

1 Impact of implementation choices on quantitative predictions of 2 cell-based computational models

3 Jochen Kursawe^{1,*}, Ruth E. Baker¹, Alexander G. Fletcher^{2,3}

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

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

8 ³ Bateson Centre, University of Sheffield, Sheffield, S10 2TN, UK

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

10 Abstract

11 ‘Cell-based’ models provide a powerful computational tool for studying the mechanisms un-
12 derlying the growth and dynamics of biological tissues in health and disease. An increasing
13 amount of quantitative data with cellular resolution has paved the way for the quantitative
14 parameterisation and validation of such models. However, the numerical implementation of
15 cell-based models remains challenging, and little work has been done to understand to what ex-
16 tent implementation choices may influence model predictions. Here, we consider the numerical
17 implementation of a popular class of cell-based models called vertex models, which are often
18 used to study epithelial tissues. In two-dimensional vertex models, a tissue is approximated
19 as a tessellation of polygons and the vertices of these polygons move due to mechanical forces
20 originating from the cells. Such models have been used extensively to study the mechanical
21 regulation of tissue topology in the literature. Here, we analyse how the model predictions
22 may be affected by numerical parameters, such as the size of the time step, and non-physical
23 model parameters, such as length thresholds for cell rearrangement. We find that vertex posi-
24 tions and summary statistics are sensitive to several of these implementation parameters. For
25 example, the predicted tissue size decreases with decreasing cell cycle durations, and cell re-
26 arrangement may be suppressed by large time steps. These findings are counter-intuitive and
27 illustrate that model predictions need to be thoroughly analysed and implementation details

28 carefully considered when applying cell-based computational models in a quantitative setting.

29 **1 Introduction**

30 Computational modelling is increasingly used in conjunction with experimental studies to un-
31 derstand the self-organisation of biological tissues [1,2]. Popular computational models include
32 ‘cell-based’ models that simulate tissue behaviour with cellular resolution. Such models natu-
33 rally capture stochastic effects and heterogeneity when only few cells are present and can be
34 used to explore tissue behaviour when complex assumptions on the cellular scale prevent us
35 from deriving continuum approximations on the tissue scale. The applications of cell-based
36 models range from embryonic development [3–7], to wound healing [8] and tumour growth [9].
37 However, the numerical solution of cell-based models remains challenging since multi-scale im-
38 plementations of such models, coupling processes at the subcellular, cellular, and tissue scales,
39 may suffer from numerical instabilities [10,11], and many such models include parameters of
40 numerical approximation or parameters that have no direct physical correlate. These issues
41 are of growing importance as cell-based models become used in an increasingly quantitative
42 way [12–14]. Thus, we need to be aware of any impacts that numerical implementation choices
43 may have on model predictions.

44 Here, we analyse a well-established class of cell-based model, the vertex model [15], to under-
45 stand to what extent choices of numerical implementation and non-physical model parameters
46 may affect model predictions. Vertex models were originally developed to study inorganic struc-
47 tures, such as foams [16] and grain boundaries [17,18], where surface tension and pressure drive
48 dynamics. They have since been modified to study epithelial tissues [19–22], one of the major
49 tissue types in animals. Epithelia form polarized sheets of cells with distinct apical (‘top’) and
50 basal (‘bottom’) surfaces, with tight lateral attachments nearer their apical surface. The growth
51 and dynamics of such sheets play a central role in morphogenesis and wound healing, as well
52 as in disease; for example, over 80% of cancers originate in epithelia [23]. In two-dimensional
53 vertex models, epithelial cell sheets are approximated by tessellations of polygons representing
54 cell apical surfaces, and vertices (where three or more cells meet) move in response to forces due
55 to growth, interfacial tension and hydrostatic pressure within each cell (figure 1A-C). Vertex
56 models typically include cell growth and proliferation. In addition, cells exchange neighbours

57 through so-called T1 transitions (figure 1D) whenever the length of a cell-cell interface falls
58 below a threshold, and any triangular cell whose area falls below a threshold is removed by a
59 so-called T2 transition (figure 1E).

60 Vertex models have been used to study a variety of processes in epithelial tissues [3–6,24–38].
61 These processes include growth of the *Drosophila* wing imaginal disc [3, 4], migration of the
62 visceral endoderm of mouse embryos [5], and tissue size control in the *Drosophila* embryonic
63 epidermis [31]. A common approach in such studies is to consider forces on vertices arising as
64 a result of minimizing the total stored energy in the tissue. The functional form for this total
65 stored energy varies between applications, but is typically chosen to reflect the effect of the
66 force-generating molecules which localise at or near the apical surface. This energy function
67 is then used either to derive forces that feed into a deterministic equation of motion for each
68 vertex, which must be integrated over time [4, 24, 28], or else minimized directly assuming the
69 tissue to be in quasistatic mechanical equilibrium at all times [3, 25]. A third approach is to
70 apply Monte Carlo algorithms to find energy minima [39, 40].

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

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

82 Importantly, previous studies such as [18] do not analyse to what extent changes in hid-
83 den model parameters, such as parameters of numerical approximation, like the size of the
84 time step, or non-physical model parameters, such as length thresholds for cell rearrangement,
85 can influence vertex configurations and other summary statistics. Here, we analyse a force-
86 propagation implementation of vertex models [48, 49] as applied to a widely studied system in

87 developmental biology, the larval wing disc of the fruit fly *Drosophila* [3, 4, 25]. We conduct
88 convergence analyses of vertex positions with respect to all numerical and non-physical model
89 parameters, and further analyse to what extent experimentally measurable summary statistics
90 of tissue morphology, such as distributions of cell neighbour numbers and areas, depend on
91 these parameters.

92 We find that vertex model predictions are sensitive to the length of cell cycle duration, the
93 time step, and the size of the edge length threshold for cell rearrangement. Specifically, vertex
94 configurations do not converge as the time step, the edge length threshold for cell rearrangement,
95 or the area threshold for cell removal are reduced. For example, reductions in the cell cycle
96 duration may promote cell removal and reduce the size of the simulated tissue by up to a factor
97 of two. We find that both the size of the time step and the size of the edge length threshold can
98 influence the rate of cell rearrangement. Counterintuitively, the rate of cell removal is robust
99 to changes in the area threshold for cell removal over multiple orders of magnitude. Further,
100 analysing the active forces within the tissue reveals that vertices are subject to stronger forces
101 during periods when cells grow and divide.

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

106 2 Methods

107 We consider a vertex model of the growing *Drosophila* wing imaginal disc, a monolayered
108 epithelial tissue that is one of the most widely used applications of vertex models. The wing
109 imaginal disc initially comprises around 30 cells, and undergoes a period of intense proliferation
110 until there are around 10,000 or more cells [3, 25]. Here, we outline the technical details of our
111 model implementation. We start by introducing the equations of motion, then describe the
112 initial and boundary conditions and implementations of cell growth and neighbour exchange.

113 **Equations of motion** In two-dimensional vertex models epithelial tissues are represented as
114 tessellations of polygons that approximate the apical cell surfaces. We propagate the position

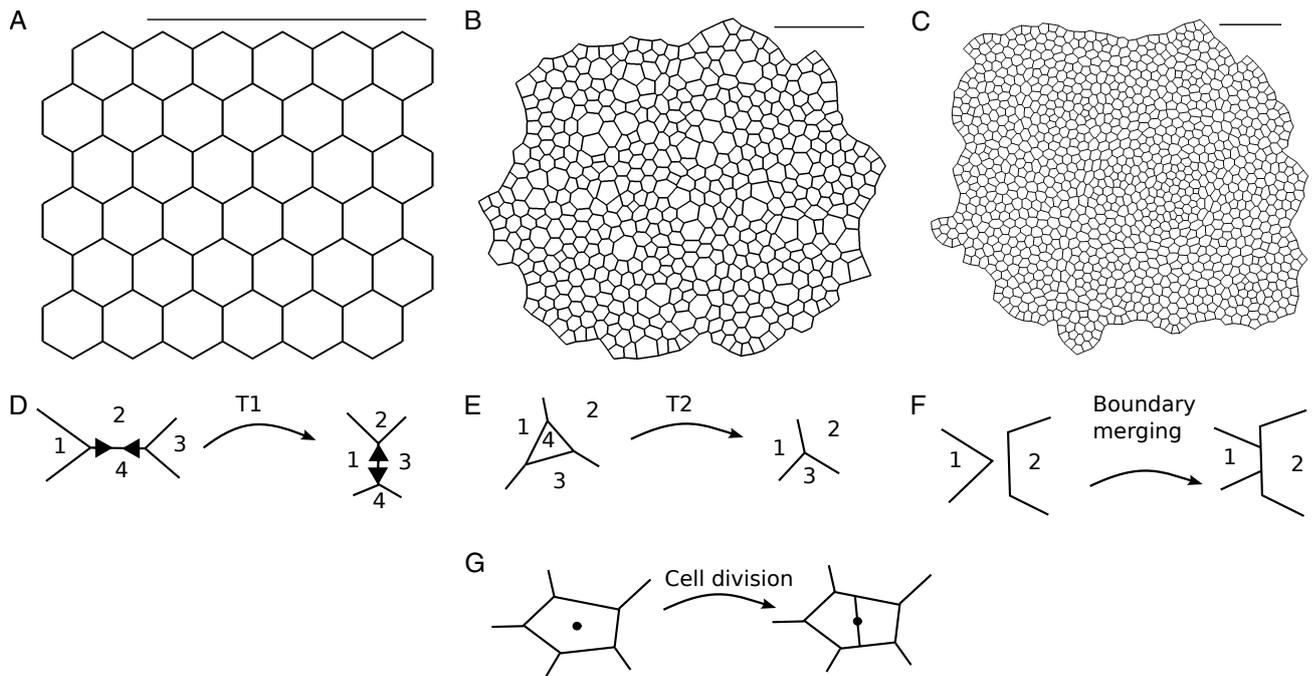


Figure 1: Two-dimensional vertex models represent cells in an epithelial tissue as polygons and allow different types of vertex rearrangement. (A-C) Snapshots of an example vertex model simulation used in our analysis. The growing *in silico* tissue undergoes five rounds of cell division. (A) The initial 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 the tissue at the end of the simulation at 27,000 time units. After the fifth (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. Parameter values 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).

115 of each vertex over time using an overdamped force equation, reflecting that cell junctions are
 116 not associated with a momentum. The force equation takes the form

$$\mu \frac{d\mathbf{x}_i}{dt} = -\nabla_i E, \quad (1)$$

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

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

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

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

$$\frac{d\mathbf{x}'_i}{dt'} = -\nabla'_i E', \quad (3)$$

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

133 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
 134 target area of cell α , the rescaled length of edge $\langle i, j \rangle$, and the rescaled cell perimeter of cell
 135 α , respectively. The symbol ∇'_i denotes the gradient with respect to the rescaled i^{th} vertex
 136 position. In the non-dimensionalised model, cell shapes are governed by the rescaled target
 137 area of each cell $A'_{0,\alpha}$ and the rescaled mechanical parameters, $\bar{\Lambda}$ and $\bar{\Gamma}$. For these parameters
 138 we use previously proposed values [3], unless stated otherwise. A complete list of parameters
 139 used in this study is provided in table 1.

