ADJOINT SENSITIVITY ANALYSIS OF A TUMOR GROWTH MODEL AND ITS APPLICATION TO SPATIOTEMPORAL RADIOTHERAPY OPTIMIZATION

Krzysztof Fujarewicz* and Krzysztof Łakomiec

Silesian University of Technology
Akademicka 16, 44-100, Gliwice, Poland

Abstract. We investigate a spatial model of growth of a tumor and its sensitivity to radiotherapy. It is assumed that the radiation dose may vary in time and space, like in intensity modulated radiotherapy (IMRT). The change of the final state of the tumor depends on local differences in the radiation dose and varies with the time and the place of these local changes. This leads to the concept of a tumor's spatiotemporal sensitivity to radiation, which is a function of time and space. We show how adjoint sensitivity analysis may be applied to calculate the spatiotemporal sensitivity of the finite difference scheme resulting from the partial differential equation describing the tumor growth. We demonstrate results of this approach to the tumor proliferation, invasion and response to radiotherapy (PIRT) model and we compare the accuracy and the computational effort of the method to the simple forward finite difference sensitivity analysis. Furthermore, we use the spatiotemporal sensitivity during the gradient-based optimization of the spatiotemporal radiation protocol and present results for different parameters of the model.

1. **Introduction.** Intensity modulated radiotherapy (IMRT) is an advanced type of radiation therapy used to treat tumors. It is a type of conformal radiation, which shapes radiation beams to closely approximate the shape of the tumor. In general, it allows to deliver a different radiation dose to different points in 3D space. The goal of IMRT is to adjust the radiation dose to the target and to reduce exposure of healthy tissue to minimize the side effects of treatment. Conformal radiation may be used in one or more sessions. Radiotherapy can also be a component of more advanced combined anticancer therapies [14].

In current clinical practice IMRT is planned based on the current state (shape and volume) of the tumor and there are very few studies which try to use mathematical modeling to predict the response to the therapy. One of the first such attempts is an approach proposed by Corwin and co-authors in [1]. They used a spatial mathematical model of tumor growth to simulate and predict the response to a dose of particular shape, and based on this information they optimized the therapy using an evolutionary optimization algorithm. It is important to stress that in [1] the mathematical model was simulated only *forward* in time, which is, of course, a typical way of using any mathematical model.

The problem of optimizing therapy is related to the problem of *sensitivity* of an objective function, measuring the quality of the response, with respect to the therapy protocol.

In this work we use adjoint sensitivity analysis, which is a sensitivity method that utilizes the adjoint system constructed based on the analyzed model and simulated *backward* in time. Adjoint sensitivity analysis, in a special structural form, is especially useful for models described by block diagrams and was originally developed for neural networks [2] and afterwards was used for different models described by ordinary differential equations [3, 4, 8, 9], systems with delays [5] and age-structured models [7]. Adjoint sensitivity analysis is especially useful in analyzing MISO (Multiple-Input Single-Output) systems due to the much lower computational complexity when compared to forward sensitivity analysis. For example, it is useful in a parameter estimation process where the system analyzed has many inputs (parameters) and usually one scalar output describing the system (objective function). In the present work adjoint sensitivity analysis is used to compute the spatiotemporal gradient of the objective function with respect to the input signal of the finite difference scheme being the result of the discretization of a given partial differential equation.

2. **Mathematical model of growth of a tumor.** In this paper we derive the spatiotemporal sensitivity to radiation for a mathematical model for tumor proliferation, invasion and response to radiotherapy (PIRT model) published previously in [1, 10, 11]. The model is formalized as the following reaction-diffusion equation:

$$\frac{\partial c}{\partial t} = \nabla \cdot (D\nabla c) + \rho c \left(1 - \frac{c}{k_t}\right) - Rc \left(1 - \frac{c}{k_t}\right) \tag{1}$$

where: D is a diffusion coefficient (tumor invasiveness), ρ means proliferation coefficient, k_t is a environment capacity for tumor cells and R is a term describing the effect of radiation on the tumor cells. In this work R is formalized as a linear-quadratic term:

$$R = 1 - e^{-\alpha d - \beta d^2} \tag{2}$$

where d is the radiation dose or, more precisely, spatiotemporal radiation dose distribution. Later in this article d is called shorter as the radiation dose distribution or simply as the dose distribution. In this work we use and optimize radiation as a continuous function of time, hence the term R has interpretation of rate constant with units $1/time$. The interpretation of the term R has been a topic of a recent discussion — see articles [12, 13] and the discussion therein. α and β are parameters of the linear-quadratic model.

