bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Impact of implementation choices on quantitative predictions of cell-based computational models

Jochen Kursawe[1,*], Ruth E. Baker[1], Alexander G. Fletcher[2,3]

[1] Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, UK

[2] School of Mathematics and Statistics, University of Sheffield, Hicks Building, Hounsfield Road, Sheffield, S3 7RH, UK

[3] Bateson Centre, University of Sheffield, Sheffield, S10 2TN, UK

[*] E-mail: kursawe@maths.ox.ac.uk

Abstract

'Cell-based' models provide a powerful computational tool for studying the mechanisms underlying the growth and dynamics of biological tissues in health and disease. An increasing amount of quantitative data with cellular resolution has paved the way for the quantitative parameterisation and validation of such models. However, the numerical implementation of cell-based models remains challenging, and little work has been done to understand to what extent implementation choices may influence model predictions. Here, we consider the numerical implementation of a popular class of cell-based models called vertex models, used extensively to simulate epithelial tissues. In two-dimensional vertex models, a tissue is approximated as a tessellation of polygons and the vertices of these polygons move due to mechanical forces originating from the cells. We analyse how the model predictions may be affected by implementation-specific parameters, such as the size of the numerical time step. We find that vertex positions and summary statistics are sensitive to multiple implementation parameters. For example, the predicted tissue size decreases with decreasing cell cycle durations, and cell rearrangement may be suppressed by large time steps. These findings are counter-intuitive and illustrate that model predictions need to be thoroughly analysed and implementation details carefully considered when applying cell-based computational models in a quantitative setting.

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

1 Introduction

Computational modelling is increasingly used in conjunction with experimental studies to understand the self-organisation of biological tissues [1, 2]. Popular computational models include 'cell-based' models that simulate tissue behaviour with cellular resolution. Such models naturally capture stochastic effects and heterogeneity when only few cells are present and can be used to explore tissue behaviour when complex assumptions on the cellular scale prevent us from deriving continuum approximations on the tissue scale. The applications of cell-based models range from embryonic development [3–7], to wound healing [8] and tumour growth [9]. However, the numerical solution of cell-based models remains challenging since multi-scale implementations of such models, coupling processes at the subcellular, cellular, and tissue scale, may suffer from numerical instabilities [10, 11] and such models may include parameters that have no direct physical correlate. These issues are of growing importance as cell-based models become used in an increasingly quantitative way [12–14]. Thus, we need to be aware of any impacts that numerical implementation choices may have on model predictions.

Here, we analyse a well-established class of cell-based model, the vertex model [15], to understand to what extent implementation choices may affect model predictions. Vertex models were originally developed to study inorganic structures, such as foams [16] and grain boundaries [17, 18], where surface tension and pressure drive dynamics. They have since been modified to study epithelial tissues [19–22], one of the major tissue types in animals. Epithelia form polarized sheets of cells with distinct apical ('top') and basal ('bottom') surfaces, with tight lateral attachments nearer their apical surface. The growth and dynamics of such sheets play a central role in morphogenesis and wound healing, as well as in disease; for example, over 80% of cancers originate in epithelia [23]. In two-dimensional vertex models, epithelial cell sheets are approximated by tessellations of polygons representing cell apical surfaces, and vertices (where three or more cells meet) move in response to forces due to growth, interfacial tension and hydrostatic pressure within each cell (figure 1A-C). Vertex models typically include cell growth and proliferation. In addition, cells exchange neighbours through so-called T1 transitions (figure 1D) whenever the length of a cell-cell interface falls below a threshold, and any triangular cell whose area falls below a threshold is removed by a so-called T2 transition (figure 1E).

Vertex models have been used to study a variety of processes in epithelial tissues [3, 5, 6,

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

24–37]. A common approach in such studies is to consider forces on vertices arising as a result of minimizing the total stored energy in the tissue. The functional form for this total stored energy varies between applications, but is typically chosen to reflect the effect of the force-generating molecules which localise at or near the apical surface. This energy function is then used either to derive forces that feed into a deterministic equation of motion for each vertex, which must be integrated over time [4, 27, 38], or else minimized directly assuming the tissue to be in quasistatic mechanical equilibrium at all times [3, 24]. A third approach is to apply a Monte Carlo stochastic simulation to find energy minima [39, 40].

Previous theoretical analyses of vertex models have elucidated ground state configurations [41], inferred bulk material properties [42–44], and introduced ways to superimpose finite-element schemes for diffusing signals with the model geometry [45]. In other work, vertex models have been compared to lattice-based cellular Potts models and other cell-based modelling frameworks [46, 47].

In the case of vertex models of grain-boundaries, the authors of [18] propose an adaptive time stepping algorithm to accurately resolve vertex rearrangements without the need of ad-hoc rearrangement thresholds and provide a numerical analysis of the simulation algorithm. However, vertex models in that context only consider energy terms linear in each grain-grain (or cell-cell) interface length, whereas energies in biophysical models typically depend non-linearly on cell areas and perimeters.

Importantly, previous studies such as [18] do not analyse to what extent changes in hidden model parameters, such as the time step and thresholds for cell rearrangement, can influence vertex configurations and other summary statistics. Here, we analyse a force-propagation implementation of vertex models [48, 49] as applied to a widely studied system in developmental biology, the larval wing disc of the fruit fly *Drosophila* [3, 4, 24]. We conduct convergence analyses of vertex positions with respect to all implementation-specific parameters of our model, and further analyse to what extent experimentally measurable summary statistics of tissue morphology, such as distributions of cell neighbour numbers and areas, depend on these parameters.

We find that vertex model predictions are sensitive to the length of cell cycle duration, the time step, and the size of the edge length threshold for cell rearrangement. Specifically, vertex configurations do not converge as the time step, the edge length threshold for cell rearrangement,

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

or the area threshold for cell removal are reduced. For example, reductions in the cell cycle duration may promote cell removal and reduce the size of the simulated tissue by up to a factor of two. We find that both the size of the time step and the size of the edge length threshold can influence the rate of cell rearrangement. Counterintuitively, the rate of cell removal is robust to changes in the area threshold for cell removal over multiple orders of magnitude. Further, analysing the active forces within the tissue reveals that vertices are subject to stronger forces during periods when cells grow and divide.

The remainder of the paper is organised as follows. In section 2, we describe our vertex model implementation of growth in the *Drosophila* larval wing disc. In section 3 we present our results. Finally, we discuss our results and draw conclusions for the use of cell-based models in quantitative biology in sections 4 and 5.

2 Methods

We consider a vertex model of the growing *Drosophila* wing imaginal disc, a monolayered epithelial tissue that is one of the most widely used applications of vertex models. The wing imaginal disc initially comprises around 30 cells, and undergoes a period of intense proliferation until there are around 10000 or more cells [3, 24]. Here, we outline the technical details of our model implementation. We start by introducing the equations of motion, then describe the initial and boundary conditions and implementations of cell growth and neighbour exchange.

Equations of motion In two-dimensional vertex models epithelial tissues are represented as tessellations of polygons that approximate the apical cell surfaces. We propagate the position of each vertex over time using an overdamped force equation, reflecting that cell junctions are not associated with a momentum. The force equation takes the form

$$\mu \frac{\mathrm{d}\mathbf{x}_i}{\mathrm{d}t} = -\nabla_i E, \tag{1}$$

where μ is the friction strength, $\mathbf{x}_i(t)$ is the position vector of vertex i at time t, and E denotes the total stored energy. The number of vertices in the system may change over time due to cell division and removal. The symbol ∇_i denotes the gradient operator with respect to the

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

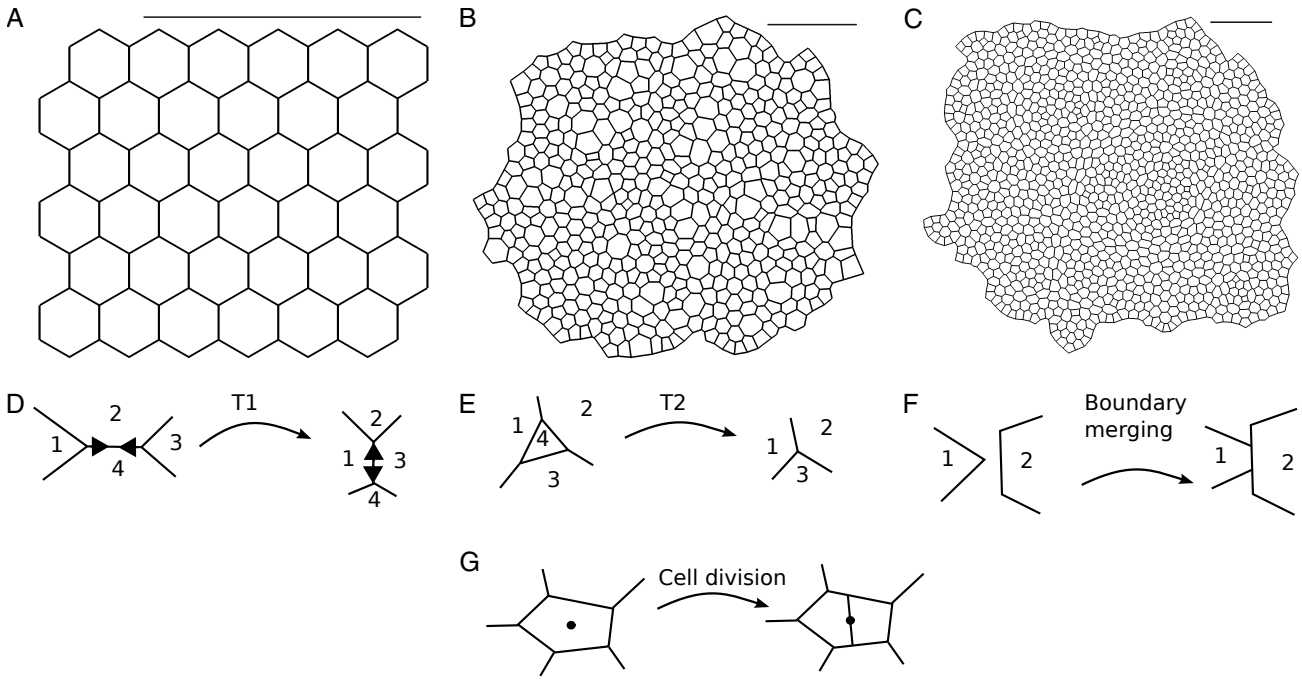

Figure 1: Two-dimensional vertex models represent cells in an epithelial tissue as polygons and allow different types of vertex rearrangement. (A-C) Snapshots of an example vertex model simulation used in our analysis. The growing *in silico* tissue undergoes four rounds of division. (A) The initial condition is a hexagonal packing of 36 cells. (B) Simulation progress after 6,750 time units at an intermediate stage of tissue growth. The tissue boundary is allowed to move freely and individual cells grow before division. (C) Snapshot of tissue at the end of the simulation at 27,000 time units. After the fourth (last) round of divisions the tissue relaxes into a stable configuration. Simulated tissues in (B-C) are rescaled to fit the view, a scale bar of fixed length is added for comparison. Simulation parameters are listed in table 1. Throughout the simulation, vertices may rearrange by T1 transitions (D), T2 transitions (E), boundary merging (F), and cell division (G).