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

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

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

142 **Initial and boundary conditions** Initially, the sheet is represented by a regular hexagonal
 143 lattice of six by six cells (figure 1A). The boundary of the lattice is allowed to move freely

144 throughout the simulation. Each cell has initial area and target area $A^s = A_0^s = 1$, respectively.

145 **Cell neighbour exchange and removal** T1 transitions (figure 1D) are executed whenever
146 the length of a given edge decreases below the threshold $l'_{T1} = 0.01$. The length of the new edge,
147 $l_{\text{new}} = \rho l_{T1}$ ($\rho = 1.5$), is chosen to be slightly longer than this threshold to avoid an immediate
148 reversion of the transition.

149 A second topological rearrangement in vertex models is a T2 transition, during which a
150 small triangular cell or void is removed from the tissue and replaced by a new vertex (figure
151 1E). In our implementation any triangular cell is removed if its area drops below the threshold
152 $A'_{T2} = 0.001$. The energy function, equation (2), in conjunction with T2 transitions can be
153 understood as a model for cell removal: cells are extruded from the sheet by a T2 transition if
154 the energy function, equation (2), leads to a sufficiently small cell. Note that in equation (2) the
155 bulk elasticity or area contribution of a cell α is finite even when the area A_α is zero, allowing
156 individual cells to become arbitrarily small if this is energetically favourable. As cells decrease
157 in area they typically also reduce their number of sides. Hence, it is sufficient to remove only
158 small triangular cells instead of cells with four or more sides [3, 4, 25].

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

165 **Cell growth and division** Unless stated otherwise the tissue is simulated for $n_d = 5$ rounds
166 of division, i.e. each cell divides exactly n_d times. To facilitate comparison with previous
167 simulations of the wing disc where vertices were propagated by minimising the energy func-
168 tion (2) [3, 41], we model each cell to have two cell cycle phases: quiescent and growing. The
169 duration of the first, quiescent, phase of the cell cycle is drawn independently from an expo-
170 nential distribution with mean $2t'_i/3$, where t'_i is the total cell cycle duration. We introduce
171 stochasticity in this phase of the cell cycle to avoid biologically unrealistic synchronous adja-
172 cent divisions; this also helps keeping the simulations in a quasistatic regime since adjacent

173 divisions are prevented from influencing each other, thus maintaining mechanical equilibrium.
174 The duration of the second, growing, phase of the cell cycle is fixed at length $t'_i/3$ for each
175 cell. During this time the target area, $A'_{0,\alpha}$, of the cell grows linearly to twice its original value.
176 Upon completion of the growth phase, the cell divides. We choose a fixed duration for the
177 growth phase to ensure gradual, quasistatic cell growth. Two-stage cell cycles with an expo-
178 nentially distributed and a fixed length contribution have previously been observed in various
179 cell cultures [50, 51] and have been applied to model growth in the *Drosophila* wing imaginal
180 disc [28].

181 The assigning of these cell cycle stages to two thirds and one third of the total cell cycle
182 duration t'_i , respectively, allows us to modify the average age of a dividing cell with a single
183 parameter. This decomposition of the cell cycle ensures that cell cycle durations are stochastic,
184 while allowing the growth phase to occupy a significant proportion of the total cell cycle dura-
185 tion, ensuring gradual, quasistatic growth. The assumption that the tissue is in a quasi-steady
186 state is common in vertex models [3, 27, 28, 34] and reflects the fact that the time scales associ-
187 ated with mechanical rearrangements (seconds to minutes) are an order of magnitude smaller
188 than typical cell cycle times (hours) [3].

189 At each cell division event, a new edge is created that separates the newly created daughter
190 cells (figure 1G). The new edge is drawn along the short axis of the polygon that represents the
191 mother cell [48]. The short axis has been shown to approximate the division direction (cleavage
192 plane) of cells in a variety of tissues [52], including the *Drosophila* wing imaginal disc [53]. The
193 short axis of a polygon crosses the centre of mass of the polygon, and it is defined as the axis
194 around which the moment of inertia of the polygon is maximised. Each daughter cell receives
195 half the target area of the mother cell upon division.

196 Applying this cell cycle model, we let the tissue grow for $n_d = 5$ generations until it contains
197 approximately 1,000 cells, making it sufficiently large to obtain summary statistics of cell pack-
198 ing. Note that the precise number of cells at the end of the simulation varies, due to variations
199 in the number of T2 transitions by which individual cells are removed from the tissue. Each
200 cell of the last generation remains in the quiescent phase of the cell cycle until the simulation
201 stops. We select the total simulation time to be $t'_{\text{tot}} = 27,000$, unless specified otherwise. This
202 duration is chosen such that the tissue can relax into its equilibrium configuration after the final

203 cell division.

204 **Computational implementation** We implement the model within Chaste, an open source
 205 C++ library that provides a systematic framework for the simulation of vertex models [48,
 206 49]. Our code is available in the supplementary material as a zip archive. Pseudocode for
 207 our implementation is provided in algorithm 1. Each time step starts by updating the cell
 208 target areas. Then, cell division, removal (T2 transitions), rearrangement (T1 transitions), and
 209 boundary merging are performed before incrementing the simulation time. The algorithm stops
 210 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.

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

Parameter	Description	Value	Reference
$\bar{\Lambda}$	Cell-cell adhesion coefficient	0.12	[3]
$\bar{\Gamma}$	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 $l'_{\text{new}} = \rho l'_{\text{T1}}$	1.5	[48]
A'^s	Initial cell area	1.0	[3]
A'_0^s	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).

211 **3 Results**

212 In this section, we analyse how model behaviour depends on numerical and non-physical model
213 parameters. Vertex models are typically used to predict summary statistics of cell packing and
214 growth, such as the distribution of cell neighbour numbers and areas [3, 25]. We analyse how
215 these summary statistics depend on simulation parameters. Specifically, we focus on the final
216 number of cells in the tissue, the total tissue area, the numbers of cell rearrangements (T1
217 transitions) and cell removals (T2 transitions), the distribution of cell neighbour numbers, and
218 the correlation between cell neighbour number and cell area. Note that we exclude cells on the
219 tissue boundary from statistics of cell neighbour numbers in order to avoid boundary artefacts,
220 which can be seen in figure 1C. In figure 1C, cell shapes along the tissue boundary differ from
221 those in the bulk of the tissue, and the cell neighbour number is poorly defined for cells along
222 the tissue boundary, since it does not coincide with the number of cell edges.

223 **Tissue size is sensitive to cell cycle duration**

224 In previous vertex model applications [3, 4, 25], experimentally measured summary statistics
225 of cell packing were reproduced using an energy minimisation implementation. Such energy
226 minimisation schemes assume quasistatic evolution of the sheet, where the tissue is in mechanical
227 equilibrium at all times. It is unclear to what extent summary statistics are preserved when
228 the tissue evolves in a dynamic regime.

229 We analyse the dependence of the summary statistics on the cell cycle duration, t'_i , in figure 2.
230 The cell number and tissue area at the end of the simulation, and the total number of cell
231 rearrangements, vary by up to a factor of two as the mean cell cycle duration increases from five
232 to 2000 non-dimensional time units (figure 2A-D). The cell number and tissue area increase with
233 the mean cell cycle duration, whereas the amount of rearrangement (T1 transitions) decreases,
234 reflecting a reduction in cell removal events (T2 transitions). The cell number and the tissue
235 area do not increase further for mean non-dimensional cell cycle durations larger than 1,000
236 time units. In this regime, the total number of rearrangements and cell removals also cease
237 decreasing. We thus identify this regime as the quasistatic regime, where the tissue maintains
238 mechanical equilibrium throughout the simulation. Note, however, that neither the total cell
239 number, nor the tissue area, the number of cell rearrangements or the number of cell removal

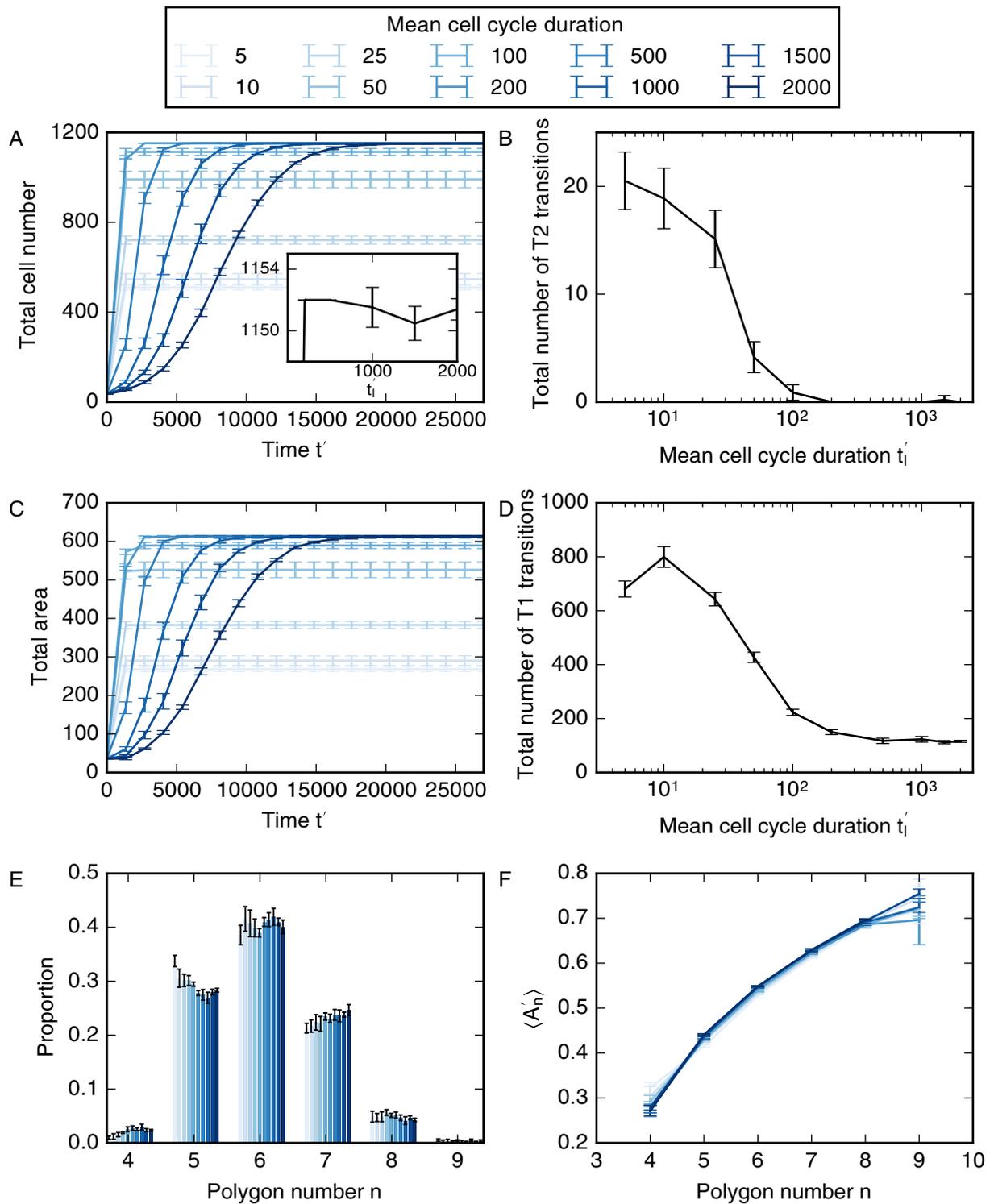


Figure 2: Variation of cell numbers (A), number of T2 transitions (B), tissue area (C), total number of T1 transitions (D), cell neighbour number distribution (E) and mean area per polygon class (F) with mean cell cycle duration. Error bars denote standard deviations across 100 simulations. All simulation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

240 converge numerically as the mean cell cycle duration increases, due to the stochastic nature of
241 the system.