We used the finite-difference method to discretize the model (1) which gives the following difference equation:

$$c_i^{j+1} = c_i^j + \Delta t \left(D\frac{c_{i-1}^j - 2c_i^j + c_{i+1}^j}{\Delta x^2} + \rho c_i^j \left(1 - \frac{c_i^j}{k_t}\right) - R_i^j c_i^j \left(1 - \frac{c_i^j}{k_t}\right)\right) \tag{3}$$

$$R_i^j = 1 - e^{-\alpha d_i^j - \beta (d_i^j)^2}$$

where: c_i^j means the density of tumor cells at spatial point i and at time j and d_i^j is the radiation dose at time j and at spatial point i. The space step Δx is equal to 1 and the final space point is 200, therefore $i_{max} = 200$. The time step Δt is 0.01 and the final time is 2 which gives $j_{max} = 200$. The parameters (in arbitrary units) of the model are: $D = 10$, $\rho = 10$, $k_t = 2.4e5$. The initial condition for the model
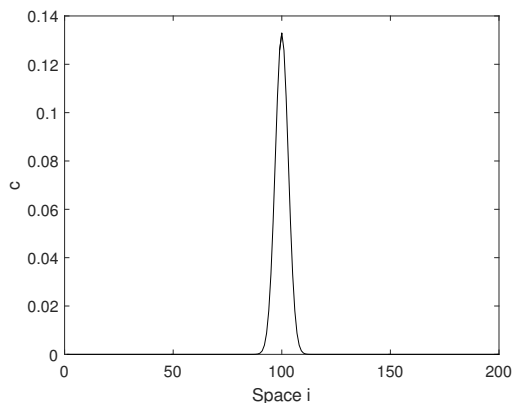
FIGURE 1. Initial condition c_i^1 for the model (3)

is shown in Fig. 1. The parameters for the linear-quadratic model of response to irradiation are: $\alpha = 1$, $\beta = 1$.

Simulations of the model without irradiation and with uniform irradiation are shown in Fig. 2.

The discretized model (3) is an example of a one dimensional finite difference scheme usually used in numerical simulations of spatiotemporal models. The simplest form of such a difference scheme is as follows:

$$y_i^{j+1} = f(y_i^j, y_{i-1}^j, y_{i+1}^j, u_i^j), \quad j = 1, 2, \ldots, j_{max} - 1, \quad i = 1, 2, \ldots, i_{max} \qquad (4)$$

where: y_i^j is a discretized scalar variable of the original PDE at the time j and at the spatial point i, u_i^j is the value of an input signal affecting the system (if such a signal exists) at the time j and at the spatial point i, $f(\cdot)$ is a general non-linear scalar function, i_{max} is the number of spatial points, and j_{max} is the final discrete time.

The difference scheme (4) is very simple: it is only one dimensional, y and $f(\cdot)$ are scalar, and y_i^{j+1} depends on the state of the difference scheme only at the previous j-th time and only in the closest neighborhood $(i-1, i, i+1)$. In general, the difference scheme resulting from the PDE may be more complicated; it may be higher dimensional, y and $f(\cdot)$ may be multidimensional, and y_i^{j+1} may depend on a longer history horizon of the difference scheme and a larger neighborhood.

The difference scheme (4) can also be viewed as a *coupled map lattice* or a special case of a *cellular automaton* with continuous state.

3. **Problem formulation.** Let us consider the difference equation (4) and let us assume that the minimized objective function is a functional and depends on the solution of this equation as follows:

$$J = \sum_{j=1}^{j_{max}} g^j(y^j) \qquad (5)$$

where y^j is the state of the difference scheme at time j

$$y^j = [y_1^j, y_2^j, \ldots, y_{i_{max}}^j]^T \qquad (6)$$

and $g^j(\cdot)$ are given non-linear scalar functions.

The whole spatiotemporal input signal U can be aggregated and presented as a matrix:

$$U = \begin{bmatrix} u_1^1 & \cdots & \cdots & \cdots & u_{i_{max}}^1 \\ \vdots & \ddots & & & \vdots \\ \vdots & & u_i^j & & \vdots \\ \vdots & & & \ddots & \vdots \\ u_1^{j_{max}} & \cdots & \cdots & \cdots & u_{i_{max}}^{j_{max}} \end{bmatrix} \tag{7}$$

The problem formulated and solved in this paper consists of finding a gradient G of an objective function J in a space of the spatiotemporal input signal:

$$G = \nabla_U J \tag{8}$$

which is a matrix of partial derivatives:

$$G = \begin{bmatrix} \frac{\partial J}{\partial u_1^1} & \cdots & \cdots & \cdots & \frac{\partial J}{\partial u_{i_{max}}^1} \\ \vdots & \ddots & & & \vdots \\ \vdots & & \frac{\partial J}{\partial u_i^j} & & \vdots \\ \vdots & & & \ddots & \vdots \\ \frac{\partial J}{\partial u_1^{j_{max}}} & \cdots & \cdots & \cdots & \frac{\partial J}{\partial u_{i_{max}}^{j_{max}}} \end{bmatrix} \tag{9}$$

Once the gradient G is obtained it may be utilized during a gradient-descent iterative minimization procedure of the objective function appearing, for example, in parameter estimation or control optimization problems.