coordinates of vertex i. The total stored energy takes the form

$$E = \sum_{\alpha} \frac{K}{2}(A_\alpha - A_{0,\alpha})^2 + \sum_{\langle i,j \rangle} \Lambda l_{i,j} + \sum_{\alpha} \frac{\Gamma}{2} P_\alpha^2. \tag{2}$$

Here, the first sum runs over every cell α in the tissue, A_α denotes the area of cell α and $A_{0,\alpha}$ is its target area. This term penalises deviations from the target area for individual cells, thus describing cellular bulk elasticity. The second sum runs over all cell edges $\langle i, j \rangle$ in the sheet and penalizes long edges (we choose $\Lambda > 0$), representing the combined effect of binding energy and contractile molecules at the interface between two cells. The third sum also runs over all cells, and P_α denotes the perimeter of cell α. This term represents a contractile acto-myosin cable along the perimeter of each cell [3]. The parameters K, Λ, and Γ together govern the strength

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

of the individual energy contributions.

Before solving the model numerically, we non-dimensionalise it to reduce the number of free parameters [3]. Rescaling space by a characteristic length scale, L, chosen to be the typical length of an individual cell, and time by the characteristic timescale, $T = \mu/KL^2$, equations (1) and (2) become

$$\frac{\mathrm{d}\mathbf{x}'_i}{\mathrm{d}t'} = -\nabla'_i E', \tag{3}$$

$$E' = \sum_\alpha \frac{1}{2}(A'_\alpha - A'_{0,\alpha})^2 + \sum_{\langle i,j\rangle} \overline{\Lambda} l'_{i,j} + \sum_\alpha \frac{\overline{\Gamma}}{2} P'^2_\alpha, \tag{4}$$

where $\mathbf{x}'_i$, A'_α, $A'_{0,\alpha}$, $l'_{i,j}$ and P'_α denote the rescaled i^{th} vertex positions, the rescaled area and target area of cell α, the rescaled length of edge $\langle i,j\rangle$, and the rescaled cell perimeter of cell α, respectively. The symbol ∇'_i denotes the gradient with respect to the rescaled i^{th} vertex position. In the non-dimensionalised model, cell shapes are governed by the rescaled target area of each cell $A'_{0,\alpha}$ and the rescaled mechanical parameters, $\overline{\Lambda}$ and $\overline{\Gamma}$. For these parameters we use previously proposed values [3], unless stated otherwise. A complete list of parameters used in this study is provided in table 1.

To solve equations (3) and (4) numerically we use a forward Euler scheme:

$$\mathbf{x}'_i(t' + \Delta t') = \mathbf{x}'_i(t') - \nabla'_i E'(t')\Delta t'. \tag{5}$$

We analyse the dependence of simulation outcomes on the size of $\Delta t'$ in the Results section.

Initial and boundary conditions Initially, the sheet is represented by a regular hexagonal lattice of six by six cells (figure 1A). The boundary of the lattice is allowed to move freely throughout the simulation. Each cell has initial area and target area $A'^s = A'^s_0 = 1$, respectively.

Cell neighbour exchange and removal T1 transitions (figure 1D) are executed whenever the length of a given edge decreases below the threshold $l'_{\text{T1}} = 0.01$. The length of the new edge, $l_{\text{new}} = \rho\, l_{\text{T1}}$ ($\rho = 1.5$), is chosen to be slightly longer than this threshold to avoid an immediate reversion of the transition.

A second topological rearrangement in vertex models is a T2 transition, during which a

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

small triangular cell or void is removed from the tissue and replaced by a new vertex (figure 1E). In our implementation any triangular cell is removed if its area drops below the threshold $A'_{\mathrm{T2}} = 0.001$. The energy function, equation (2), in conjunction with T2 transitions can be understood as a model for cell removal: cells are extruded from the sheet by a T2 transition if the energy function, equation (2), leads to a sufficiently small cell. Note that in equation (2) the bulk elasticity or area contribution of a cell α is finite even when the area A_α is zero, allowing individual cells to become arbitrarily small if this is energetically favourable. As cells decrease in area they typically also reduce their number of sides. Hence it is sufficient to remove only small triangular cells instead of cells with four or more sides [3, 4, 24].

We further model the merging of overlapping tissue boundaries (figure 1F). Whenever two boundary cells overlap, a new edge of length l_{new} is created that is shared by the overlapping cells. In cases where the cells overlap by multiple vertices, or if the same cells overlap again after a previous merging of edges, the implementation ensures that two adjacent polygons never share more than one edge by removing obsolete vertices. The merging of boundary edges is discussed in further detail in [48].

Cell growth and division Unless stated otherwise the tissue is simulated for $n_d = 4$ rounds of division, i.e. each cell divides exactly n_d times. To facilitate comparison with previous simulations of the wing disc where vertices were propagated by minimising the energy function (2), we model each cell to have two cell cycle phases: quiescent and growing. The duration of the first, quiescent, phase of the cell cycle is drawn independently from an exponential distribution with mean $2t'_l/3$, where t'_l is the total cell cycle duration. We choose stochasticity in this phase of the cell cycle to avoid biologically unrealistic synchronous adjacent divisions; this also helps keeping the simulations in a quasistatic regime. The duration of the second, growing, phase of the cell cycle is fixed at length $t'_l/3$ for each cell. During this time the target area, $A'_{0,\alpha}$, of the cell grows linearly to twice its original value. Upon completion of the growth phase, the cell divides. We choose a fixed duration for the growth phase to ensure gradual, quasistatic cell growth. Two-stage cell cycles with an exponentially distributed and a fixed length contribution have previously been observed in various cell cultures [50, 51] and have been applied to model growth in the *Drosophila* wing imaginal disc [27].

At each cell division event, a new edge is created that separates the newly created daughter

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

cells. The new edge is drawn along the short axis of the polygon that represents the mother cell [48]. The short axis has been shown to approximate the division direction (cleavage plane) of cells in a variety of tissues [52], including the *Drosophila* wing imaginal disc [53]. The short axis of a polygon crosses the centre of mass of the polygon, and it is defined as the axis around which the moment of inertia of the polygon is maximised. Each daughter cell receives half the target area of the mother cell upon division.

Applying this cell cycle model, we let the tissue grow for $n_d = 4$ generations. Each cell of the last generation remains in the quiescent phase of the cell cycle until the simulation stops. We select the total simulation time $t'_{\text{tot}} = 27,000$, unless specified otherwise. This duration is chosen such that the tissue can relax into its equilibrium configuration after the final cell division.

Computational implementation We implement the model within Chaste, an open source C++ library that provides a systematic framework for the simulation of vertex models [48, 49]. Our code is available in the supplementary material as a zip archive. Pseudocode for our implementation is provided in algorithm 1. Each time step starts by updating the cell target areas. Then, cell division, removal (T2 transitions), rearrangement (T1 transitions), and boundary merging are performed before incrementing the simulation time. The algorithm stops when the end time of the simulation is reached.

Initialize time $t' = 0$;
Generate initial configuration;
while $t' < t'_{\text{tot}}$ **do**
1. Update cell target areas;
2. Perform cell division on cells that have reached the end of their cell cycle;
3. Perform any T2 transitions;
4. Perform any T1 transitions;
5. Perform boundary merging;
6. Propagate vertex positions using equation (3);
7. Increment time by $\Delta t'$;

end

Algorithm 1: Pseudocode of the simulation algorithm.

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Table 1: **Description of parameter values used in our simulations.**

Parameter	Description	Value	Reference
$\overline{\Lambda}$	Cell-cell adhesion coefficient	0.12	[3]
$\overline{\Gamma}$	Cortical contractility coefficient	0.04	[3]
$\Delta t'$	Time step	0.01	[48]
A'_{min}	T2 transition area threshold	0.001	[48]
l'_{T1}	T1 transition length threshold	0.01	[48]
ρ	New edges after a T1 transition have the length $l'_{\mathrm{new}} = \rho l'_{\mathrm{T1}}$	1.5	[48]
A'^s	Initial cell area	1.0	[3]
A'^s_0	Initial cell target area	1.0	[3]
N^s	Initial cell number	36	[3]
t'_l	Mean cell cycle duration	1,750	–
t'_{tot}	Simulation duration	27,000	–
n_d	Total number of divisions per cell	4	–

For parameter values for which no reference is given, please see main text for details on how these values were estimated. Spatial and temporal parameters are non-dimensionalised (see section 2 for details).

3 Results

In this section, we analyse how the model behaviour depends on implementation parameters. Vertex models are typically used to predict summary statistics of cellular packing and growth, such as the distribution of cell neighbour numbers and areas [3, 24]. We analyse how these summary statistics may depend on simulation parameters. Specifically, we focus on the final cell number of the simulation, the total tissue area, the numbers of cell rearrangements (T1 transitions) and cell removals (T2 transitions), the distribution of cell neighbour numbers, and the correlation between cell neighbour number and cell area. Note that we exclude cells on the tissue boundary from statistics of cell neighbour numbers in order to avoid boundary artefacts.