242 The cell neighbour number distribution depends on the cell cycle duration in a non-linear
243 fashion (figure 2E). For example, the number of hexagons peaks at cell cycle durations of 10
244 as well as 1,000 time units. For cell cycle durations longer than 1,000 time units the numbers
245 of pentagons and heptagons increase as the cell cycle duration increases, while the number
246 of hexagons decreases. We interpret this non-linear dependence as resulting from changes in
247 cell neighbour numbers due to cell division and due to cell neighbour exchanges. As the cell
248 cycle duration exceeds $t'_l = 10$, a decrease in the number of cell removal events leads to an
249 increase in cell division events which, in turn, drives the polygon distribution away from its
250 hexagonal initial condition. As the number of cell divisions ceases to increase the number of cell
251 rearrangements drops as well, and the number of hexagons reaches a second peak. Increasing
252 the time between cell divisions further decreases the number of hexagons. Note that none of the
253 simulated polygon histograms coincide with previously reported histograms in which pentagons
254 outweigh hexagons [3, 25], despite choosing identical parameters in energy equation (2). We
255 discuss possible reasons for this difference in section 4.

256 Another common summary statistic of cell packing is the mean area of cells of each polygon
257 number $\langle A'_n \rangle$, where $\langle \cdot \rangle$ denotes an average across all cells in the tissue that are not on the tissue
258 boundary, A' is the rescaled cell area, and n is the polygon number, i.e. the number of neighbours
259 that each cell has. This summary statistic is often used to characterise epithelia [3, 26, 54, 55].
260 We find that the mean cell area for each polygon number is not sensitive to changes in cell cycle
261 length and increases monotonically with polygon number (figure 2F).

262 We interpret the data in figure 2 as follows. Differences in tissue size and cell packing arise
263 due to a sensitive interplay between the cell cycle duration and the timescale for mechanical
264 relaxation of the tissue, T . Growing cells push against their neighbours, leading to tissue
265 growth. This outward movement is counteracted by the friction term in the force equation (1).
266 As cells grow more quickly, i.e. with smaller cell cycle durations, the force required to push the
267 surrounding cells outward increases. For sufficiently small cell cycle durations, the forces may
268 become strong enough to cause cell extrusion. This finding is may not be biologically relevant
269 when studying growth in the *Drosophila* wing imaginal disc, since in this system the time scales

270 for mechanical rearrangement are orders of magnitude smaller than the time scales associated
271 with growth and proliferation [3]. However our results suggest that, in other systems, where
272 cells divide on the time scales of minutes rather than hours, such as the *Drosophila* embryonic
273 epidermis, cell extrusion may be induced during periods of fast tissue growth.

274 Cell growth and division increase forces within the tissue

275 The energy expression (4) leads to three different force contributions on each vertex: an area
276 force; an edge force; and a perimeter force. In figure 3 we analyse the magnitude of these
277 contributions for a simulation with mean cell cycle duration $t'_i = 2000$. The solid line represents
278 the average magnitudes for the individual contributions for all forces in the tissue, and the
279 shaded areas mark one standard deviation. The strongest force contribution is the area force
280 (figure 3A), whereas the weakest is the edge force (figure 3B). This relationship is intuitive
281 if one considers the directions of the individual force contributions when both $\bar{\Lambda}$ and $\bar{\Gamma}$ are
282 positive: Most cells in the tissue have areas smaller than their target area of 1.0 (compare with
283 figure 2F), hence for an individual cell, the area force contribution points outwards from the
284 cell. The edge contribution and perimeter contribution (figure 3C) point inwards for individual
285 cells, thus counteracting the area force. It follows that the area contribution is strongest since,
286 in mechanical equilibrium, it counteracts the sum of the edge and perimeter contributions. The
287 variation of each force contribution has the same order of magnitude as their mean values,
288 illustrating that the forces on vertices can vary strongly across the tissue. The force magnitudes
289 change throughout the simulation, and they peak at a value that is 50% higher than the final
290 values. For times larger than 15000 time units, the forces do not change with time in figure 3.
291 At this time cells stop dividing and the final cell number is reached, illustrating that the forces
292 are largest when the tissue size is increasing most rapidly. This transient rise in forces emerges
293 because cells in the interior of the simulated tissue push on their neighbours as they grow before
294 division. These observations enable us to predict that cells undergoing active processes, such
295 as growth and division, are subject to significantly higher forces than cells in quiescent tissues.

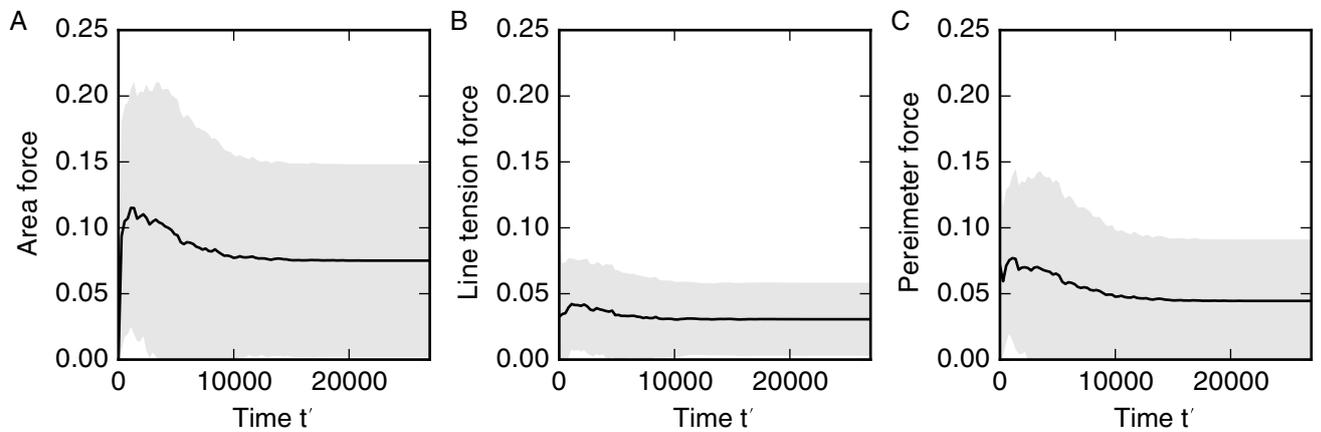


Figure 3: Magnitude of area (A), edge (B), and perimeter force (C) contributions over time. The solid lines represent the average of force contribution magnitudes across all vertices of one simulation. The shaded regions represent one standard deviation of the force contribution magnitudes across the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

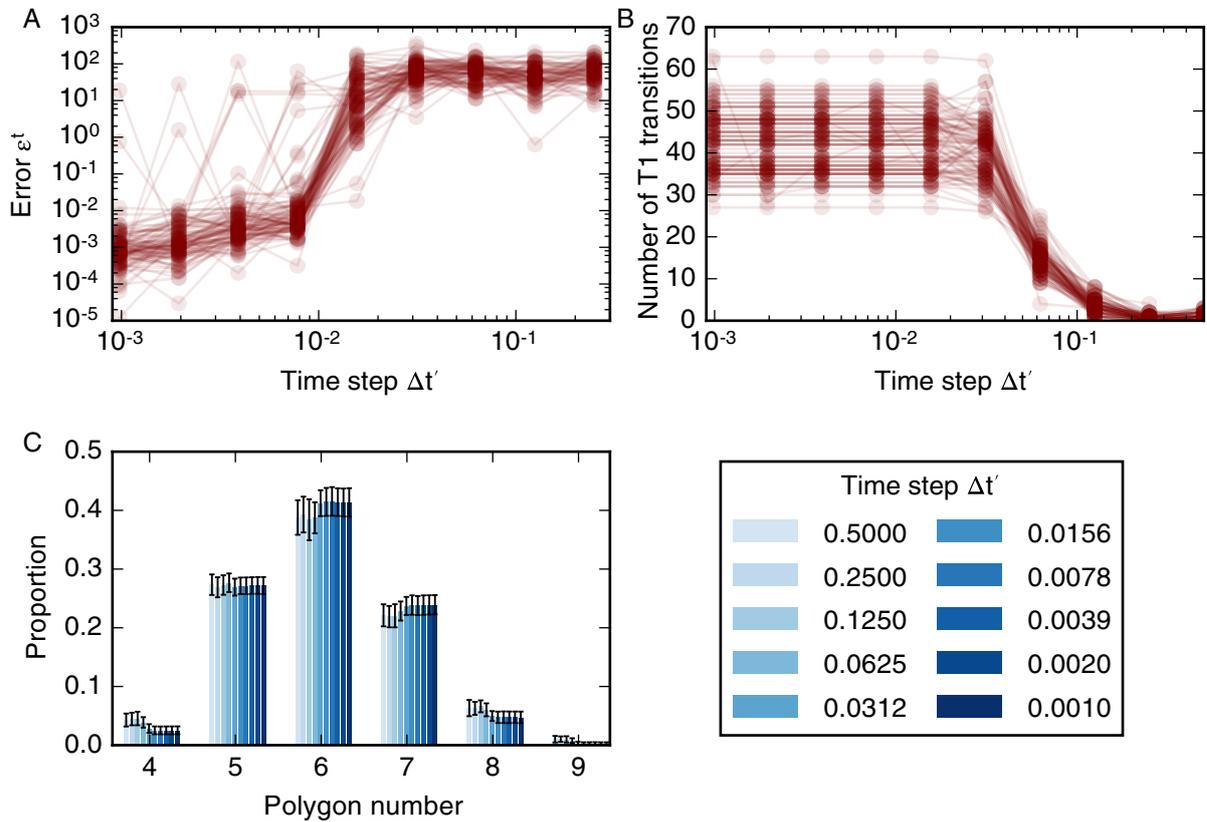


Figure 4: Variation in simulation result with the time step. (A) The error function (6) for 100 different realisations of the model plotted as overlapping, opaque curves. The error function decreases as the time step is decreased, but does not converge for all simulations. (B) The dependence of the number of T1 transitions on the time step for 100 model realisations. The number of T1 transitions in the simulations is stable for time steps smaller than 0.02 and decreases with time steps greater than 0.002. (C) 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 $n_d = 4$ rounds of division, and the total simulation time is $t'_{\text{tot}} = 21,000$. All other parameter values are listed in table 1. Error bars denote standard deviations across 100 simulations.

