Large Scale Parallel Simulations of 3-D Cell Colony Dynamics

Maciej Cytowski, Zuzanna Szymańska[*]

Abstract

Biological processes are inherently very complex and involve many unknown relationships and mechanisms at different scales. Despite many efforts, one still cannot explain all the observed phenomena and, if necessary, make any desirable changes in the dynamics. Recently, it has become apparent that the opportunity lies in complementing the traditional, heuristic experimental approach with mathematical modelling and computer simulations. Achieving a simulation scale that corresponds for instance to clinically detectable tumour sizes is still a huge challenge, however it is necessary to understand and control complex biological processes. In this paper we present a novel high performance computational approach allowing simulations of 3D cell colony dynamics in previously unavailable tissue scale. Due to the high parallel scalability we are able to simulate cell colonies composed of 10^9 cells, which allows for instance to describe tumour growth in its early clinical stage.

Keywords: mathematical modelling, cellular processes, parallel computing.

1   Introduction

Experimental methods alone are very often not sufficient to build a consistent, systematic theory which is capable of describing biological phenomena. Exploring all the complexities of biological systems may require many experiments to be performed. Moreover, a full understanding of the governing mechanisms with the use of experimental method alone is nearly impossible due to their highly nonlinear nature. One of the tools extensively used to complement the traditional, experimental approach is mathematical modelling, which aims at the mathematical representation of biological processes using a variety of analytical and computational techniques. Such models have proved to be very successful in many applications e.g. systems biology, bioinformatics, molecular dynamics, neurobiology, biological tissue and cell modelling. In particular, numerous mathematical models have been proposed to study biological processes related to cell

[*]M. Cytowski and Z. Szymańska are with the Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, Warsaw, Poland e-mail: m.cytowski@icm.edu.pl, z.szymanska@icm.edu.pl

colonies growth from discrete individual based methods simulating intercellular dynamics through continuum models describing global population mechanisms, to hybrid models combining both individual based and continuum approaches. Discrete or hybrid models have been widely used in mathematical oncology, many aspects of tumour growth such as early avascular tumour growth, tumour invasion or tumour angiogenesis have been addressed (see an excellent review of Lowengrub et al. [1]). Other biological processes that have been modelled with the use of individual-based models are chemotactic cell movement, blood clotting, tissue formation and morphogenesis [2].

Biological cells are discrete entities and their precise spatial location within a cell colony has an impact on whether they grow and divide. They are influenced by signals from other cells, external factors such as nutrient concentration, stress and of course their own internal signalling pathways. Individual based models proved to be very successful in modelling many biological phenomena related to cellular processes, since they analyse spatial and temporal dynamics at the level of the cell, linking individual behaviour with the macroscopic one. Individual based approach provides scope to include different cell types, sub-cellular structures, intracellular dynamics and the spatial location of individual cells. Moreover, just like actual biological cells, each computational cell can have a stochastic nature.

At the cellular level, the key processes that are modelled are division, differentiation, apoptosis and interactions between cells. These processes are regulated by cascades of biochemical reactions called signalling pathways. Regulatory proteins, whose production is triggered by signalling pathways, initiate or modify the processes of cell division and death. For instance, the loss of control over cell division is linked to mutations of genes encoding proteins involved in the regulation of the cell cycle. As a result of these mutations, intracellular signalling pathways act in an altered way leading to further pathological changes in the organism (these pathways may become a therapeutic goal in the future). In turn, cell differentiation, both normal and pathological, influences the signalling pathway dynamics, which leads to subsequent changes at the cellular level. Taking into account the connections of the models at the cellular and sub-cellular scale is necessary in order to describe the function of cells.

Individual-based models can be on- or off-lattice. Off-lattice models however, being more biologically realistic (since the cells are not constrained by a predefined underlying grid), are much more computationally expensive.

Although mathematical modelling in biology led to the development of many computer simulation tools, none of those tools is capable of simulating cellular systems on the clinically detectable scale. Modelling complex and multiscale biological processes that involve many unknown relations and mechanisms on such scales requires the development of high performance computational (HPC) tools. It is especially important for biological systems in which the properties of individual cells considered at extremely high resolutions can influence the dynamics and geometry of a large population of cells. The increasing need for accuracy has led to the development of highly complex and computationally demanding models and we believe that computational methodology described

in this paper can be used for their further development.