Tissue size is sensitive to cell cycle duration

In previous vertex model applications [3, 4, 24], experimentally measured summary statistics of cellular packing were reproduced using an energy-minimisation implementation. Such energy minimisation schemes assume quasistatic evolution of the sheet. It is unclear to what extent summary statistics are preserved when the tissue evolves in a dynamic regime.

We analyse the dependence of the summary statistics on the cell cycle duration, t'_l, in figure 2. The cell number and tissue area at the end of the simulation, and the total number

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

of cell rearrangements, vary by up to a factor of two as the mean cell cycle duration increases from five to 2000. The cell number and tissue area increase with the mean cell cycle duration, whereas the amount of rearrangement decreases, reflecting a reduction in cell removal events. The cell number, tissue area, and the total number of cell rearrangement converge for mean cell cycle durations of 1500 or longer.

The cell neighbour number distribution depends on the cell cycle duration in a non-linear fashion. For example, the number of hexagons peaks at cell cycle durations of 10 as well as at 1000 time units. For cell cycle durations longer than 1000 rescaled time units the numbers of pentagons and heptagons increase as the cell cycle duration increases, while the number of hexagons decreases. We interpret this non-linear dependence as resulting from changes in cell neighbour numbers due to cell division and due to cell neighbour exchanges. As the cell cycle duration exceeds $t'_l = 10$, a decrease in the number of cell removal events leads to an increase in cell division events which, in turn, drives the polygon distribution away from its hexagonal initial condition. As the number of cell divisions ceases to increase the number of cell rearrangements drops as well, and the number of hexagons reach a second peak. Increasing the time between cell divisions further decreases the number of hexagons. Note that none of the simulated polygon histograms coincide with previously reported histograms in which pentagons outweigh hexagons [3, 24], despite choosing identical parameters in energy equation (2). We discuss possible reasons for this difference in section 4.

A further summary statistic of cellular packing is the mean area for cells of each polygon number. This summary statistic often used to characterise epithelia [3, 25, 54, 55]. We find that the mean cell area for each polygon number is not sensitive to changes in cell cycle length and increases monotonically with polygon number.

Differences in tissue size and cellular packing arise due to a sensitive interplay between the cell cycle duration and the timescale for mechanical relaxation of the tissue, T. Growing cells push against their neighbours, leading to tissue growth. This outward movement is counteracted by the friction term in the force equation (1). As cells grow more quickly, i.e. with smaller cell cycle durations, the force required to push the surrounding cells outward increases. For sufficiently small cell cycle durations, the forces may become strong enough to cause cell extrusion.

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

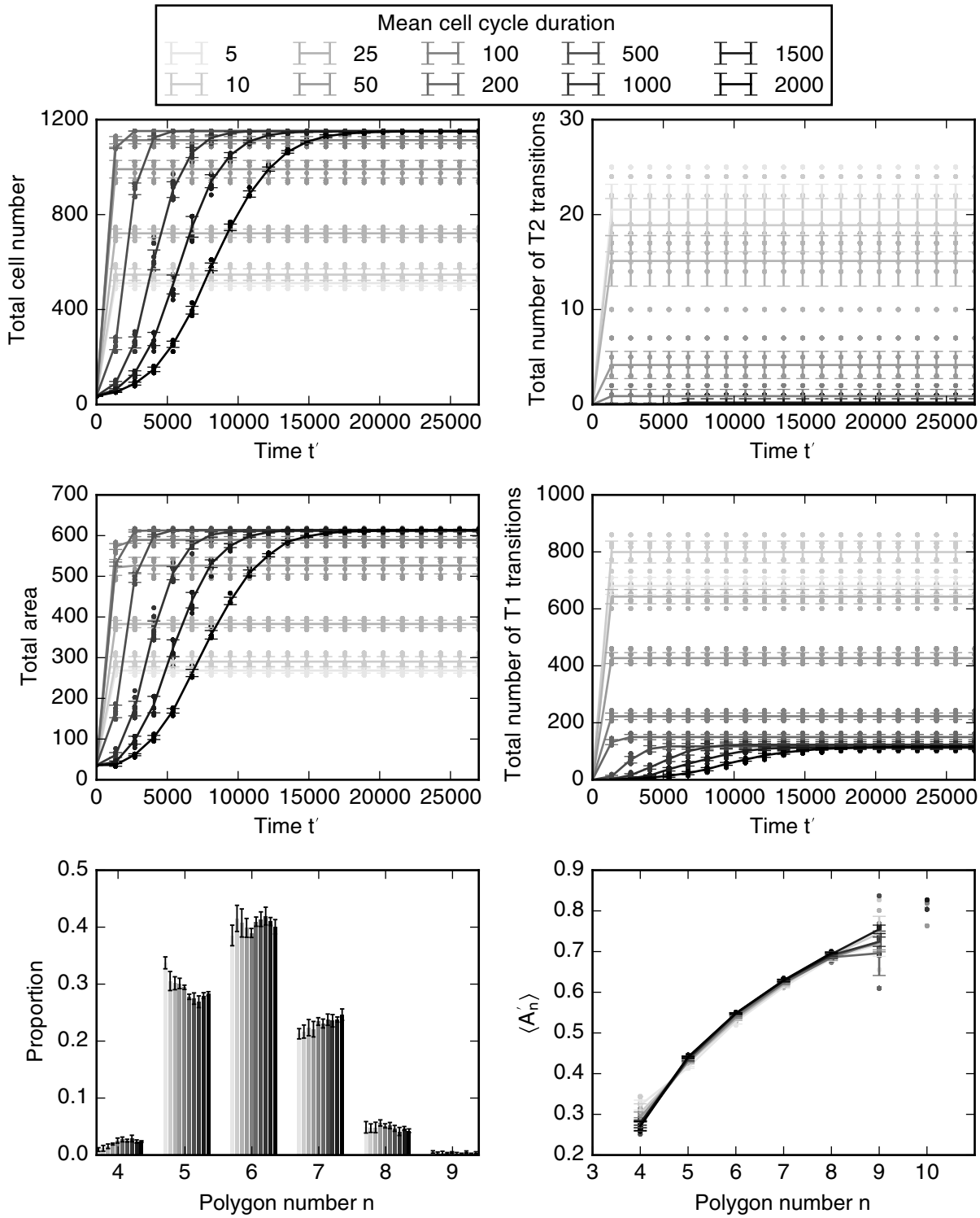

Figure 2: Variation of cell numbers, number of T2 transitions, tissue area, total number of T1 transitions, cell neighbour number distribution and mean area per polygon class with mean cell cycle duration. Error bars denote standard deviations across 100 simulations. All simulation parameters are provided in table 1.

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Cell growth and division increase forces within the tissue

The energy expression (2) leads to three different force contributions on each node: an area force; an edge force; and a perimeter force. In figure 3 we analyse the magnitude of these contributions for a simulation with mean cell cycle duration $t'_l = 2000$. The solid line represents the average magnitudes for the individual contributions for all forces in the tissue, and the shaded areas mark one standard deviation. The strongest force contribution is the area force, whereas the weakest is the edge force. This relationship is intuitive. Most cells in the tissue have areas smaller than their target area of 1.0 (compare with figure 2). Hence, for an individual cell the area force contribution points outwards from the cell. The edge contribution and perimeter contribution point inwards for individual cells, thus counteracting the area force. The variation of each force contribution has the same order of magnitude as their mean values, illustrating that the forces on nodes can vary strongly across the tissue. The force magnitudes change throughout the simulation, and they peak at a value that is 50% higher than the final values. For times larger than 15000 time units, the forces do not change with time in figure 3. At this time cells stop dividing and the final cell number is reached, illustrating that the forces are largest when the tissue size is increasing most rapidly. This transient rise in forces emerges as cells in the interior of the simulated tissue push on their neighbours as they grow before division. This predicts that cells undergoing active processes, such as growth and division, are subject to significantly higher forces than cells in quiescent tissues.

Large time steps suppress cell rearrangement

When using an explicit Euler method, such as as in equation (5), the time step must be chosen sufficiently small to provide a stable and accurate numerical approximation of the model dynamics. To this end, we conduct a convergence analysis. To reduce simulation times, we conduct the convergence analysis on sample simulations in which each cell divides $n_d = 3$ times instead of four, and with total simulation time $t'_{\text{tot}} = 21000$. We choose a series of decreasing time steps $\Delta t'_i$ and define the error function

$$\epsilon_i^t = \left\| \sum_j \mathbf{x}_j^i - \sum_j \mathbf{x}_j^{i-1} \right\|, \tag{6}$$

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

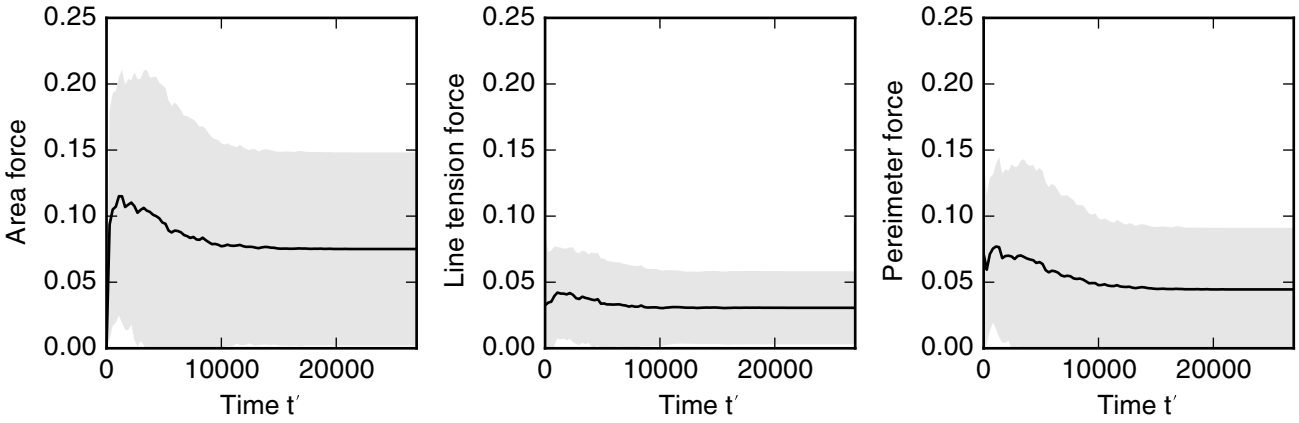

Figure 3: Magnitude of area, edge, and perimeter force contributions over time. The solid lines represent the average of force contribution magnitudes across all nodes for one simulation. The shaded regions represent one standard deviation of the force contribution magnitudes across the tissue. A cell cycle duration of $t'_l = 2000$ is used. All other parameters are listed in table 1.