296 Large time steps suppress cell rearrangement

297 When using an explicit Euler method to propagate the model forward in time, such as in equa-
298 tion (5), the time step should be chosen sufficiently small to provide a stable and accurate
299 numerical approximation of the model dynamics. To this end, we conduct a convergence anal-
300 ysis. To reduce simulation times, we conduct the convergence analysis on sample simulations
301 in which each cell divides $n_d = 4$ times instead of five, and set the total simulation time as
302 $t'_{\text{tot}} = 21,000$. We choose a series of decreasing time steps, $\Delta t'_k$, and define the error function

$$\epsilon_k^t = \left\| \sum_j \mathbf{x}_j^k - \sum_j \mathbf{x}_j^{k-1} \right\|, \quad (6)$$

304 where the sums run over all vertex positions, \mathbf{x}_j^k , at the end of the simulation with time steps
305 $\Delta t'_k$ and $\Delta t'_{k-1}$. The error function (6) evaluates the differences between the sums of final
306 vertex positions at decreasing values of the time step. To ensure that simulations with consec-
307 utive values of the time step follow identical dynamics we generate fixed series of exponentially
308 distributed random variates from which we calculate the cell cycle durations.

309 We plot results of our analysis of the convergence of the vertex positions with the time
310 step $\Delta t'$ in figure 4. In general, the error function does not converge. However, for most
311 simulations the error function (6) assumes values smaller than 10^{-1} for time steps smaller than
312 10^{-2} (figure 4A). Note that this time step is five orders of magnitude smaller than the average
313 cell cycle duration. When the time step is larger than 10^{-2} the error function (6) is larger
314 than one since a significant number of T1 transitions are suppressed. On rare occasions, for
315 less than five examples out of 100, the error function may be non-negligible even if the time
316 step is smaller than 10^{-2} . These large values of the error function (6) reflect changes in the
317 number of T1 transitions as the time step decreases (figure 4B). When the time step is smaller
318 than 10^{-2} summary statistics of cell packing, such as the distribution of cell neighbour numbers
319 (figure 4C) or the total number of cells, do not change as the time step is decreased further.
320 Note that the distribution of cell neighbour numbers in figure 4C differs from those in figure 2
321 due to the decreased number of divisions per cell, n_d . Further, we conclude from our analysis in
322 figure 4 that it is necessary to use a time step smaller than 0.01 in order to arrive at physically
323 meaningful solutions of the vertex model, since otherwise the amount of cell rearrangement
324 and summary statistics of cell packing will be affected by the numerical implementation of the

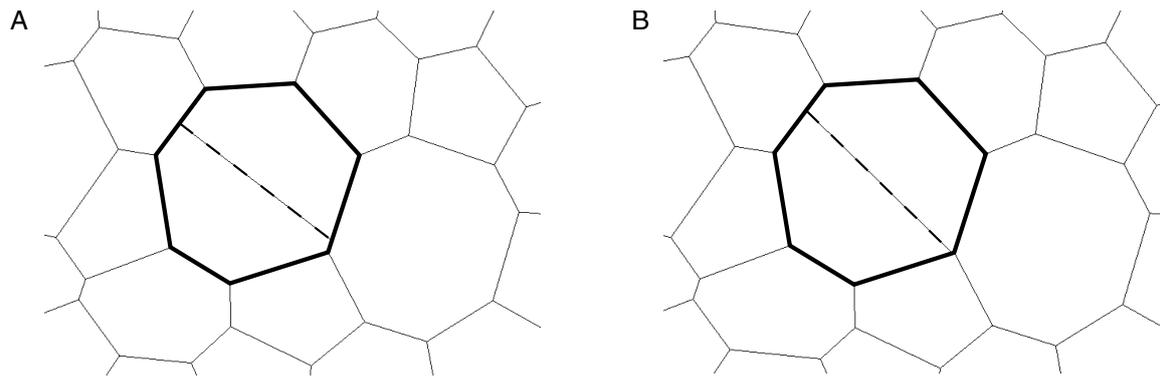


Figure 5: Differences in vertex configurations can arise in simulations run with different temporal resolution. A dividing cell in simulations run with time steps $\Delta t' = 0.004$ (A) and $\Delta t' = 0.002$ (B) is shown in bold. During the cell division, a new cell-cell interface (dashed line) is 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.

325 model.

326 An example of how differences in the number of T1 transitions and final vertex positions
327 can emerge when the time step is smaller than 0.01 is shown in figure 5. In this figure, a cell
328 division occurs in two simulations using a time step of 0.004 (figure 5A) and a time step of 0.002
329 (figure 5B). Both simulations use the same, fixed, series of cell cycle times, and vertex positions
330 in both simulations are similar over time up until the illustrated division. Here, and throughout,
331 cells divide along their short axis. In this example, the short axis of the cell intersects the cell
332 boundary close to an existing vertex. Due to differences in the vertex positions of the cell,
333 the new vertex is created on different cell-cell interfaces as the size of the time step varies. As
334 the simulation progresses, these different vertex configurations propagate towards different final
335 tissue configurations, leading to differences in the total number of T1 transitions and the error
336 function. In figure 4, differences in final vertex positions are observed for all considered values
337 of the time step. However, such differences in vertex positions do not propagate through to
338 tissue-level summary statistics such as the distribution of cell neighbour numbers or areas.

339 **Model convergence with time step is not improved if higher-order numerical** 340 **methods are used**

341 The results in figures 4 and 5 were generated by propagating the vertex positions using a forward
342 Euler time-stepping scheme. The choice of a forward Euler scheme over more accurate numerical

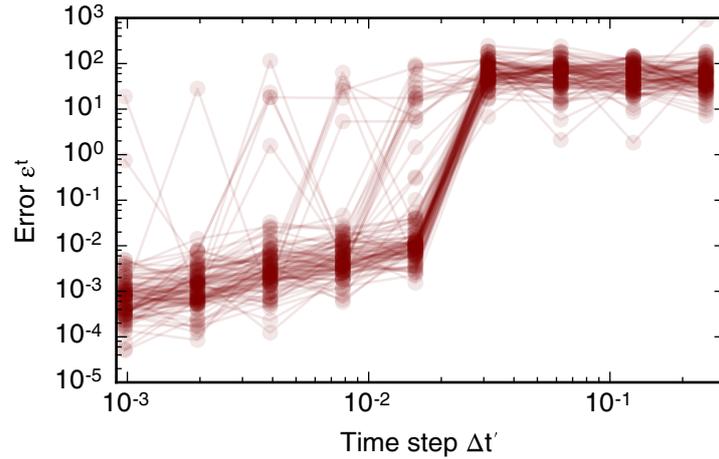


Figure 6: Variation in simulation result with the time step if a fourth-order Runge-Kutta scheme is used. The error function (6) for 100 different realisations of the model, evaluated using a fourth-order Runge-Kutta scheme, is plotted as overlapping, opaque curves. The error function decreases as the time step is decreased, but does not converge for all simulations. This result is similar for simulations run with a forward Euler scheme in figure 4A.

343 methods is common in vertex models. For example, in a previous application where a tissue
344 was relaxed starting from a random initial condition, it was shown that, in order to accurately
345 resolve all T1 transitions, sufficiently small time steps had to be chosen that the benefits of higher
346 order numerical methods were negligible [56]. However, in figures 4 and 5 vertex positions do
347 not converge as the time step is decreased due to differences in T1 transitions and cell divisions
348 for varying values of the time step, suggesting that convergence might be achieved if higher-
349 order numerical methods were used. We test this hypothesis in figure 6, where we record the
350 error function (6) when propagating the vertex model with a fourth-order Runge-Kutta time-
351 stepping scheme as follows. First, all vertices are accumulated into the vertex vector \mathbf{x}' , such
352 that if there are N vertices at time t' then the vector $\mathbf{x}'(t')$ has $2N$ components. We propagate

353 the vertex vector using

$$\mathbf{x}'(t' + \Delta t') = \mathbf{x}'(t') + \frac{\Delta t'}{6} (\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4), \quad (7)$$

$$\mathbf{k}_1 = -\nabla' E'(t', \mathbf{x}'(t')), \quad (8)$$

$$\mathbf{k}_2 = -\nabla' E'(t' + \frac{\Delta t'}{2}, \mathbf{x}'(t') + \frac{\Delta t'}{2} \mathbf{k}_1), \quad (9)$$

$$\mathbf{k}_3 = -\nabla' E'(t' + \frac{\Delta t'}{2}, \mathbf{x}'(t') + \frac{\Delta t'}{2} \mathbf{k}_2), \quad (10)$$

$$\mathbf{k}_4 = -\nabla' E'(t' + \Delta t', \mathbf{x}'(t') + \Delta t' \mathbf{k}_3). \quad (11)$$

354 Here, ∇' denotes the gradient with respect to the vector \mathbf{x} .

355 Similar to the error function obtained using a forward Euler numerical scheme in figure
 356 4A, the error function obtained using a fourth-order Runge-Kutta numerical scheme in figure
 357 6 assumes values smaller than one for time steps below 0.01, but does not converge as the
 358 time step is decreased further. Comparing figures 4A and 6 we conclude that a higher-order
 359 time-stepping scheme does not improve the accuracy of vertex model propagation, since both
 360 the forward Euler and the fourth-order Runge-Kutta scheme require time steps smaller than
 361 roughly 0.01 in order for the error function (6) to assume values smaller than one on average,
 362 while exhibiting a similar degree of variability across all simulations.

363 Occurrence of cell rearrangements is regulated by rearrangement threshold

364 We further analyse the dependence of vertex positions and summary statistics on the T1 transi-
 365 tion threshold, l'_{T1} . Similar to the time step convergence analysis, we define a series of decreasing
 366 values of $l'_{T1,k}$ and the error function

$$\epsilon_k^{T1} = \left\| \sum_j \mathbf{x}_j^k - \sum_j \mathbf{x}_j^{k-1} \right\|, \quad (12)$$

367 which measures the difference between the final vertex positions of simulations with decreasing
 368 values of the T1 transition threshold, $l'_{T1,k}$. The variation of the error function with decreasing
 369 values of $l'_{T1,k}$ is shown in figure 8A. For all considered values of l'_{T1} the error function does
 370 not converge and varies between values of 1 and 10^3 . Only for $l'_{T1} < 10^{-3}$ is the error function
 371 (12) smaller than one for some simulations. However, for such small values of l'_{T1} , many simula-

372 tions fail as the simulation algorithm encounters situations that it cannot resolve, for example
373 configurations including overlapping cells (figure 8B).

374 A large T1 transition threshold of 0.2 length units leads to a large number of T1 transitions,
375 whereas T1 transitions are suppressed for thresholds of 0.003 length units or smaller (figure 8C).
376 This variation in the number of cell rearrangements influences summary statistics of cell pack-
377 ing, for example leading to variations in the cell neighbour number distribution. For large
378 rearrangement thresholds, e.g. $l'_{T1} = 0.2$, the number of cell rearrangements is high, leading
379 to a high proportion of hexagons (around 0.6), whereas suppression of cell rearrangements for
380 small cell rearrangement thresholds, for example $l'_{T1} = 0.2$, leads to a wider distribution of cell
381 neighbour numbers with a proportion of hexagons below 0.4. The number of cell rearrangements
382 is stable between T1 transition thresholds of 0.02 and 0.003. In this regime, the proportion of
383 hexagons varies slightly between 0.425 and 0.409 (figure 8D). Despite the stable number of T1
384 transitions across this parameter regime between 0.02 and 0.003 the final vertex positions differ
385 for any two values of the T1 transition threshold, as reflected in values of the error function.