Most of the published models focus on theoretical modelling issues so to speak, neglecting the computational implementation corresponding to the real scale of the problem. Here, we do the opposite, i.e. we place the emphasis on the efficient implementation of the generic off-lattice individual-based model of cell colony dynamics rather then modelling particular biological processes. The ongoing research will be carried out on an HPC systems. The computer code was developed for optimal performance on the IBM Blue Gene/Q supercomputer [3] and can be used to simulate biological processes on previously unachievable scales. Our goal is to enable accurate and complex simulations of biological processes in cell colonies consisting of 10^9 and more cells. This large-scale computational approach will allow for simulations to be carried out over spatial scales up to 1cm in size i.e. the tissue scale. The proposed computational model is a general framework for large scale simulations of various biological processes occurring in cellular colonies, that can be further adapted to simulate particular processes such as biofilm development, vasculogenesis or solid tumour growth.

Here, we first describe the mathematical model of cell colony dynamics. We then discuss computational methods and present details of the corresponding parallel implementation. Finally, we share results from our recent applications to tumour growth invasion.

2   Description of the model

In this section, we describe the basic assumptions of our model. Thanks to the continuous development and enhancement of HPC we were able to address the complexity and dynamics of very large cell colonies with the use of off-lattice individual-based models.

2.1   Cell description

We consider cells as free objects that reside in three-dimensional space. The number of cells in the model is referred to as N. Each cell c_i position is described by its Cartesian coordinates $(x_{c_i}, y_{c_i}, z_{c_i})$. Each cell size and shape is described by a sphere centered at $(x_{c_i}, y_{c_i}, z_{c_i})$ with radius r_{c_i}. Both the position and size of a cell can change in time. Cells interact with each other. For every cell c_i its maximum intercellular interaction distance is defined by a neighbourhood sphere $B_{\epsilon_{ci}}(c_i)$ centered on $(x_{c_i}, y_{c_i}, z_{c_i})$ with a radius of ϵ_{ci}.

2.2   Cell cycle and cell mitosis

The growth of a population of cells is based on the growth and division of individual cells. In order to proliferate, a cell must undergo an ordered series of reactions that allows for the duplication of its contents and finally the splitting into two new daughter cells. These processes of duplication and division are called, collectively, the cell division cycle (or simply the cell cycle) [4]. The eukaryotic cell cycle is divided into four phases: M-phase, when the actual

division of cell occurs, G_1-phase, which is a phase of cell growth that occurs after the end of the division, and prior to DNA duplication, S-phase, when nuclear DNA is duplicated, and G_2-phase, when a cell completes preparation for cell division (see Fig. 1).
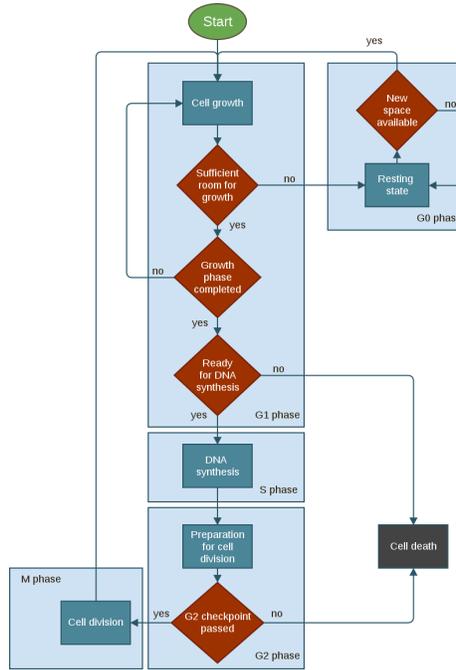

Fig. 1: Sequence of events leading to cell division. Decision points: *"Sufficient room for growth"*, *"Growth phase completed"* and *"Ready for DNA synthesis"* corresponds to the G_1/S checkpoint, decision point *"G2 checkpoint passed"* corresponds to the G_2/M checkpoint. Decision Point *"New space available"* corresponds to the cell decision whether it can re-enter the G_1 phase.

The G_1-phase allows additional time for a cell to grow and to monitor its environment. If conditions are particularly unfavourable, instead of entering S-phase a cell can enter a resting state - G_0-phase, where it remains until conditions improve and it continues the cell cycle. The first cell cycle checkpoint, so called G_1/S checkpoint, is located at the end of the G_1-phase, when a cell takes a decisive step to enter S-phase and to initiate DNA duplication, or to undergo apoptosis (i.e. to die).