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

where the sums run over all vertex positions, $\mathbf{x}_j^i$, at the end of the simulation with time steps $\Delta t'_i$ and $\Delta t'_{i-1}$. The error function (6) evaluates the differences between the sums of final vertex positions at decreasing values of the time step. In order to ensure that simulations with consecutive values of the time step follow identical dynamics we generate fixed series of exponentially distributed random variates from which we calculate the cell cycle durations. We analyse the convergence of the vertex positions in the time step $\Delta t'$ in figure 4. In general, the error function does not converge. However, for most simulations the error function (6) assumes values smaller than 10^{-1} for time steps smaller than 10^{-2}. Note that this time step is five orders of magnitudes smaller than the average cell cycle duration. When the time step is larger than 10^{-2} the error function (6) is large since a significant number of T1 transitions are suppressed. On rare occasions, the error function may be non-negligible even if the time step is smaller than 10^{-2}. These large values of the error function (6) reflect changes in the number of T1 transitions as the time step decreases. When the time step is smaller than 10^{-2} summary statistics of cellular packing, such as the distribution of cell neighbour numbers or the total number of cells, do not change as the time step is decreased further.

An example of how differences in the number of T1 transitions and final vertex positions can emerge when the time step is smaller than 10^{-2} is shown in figure 5. In this figure, a cell division occurs in simulations using a time step of 0.004 and 0.002. Both simulation use the same, fixed, series of cell cycle times and vertex positions in both simulations are similar over time up to the time point of the illustrated division. Here, and throughout, cells divide along their short axis. In this example, the short axis of the cell intersects the cell boundary close to an existing vertex. Due to differences in the vertex positions of the cell, the new vertex is created on different cell-cell interfaces as the the size of the time step varies. As the simulation progresses, these different vertex configurations propagate towards different final tissue configurations, leading to differences in the total number of T1 transitions and the error function. In figure 4, differences in final vertex positions are observed for all considered values of the time step. However, such differences in vertex positions do not propagate through to tissue level summary statistics such as the distribution of cell neighbour numbers or areas.

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

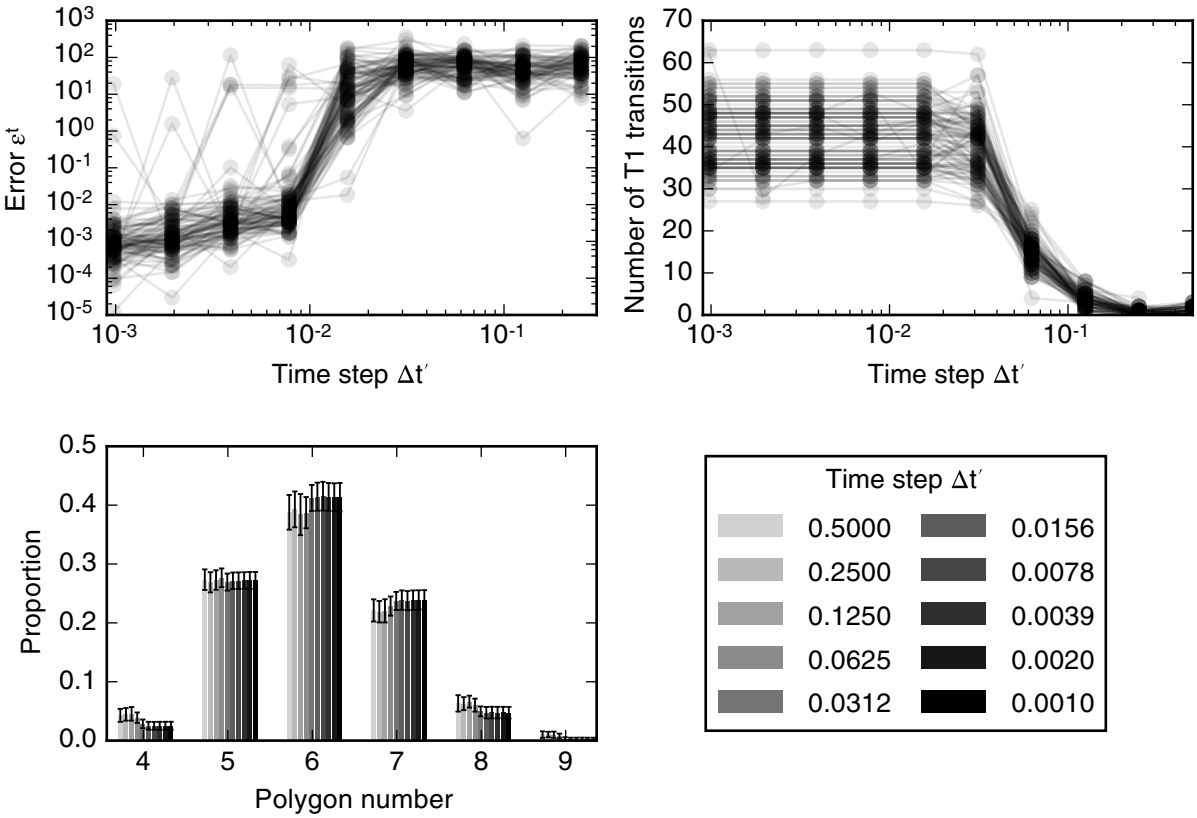

Figure 4: Variation in simulation result with the time step. The error function decreases as the time step is decreased. The number of T1 transitions in the simulations is stable for time steps smaller than 0.02 and decreases with time steps longer than 0.002. For time steps $\Delta t' < 0.02$ the cell neighbour number distribution is stable; the means of individual polygon class proportions vary by less than 0.01. In these simulations, cells undergo $n_d = 3$ rounds of division, and the total simulation time is $t'_{\text{tot}} = 21000$. All other parameter values are listed in table 1. Error bars denote standard deviations across 100 simulations.

Occurrence of cell rearrangements is regulated by rearrangement threshold

We further analyse the dependence of vertex positions and summary statistics on the T1 transition threshold, l'_{T1}. Similar to the time step convergence analysis we define a series of decreasing values of $l'_{\text{T1},i}$ and the error function

$$\epsilon_i^{T1} = \left\| \sum_j \mathbf{x}_j^i - \sum_j \mathbf{x}_j^{i-1} \right\|, \tag{7}$$

which measures the difference between the final vertex positions of simulations with decreasing values of the T1 transition threshold, $l'_{\text{T1},i}$. The variation of the error function with decreasing values of $l'_{\text{T1},i}$ is shown in figure 8. For all considered values of l'_{T1} the error function does not

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

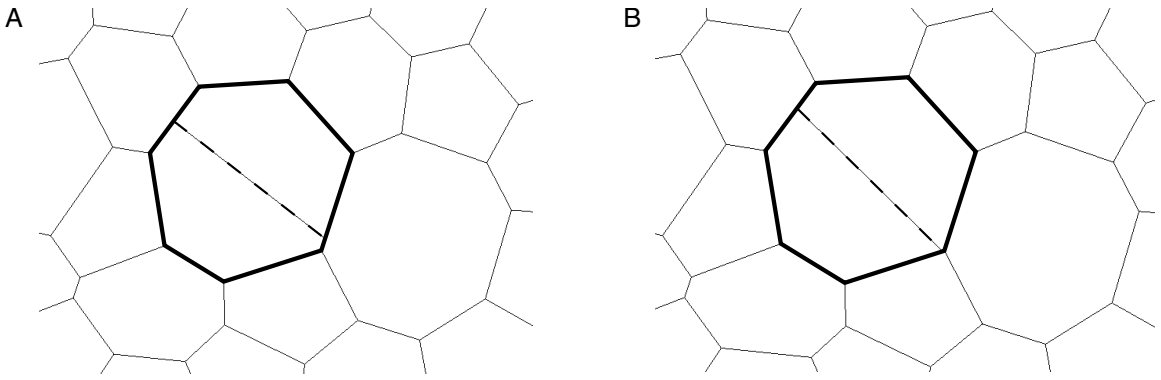

Figure 5: Differences in vertex configurations can arise from simulations run with different temporal resolution. A dividing cell in simulations run with time steps $\Delta t = 0.004$ (A) and $\Delta t = 0.002$ (B) is shown in bold. During the cell division, a new cell-cell interface (dashed line) is created along the short axis of the dividing cell by creating new vertices (see Methods section for details). The daughter cells of the dividing cell contain different vertices in the configurations corresponding to the two time steps. This leads to different vertex configurations at the end of the simulations.

converge and varies between values of 1 and 10^3. Only for $l'_{\mathrm{T1}} < 10^{-3}$ is the error function (7) smaller than one for some simulations. However, for such small values of l'_{T1}, many simulations fail as the simulation algorithm encounters situations that it cannot resolve, for example configurations including overlapping cells.

A large T1 transition threshold of 0.2 length units leads to a large number of T1 transitions, whereas T1 transitions are suppressed for thresholds of 0.003 length units or smaller. This variation in the amount of cell rearrangement influences summary statistics of cellular packing, for example leading to variations in the cell neighbour number distribution. For large rearrangement thresholds, e.g. $l'_{\mathrm{T1}} = 0.2$, the number of cell rearrangements is high, leading to a high proportion of hexagons (around 0.6), whereas suppression of cell rearrangements for small cell rearrangement thresholds, for example $l'_{\mathrm{T1}} = 0.2$, leads to a wider distribution of cell neighbour numbers with a proportion of hexagons (below 0.4). The number of cell rearrangements is stable between T1 transition thresholds of 0.02 and 0.003. In this regime, the proportion of hexagons varies slightly between 0.425 and 0.409. Despite the stable number of T1 transitions across this parameter regime between 0.02 and 0.003 the final vertex positions differ for any two values of the T1 transition threshold, as reflected in values of the error function.