386 As illustrated in figure 8B, if the T1 transition threshold is smaller than 0.001, simulations
387 fail to complete as the simulation algorithm encounters situations that it cannot resolve, for
388 example due to overlapping or self-intersecting cells. An example of how a simulation can fail
389 due to a small value of the T1 transition threshold is provided in figure 7. A snapshot is taken
390 of the simulation at the last two time steps before simulation failure. Due to a short edge two
391 boundary vertices in the tissue appear merged (arrow in figure 7A). This short edge is magnified
392 for the penultimate (figure 7B) and last time steps (figure 7C) before simulation failure. At
393 this last time step, one of the boundary cells becomes concave. The simulation then fails since
394 our vertex model implementation cannot resolve this configuration. When two boundary cells
395 overlap, the simulation procedure attempts to merge the vertex with its closest cell boundary.
396 This procedure fails because the identified boundary is internal to the tissue rather than a
397 boundary interface.

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

399 When cells exchange neighbours by way of T1 transitions, new edges are formed. Each new
400 edge has length $l'_{\text{new}} = \rho l'_{T1}$. In order to investigate the extent to which changes in the length

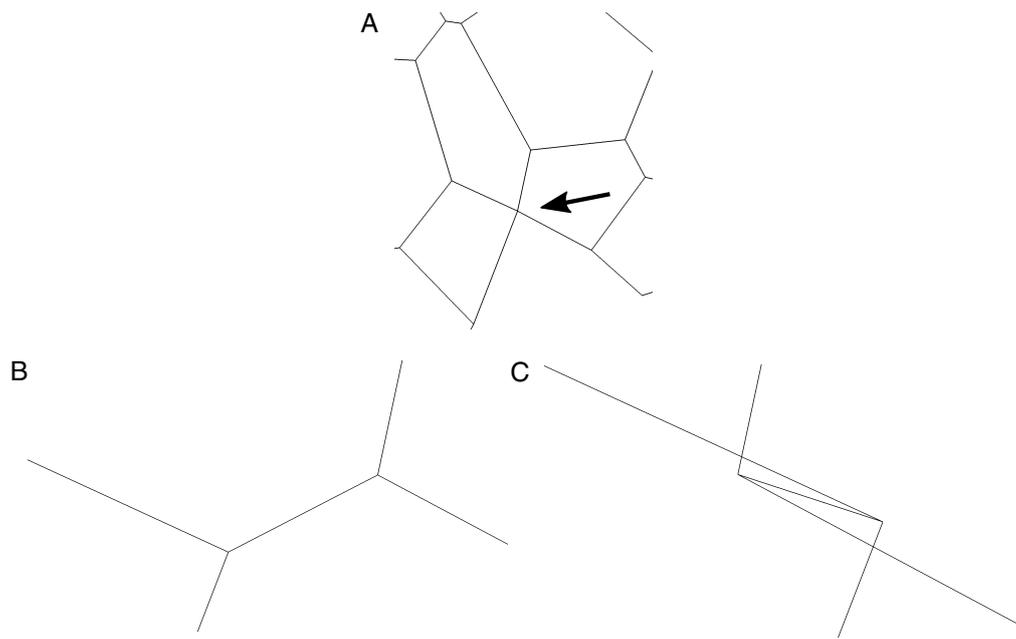


Figure 7: Small values of the T1 transition threshold, $l'_{T1} < 10^{-3}$, suppress rearrangement and lead to failure of the simulation algorithm. One of the failing simulations in figure 8 is analysed. The tissue configuration in the last time step before simulation failure contains two vertices that 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 completed time step (C) of the simulation. Since 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.

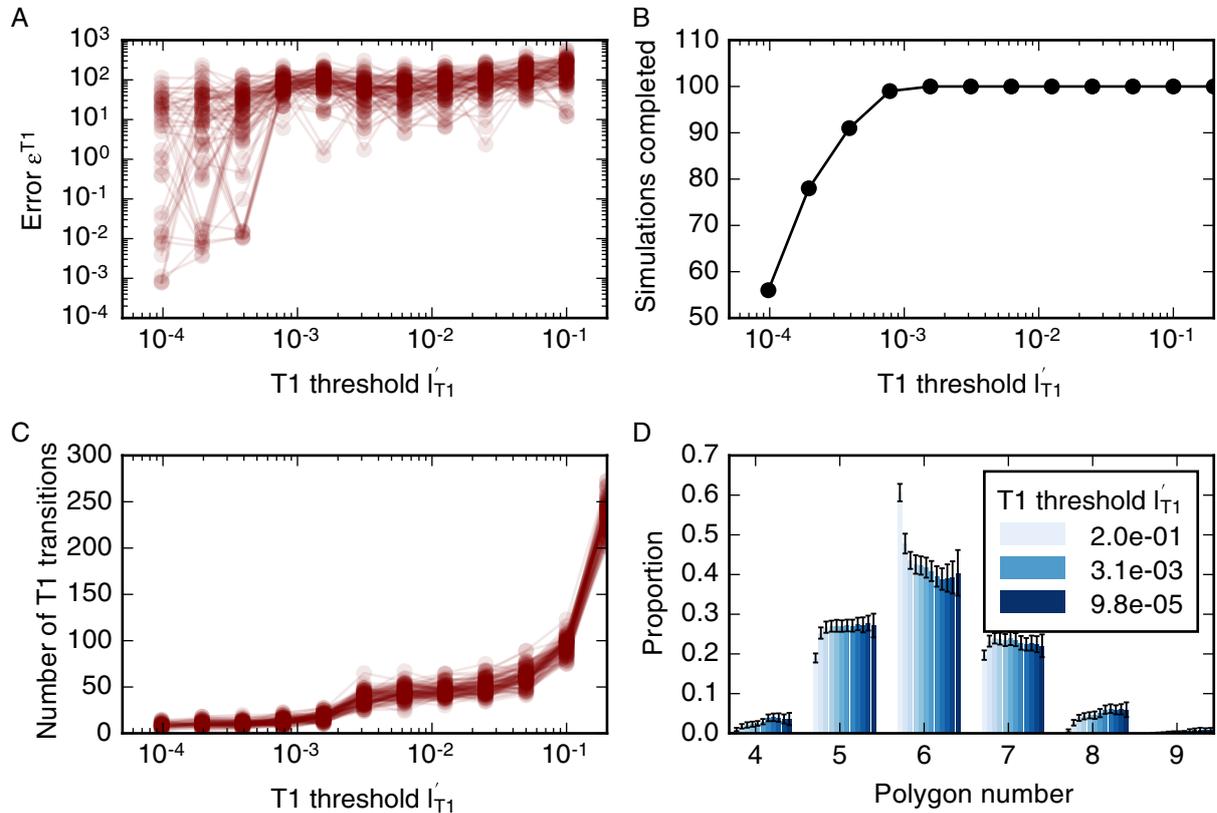


Figure 8: Variation of simulation result with size of the T1 transition threshold, l'_{T1} . (A) The dependence of the error function on l'_{T1} for 100 model realisations. The error function (12) does not converge as l'_{T1} decreases. (B) For small values of the T1 transition threshold, some simulations fail to complete (see main text). (C) The dependence of the number of cell rearrangements on l'_{T1} for 100 model realisations. The number of cell rearrangements is larger than 100 for a large value of the rearrangement threshold, $l'_{T1} > 0.1$, whereas cell rearrangements are suppressed for small values of the rearrangement threshold, $l'_{T1} < 0.001$, with cell rearrangement numbers less than 30. (D) Varying amounts of cell rearrangement lead to different distributions in cell neighbour numbers. Parameter values are listed in table 1. Error bars denote standard deviations across 100 simulations.

401 of newly formed edges can affect simulation results we define a series of increasing values for ρ^k
402 and the error function

$$\epsilon_k^\rho = \left\| \sum_j \mathbf{x}_j^k - \sum_j \mathbf{x}_j^0 \right\|, \quad (13)$$

404 which measures the difference in vertex positions relative to simulations with $\rho^0 = 1.05$. As
405 shown in figure 9, individual simulations may result in different final tissue configurations than
406 the reference configuration if newly formed edges are twice as long as the rearrangement thresh-
407 old or longer. Such differences in configuration were observed for three out of 100 simulations,
408 illustrating the robustness of simulation results to the length of newly formed edges.

409 **Rate of T2 transitions is robust to variation in the T2 transition threshold** 410 **over five orders of magnitude**

411 Next, we turn to the value of the T2 transition threshold. We define a series of decreasing values
412 of A_{T2}^k and the error function

$$\epsilon_k^{T2} = \left\| \sum_j \mathbf{x}_j^k - \sum_j \mathbf{x}_j^{k-1} \right\|, \quad (14)$$

413 which measures the difference between the final vertex positions of simulations with decreasing
414 values of the T2 transition threshold, A_{T2}^k . To analyse the value of the error function (14) in a
415 simulation with a significant amount of cell rearrangement and removal we run simulations with
416 $n_d = 8$ generations, a cell cycle duration of $t'_l = 700$, and total simulation time $t'_{tot} = 19600$.
417 All other parameter values are listed in table 1.

418 The value of the error function, on average, is small (figure 10A). However, the error function
419 does not converge for individual simulations and may be large between consecutive values of
420 the threshold. In particular, the error function does not converge to zero. As the threshold
421 decreases, the overall number of T2 transitions in the simulations is stable at approximately
422 150 T2 transitions per simulation (figure 10B). However, for individual simulations, the total
423 number of T2 transitions may vary by up to 10 as the threshold A'_{T2} is decreased. The overall
424 number of T2 transitions does not change over a large range of T2 transition thresholds that
425 covers multiple orders of magnitude, and all simulations complete without errors even if the
426 T2 transition threshold is smaller than 10^{-6} , which is three orders of magnitudes smaller than

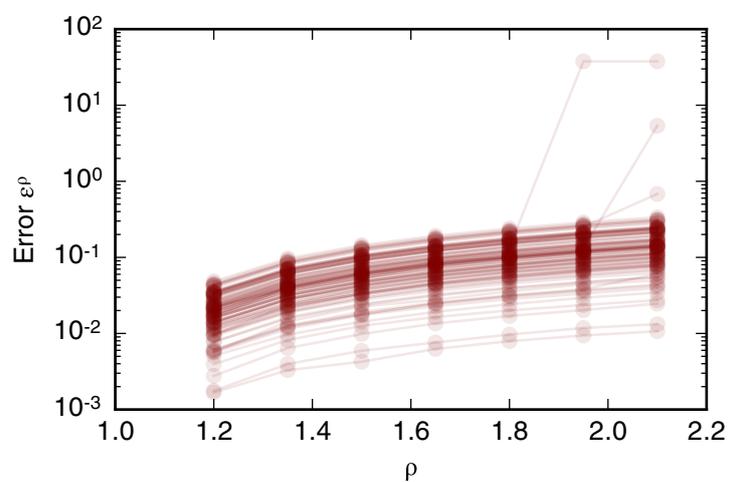


Figure 9: Dependence of simulation results on the length of edges created by T1 transitions, $l'_{\text{new}} = \rho l'_{\text{T1}}$. The error function (13) is recorded for 100 simulations. All simulation parameters are listed in table 1. The error function is smaller than one for $\rho < 2.0$.