During G_2-phase the synthesis of proteins needed to carry out cell division occurs. This also provides a safety gap, allowing the cell to ensure that DNA duplication is completed before mitosis [5]. The second checkpoint, so called G_2/M checkpoint, is located in the G_2-phase. Passing through this checkpoint, a cell decides whether it is ready to enter M-phase and if not it undergoes
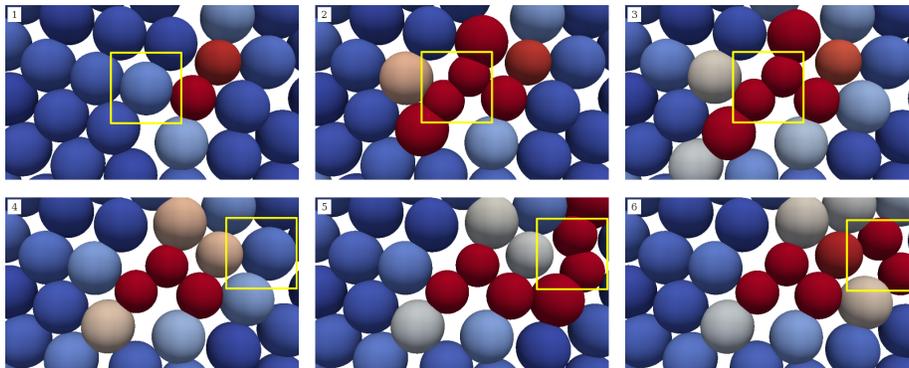
Fig. 2: Figures showing in detail the process of mitosis in two given cells (indicated by a yellow frame) each dividing into two daughter cells in a cellular colony.

apoptosis. During M-phase, at first, the cell's nucleus divides into two (a process known as mitosis) and later the whole cell splits into two (a process known as cytokinesis).

   We assume that during cytokinesis the parent cell c_i of radius r_{c_i} splits into two daughter cells of radius $\frac{1}{\sqrt[3]{2}} r_{c_i}$ (symmetric division), chosen to preserve mass conservation. Daughter cells are located at a distance of $\frac{r_{c_i}}{2}$ from the center of the parent cell in a randomly chosen direction (see Fig. 2). At any time, each cell of the model can be in any one of the different states: G_1-phase, S-phase, G_2-phase, M-phase or G_0-phase. The exact length of each cell cycle phase for a respective cell is taken randomly from the interval $[T \cdot (1-V), T \cdot (1+V)]$, where T is the average length of the phase, and V is the variance. Each cell during the simulation has a certain probability P_{c_i} of being marked for programmed cell death (apoptosis) at the nearest checkpoint. The exact value of P_{c_i} depends on the particular cell type, e.g. in the example simulation presented in this paper we assume that tumour cells have down-regulated their apoptotic pathways (i.e. they are immortal).

2.3   Cell-cell interactions

The interactions between cells in the system are described by a modified Hertz model as proposed in [6]. A decreasing distance between the centers of cells results in an attractive interaction related to adhesive forces. Furthermore, experiments suggest that cells have only limited compressibility which give rise to repulsive interaction. The potential $V_{c_i c_j}$ between two adjacent cells c_i and c_j of radius r_{c_i} and r_{c_j} in this model is a combination of repulsive and attractive forces and is given by:

$$V_{c_i c_j} = (r_{c_i} + r_{c_j} - d_{c_i c_j})^{\frac{5}{2}} \frac{1}{5 E_{c_i c_j}} \sqrt{\frac{r_{c_i} r_{c_j}}{r_{c_i} + r_{c_j}}} + A_{c_i c_j}.$$

The first term of the above formula relates to repulsive interactions modelled with the Hertz formula. Here $d_{c_i c_j}$ is the distance between two cells and $E_{c_i c_j}$ is computed with the use of Young moduli E_{c_i} and E_{c_j} and the Poisson ratios ν_{c_i} and ν_{c_j}. The parameter values of Young moduli for each cell were randomly chosen from range of 2100 ± 100 Pa in accordance with [7]. The Poisson ratio is assumed to be equal to 0.33 as suggested in [8].

The second term $A_{c_i c_j}$ of the potential formula relates to the adhesive forces and is given by:

$$A_{c_i c_j} = \rho_m D_{c_i c_j} 25 k_B T,$$

where k_B is the Boltzmann constant, T is the temperature, $D_{c_i c_j}$ is the contact area between cells c_i, c_j and ρ_m is the density of surface adhesion molecules in the contact zone [9], which we assume is given for a specific cell type.

The potential function value for cell c_i is given by the sum:

$$V_{c_i} = \sum_{c_j \in B_{\epsilon_{ci}}(c_i)} V_{c_i c_j}.$$

2.4 Cell movements

In this paper we discuss a simplified model where the motion D_{c_i} of cell c_i is based only on the potential function. The aforementioned formulation of the potential function allows us to model two very important types of cell-cell interactions. First of all adhesive forces which allow cells to be bound to other cells or extracellular components of tissue (so called extracellular matrix - ECM). Secondly repulsive forces which appear when cell cytoskeletons and membranes are significantly stressed due to attractive forces and the decreasing distance between cells' centers.