If the T1 transition threshold is smaller than 0.001 simulations fail to complete. An example of how a simulation can fail due to a small value of the T1 transition threshold is provided in figure 7. A snapshot is taken of the simulation at the last two time steps before simulation

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

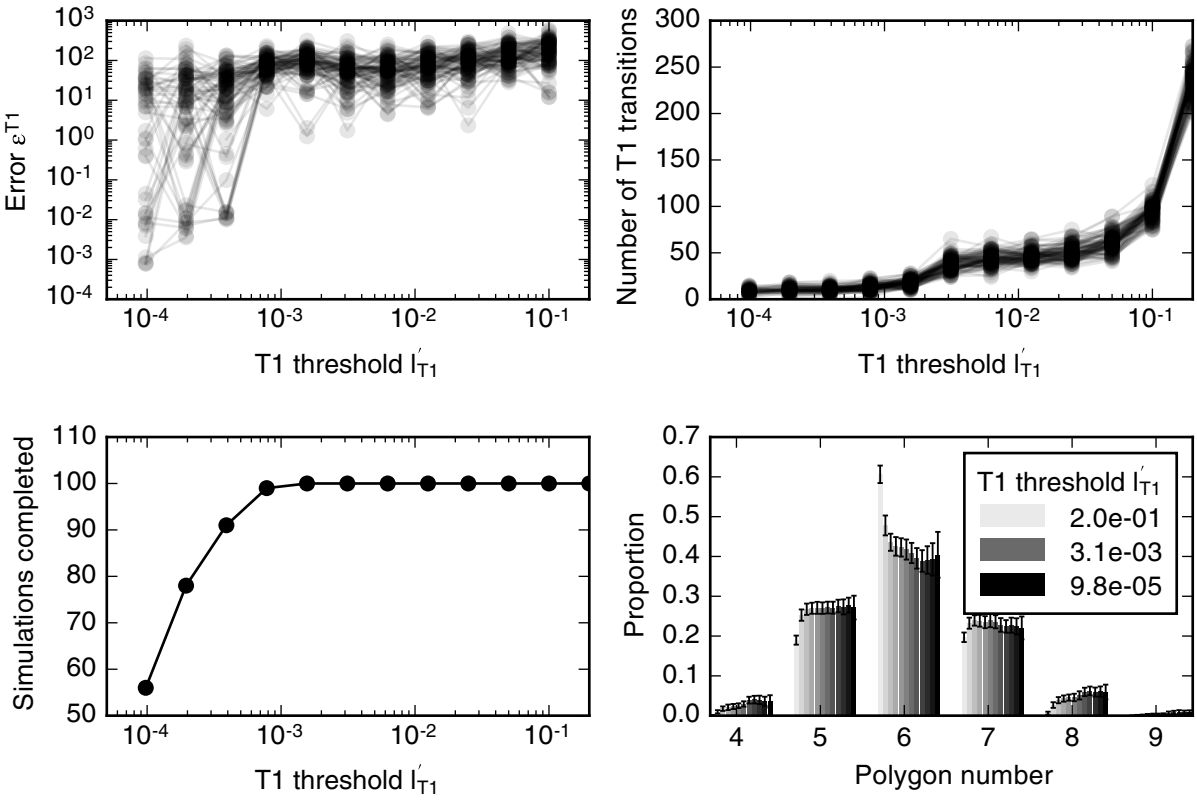

Figure 6: Variation of simulation result with size of the T1 transition threshold, l'_{T1}. The error function (7) does not converge as l'_{T1} decreases. The number of cell rearrangements is larger than 100 for a large value of the rearrangement threshold $l'_{T1} > 0.1$, whereas cell rearrangement is suppressed for small values of the rearrangement threshold $l'_{T1} < 0.001$, with cell rearrangement numbers less than 30. When cell rearrangement is suppressed, simulations fail to complete. Varying amounts of cell rearrangement lead to different distributions of cell neighbour numbers. All simulation parameters are listed in table 1. Error bars denote standard deviations across 100 simulations.

failure. Due to a short edge two boundary vertices in the tissue appear merged (arrow in figure 7A). This short edge is magnified for the penultimate (figure 7B) and last time step (figure 7C) before simulation failure. At the last time step, one of the boundary cells becomes concave. The simulation then fails since our vertex model implementation cannot resolve this configuration. When two boundary cells overlap, the simulation procedure attempts to merge the vertex with its closest cell boundary. This procedure fails because the identified boundary is internal to the tissue rather than a boundary interface.

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

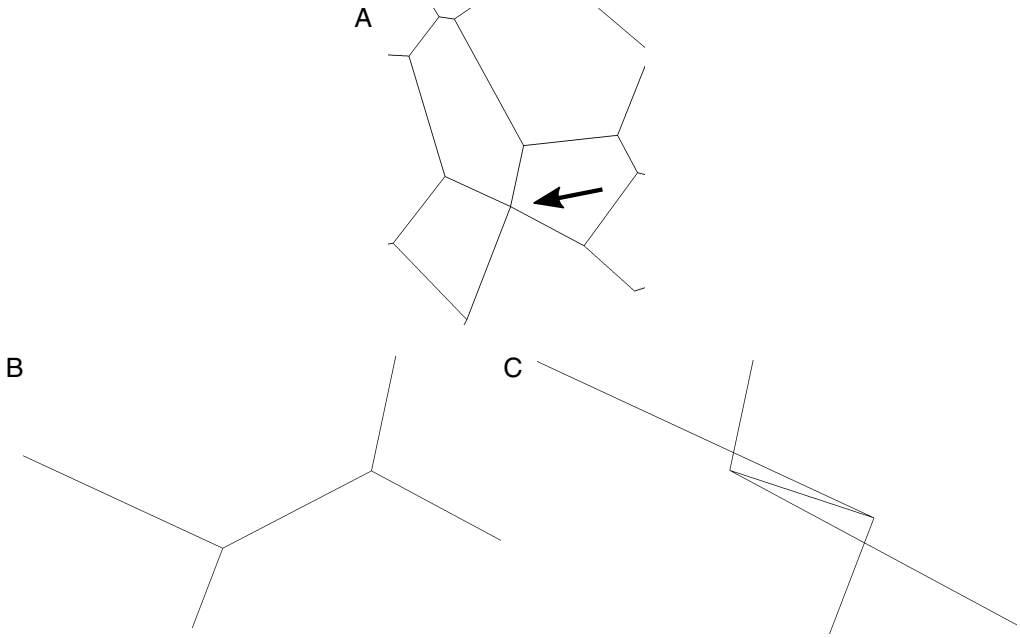

Figure 7: Small values of the T1 transition threshold, $l'_{\mathrm{T1}} < 10^{-3}$, suppress rearrangement and lead to failure of the simulation algorithm. One of the failing simulations in figure 8 is analysed. The tissue configuration in the last time step before simulation failure contains two vertices that appear to be merged due to a short edge on the tissue boundary. The short edge is indicated by an arrow (A) and magnified for the penultimate (B) and final time step (C) of the simulation. Since the the short edge in the penultimate time step is prevented from rearranging, the two adjacent boundary cells intersect each other, leading to failure of the simulation.

Simulation results are robust to variation in length of newly formed edges.

When cells exchange neighbours by way of T1 transitions, new edges are formed. New edges have length $l'_{\mathrm{new}} = \rho l'_{\mathrm{T1}}$. In order to investigate the extent to which changes in the length of newly formed edges can affect simulation results we define a series of increasing values for ρ^j and the error function

$$\epsilon_i^\rho = \left\| \sum_j \mathbf{x}_j^i - \sum_j \mathbf{x}_j^0 \right\|, \tag{8}$$

which measures the difference in vertex positions relative to simulations with $\rho^0 = 1.05$. We find that individual simulations may result in different final tissue configurations than the reference configuration if newly formed edges are twice as long as the rearrangement threshold or longer. Such differences in configuration were observed for three out of 100 simulations, illustrating the robustness of simulation results to the length of newly formed edges.

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

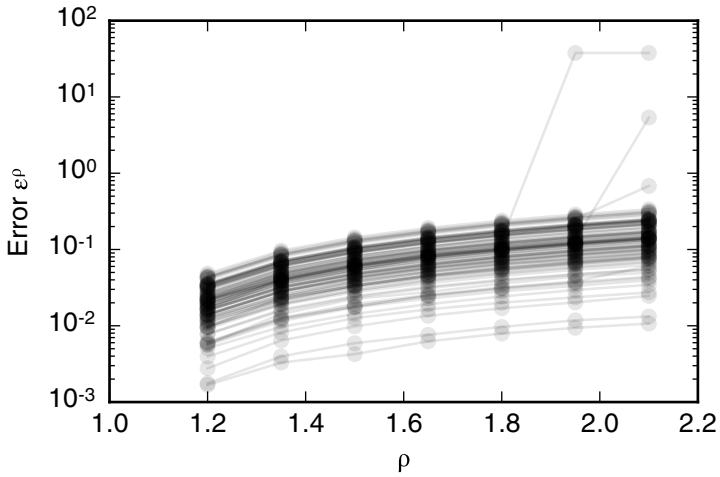

Figure 8: Dependence of simulation results on the length of edges created by T1 transitions, $l'_{\text{new}} = \rho l'_{\text{T1}}$. The error function (8) is recorded for 100 simulations. All simulation parameters are listed in table 1. The error function is smaller than 1 for for $\rho < 2.0$.

Rate of T2 transitions is robust to variation in the T2 transition threshold over five orders of magnitude

Next, we turn to the value of the T2 transition threshold. We define a series of decreasing values of A'^{i}_{T2} and the error function

$$\epsilon_i^{T2} = \left\| \sum_j \mathbf{x}_j^i - \sum_j \mathbf{x}_j^{i-1} \right\|, \tag{9}$$

which measures the difference between the final vertex positions of simulations with decreasing values of the T2 transition threshold, A'^{i}_{T2}. In order to analyse the value of the error function (9) in a simulation with a significant amount of cell rearrangement we run simulations with $n_d = 7$ generations, a cell cycle duration of $t'_l = 700$, and total simulation time $t'_{\text{tot}} = 19600$. All further simulation parameters are listed in table 1.