427 the standard value for this parameter in our simulations. The independence of the number of
428 T2 transitions of the threshold A_{T2}^k is reflected in tissue-level summary statistics, such as the
429 distributions of cell neighbour numbers, which are unaffected by changes in the T2 transition
430 threshold (figure 10C).

431 **Dependence of the simulation results on the update ordering in each time** 432 **step**

433 Finally, we investigate whether the update ordering within algorithm 1 may affect simulation
434 results. To this end, we randomise the order in which T1 transitions are conducted during one
435 time step. We find that the update order does not lead to differences in final vertex positions
436 in 100 simulations. This is intuitive, considering that the order in which individual events are
437 conducted is most likely to be relevant in situations where events happen directly adjacent to
438 each other, for example if two adjacent edges undergo T1 transitions at the same time step, if
439 there are two adjacent divisions, or if a dividing cell also participates in cell rearrangement. In
440 these examples, the order in which these events occur during one time step may have an impact
441 on simulation outcomes. Our results imply that no adjacent two edges undergo T1 transitions
442 in 100 sample simulations.

443 **4 Discussion**

444 Cell-based models have the potential to help unravel fundamental biophysical mechanisms un-
445 derlying the growth and dynamics of biological tissues. However, the numerical implementation
446 of such models is rarely analysed and the dependence of model predictions on implementation
447 details often remains unexplored. Here, we analyse a widely applied class of cell-based models, a
448 vertex model, and probe to what extent experimentally relevant summary statistics can depend
449 on implementation details, such as the choice of numerical or non-physical model parameters.

450 For example, we find that the speed at which cells grow and divide relative to the speed
451 of tissue relaxation can significantly alter *in silico* tissue behaviour. The total number cells in
452 the tissue, as well as the tissue area and the number of cell rearrangements, varies by up to a
453 factor of two as the mean cell cycle duration is changed. Summary statistics of cell packing,
454 such as the distribution of cell neighbour numbers, or the correlation between cell neighbour
455 number and area, are less strongly affected by the exact choice of timescale; the main features

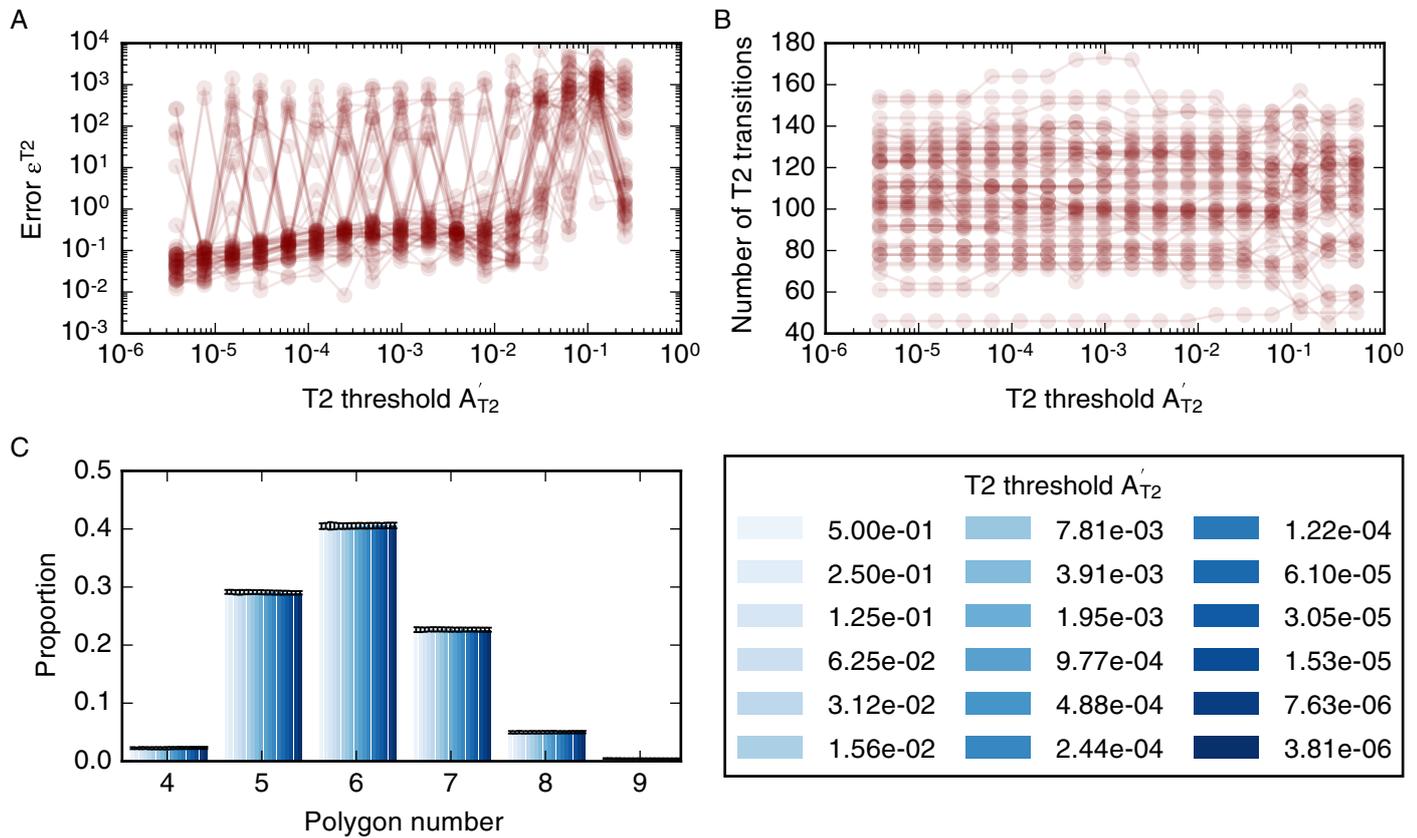


Figure 10: Dependence of simulation results on the T2 transition threshold, A'_{T2} . (A) The dependence of the error function (14) on the T2 transition threshold for 50 model realisations. The error function assumes values less than one for $A_{T2} < 10^{-2}$ but does not converge. (B) The total number of T2 transitions for 50 model realisations is stable for all observed values of A_{T2} . (C) 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 $n_d = 8$ rounds of division, a cell cycle duration of $t'_l = 700$, and total simulation time $t'_{tot} = 19600$. All further simulation parameters are listed in table 1.

456 of these statistics are preserved in all cases. This finding that the total cell number and tissue
457 area depend on the mean cell cycle duration suggests that cell extrusion may be induced in
458 fast-growing tissues.

459 The distribution of cell numbers for the case of quasistatic simulations, identified as simu-
460 lations where increases in the cell cycle duration would not lead to an overall increase in tissue
461 area or cell number, differs from previously reported results [3]. Specifically, we observe fewer
462 pentagons than hexagons. This discrepancy might arise from a difference in how equation (2) is
463 used to evolve the tissue. For example, our implementation of the cell cycle differs from other
464 implementations where the cell cycle duration varies spatially in the tissue [4, 24, 28]. Further,
465 in [3], a global energy minimisation scheme is used to propagate vertex positions, whereas a more
466 accurate force-based approach is used here. A major difference between the two approaches is
467 the fraction of cells in the tissue that are allowed to grow and divide concurrently. In our
468 implementation, up to one third of the cells undergo cell-growth at any given time, whereas
469 in other implementations all cells grow and divide sequentially. Further analysis is required to
470 understand to what extent synchronous growth and division can affect cell packing in epithe-
471 lial tissues. Milan et al. report that up to 1.7% of cells in the early wing disc are mitotic at
472 any given time [57]. However, mitosis and cell growth may not happen consecutively, hence
473 the optimal choice of the duration of the growth phase in our simulations is unclear. Overall,
474 it is unclear to what extent different choices for the cell cycle model may influence summary
475 statistics of cell packing.

476 Our analysis of forces throughout simulations, presented in figure 3, reveals that, on average,
477 the area force contribution is stronger than the edge force contribution and the perimeter force
478 contribution on a given vertex. Further, forces on cells increase during phases of proliferation
479 and growth. Our findings may be of relevance in force-inference approaches that estimate forces
480 using segmented microscopy images of epithelial tissues [58–60]. Force-inference methods often
481 assume that the measured configuration of cells is in equilibrium and it is unclear to what extent
482 force-inference approaches introduce errors if this is not the case. In our simulations, forces are
483 up to 50% higher when simulations are run in a dynamic regime, where cells grow and divide,
484 than in the static regime at the end of the simulation, where cells are relaxed into a static
485 configuration.

486 The vertex positions, as well as simulation summary statistics, vary as the time step is
487 changed, and differences in vertex positions decrease with the time step. Counterintuitively,
488 large time steps can suppress cell rearrangement in vertex simulations. This may be explained by
489 considering that, for large time steps, vertex positions move further than the length threshold
490 for cell rearrangements, and instances when the lengths of cell-cell interfaces fall below this
491 threshold may not be resolved. Importantly, in order for differences in simulation results to
492 be negligibly small, a time step has to be chosen that is five orders of magnitude smaller
493 than the average cell cycle duration in our simulation, and six orders of magnitude smaller
494 than the simulation time. For individual simulations, simulation outcomes may change if a
495 smaller time step is chosen, an effect that is preserved even when a higher-order numerical
496 scheme, such as fourth-order Runge-Kutta, is used. The latter finding confirms that, for vertex
497 model implementations with ad-hoc rules for cell rearrangement and division, such as in this
498 study, the benefits of higher-order numerical schemes diminish, and it is beneficial to reduce
499 the computational cost of the algorithm by using a simpler numerical scheme, such as forward
500 Euler. A forward Euler scheme is more computationally efficient than a fourth-order Runge-
501 Kutta scheme since it requires fewer floating point operations per time step. In our simulations,
502 differences in simulation outcomes with decreasing time steps occurred at all observed choices
503 of the time step for both numerical schemes investigated. More research is required to analyse
504 the extent to which further decreases in the time step can lead to convergence of the simulation
505 results. Here, we stopped investigating the effects of further decreasing the time step due to
506 prohibitive increases in calculation times as the time step is decreased. In previous studies,
507 vertex models have been reported to converge as the time step is decreased [45, 56]. Our
508 analysis differs from these previous studies by considering a tissue undergoing cell division and
509 rearrangement rather than relaxation from an initial condition.

510 The simulation results are sensitive to the T1 transition threshold chosen in the simulation.
511 The size of the T1 transition threshold can be used to regulate the extent to which the simulated
512 tissue is allowed to rearrange in order to minimise energy. Literature values for this quantity
513 span a range from 0.1 [4, 48] to 0.01 [31]. Final vertex positions of individual simulations change
514 with the value for the T1 transition threshold and do not converge as the threshold is decreased.