Cells can change their position after each simulation step. The displacement velocity vector is computed by deriving the potential function, i.e.:

$$D_{c_i} = -\nabla V_{c_i}.$$

The other mechanisms which might affect the dynamics of cell colonies (e.g. chemotaxis, haptotaxis) are not currently treated within the present model. In the general case cells will additionally tend to move with respect to gradient of substances in their environment, see [10] for 2-D example.

3 Computational methods

This section presents details of the computational methods that we have used to translate mathematical model into computer language. Large scale simulations at tissue scale are both highly memory and computational demanding. Therefore proposed computational framework was designed to work on novel supercomputing architectures and parallelized from scratch. Computational algorithms have been carefully selected to ensure high accuracy on the one hand and efficient parallelization and high scalability on the other. Performance of

the resulting computer model was tested in various setups, e.g. for different domain decomposition scenarios. Consecutive steps of implementation, performance analysis and optimization led us to the development of highly scalable simulation environment on the IBM Blue Gene/Q supercomputer.

A schematic simulation diagram is presented in Scheme 1. Simulations are made up of successive iterations, which computational nature will be described next.

Scheme 1 Computational scheme of the simulation

$iter \leftarrow 0$
while $iter \leq max\_iter$ **do**
   Step 1: Perform domain decomposition
   Step 2: Build tree
   Step 3a: Find exchange regions and initiate data exchange
   for all local cells **do**
     Step 4a: Find cell's neighbours $\leftarrow$ local data
     Step 5a: Compute potential and density functions $\leftarrow$ local data
   end for
   Step 3b: Wait until data exchange is finished
   for all local cells **do**
     Step 4b: Find cell's neighbours $\leftarrow$ remote data
     Step 5b: Compute potential and density functions $\leftarrow$ remote data
   end for
   for all local cells **do**
     Step 6: Update cells' cycle
     Step 7: Compute forces and move cells to their new positions
   end for
end while

3.1   Domain decomposition

Each iteration of the simulation begins with domain decomposition step (Step 1) which distributes cells accross available parallel processes. We have selected the most appropriate decomposition method based on the analysis of resulting load balancing. We decided to use the dynamical decomposition algorithm based on the concept of Peano-Hilbert space filling curves (PHSFC), which is particularly useful in the case of systems of free moving objects.

One of the most important features of PHSFC method is that decomposition domains assigned to parallel processes are topologically connected. The amount of communication between adjacent processes is thereby minimized and most of the calculations can be performed independently. Data exchange regions for parallel processes, which can also be understood as borders of domain fragments, are illustrated in Fig. 3.

3.2   Nearest neighbours searching

In consecutive iterations each cell sample the environment by identifying it's neighbours. This process can be very time consuming, especially when a naive algorithmic approach is used. For instance the computational complexity of the
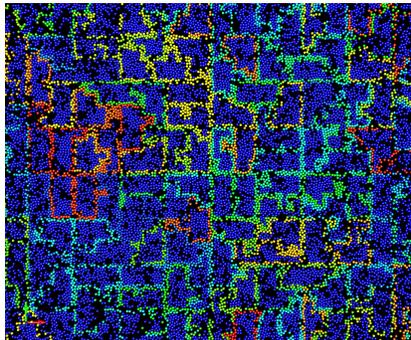
Fig. 3: Halo exchange areas shown with colours in a 2-D simulation of a cellular colony on 256 parallel MPI processes. Local cells of each process are denoted by a navy blue colour. The colours of halo exchange cells have been randomly chosen for each process.

Particle-Particle method is known to be of the order of $O(N^2)$ where N indicates the total number of cells. Much more efficient methods are those based on tree algorithms, with complexity of the order of $O(NlogN)$. These methods have been widely used in research areas such as computational cosmology [11] and molecular dynamics [12]. We have used similar approach in this work. The tree is built on each parallel process by an iterative procedure (Step 2). It starts from a cubical root node containing all cells assigned to the process. This root node is repeatedly subdivided into eight daughter nodes of half the side length each, until one ends up with leaf nodes containing single cells. We consider only short-distance interactions since each cell have finite neighbourhood sphere $B_{\epsilon_{ci}}(c_i)$. Neighbours of each cell are found by traversing the tree starting from the leaf upwards. In this way we are able to compute the most important characteristics of the simulation: potential and density functions.

3.3 Communication and data exchange

One of the consequences of parallel realization is the need of communication and data exchange, since neighbours of boundary cells might be stored on different processes. As can be seen from Scheme 1, potential and density function computations, which are based on tree traversing algorithm, are divided into two stages. Data exchange between processes is initiated (Step 3a) before the first stage which is calculated only with the use of locally available data (Steps 4a and 5a). The second stage (Steps 4b and 5b) is computed when the remote data from other processes is received which is ensured in Step 3b. With this approach, we are able to overlap computations and communication, which have an important impact on final scalability of the simulator.