The value of the error function, on average, is small in figure 9. However, the error function does not converge for individual simulations and may be large between consecutive values of the threshold. In particular, the error function does not converge to zero. As the threshold decreases, the overall number of T2 transitions in the simulations is stable at approximately 150 T2 transitions per simulation. However, for individual simulations the total number of T2 transitions may vary by up to 10 as the threshold A'_{T2} is decreased. The overall number of T2

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

transitions does not change over a large range of T2 transition thresholds that covers multiple orders of magnitude, and all simulations complete without errors even if the T2 transition threshold is smaller than 10^{-6}, which is three orders of magnitudes smaller than the standard value for this parameter in our simulations. The independence of the number of T2 transitions of the threshold A'^{i}_{T2} is reflected in tissue-level summary statistics, such as the distributions of cell neighbour numbers numbers, which are unaffected by changes in the T2 transition threshold.

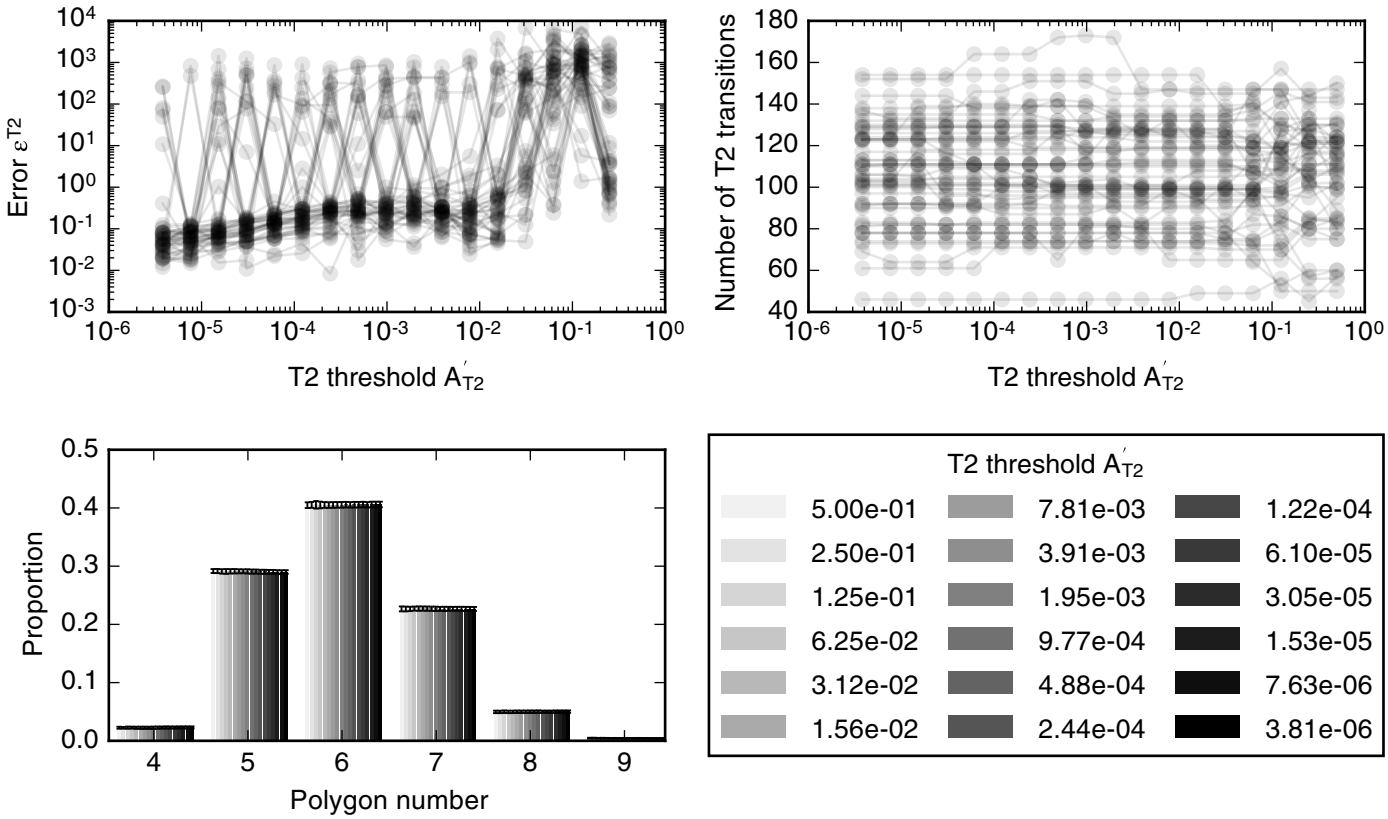

Figure 9: Dependence of simulation results on the T2 transition threshold, A'_{T2}. The error function (9) assumes values less than one for $A_{T2} < 10^{-2}$. The total number of T2 transitions per simulation is stable for all observed values of A_{T2}. Tissue-level summary statistics such as the cell neighbour number distribution are not affected by changes in the threshold. Error bars denote standard deviations across 50 simulations. Simulations are run with $n_d = 7$ rounds of division, a cell cycle duration of $t'_l = 700$, and total simulation time $t'_{\mathrm{tot}} = 19600$. All further simulation parameters are listed in table 1.

Finally, we investigate whether the update ordering within algorithm 1 may affect simulation results. To this end, we randomise the order in which T1 transitions are conducted during one time step. We find that the update order does not lead to differences in final vertex positions in 100 simulations. This is intuitive, considering that the order in which individual events are

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

conducted is most likely to be relevant in situations where events happen directly adjacent to each other, for example if two adjacent edges undergo T1 transitions at the same time step, if there are two adjacent divisions, or if a dividing cell also participates in cell rearrangement. In these examples, the order in which these events occur during one time step may have an impact on simulation outcomes. Our results imply that no adjacent two edges undergo T1 transitions in 100 sample simulations.

4 Discussion

Cell-based models have the potential to help unravel fundamental biophysical mechanisms underlying the growth and dynamics of biological tissues. However, the numerical implementation of such models is rarely analysed and the dependence of model predictions on implementation details often remains unknown. Here, we analyse a widely applied class of cell-based models called a vertex model, and probe to what extent experimentally relevant summary statistics can depend on implementation details, such as choice of timescales or numerical implementation parameters.

For example, we find that the speed with which cells grow and divide relative to the timescale of tissue relaxation can significantly alter the *in silico* tissue behaviour. The total number cells in the tissue, as well as the tissue area and the amount of cell rearrangement, vary by up to a factor of two as the mean cell cycle duration is changed. Summary statistics of cellular packing, such as the distribution of cell neighbour numbers, or the correlation between cell neighbour number and area, are less strongly affected by the exact choice of timescale; the main features of these statistics are preserved in all cases.

The distribution of cell numbers for the case of quasistatic simulations differs from previously reported results [3]. Specifically, we observe fewer pentagons than hexagons. This discrepancy might arise from a difference in how equation (2) is used to evolve the tissue. For example, our implementation of the cell cycle differs from other implementations where the cell cycle duration varies spatially in the tissue [4, 27, 38]. Further, in [3], a global energy minimisation scheme is used to propagate vertex positions, whereas a force-based approach is used here. A major difference between the two approaches is the fraction of cells in the tissue that are allowed to divide concurrently. In our implementation, up to one third of the cells undergo cell-growth

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

at any given time, whereas in previous implementations all cells grow and divide sequentially. Further analysis is required to understand to what extent synchronous growth and division can affect cellular packing in epithelial tissues. Milan et al. report that up to 1.7% of cells in the early wing disc are mitotic at any given time [56]. However, mitosis and cell growth may not happen consecutively, hence the optimal choice of the duration of the growth phase in our simulations is unclear. Overall, it is unclear to what extent different choices for the cell cycle model may influence summary statistics of cellular packing.

Our analysis of forces throughout our simulations in figure 3 reveals that, on average, the area force contribution is stronger than the edge force contribution and the perimeter force contribution on a given node. Further, forces on cells increase during phases of proliferation and growth. Our findings may be of relevance to force-inference approaches that estimate forces using segmented microscopy images of epithelial tissues [57–59]. Force-inference methods often assume that the measured configuration of cells is assumed to be in equilibrium. It is unclear to what extent force-inference approaches introduce errors if this is not the case. In our simulations, forces are up to 50% higher when simulations are run in a dynamic regime, where cells grow and divide, than in the static regime at the end of the simulation, where cells are relaxed into a static configuration.

The vertex positions, as well as simulation summary statistics, vary as the time step of the simulation is changed, and differences in vertex positions decrease with the time step. Counterintuitively, large time steps can suppress cell rearrangement in vertex simulations. This may be explained by considering that, for large time steps, vertex positions move further than the length threshold for cell rearrangements, and instances when the lengths of cell-cell interfaces fall below this threshold may not be resolved. Importantly, in order for differences in simulation results to be negligibly small, a time step has to be chosen that is five orders of magnitude smaller than the average cell cycle duration in our simulation, and six orders of magnitude smaller than the simulation time. For individual simulations, simulation outcomes may change if a smaller time step is chosen. In our simulations, differences in simulation outcomes with decreasing time steps occurred at all observed choices of the time step. More research is required to analyse the extent to which further decreases in the time step can lead to convergence of the simulation results. Here, we stopped investigating the effects of further decreasing the time

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

step due to prohibitive increases in calculation times as the time step is decreased. In previous studies, vertex models have been reported to converge as the time step is decreased [45, 60]. Our analysis differs from these previous studies by considering a tissue undergoing cell division and rearrangement rather than relaxation from an initial condition.

The simulation results are sensitive to the T1 transition threshold chosen in the simulation. The size of the T1 transition threshold can be used to regulate the extent to which the simulated tissue is allowed to rearrange in order to minimise energy. Literature values for this quantity span a range from 0.1 [4] to 0.01 [27]. Final vertex positions of individual simulations change with the value for the T1 transition threshold and do not converge as the threshold is decreased.

