murray JD. Mathematical Biology: II: Spatial Models and Biomedical Applications. vol. 3. Springer;
505 2003.
- 506 [21] Albi G, Pareschi L. Modeling of self-organized systems interacting with a few individuals: from micro-
507scopic to macroscopic dynamics. *Appl Math Lett*. 2013;26(4):397-401.

- 508 [22] Edelstein-Keshet L, Ermentrout GB. Models for contact-mediated pattern formation: cells that form
509 parallel arrays. *J Math Biol.* 1990;29:33-58.
- 510 [23] Großmann R, Peruani F, Bär M. Mesoscale pattern formation of self-propelled rods with velocity
511 reversal. *Phys Rev E.* 2016;94(5):050602.
- 512 [24] He S, Green Y, Saeidi N, Li X, Fredberg JJ, Ji B, et al. A theoretical model of collective cell polarization
513 and alignment. *J Mech Phys Solids.* 2020;137:103860.
- 514 [25] Manhart A. Counter-propagating wave patterns in a swarm model with memory. *J Math Biol.*
515 2019;78(3):655-82.
- 516 [26] Markham D, Baker R, Maini P. Modelling collective cell behaviour. *Disc Cont Dyn Sys.*
517 2014;34(12):5123-33.
- 518 [27] Mogilner A, Edelstein-Keshet L. Selecting a common direction: I. How orientational order can arise
519 from simple contact responses between interacting cells. *J Math Biol.* 1995;33:619-60.
- 520 [28] Mogilner A, Manhart A. Agent-based modeling: case study in cleavage furrow models. *Mol Biol Cell.*
521 2016;27(22):3379-84.
- 522 [29] Yates C, Baker R, Erban R, Maini P. Refining self-propelled particle models for collective behaviour.
523 *Can Appl Maths Quat.* 2010;18(3).
- 524 [30] Cucker F, Smale S. Emergent behavior in flocks. *IEEE Trans Automat Contr.* 2007;52(5):852-62.
- 525 [31] D'Orsogna MR, Chuang YL, Bertozzi AL, Chayes LS. Self-propelled particles with soft-core interactions:
526 patterns, stability, and collapse. *Phys Rev Lett.* 2006;96(10):104302.
- 527 [32] Vicsek T, Czirók A, Ben-Jacob E, Cohen I, Shochet O. Novel type of phase transition in a system of
528 self-driven particles. *Phys Rev Lett.* 1995;75(6):1226.
- 529 [33] Park D, Wershof E, Boeing S, Labernadie A, Jenkins RP, George S, et al. Extracellular matrix anisotropy
530 is determined by TFAP2C-dependent regulation of cell collisions. *Nat Mater.* 2020;19(2):227-38.
- 531 [34] Degond P, Manhart A, Yu H. An age-structured continuum model for myxobacteria. *Math Models
532 Methods Appl Sci.* 2018;28(09):1737-70.
- 533 [35] Moure A, Gomez H. Phase-field modeling of individual and collective cell migration. *Arch Comput
534 Methods Eng.* 2021;28:311-44.
- 535 [36] Najem S, Grant M. Phase-field model for collective cell migration. *Phys Rev E.* 2016;93(5):052405.
- 536 [37] Nonomura M. Study on multicellular systems using a phase field model. *PLoS One.* 2012;7(4):e33501.

- 537 [38] Löber J, Ziebert F, Aranson IS. Collisions of deformable cells lead to collective migration. *Sci Rep.*
538 2015;5(1):9172.
- 539 [39] Shao D, Rappel WJ, Levine H. Computational model for cell morphodynamics. *Phys Rev Lett.*
540 2010;105(10):108104.
- 541 [40] Peskin C. The immersed boundary method. *Acta numerica.* 2002;11:479-517.
- 542 [41] Strychalski W, Copos CA, Lewis OL, Guy RD. A poroelastic immersed boundary method with appli-
543 cations to cell biology. *J Comput Phys.* 2015;282:77-97.
- 544 [42] Cooper FR, Baker RE, Fletcher AG. Numerical analysis of the immersed boundary method for cell-based
545 simulation. *SIAM J Sci Comput.* 2017;39(5):B943-67.
- 546 [43] Graner F, Glazier J. Simulation of biological cell sorting using a two-dimensional extended Potts model.
547 *Phys Rev Lett.* 1992;69(13):2013.
- 548 [44] Matsushita K. Cell-alignment patterns in the collective migration of cells with polarized adhesion. *Phys*
549 *Rev E.* 2017;95(3):032415.
- 550 [45] Debets VE, Janssen LM, Storm C. Enhanced persistence and collective migration in cooperatively
551 aligning cell clusters. *Biophys J.* 2021;120(8):1483-97.
- 552 [46] Müller R, Boutillon A, Jahn D, Starruß J, David NB, Brusch L. Collective cell migration due to
553 guidance-by-followers is robust to multiple stimuli. *Front Appl Math Stat.* 2023;9:1163583.
- 554 [47] Alt S, Ganguly P, Salbreux G. Vertex models: from cell mechanics to tissue morphogenesis. *Philos*
555 *Trans R Soc Lond B Biol Sci.* 2017;372(1720):20150520.
- 556 [48] Barton DL, Henkes S, Weijer CJ, Sknepnek R. Active vertex model for cell-resolution description of
557 epithelial tissue mechanics. *PLoS Comput Biol.* 2017;13(6):e1005569.
- 558 [49] Koride S, Loza AJ, Sun SX. Epithelial vertex models with active biochemical regulation of contractility
559 can explain organized collective cell motility. *APL Bioeng.* 2018;2(3).
- 560 [50] Li B, Sun SX. Coherent motions in confluent cell monolayer sheets. *Biophys J.* 2014;107(7):1532-41.
- 561 [51] Starruß J, De Back W, Brusch L, Deutsch A. Morpheus: a user-friendly modeling environment for
562 multiscale and multicellular systems biology. *Bioinformatics.* 2014;30(9):1331-2.
- 563 [52] Lee HG, Lee KJ. Neighbor-enhanced diffusivity in dense, cohesive cell populations. *PLoS Comput Biol.*
564 2021;17(9):e1009447.
- 565 [53] Palsson E, Othmer HG. A model for individual and collective cell movement in Dictyostelium dis-
566 coideum. *PNAS.* 2000;97(19):10448-53.