3.4   Potential and density functions

Having the information about the neighbours of each cell, the value of the potential function can be easily determined. At this stage of the iteration we are also able to calculate the approximate value of the density function for each cell. This is achieved by adopting the concept of Smooth Particle Hydrodynamics (SPH) [13] and it's smoothing lengths. The SPH algorithm was originally developed for fluid dynamics problems.

We assume that the value of the density function $\tilde{\rho}$ for given position $\vec{p}$ in three-dimensional space can be computed as a sum over all cells c_i in the system with the use of following formula:

$$\tilde{\rho}(\vec{p}) \simeq \sum_{i=1,N} m_{c_i} W(\vec{p} - \vec{p_{c_i}}, h),$$

where N denotes the number of cells in the system, m_{c_i} is the mass of cell c_i, $\vec{p_{c_i}}$ is its location and h is the parameter that defines the radius of the region of influence for the approximation of the density function. Kernel function used in our model is of the following form:

$$W(z, h) = \frac{8}{\pi h^3} \begin{cases} 1 - 6(\frac{z}{h})^2 + 6(\frac{z}{h})^3, & 0 \leq \frac{z}{h} \leq 0.5 \\ 2(1 - \frac{z}{h})^3, & 0.5 \leq \frac{z}{h} \leq 1 \\ 0, & \frac{z}{h} > 1 \end{cases}.$$

Values of the kernel for cells further away than a distance of h from $\vec{p}$ are exactly zero. In our approach we assume that h is equal to the radius of cells' neighbourhood r. With this assumption we are able to calculate the value of the density function at the location of each cell with the use of previously determined neighbourhood. An example mapping of the density function on spheres representing cells is shown on Fig. 4.

3.5   Cell cycle update and motion of cells

The development and dynamics of cell colony in our model is driven by two previously described characteristics: potential and density functions. Each cell undergo cell cycle (Step 6) which is controlled by density function. It means that cells do not divide or grow if locally there is no sufficient space, i.e. if the cell has no place to divide and is unable to push the neighbours, so as to generate a needed place. In the particular case when cell division occurs, the number of cells in the cellular system is increased accordingly. The next iteration of the simulation starts with a new cell number. The motion of each cell is on the other hand driven by the gradient of the potential function. Cells move towards the direction computed from repulsive and adhesive interactions with their neighbours (Step 7). In our model, we assume that all cells involved in the simulation reside in a 3D finite computational box. If any cell moves out of the box, it is removed from the simulation and not considered further.
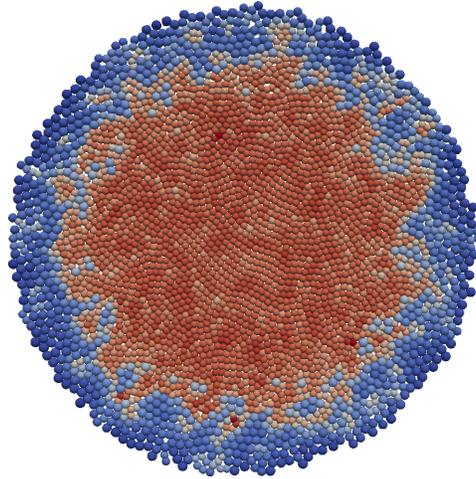
Fig. 4: Density of cells in a colony shown for clarity in 2-D (red - higher density, blue - lower density). Proliferation occurs mainly in lower density areas.

4   Details of parallel implementation

Developing powerful applications for today's largest supercomputing systems, equipped with thousands and millions of computational cores, requires the use of advanced parallel programming techniques and tools. In this section we describe the effort we have made to optimize performance of described biological computational framework on the IBM Blue Gene/Q system.

Maintaining extremely large number of processes with the use of flat MPI programming model results in a very large memory overhead and significantly reduces the scalability of applications. One of the most popular ways to overcome these issues is to use a multi-level or hybrid parallelization model. In this case we are using mixed MPI and OpenMP model. Depending on the size of the biological system under consideration we can use from one to sixteen MPI processes on each node of BG/Q. Each MPI process is further parallelized with the use of OpenMP. Total number of threads per node in each simulation run is equal to 64. In this way the SMT hardware support available on PowerPC A2 chip is fully utilized.