Our results that both the time step and the cell rearrangement threshold may influence the rate of T1 transitions illustrates that these parameters are interconnected. When the time step is chosen sufficiently large such that vertices move further than the cell rearrangement threshold between time steps, cell rearrangement is suppressed. This means that if a small cell rearrangement threshold is chosen, a sufficiently small time step needs to be chosen. A careful choice of time steps and cell rearrangement threshold is crucial since an incorrect choice may lead to failure of the simulation algorithm. For vertex models designed to simulate polycrystalline materials an adaptive time stepping scheme has been developed that resolves the exact time at which the end points of a short edge meet, and a T1 transition is performed whenever this happens [18]. More work is required to understand how rates of T1 transitions differ if different conditions for rearrangement are implemented, such as the shortening of an edge to a given threshold or the shrinking edge of an edge to a point. Ultimately, the optimal algorithm to simulate cell rearrangement in epithelial tissues can only be chosen through comparison with experimental results.

While simulated vertex model configurations are sensitive to the size of the time step and thresholds for cell rearrangement, they are less sensitive to the length of newly formed edges, and to thresholds for cell removal. We find that the length of newly formed edges may be up to twice as long as the threshold for T1 transitions without affecting final vertex configurations. Note that this factor by which the length of newly formed edges can be varied without affecting simulation outcomes may change in other parameter regimes, for example if larger values for the cell rearrangement threshold are chosen.

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The size of the area threshold for cell removal may be varied over six orders of magnitude without impacting tissue level summary statistics, even though the exact number of T2 transitions may differ for any two values of the area threshold. In particular, it seems to be possible to choose arbitrarily small values for the T2 transition threshold without causing the algorithm to fail. There are three effects that may contribute to the stability of small elements in our simulations. (i) Since small cells with areas close to the threshold for cell removal are far away from their preferred area in our simulations ($A_{0,\alpha} > 1.0$), their area force is larger than that of adjacent neighbours. This makes the cells stiff and prevents them from becoming inverted or otherwise misshapen. (ii) Additionally, the relationship between area and cell neighbour numbers number in figure 2 shows that small elements are most likely to be triangular. Our simulation algorithm does not permit T1 transitions if the short edge is part of a triangular cell in order to prevent triangular elements from becoming inverted and thus the algorithm from failure. (iii) This relationship between cell area and cell neighbour number may also contribute to the stability of the algorithm when the area threshold is large, for example 0.2. In this case, individual cells may be smaller than the area threshold without undergoing T2 transitions if they are not triangular.

The energy equation (2) provides a geometrical hyphothesis for the removal of cells from epithelia, in which cells are removed from the tissue if this is energetically favourable. Mechanical effects of cell death are an area of increasing biophysical interest [61], and it is subject of future work to design vertex models that allow alternative hypothesis for cell death to be tested.

Here, we analysed how non-biophysical parameters can influence experimentally measurable summary statistics in cell-based models by examining a force-propagation-based implementation of vertex models. Individual results may be relevant to other implementation choices. For example, our finding that the duration of the cell cycle in our model influences simulation outcomes may mean that parameters that control the rate of energy-minimisation may influence results in other vertex model implementations [3, 24, 62]. In general, further work is required to understand how other choices of implementation schemes may impact computational model predictions. For example, the noise strength in a Monte-Carlo vertex propagation scheme [39,40] or the choice of energy-minimisation algorithm may influence vertex model behaviour.

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

5 Conclusions

Our results illustrate that care needs to be taken when drawing predictions using cell-based computational models because implementation details such as the size of the time step may influence model predictions significantly. With the rise of quantitative analysis and quantitative model-data comparison in biophysical applications, choices of model implementation become increasingly relevant. To enable the use of cell-based models in quantitative settings, it is important to be aware of any influences that implementation parameters may have on model predictions when analysing a specific biophysical phenomenon. Understanding model behaviour in detail is crucial to prevent modelling artefacts from influencing experimental predictions and clouding our biophysical understanding and, as such, our findings emphasise the need to fully document algorithms for simulating cell-based models. Close attention to implementation details is required in order to unravel the full predictive power of cell-based models.

Acknowledgements

J.K. acknowledges funding from the Engineering and Physical Sciences Research Council through a studentship. The authors would like to acknowledge the use of the University of Oxford Advanced Research Computing (ARC) facility in carrying out this work.

References

[1] Kitano, H. Systems biology: A brief overview. Science, 295(5560):1662–1664, 2002. doi: 10.1126/science.1069492.

[2] Fisher, J. and Henzinger, T. A. Executable cell biology. Nature Biotechnol., 25(11):1239–49, 2007. doi:10.1038/nbt1356.

[3] Farhadifar, R., Röper, J.-C., Aigouy, B., Eaton, S., and Jülicher, F. The influence of cell mechanics, cell-cell interactions, and proliferation on epithelial packing. Curr. Biol., 17(24):2095–2104, 2007. doi:10.1016/j.cub.2007.11.049.

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

[4] Mao, Y., Tournier, A. L., Bates, P. A., Gale, J. E., Tapon, N., and Thompson, B. J. Planar polarization of the atypical myosin Dachs orients cell divisions in *Drosophila*. Genes Dev., 25(2):131–136, 2011. doi:10.1101/gad.610511.

[5] Trichas, G., Smith, A. M., White, N., Wilkins, V., Watanabe, T., Moore, A., Joyce, B., Sugnaseelan, J., Rodriguez, T. A., Kay, D., et al. Multi-cellular rosettes in the mouse visceral endoderm facilitate the ordered migration of anterior visceral endoderm cells. PLoS Biol., 10(2):e1001256, 2012. doi:10.1371/journal.pbio.1001256.

[6] Monier, B., Gettings, M., Gay, G., Mangeat, T., Schott, S., Guarner, A., and Suzanne, M. Apico-basal forces exerted by apoptotic cells drive epithelium folding. Nature, 518(7538):245–248, 2015. doi:10.1038/nature14152.

[7] Atwell, K., Qin, Z., Gavaghan, D., Kugler, H., Hubbard, E. J., and Osborne, J. M. Mechano-logical model of C. elegans germ line suggests feedback on the cell cycle. Development, 142(22):3902–3911, 2015. doi:10.1242/dev.126359.

[8] Walker, D. C., Hill, G., Wood, S. M., Smallwood, R. H., and Southgate, J. Agent-based computational modeling of wounded epithelial cell monolayers. IEEE Trans. Nanobioscience, 3(3):153–163, 2004.

[9] Anderson, A. R. A., Weaver, A. M., Cummings, P. T., and Quaranta, V. Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment. Cell, 127(5):905–915, 2006. doi:10.1016/j.cell.2006.09.042.

[10] Weinan, E. and Engquist, B. Multiscale modeling and computation. Notices Amer. Math. Soc., 50(9):1062–1070, 2003.

[11] Miller, R. E. and Tadmor, E. B. A unified framework and performance benchmark of fourteen multiscale atomistic/continuum coupling methods. Model. Simul. Mater. Sci. Eng., 17(5):053001, 2009. doi:10.1088/0965-0393/17/5/053001.

[12] Pantazis, P. and Supatto, W. Advances in whole-embryo imaging: a quantitative transition is underway. Nat. Rev. Mol. Cell. Biol., 15(5):327–339, 2014. doi:10.1038/nrm3786.

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

[13] Yu, J. C. and Fernandez-Gonzalez, R. Quantitative modelling of epithelial morphogenesis: integrating cell mechanics and molecular dynamics. Sem. Cell Dev. Biol., 2016. doi:10.1016/j.semcdb.2016.07.030.

[14] Pargett, M. and Umulis, D. M. Quantitative model analysis with diverse biological data: Applications in developmental pattern formation. Methods, 62(1):56–67, 2013. doi:10.1016/j.ymeth.2013.03.024.

[15] Fletcher, A. G., Osterfield, M., Baker, R. E., and Shvartsman, S. Y. Vertex models of epithelial morphogenesis. Biophys. J., 106(11):2291–2304, 2014. doi:10.1016/j.bpj.2013.11.4498.

[16] Okuzono, T. and Kawasaki, K. Intermittent flow behavior of random foams: A computer experiment on foam rheology. Phys. Rev. E, 51(2):1246–1253, 1995. doi:10.1103/PhysRevE.51.1246.

[17] Kawasaki, K., Nagai, T., and Nakashima, K. Vertex models for two-dimensional grain growth. Philos. Mag. B, 60(3):399–421, 1989. doi:10.1080/13642818908205916.

[18] Torres, C. E., Emelianenko, M., Golovaty, D., Kinderlehrer, D., and Ta'asan, S. Numerical analysis of the vertex models for simulating grain boundary networks. SIAM J. Appl. Math., 75(2):762–786, 2015. doi:10.1137/140999232.

[19] Honda, H. Description of cellular patterns by Dirichlet domains: The two-dimensional case. J. Theor. Biol., 72(3):523–543, 1978. doi:10.1016/0022-5193(78)90315-6.

[20] Nagai, T., Kawasaki, K., and Nakamura, K. Vertex dynamics of two-dimensional cellular patterns. J. Phys. Soc. Jpn., 57(7):2221–2224, 1988. doi:10.1143/JPSJ.57.2221.

[21] Weliky, M. and Oster, G. The mechanical basis of cell rearrangement. I. Epithelial morphogenesis during *Fundulus* epiboly. Development, 109(2):373–386, 1990.

[22] Nagai, T. and Honda, H. A dynamic cell model for the formation of epithelial tissues. Philos. Mag. B, 81(7):699–719, 2001. doi:10.1080/13642810108205772.

[23] Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. Molecular Biology of the Cell, 5th Ed. Garland Science, 2008.

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

[24] Aegerter-Wilmsen, T., Smith, A., Christen, A., Aegerter, C., Hafen, E., and Basler, K. Exploring the effects of mechanical feedback on epithelial topology. Development, 137(3):499–506, 2010. doi:10.1242/dev.041731.

[25] Patel, A. B., Gibson, W. T., Gibson, M. C., and Nagpal, R. Modeling and inferring cleavage patterns in proliferating epithelia. PLoS Comput. Biol., 5(6):e1000412, 2009. doi:10.1371/journal.pcbi.1000412.