515 Our results that both the time step and the cell rearrangement threshold may influence

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

529 While simulated vertex model configurations are sensitive to the size of the time step and
530 thresholds for cell rearrangement, they are less sensitive to the length of newly formed edges,
531 and to thresholds for cell removal. We find that the length of newly formed edges may be up to
532 twice as long as the threshold for T1 transitions without affecting final vertex configurations.
533 However, this may change in other parameter regimes, for example if larger values for the cell
534 rearrangement threshold are chosen.

535 The size of the area threshold for cell removal may be varied over six orders of magnitude
536 without impacting tissue-level summary statistics, even though the exact number of T2 transi-
537 tions may differ for any two values of the area threshold. In particular, it seems to be possible
538 to choose arbitrarily small values for the T2 transition threshold without causing the algorithm
539 to fail. There are three effects that may contribute to the stability of small elements in our
540 simulations. First, since small cells with areas close to the threshold for cell removal are far away
541 from their preferred area in our simulations ($A_{0,\alpha} > 1.0$), their area force is larger than that
542 of adjacent neighbours. This makes the cells stiff and prevents them from becoming inverted
543 or otherwise misshapen. Second, the relationship between area and cell neighbour numbers
544 presented in figure 2 shows that small elements are most likely to be triangular. Our simulation
545 algorithm does not permit T1 transitions if the short edge is part of a triangular cell in order

546 to prevent triangular elements from becoming inverted and thus the algorithm from failure.
547 Third, this relationship between cell area and cell neighbour number may also contribute to
548 the stability of the algorithm when the area threshold is large, for example 0.2. In this case,
549 individual cells may be smaller than the area threshold without undergoing T2 transitions if
550 they are not triangular.

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

556 Here, we analysed how numerical and non-physical parameters can influence experimentally
557 measurable summary statistics in cell-based models by examining a force-propagation-based
558 implementation of vertex models. Individual results may be relevant to other implementation
559 choices. For example, our finding that the duration of the cell cycle in our model influences
560 simulation outcomes may mean that parameters that control the rate of energy-minimisation
561 may influence results in other vertex model implementations [3,25,62]. In general, further work is
562 required to understand how other choices of implementation schemes may impact computational
563 model predictions. For example, the noise strength in a Monte Carlo vertex propagation scheme
564 [39,40] or the choice of energy-minimisation algorithm may influence vertex model behaviour.

565 While most of our findings are of a numerical nature, some have explicit biological relevance.
566 Our analysis of the dependence of tissue properties and forces on the mean cell cycle duration
567 reveals that the vertex model predicts increased forces in tissues undergoing growth and prolif-
568 eration, and that fast tissue growth may induce cell extrusion. Our findings further suggest that
569 statistics of cell packing may depend on the nature of the cell cycle or the boundary condition
570 of the tissue. Note that findings that do not make explicit biological predictions, such as the
571 robustness of the vertex model to changes in the area threshold for cell removal, or its sensitivity
572 to changes in the length threshold for cell rearrangement, are nonetheless highly relevant, since
573 these findings highlight that choices of model design and implementation have to be carefully
574 considered when applying vertex models quantitatively.

575 Throughout the manuscript we use non-dimensional parameters that arise when rescaling

576 time and space by the characteristic length and time scales of the model. The use of such
577 rescaled parameters is beneficial in this case since it allows, for example, the comparison of our
578 model parameters to previously used values [3, 4, 28]. Further, we identify reference parameter
579 values for which our simulations are physically reasonable. By providing non-dimensional values
580 for these parameters we facilitate their reuse in other applications where the physical values of
581 the characteristic length or time scales may be different.

582 **5 Conclusions**

583 Our results illustrate that care needs to be taken when drawing predictions using cell-based
584 computational models because implementation details such as the size of the time step or non-
585 physical parameters, such as length thresholds for cell rearrangement, may influence model
586 predictions significantly. With the rise of quantitative analysis and quantitative model-data
587 comparison in biophysical applications, choices of model implementation become increasingly
588 relevant. To enable the use of cell-based models in quantitative settings, it is important to
589 be aware of any influences that implementation choices may have on model predictions when
590 analysing a specific biophysical phenomenon. Understanding model behaviour in detail is cru-
591 cial to prevent modelling artefacts from influencing experimental predictions and clouding our
592 biophysical understanding and, as such, our findings emphasise the need to fully document al-
593 gorithms for simulating cell-based models. Close attention to implementation details is required
594 in order to unravel the full predictive power of cell-based models.

595 **Acknowledgements**

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

599 **References**

600 [1] Kitano, H. Systems biology: A brief overview. *Science*, 295(5560):1662–1664, 2002. doi:
601 10.1126/science.1069492.

- 602 [2] Fisher, J. and Henzinger, T. A. Executable cell biology. *Nature Biotechnol.*, 25(11):1239–
603 49, 2007. doi:10.1038/nbt1356.
- 604 [3] Farhadifar, R., Röper, J.-C., Aigouy, B., Eaton, S., and Jülicher, F. The influence of
605 cell mechanics, cell-cell interactions, and proliferation on epithelial packing. *Curr. Biol.*,
606 17(24):2095–2104, 2007. doi:10.1016/j.cub.2007.11.049.
- 607 [4] Mao, Y., Tournier, A. L., Bates, P. A., Gale, J. E., Tapon, N., and Thompson, B. J. Planar
608 polarization of the atypical myosin Dachs orients cell divisions in *Drosophila*. *Genes Dev.*,
609 25(2):131–136, 2011. doi:10.1101/gad.610511.
- 610 [5] Trichas, G., Smith, A. M., White, N., Wilkins, V., Watanabe, T., Moore, A., Joyce,
611 B., Sugnaseelan, J., Rodriguez, T. A., Kay, D., et al. Multi-cellular rosettes in the mouse
612 visceral endoderm facilitate the ordered migration of anterior visceral endoderm cells. *PLoS*
613 *Biol.*, 10(2):e1001256, 2012. doi:10.1371/journal.pbio.1001256.
- 614 [6] Monier, B., Gettings, M., Gay, G., Mangeat, T., Schott, S., Guarner, A., and Suzanne,
615 M. Apico-basal forces exerted by apoptotic cells drive epithelium folding. *Nature*,
616 518(7538):245–248, 2015. doi:10.1038/nature14152.
- 617 [7] Atwell, K., Qin, Z., Gavaghan, D., Kugler, H., Hubbard, E. J., and Osborne, J. M.
618 Mechano-logical model of *C. elegans* germ line suggests feedback on the cell cycle. *De-*
619 *velopment*, 142(22):3902–3911, 2015. doi:10.1242/dev.126359.
- 620 [8] Walker, D. C., Hill, G., Wood, S. M., Smallwood, R. H., and Southgate, J. Agent-based
621 computational modeling of wounded epithelial cell monolayers. *IEEE Trans. Nanobio-*
622 *science*, 3(3):153–163, 2004.
- 623 [9] Anderson, A. R. A., Weaver, A. M., Cummings, P. T., and Quaranta, V. Tumor morphology
624 and phenotypic evolution driven by selective pressure from the microenvironment. *Cell*,
625 127(5):905–915, 2006. doi:10.1016/j.cell.2006.09.042.
- 626 [10] Weinan, E. and Engquist, B. Multiscale modeling and computation. *Notices Amer. Math.*
627 *Soc.*, 50(9):1062–1070, 2003.

- 628 [11] Miller, R. E. and Tadmor, E. B. A unified framework and performance benchmark of
629 fourteen multiscale atomistic/continuum coupling methods. *Model. Simul. Mater. Sci.*
630 *Eng.*, 17(5):053001, 2009. doi:10.1088/0965-0393/17/5/053001.
- 631 [12] Pantazis, P. and Supatto, W. Advances in whole-embryo imaging: a quantitative transition
632 is underway. *Nat. Rev. Mol. Cell. Biol.*, 15(5):327–339, 2014. doi:10.1038/nrm3786.
- 633 [13] Yu, J. C. and Fernandez-Gonzalez, R. Quantitative modelling of epithelial morphogenesis:
634 integrating cell mechanics and molecular dynamics. *Sem. Cell Dev. Biol.*, 2016. doi:10.
635 1016/j.semcdb.2016.07.030.
- 636 [14] Pargett, M. and Umulis, D. M. Quantitative model analysis with diverse biological data:
637 Applications in developmental pattern formation. *Methods*, 62(1):56–67, 2013. doi:10.
638 1016/j.ymeth.2013.03.024.
- 639 [15] Fletcher, A. G., Osterfield, M., Baker, R. E., and Shvartsman, S. Y. Vertex models of
640 epithelial morphogenesis. *Biophys. J.*, 106(11):2291–2304, 2014. doi:10.1016/j.bpj.2013.11.
641 4498.
- 642 [16] Okuzono, T. and Kawasaki, K. Intermittent flow behavior of random foams: A computer
643 experiment on foam rheology. *Phys. Rev. E*, 51(2):1246–1253, 1995. doi:10.1103/PhysRevE.
644 51.1246.
- 645 [17] Kawasaki, K., Nagai, T., and Nakashima, K. Vertex models for two-dimensional grain
646 growth. *Philos. Mag. B*, 60(3):399–421, 1989. doi:10.1080/13642818908205916.
- 647 [18] Torres, C. E., Emelianenko, M., Golovaty, D., Kinderlehrer, D., and Ta’asan, S. Numerical
648 analysis of the vertex models for simulating grain boundary networks. *SIAM J. Appl.*
649 *Math.*, 75(2):762–786, 2015. doi:10.1137/140999232.
- 650 [19] Honda, H. Description of cellular patterns by Dirichlet domains: The two-dimensional
651 case. *J. Theor. Biol.*, 72(3):523–543, 1978. doi:10.1016/0022-5193(78)90315-6.
- 652 [20] Nagai, T., Kawasaki, K., and Nakamura, K. Vertex dynamics of two-dimensional cellular
653 patterns. *J. Phys. Soc. Jpn.*, 57(7):2221–2224, 1988. doi:10.1143/JPSJ.57.2221.