Performance optimization on the IBM Blue Gene/Q system was carried out iteratively. We have designed and executed a large scale benchmark run. We have collected MPI trace data with the use of TAU performance analysis tool [14]. Results of the analysis showed an unexpected imbalance between parallel processes which was identified as the main bottleneck limiting code's scalability. Fig. 5 shows a detailed look at the computer program activities in a single iteration step of the simulation, before and after optimization. Computational activities are depicted in pink whereas the standby mode when processes are idle is depicted in purple. At the end of each iteration step there is a synchronization
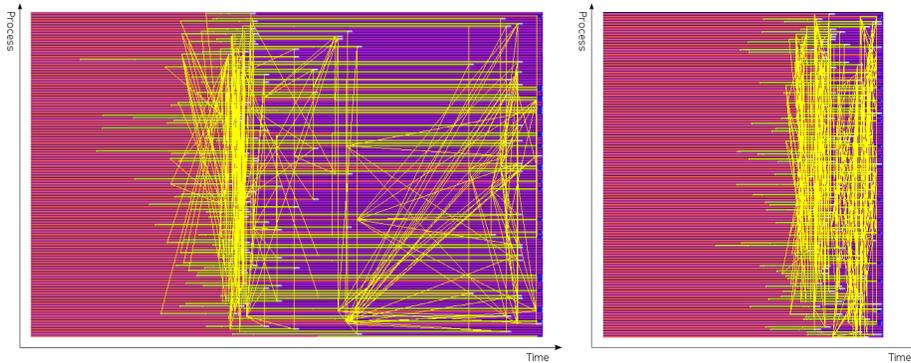
Fig. 5: Visualization of the trace data collected during one simulation step. The imbalance between working processes in the case of non-optimal domain partitioning (left) is confronted with optimized domain partitioning (right). Measurements and plots were obtained with the use of TAU performance analysis tool [14].

barrier which is preceded by an exchange of data between processes (depicted with yellow lines). As can be seen on the non-optimal version (left) some of the processes are waiting on the barrier for other to finish their work.

We have identified that non-optimal load balancing between working processes was the main performance and scalability bottleneck. The main observation was that computing potential function values is more time consuming for cells which reside in high-density areas (e.g. inside of a developing solid tumour). Such cells have more interactions with their neighbours and therefore tree traversal is more complex. To solve this problem we have used weighted PHSFC partitioning. Now each cell has a weight assigned to itself, which is equal to the value of the density function computed in cell's center. The trace analysis of the optimized code is shown in Fig. 5 (right). As can be seen the imbalance between processes was significantly reduced.

Fig. 6 presents applications scalability on the IBM Blue Gene/Q system. The maximal computational partition used in this measurement consisted of 512 nodes (which is equivalent to approximately 100 TFlop/s). As can be seen we achieved nearly linear speedup and we expect that this behaviour will remain on larger computational partitions for bigger benchmark sizes.

5   Example of application

In this section we present an example simulation of early tumour growth obtained with our model. Solid tumours are believed to start from a single mutated cell that, due to subsequent divisions, becomes a multi-cellular spheroid (MCS) [15]. In the subsequent development tumour adapts more complex structure and looses its spherical symmetry. Here, the simulation size was reduced to
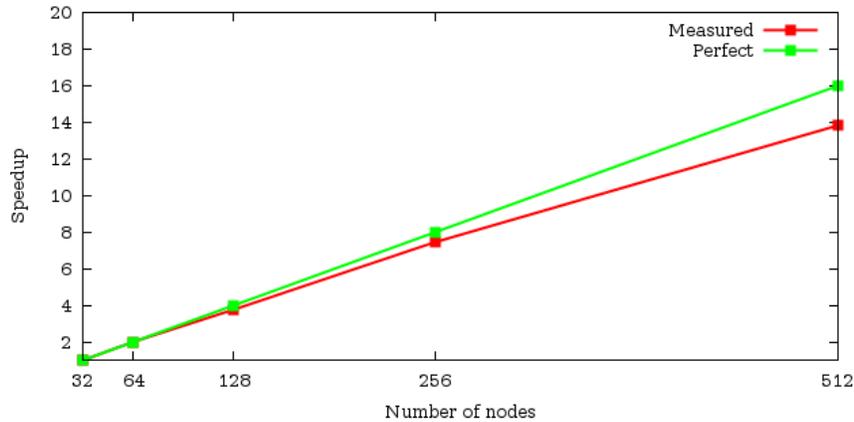
Fig. 6: Scalability results: simulation of a large scale cellular colony on the IBM
         Blue Gene/Q.

10^6. Thanks to this we were able to present visualizations of the underlying phenomena. Visualizations of 10^9 size simulations are actually limited to the analysis of specyfied smaller subregions of the tissue.