[26] Landsberg, K. P., Farhadifar, R., Ranft, J., Umetsu, D., Widmann, T. J., Bittig, T., Said, A., Jülicher, F., and Dahmann, C. Increased cell bond tension governs cell sorting at the *Drosophila* anteroposterior compartment boundary. Curr. Biol., 19(22):1950–1955, 2009. doi:10.1016/j.cub.2009.10.021.

[27] Canela-Xandri, O., Sagués, F., Casademunt, J., and Buceta, J. Dynamics and mechanical stability of the developing dorsoventral organizer of the wing imaginal disc. PLoS Comput. Biol., 7(9):e1002153, 2011. doi:10.1371/journal.pcbi.1002153.

[28] Salbreux, G., Barthel, L. K., Raymond, P. A., and Lubensky, D. K. Coupling mechanical deformations and planar cell polarity to create regular patterns in the zebrafish retina. PLoS Comput. Biol., 8(8):e1002618, 2012. doi:10.1371/journal.pcbi.1002618.

[29] Sugimura, K. and Ishihara, S. The mechanical anisotropy in a tissue promotes ordering in hexagonal cell packing. Development, 140(19):4091–4101, 2013. doi:10.1242/dev.094060.

[30] Kursawe, J., Brodskiy, P. A., Zartman, J. J., Baker, R. E., and Fletcher, A. G. Capabilities and limitations of tissue size control through passive mechanical forces. PLoS Comput. Biol., 11(12):e1004679, 2015. doi:10.1371/journal.pcbi.1004679.

[31] Wartlick, O., Mumcu, P., Jülicher, F., and González-Gaitán, M. Understanding morphogenetic growth control - lessons from flies. Nat. Rev. Mol. Cell Biol., 12(9):594–604, 2011. doi:10.1038/nrm3169.

[32] Bi, D., Lopez, J. H., Schwarz, J., and Manning, M. L. Energy barriers and cell migration in densely packed tissues. Soft Matter, 10:1885–1890, 2014. doi:10.1039/C3SM52893F.

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

[33] Marinari, E., Mehonic, A., Curran, S., Gale, J., Duke, T., and Baum, B. Live-cell delamination counterbalances epithelial growth to limit tissue overcrowding. Nature, 484(7395):542–545, 2012. doi:10.1038/nature10984.

[34] Nagai, T. and Honda, H. Computer simulation of wound closure in epithelial tissues: Cell-basal-lamina adhesion. Phys. Rev. E, 80:061903, 2009. doi:10.1103/PhysRevE.80.061903.

[35] Rauzi, M., Verant, P., Lecuit, T., and Lenne, P.-F. Nature and anisotropy of cortical forces orienting *Drosophila* tissue morphogenesis. Nat. Cell Biol., 10(12):1401–1410, 2008. doi:10.1038/ncb1798.

[36] Collinet, C., Rauzi, M., Lenne, P.-F., and Lecuit, T. Local and tissue-scale forces drive oriented junction growth during tissue extension. Nat. Cell Biol., 17(10):1247–1258, 2015. doi:10.1038/ncb3226.

[37] Tetley, R. J., Blanchard, G. B., Fletcher, A. G., Adams, R. J., and Sanson, B. Unipolar distributions of junctional myosin ii identify cell stripe boundaries that drive cell intercalation throughout *Drosophila* axis extension. eLife, 5:1–35, 2016. doi:10.7554/eLife.12094.

[38] Mao, Y., Tournier, A. L., Hoppe, A., Kester, L., Thompson, B. J., and Tapon, N. Differential proliferation rates generate patterns of mechanical tension that orient tissue growth. EMBO J., 32(21):2790–2803, 2013. doi:10.1038/emboj.2013.197.

[39] Etournay, R., Popovi, M., Merkel, M., Nandi, A., Blasse, C., Aigouy, B., Brandl, H., Myers, G., Salbreux, G., Jlicher, F., et al. Interplay of cell dynamics and epithelial tension during morphogenesis of the *Drosophila* pupal wing. eLife, 4:e07090, 2015. doi:10.7554/eLife.07090.

[40] Ray, R. P., Matamoro-Vidal, A., Ribeiro, P. S., Tapon, N., Houle, D., Salazar-Ciudad, I., and Thompson, B. J. Patterned anchorage to the apical extracellular matrix defines tissue shape in the developing appendages of *Drosophila*. Dev. Cell, 34(3):310–322, 2015. doi:10.1016/j.devcel.2015.06.019.

[41] Staple, D., Farhadifar, R., Röper, J.-C., Aigouy, B., Eaton, S., and Jülicher, F. Mechanics and remodelling of cell packings in epithelia. Eur. Phys. J. E, 33:117–127, 2010. doi:10.1140/epje/i2010-10677-0.

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

[42] Merzouki, A., Malaspinas, O., and Chopard, B. The mechanical properties of a cell-based numerical model of epithelium. Soft Matter, 12(21):4745–4754, 2016. doi:10.1039/C6SM00106H.

[43] Xu, G.-K., Liu, Y., and Li, B. How do changes at the cell level affect the mechanical properties of epithelial monolayers? Soft Matter, 11(45):8782–8788, 2015. doi:10.1039/C5SM01966D.

[44] Davit, Y., Osborne, J., Byrne, H., Gavaghan, D., and Pitt-Francis, J. Validity of the cauchy-born rule applied to discrete cellular-scale models of biological tissues. Phys. Rev. E, 87(4):042724, 2013. doi:10.1103/PhysRevE.87.042724.

[45] Smith, A. M., Baker, R. E., Kay, D., and Maini, P. K. Incorporating chemical signalling factors into cell-based models of growing epithelial tissues. J. Math. Biol., 65:441–463, 2012. doi:10.1007/s00285-011-0464-y.

[46] Osborne, J. M., Fletcher, A. G., Pitt-Francis, J. M., Maini, P. K., and Gavaghan, D. J. Comparing individual-based approaches to modelling the self-organization of multicellular tissues. bioRxiv, 2016. doi:10.1101/074351.

[47] Magno, R., Grieneisen, V. A., and Marée, A. F. M. The biophysical nature of cells: potential cell behaviours revealed by analytical and computational studies of cell surface mechanics. BMC Biophys., 8(1):8, 2015. doi:10.1186/s13628-015-0022-x.

[48] Fletcher, A. G., Osborne, J. M., Maini, P. K., and Gavaghan, D. J. Implementing vertex dynamics models of cell populations in biology within a consistent computational framework. Prog. Biophys. Mol. Biol., 113(2):299 –326, 2013. doi:10.1016/j.pbiomolbio.2013.09.003.

[49] Mirams, G. R., Arthurs, C. J., Bernabeu, M. O., Bordas, R., Cooper, J., Corrias, A., Davit, Y., Dunn, S.-J., Fletcher, A. G., Harvey, D. G., et al. Chaste: An open source C++ library for computational physiology and biology. PLoS Comput. Biol., 9(3):e1002970, 2013. doi:10.1371/journal.pcbi.1002970.

[50] Smith, J. A. and Martin, L. Do cells cycle? Proc. Natl. Acad. Sci. USA, 70(4):1263–1267, 1973. doi:10.1073/pnas.70.4.1263.

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

[51] Shields, R. Further evidence for a random transition in the cell cycle. Nature, 273:755–758, 1978. doi:10.1038/273755a0.

[52] Hofmeister, W. Zusätze und Berichtigungen zu den 1851 veröffentlichen Untersuchungengen der Entwicklung höherer Kryptogamen. Jahrb. wiss. Bot., 3:259–293, 1863.

[53] Gibson, W. T., Veldhuis, J. H., Rubinstein, B., Cartwright, H. N., Perrimon, N., Brodland, G. W., Nagpal, R., and Gibson, M. C. Control of the mitotic cleavage plane by local epithelial topology. Cell, 144(3):427–438, 2011. doi:10.1016/j.cell.2010.12.035.

[54] Lewis, F. T. The correlation between cell division and the shapes and sizes of prismatic cells in the epidermis of cucumis. Anat. Rec., 38(3):341–376, 1928. doi:10.1002/ar.1090380305.

[55] Kim, S., Cai, M., and Hilgenfeldt, S. Lewis' law revisited: the role of anisotropy in sizetopology correlations. New J. Phys., 16(1):015024, 2014. doi:10.1088/1367-2630/16/1/015024.

[56] Milán, M., Campuzano, S., and García-Bellido, A. Cell cycling and patterned cell proliferation in the wing primordium of *Drosophila*. Proc. Natl. Acad. Sci. USA, 93:640–645, 1996. doi:10.1073/pnas.93.2.640.

[57] Ishihara, S. and Sugimura, K. Bayesian inference of force dynamics during morphogenesis. J. Theor. Biol., 313(0):201–211, 2012. doi:10.1016/j.jtbi.2012.08.017.

[58] Ishihara, S., Sugimura, K., Cox, S., Bonnet, I., Bellaïche, Y., and Graner, F. Comparative study of non-invasive force and stress inference methods in tissue. Eur. Phys. J. E, 36(4):1–13, 2013. doi:10.1140/epje/i2013-13045-8.

[59] Chiou, K. K., Hufnagel, L., and Shraiman, B. I. Mechanical stress inference for two-dimensional cell arrays. PLoS Comput. Biol., 8(5):e1002512, 2012. doi:10.1371/journal.pcbi.1002512.

[60] Smith, A. M. Vertex model approaches to epithelial tissues in developmental systems. Ph.D. thesis, University of Oxford, 2011.

[61] Ambrosini, A., Gracia, M., Proag, A., Rayer, M., Monier, B., and Suzanne, M. Apoptotic forces in tissue morphogenesis. Mech. Dev., 2016. doi:10.1016/j.mod.2016.10.001.

bioRxiv preprint doi: https://doi.org/10.1101/092924; this version posted December 9, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

[62] Aegerter-Wilmsen, T., Heimlicher, M., Smith, A., de Reuille, P., Smith, R., Aegerter, C., and Basler, K. Integrating force-sensing and signaling pathways in a model for the regulation of wing imaginal disc size. Development, 139(17):3221–3231, 2012. doi:10.1242/dev.082800.