- 567 [54] Palsson E. A 3-D model used to explore how cell adhesion and stiffness affect cell sorting and movement
568 in multicellular systems. *J Theor Biol.* 2008;254(1):1-13.
- 569 [55] Kenny FN, Marcotti S, De Freitas DB, Drudi EM, Leech V, Bell RE, et al. Autocrine IL-6 drives cell
570 and extracellular matrix anisotropy in scar fibroblasts. *Matrix Biol.* 2023;123:1-16.
- 571 [56] Chen WT. Mechanism of retraction of the trailing edge during fibroblast movement. *J Cell Biol.*
572 1981;90(1):187-200.
- 573 [57] Felder S, Elson EL. Mechanics of fibroblast locomotion: quantitative analysis of forces and motions at
574 the leading lamellas of fibroblasts. *J Cell Biol.* 1990;111(6):2513-26.
- 575 [58] Escaff D, Toral R, Van den Broeck C, Lindenberg K. A continuous-time persistent random walk model
576 for flocking. *Chaos.* 2018;28(7).
- 577 [59] Wu PH, Giri A, Wirtz D. Statistical analysis of cell migration in 3D using the anisotropic persistent
578 random walk model. *Nat Protoc.* 2015;10(3):517-27.
- 579 [60] Ilkanaiv B, Kearns DB, Ariel G, Be'er A. Effect of cell aspect ratio on swarming bacteria. *Phys Rev*
580 *Lett.* 2017;118(15):158002.
- 581 [61] Bera P, Wasim A, Mondal J, Ghosh P. Mechanistic underpinning of cell aspect ratio-dependent emergent
582 collective motions in swarming bacteria. *Soft matter.* 2021;17(31):7322-31.
- 583 [62] Jayaram A, Fischer A, Speck T. From scalar to polar active matter: Connecting simulations with
584 mean-field theory. *Phys Rev E.* 2020;101(2):022602.
- 585 [63] Narayan V, Ramaswamy S, Menon N. Long-lived giant number fluctuations in a swarming granular
586 nematic. *Science.* 2007;317(5834):105-8.
- 587 [64] Palachanis D, Szabó A, Merks RM. Particle-based simulation of ellipse-shaped particle aggregation as
588 a model for vascular network formation. *Comput Part Mech.* 2015;2:371-9.
- 589 [65] Shi X, Ma Y. Topological structure dynamics revealing collective evolution in active nematics. *Nat*
590 *Commun.* 2013;4(1):3013.
- 591 [66] Kraikivski P, Lipowsky R, Kierfeld J. Enhanced ordering of interacting filaments by molecular motors.
592 *Physical review letters.* 2006;96(25):258103.
- 593 [67] Peruani F, Deutsch A, Bär M. Nonequilibrium clustering of self-propelled rods. *Phys Rev E.*
594 2006;74(3):030904.
- 595 [68] Yang Y, Marceau V, Gompper G. Swarm behavior of self-propelled rods and swimming flagella. *Physical*
596 *Review E—Statistical, Nonlinear, and Soft Matter Physics.* 2010;82(3):031904.

- 597 [69] Henkes S, Fily Y, Marchetti MC. Active jamming: Self-propelled soft particles at high density. Phys
598 Rev E. 2011;84(4):040301.
- 599 [70] Baskaran A, Marchetti MC. Enhanced diffusion and ordering of self-propelled rods. Physical review
600 letters. 2008;101(26):268101.
- 601 [71] Volkmann N, DeRosier D, Matsudaira P, Hanein D. An atomic model of actin filaments cross-linked by
602 fimbrin and its implications for bundle assembly and function. J Cell Biol. 2001;153(5):947-56.
- 603 [72] Bathe M, Heussinger C, Claessens MM, Bausch AR, Frey E. Cytoskeletal bundle mechanics. Biophys
604 J. 2008;94(8):2955-64.