Prior to starting the simulation we develop a large colony consisting of 10^6 cells that represents a healthy tissue. At the beggining of simulation healthy cells start to undergo the G_1/S and G_2/M checkpoints. We have chosen the probability P_{c_i} of being marked for programmed cell death (apoptosis) to be equal to 0.5. This ensures the homeostasis of the whole cell population, which basically means that we obtain the stabilization of the cell number in the colony. Also, at this point of the simulation we place a single tumour cell in G_1 phase in the middle of the tissue.

This initial tumour cell divides repeatedly following a set of rules and produces a cluster of cells. The system is updated repeatedly as the program runs through a loop. During one time step, for each cell its phase is checked and, if necessary, updated. A cell divides if it has sufficient space around it to place its daughter cell. If there is no free place nearby, the cell can "push" up some neighbouring cells in order to create an empty space for its daughter cell. The exact spatial location for the new cell is computed using a potential function.

There is a significant differentiation between cancer and healthy cells in presented simulation. Firstly, for healthy cells average cell cycle phases lengths are: 11, 8, 4 and 1 hour for G_1, S, G_2 and M phases respectively. Cancer cells have shorter G_1 phase which is only 6 hours in average. Secondly, cancer cells have greater ability to grow and divide in high-density neighbourhood than healthy cells. Finally, the diseased cells produce slightly weaker adhesion forces. We have been able to incorporate all those differences to the life cycles of single cells thanks to the individual-based approach of our model. Chosen steps of the simulation together with the description of their main characteristics are
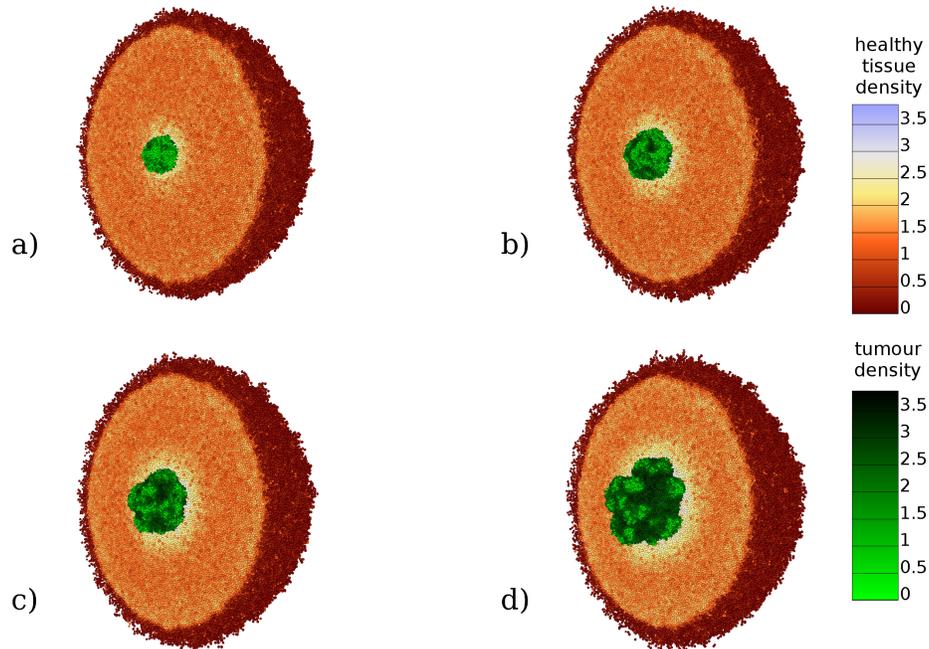
Fig. 7: Plots showing the simulation of growing solid tumour inside a healthy
tissue. Plot a) corresponds to the growing tumour after 80 days, with
2742 cancer cells. Plot b) corresponds to the growing tumour after 82.5
days, with 8091 cancer cells. Plot c) corresponds to the growing tumour
after 85 days, with 15626 cancer cells. Finally, plot d) corresponds to
the growing tumour after 90 days, with 37681 cancer cells. The whole
cell population in all time steps consisted of approximately 10^6 cells.

presented on Fig. 7.

6   Summary and future work

In this paper we presented a new powerful computational tool with a very wide
range of applications in the biomedical sciences. The proposed model can be
adapted to simulate many biological processes, such as: i) tumour development;
ii) vasculogenesis which is the process of blood vessel formation occurring by a
de novo production of endothelial cells, iii) angiogenesis which is the process of
new blood vessels formation from pre-existing vessels; iv) wound healing; v) the
development of biofilms, the complex structures of bacteria and other organisms
surrounded by a layer of organic and inorganic substances produced by these
micro-organisms; vi) tissue regeneration and many others.

   In the future the proposed model will be extended by taking into account the
environment described in a continuous manner by partial differential equations.