- 654 [21] Weliky, M. and Oster, G. The mechanical basis of cell rearrangement. I. Epithelial mor-
655 phogenesis during *Fundulus* epiboly. *Development*, 109(2):373–386, 1990.
- 656 [22] Nagai, T. and Honda, H. A dynamic cell model for the formation of epithelial tissues.
657 *Philos. Mag. B*, 81(7):699–719, 2001. doi:10.1080/13642810108205772.
- 658 [23] Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. *Molecular Biology*
659 *of the Cell*, 5th Ed. Garland Science, 2008.
- 660 [24] Mao, Y., Tournier, A. L., Hoppe, A., Kester, L., Thompson, B. J., and Tapon, N. Differ-
661 ential proliferation rates generate patterns of mechanical tension that orient tissue growth.
662 *EMBO J.*, 32(21):2790–2803, 2013. doi:10.1038/emboj.2013.197.
- 663 [25] Aegerter-Wilmsen, T., Smith, A., Christen, A., Aegerter, C., Hafen, E., and Basler, K. Ex-
664 ploring the effects of mechanical feedback on epithelial topology. *Development*, 137(3):499–
665 506, 2010. doi:10.1242/dev.041731.
- 666 [26] Patel, A. B., Gibson, W. T., Gibson, M. C., and Nagpal, R. Modeling and inferring
667 cleavage patterns in proliferating epithelia. *PLoS Comput. Biol.*, 5(6):e1000412, 2009.
668 doi:10.1371/journal.pcbi.1000412.
- 669 [27] Landsberg, K. P., Farhadifar, R., Ranft, J., Umetsu, D., Widmann, T. J., Bittig, T., Said,
670 A., Jülicher, F., and Dahmann, C. Increased cell bond tension governs cell sorting at the
671 *Drosophila* anteroposterior compartment boundary. *Curr. Biol.*, 19(22):1950–1955, 2009.
672 doi:10.1016/j.cub.2009.10.021.
- 673 [28] Canela-Xandri, O., Sagués, F., Casademunt, J., and Buceta, J. Dynamics and mechanical
674 stability of the developing dorsoventral organizer of the wing imaginal disc. *PLoS Comput.*
675 *Biol.*, 7(9):e1002153, 2011. doi:10.1371/journal.pcbi.1002153.
- 676 [29] Salbreux, G., Barthel, L. K., Raymond, P. A., and Lubensky, D. K. Coupling mechanical
677 deformations and planar cell polarity to create regular patterns in the zebrafish retina.
678 *PLoS Comput. Biol.*, 8(8):e1002618, 2012. doi:10.1371/journal.pcbi.1002618.
- 679 [30] Sugimura, K. and Ishihara, S. The mechanical anisotropy in a tissue promotes ordering in
680 hexagonal cell packing. *Development*, 140(19):4091–4101, 2013. doi:10.1242/dev.094060.

- 681 [31] Kursawe, J., Brodskiy, P. A., Zartman, J. J., Baker, R. E., and Fletcher, A. G. Capabilities
682 and limitations of tissue size control through passive mechanical forces. *PLoS Comput.*
683 *Biol.*, 11(12):e1004679, 2015. doi:10.1371/journal.pcbi.1004679.
- 684 [32] Wartlick, O., Mumcu, P., Jülicher, F., and González-Gaitán, M. Understanding morpho-
685 genetic growth control - lessons from flies. *Nat. Rev. Mol. Cell Biol.*, 12(9):594–604, 2011.
686 doi:10.1038/nrm3169.
- 687 [33] Bi, D., Lopez, J. H., Schwarz, J., and Manning, M. L. Energy barriers and cell migration
688 in densely packed tissues. *Soft Matter*, 10:1885–1890, 2014. doi:10.1039/C3SM52893F.
- 689 [34] Marinari, E., Mehonic, A., Curran, S., Gale, J., Duke, T., and Baum, B. Live-cell delamina-
690 tion counterbalances epithelial growth to limit tissue overcrowding. *Nature*, 484(7395):542–
691 545, 2012. doi:10.1038/nature10984.
- 692 [35] Nagai, T. and Honda, H. Computer simulation of wound closure in epithelial tissues: Cell-
693 basal-lamina adhesion. *Phys. Rev. E*, 80:061903, 2009. doi:10.1103/PhysRevE.80.061903.
- 694 [36] Rauzi, M., Verant, P., Lecuit, T., and Lenne, P.-F. Nature and anisotropy of cortical
695 forces orienting *Drosophila* tissue morphogenesis. *Nat. Cell Biol.*, 10(12):1401–1410, 2008.
696 doi:10.1038/ncb1798.
- 697 [37] Collinet, C., Rauzi, M., Lenne, P.-F., and Lecuit, T. Local and tissue-scale forces drive
698 oriented junction growth during tissue extension. *Nat. Cell Biol.*, 17(10):1247–1258, 2015.
699 doi:10.1038/ncb3226.
- 700 [38] Tetley, R. J., Blanchard, G. B., Fletcher, A. G., Adams, R. J., and Sanson, B. Unipolar dis-
701 tributions of junctional myosin ii identify cell stripe boundaries that drive cell intercalation
702 throughout *Drosophila* axis extension. *eLife*, 5:1–35, 2016. doi:10.7554/eLife.12094.
- 703 [39] Etournay, R., Popovi, M., Merkel, M., Nandi, A., Blasse, C., Aigouy, B., Brandl, H., Myers,
704 G., Salbreux, G., Jlicher, F., et al. Interplay of cell dynamics and epithelial tension during
705 morphogenesis of the *Drosophila* pupal wing. *eLife*, 4:e07090, 2015. doi:10.7554/eLife.
706 07090.
- 707 [40] Ray, R. P., Matamoro-Vidal, A., Ribeiro, P. S., Tapon, N., Houle, D., Salazar-Ciudad,
708 I., and Thompson, B. J. Patterned anchorage to the apical extracellular matrix defines

- 709 tissue shape in the developing appendages of *Drosophila*. *Dev. Cell*, 34(3):310–322, 2015.
710 doi:10.1016/j.devcel.2015.06.019.
- 711 [41] Staple, D., Farhadifar, R., Röper, J.-C., Aigouy, B., Eaton, S., and Jülicher, F. Mechanics
712 and remodelling of cell packings in epithelia. *Eur. Phys. J. E*, 33:117–127, 2010. doi:
713 10.1140/epje/i2010-10677-0.
- 714 [42] Merzouki, A., Malaspinas, O., and Chopard, B. The mechanical properties of a cell-
715 based numerical model of epithelium. *Soft Matter*, 12(21):4745–4754, 2016. doi:10.1039/
716 C6SM00106H.
- 717 [43] Xu, G.-K., Liu, Y., and Li, B. How do changes at the cell level affect the mechanical
718 properties of epithelial monolayers? *Soft Matter*, 11(45):8782–8788, 2015. doi:10.1039/
719 C5SM01966D.
- 720 [44] Davit, Y., Osborne, J., Byrne, H., Gavaghan, D., and Pitt-Francis, J. Validity of the
721 cauchy-born rule applied to discrete cellular-scale models of biological tissues. *Phys. Rev.*
722 *E*, 87(4):042724, 2013. doi:10.1103/PhysRevE.87.042724.
- 723 [45] Smith, A. M., Baker, R. E., Kay, D., and Maini, P. K. Incorporating chemical signalling
724 factors into cell-based models of growing epithelial tissues. *J. Math. Biol.*, 65:441–463,
725 2012. doi:10.1007/s00285-011-0464-y.
- 726 [46] Osborne, J. M., Fletcher, A. G., Pitt-Francis, J. M., Maini, P. K., and Gavaghan, D. J.
727 Comparing individual-based approaches to modelling the self-organization of multicellular
728 tissues. *bioRxiv*, 2016. doi:10.1101/074351.
- 729 [47] Magno, R., Grieneisen, V. A., and Marée, A. F. M. The biophysical nature of cells:
730 potential cell behaviours revealed by analytical and computational studies of cell surface
731 mechanics. *BMC Biophys.*, 8(1):8, 2015. doi:10.1186/s13628-015-0022-x.
- 732 [48] Fletcher, A. G., Osborne, J. M., Maini, P. K., and Gavaghan, D. J. Implementing vertex dy-
733 namics models of cell populations in biology within a consistent computational framework.
734 *Prog. Biophys. Mol. Biol.*, 113(2):299–326, 2013. doi:10.1016/j.pbiomolbio.2013.09.003.
- 735 [49] Mirams, G. R., Arthurs, C. J., Bernabeu, M. O., Bordas, R., Cooper, J., Corrias, A.,
736 Davit, Y., Dunn, S.-J., Fletcher, A. G., Harvey, D. G., et al. Chaste: An open source C++

- 737 library for computational physiology and biology. PLoS Comput. Biol., 9(3):e1002970,
738 2013. doi:10.1371/journal.pcbi.1002970.
- 739 [50] Smith, J. A. and Martin, L. Do cells cycle? Proc. Natl. Acad. Sci. USA, 70(4):1263–1267,
740 1973. doi:10.1073/pnas.70.4.1263.
- 741 [51] Shields, R. Further evidence for a random transition in the cell cycle. Nature, 273:755–758,
742 1978. doi:10.1038/273755a0.
- 743 [52] Hofmeister, W. Zusätze und Berichtigungen zu den 1851 veröffentlichten Untersuchungen
744 der Entwicklung höherer Kryptogamen. Jahrb. wiss. Bot., 3:259–293, 1863.
- 745 [53] Gibson, W. T., Veldhuis, J. H., Rubinstein, B., Cartwright, H. N., Perrimon, N., Brodland,
746 G. W., Nagpal, R., and Gibson, M. C. Control of the mitotic cleavage plane by local
747 epithelial topology. Cell, 144(3):427–438, 2011. doi:10.1016/j.cell.2010.12.035.
- 748 [54] Lewis, F. T. The correlation between cell division and the shapes and sizes of prismatic cells
749 in the epidermis of cucumis. Anat. Rec., 38(3):341–376, 1928. doi:10.1002/ar.1090380305.
- 750 [55] Kim, S., Cai, M., and Hilgenfeldt, S. Lewis’ law revisited: the role of anisotropy in
751 sizetopology correlations. New J. Phys., 16(1):015024, 2014. doi:10.1088/1367-2630/16/1/
752 015024.
- 753 [56] Smith, A. M. Vertex model approaches to epithelial tissues in developmental systems.
754 Ph.D. thesis, University of Oxford, 2011.
- 755 [57] Milán, M., Campuzano, S., and García-Bellido, A. Cell cycling and patterned cell prolif-
756 eration in the wing primordium of *Drosophila*. Proc. Natl. Acad. Sci. USA, 93:640–645,
757 1996. doi:10.1073/pnas.93.2.640.
- 758 [58] Ishihara, S. and Sugimura, K. Bayesian inference of force dynamics during morphogenesis.
759 J. Theor. Biol., 313(0):201–211, 2012. doi:10.1016/j.jtbi.2012.08.017.
- 760 [59] Ishihara, S., Sugimura, K., Cox, S., Bonnet, I., Bellaïche, Y., and Graner, F. Comparative
761 study of non-invasive force and stress inference methods in tissue. Eur. Phys. J. E, 36(4):1–
762 13, 2013. doi:10.1140/epje/i2013-13045-8.

- 763 [60] Chiou, K. K., Hufnagel, L., and Shraiman, B. I. Mechanical stress inference for two-
764 dimensional cell arrays. *PLoS Comput. Biol.*, 8(5):e1002512, 2012. doi:10.1371/journal.
765 pcbi.1002512.
- 766 [61] Ambrosini, A., Gracia, M., Proag, A., Rayer, M., Monier, B., and Suzanne, M. Apoptotic
767 forces in tissue morphogenesis. *Mech. Dev.*, 2016. doi:10.1016/j.mod.2016.10.001.
- 768 [62] Aegerter-Wilmsen, T., Heimlicher, M., Smith, A., de Reuille, P., Smith, R., Aegerter,
769 C., and Basler, K. Integrating force-sensing and signaling pathways in a model for the
770 regulation of wing imaginal disc size. *Development*, 139(17):3221–3231, 2012. doi:10.1242/
771 dev.082800.