Further developments will include introduction of some intracellular kinetics into each individual cell. Intracellular dynamics will be described with systems of ODE and the whole model will gain genuine multi-scale character. First work of that type, where authors describe intracellular dynamics of proteins involved in regulation of adhesion one can find in [16]. Another extension of the model could be the consideration of fibres of the extracellular matrix. It is particularly important in case of solid tumour growth, cancer invasion and cellular motility phenomena. Preliminary works on this topic can be found in [17]. Such methodology will provide a new insight into processes occurring in living organisms and multicellular systems enabling its better understanding and control. It will make possible to simulate not only selected processes but the whole phenomena. In particular, in case of cancer disease it will allow not only the simulation of selected processes such as invasion and angiogenesis but it will make possible to combine them into a one large multi-scale model of cancer growth. Such virtual solid tumour may have a genuine impact on cancer treatment.

Acknowledgements

References

[1] J. S. Lowengrub, H. B. Frieboes, F. Jin, Y-L. Chuang, X. Li, P. Macklin, S. M. Wise and V. Cristini. Nonlinear modelling of cancer: bridging the gap between cells and tumours. *Nonlinearity*, 23:R1–R9, 2010.

[2] Single-cell-based Models in Biology and Medicine, eds. A. Anderson, M. A. J. Chaplain and K. Rejniak, *Birkhauser Verlag*, Basel. 2007.

[3] P. W. Coteus Packaging the IBM Blue Gene/Q supercomputer. *IBM Journal of Research and Development*, 57: 2:1–2:13, 2013.

[4] M. Cytowski, Z. Szymańska, B. Borucki. A 2-D large-scale individual-based model of solid tumour growth. *GAKUTO International Series Mathematical Sciences and Applications, Nonlinear Analysis in Interdisciplinary Sciences*, 36:43-56, 2013.

[5] M. D. Garrett Cell cycle control and cancer. *Curr. Sci.*, 81:515–522, 2001.

[6] J. Galle, M. Loeffer and D. Drasdo, Modelling the Effect of Deregulated Proliferation and Apoptosis on the Growth Dynamics of Epithelial Cell Populations In Vitro, em Biophysical Journal, 88, 62–75 2005.

[7] R. E. Mahaffy, C.K. Shih, F. C. MacKintosh, and J. Kas. Scanning Probe-Based Frequency-Dependent Microrheology of Polymer Gels and Biological Cells. *Physical Review Letters*, 85(4):880–883, 2000.

[8] A. J. Maniotis, C. S. Chen, and D. E. Ingber. Demonstration of mechanical connections between integrins, cytoskeletal filaments, and nucleoplasm that stabilize nuclear structure. *Cell Biology*, 94:849–854, 1997.

[9] I. Ramis-Conde, D. Drasdo, A. R. A. Anderson, and M. A. J. Chaplain. Modelling the the influence of the E-Cadherin-β-catenin pathway in cancer cell invasion: A multi-scale approach. *Biophys. J.*, 95:155-165, 2008.

[10] I. Ramis-Conde, M. A. J. Chaplain, and A. R. A. Anderson. Mathematical modelling of cancer cell invasion of tissue. *Math. Comput. Modelling*, 47:533-545, 2008.

[11] V. Springel. The cosmological simulation code GADGET-2. *Mon. Not. R. Astron. Soc.*, 364:1105–1134, 2005.

[12] K. Misoo, Y. Akiyama, Y. Shizawa and M. Saito Development of molecular dynamics programs for proteins with a parallelized Barnes-Hut tree code. *High Performance Computing in the Asia-Pacific Region, 2000. Proceedings.*, 2, pp.1103 – 1111, 2000.

[13] J. J. Monaghan Smoothed particle hydrodynamics. *Annu. Rev. Astron. Astrophys.*, 30:543–574, 1992.

[14] S. Shende and A. D. Malony. The TAU Parallel Performance System. *International Journal of High Performance Computing Applications, SAGE Publications*, 20:287–331, 2006.

[15] S. Kruś Anatomia patologiczna, *PZWL*, Warszawa, 2000.

[16] I. Ramis-Conde, M. A. J. Chaplain, A. R. A. Anderson, and D. Drasdo. Multiscale modelling of cancer cell intravasation: the role of cadherins in metastasis. *Phys. Biol.*, 6:016008, 2009.

[17] Schlüter, D.K., Ramis-Conde, I., Chaplain, M.A.J. Computational modeling of single cell migration: The leading role of extracellular matrix fibbers. Biophys. J., 103, 1141-1151, 2012.

[18] VisNow, a generic visualization framework. *http://visnow.icm.edu.pl*