4. **Problem solution.** To calculate the spatiotemporal gradient of the objective function J with respect to the whole input signal U we will utilize the concept of the adjoint system and adjoint variables known from the discrete formulation of the Pontryagin maximum principle.

Let us introduce a scalar function called Hamiltonian:

$$H(y^j, u^j, p^{j+1}) = g^j(y^j) + \sum_{i=1}^{i_{max}} p_i^{j+1} f(y_i^j, y_{i-1}^j, y_{i+1}^j, u_i^j) \tag{10}$$

where $p_i^j, i = 1, 2, \ldots, i_{max}, j = 1, 2, \ldots, j_{max}$ are adjoint variables (or Lagrange multipliers), each corresponding to one scalar constraint (4) and satisfying the following equations of the adjoint system:

$$p_i^j = \frac{\partial H(\cdot)}{\partial y_i^j}, \quad j = 1, 2, \ldots, j_{max}, \quad i = 1, 2, \ldots, i_{max} \tag{11}$$

Now, let us calculate p_i^j based on (10) and taking into account that y_i^j appears under the sum symbol in (10) three times: for i as y_i^j, for $i-1$ as y_{i+1}^j and for $i+1$ as y_{i-1}^j. This observation leads to the following equation of the adjoint system

$$p_i^j = \frac{\partial g^j(\cdot)}{\partial y_i^j} + p_i^{j+1} \cdot \left(\frac{\partial f(\cdot)}{\partial y_i^j}\right)_i^j + p_{i-1}^{j+1} \cdot \left(\frac{\partial f(\cdot)}{\partial y_{i+1}^j}\right)_{i-1}^j + p_{i+1}^{j+1} \cdot \left(\frac{\partial f(\cdot)}{\partial y_{i-1}^j}\right)_{i+1}^j . \tag{12}$$

The adjoint system (12) should be solved backward in time with the initial condition

$$p_i^{j_{max}} = \frac{\partial g^{j_{max}}(\cdot)}{\partial y_i^{j_{max}}}, \quad i = 1, 2, \ldots, i_{max} \tag{13}$$

Then the elements of the searched gradient (9) can be calculated as partial derivatives of the Hamiltonian with respect to u_i^j:

$$\frac{\partial J}{\partial u_i^j} = \frac{\partial H(\cdot)}{\partial u_i^j} = p_i^{j+1} \frac{\partial f(\cdot)}{\partial u_i^j}. \tag{14}$$

5. **Results of gradient calculation — spatiotemporal sensitivity.** The objective function J is defined as the total density of tumor cells at the final time j_{max}:

$$J = \sum_{i=1}^{i_{max}} c_i^{j_{max}} \tag{15}$$

where $j_{max} = 200$, $i_{max} = 200$. Acquired from adjoint sensitivity analysis, the spatiotemporal gradient of the objective function (15) with respect to the radiation dose distribution d for model (3) is shown in Fig. 2. We have calculated this gradient twice, for two nominal spatiotemporal irradiation protocols: for zero irradiation (left panels) and for uniform irradiation $d \equiv 1$ (right panels). The results of simulation of the model compared to the antigradient (the gradient multiplied by -1) for particular time moments $j = 1, 100$, and 200 are shown in Fig. 3.

In this Figure we can observe that the shape of the spatiotemporal sensitivity of J is different from the shape of the growing tumor. This is in contradiction, to some extent, with our intuition that J should be more sensitive for a higher tumor density. This phenomena and its possible reasons will be discussed further in next sections where the optimal spatiotemporal radiation will be presented.

In order to confirm that the gradient is properly calculated we compared it with the gradient computed by the finite difference method. In this method we change the radiation dose distribution only at one time j and one space point i by Δd_i^j and check the resulting variation of the objective function ΔJ. Then the gradient at time j and space point i can be approximated by the following formula:

$$(\nabla_d J)_i^j \approx \frac{\Delta J}{\Delta d_i^j}. \tag{16}$$

This has to be repeated $i_{max} \times j_{max}$ times to obtain the approximation of the whole gradient. Comparisons of the gradient calculated by the finite difference method and the adjoint sensitivity analysis for time $j = 150$ are shown in Figs. 4 and 5.

We can also compare the time needed to execute both methods. On a laptop with Intel Core i7-4700MQ CPU and with implementation of the algorithms in MATLAB, the execution time that is needed to calculate the spatiotemporal gradient is about 2 minutes when using the finite difference method. Using the adjoint sensitivity analysis described in this paper we can calculate the whole spatiotemporal gradient in about 0.3 second. One can see that the adjoint sensitivity approach is especially useful for this type of task, especially when we take into account that during gradient-based therapy optimization the gradient has to be calculated many times.

The MATLAB code of the adjoint sensitivity analysis and the finite difference method used in this paper can be downloaded from the MathWorks File Exchange site.
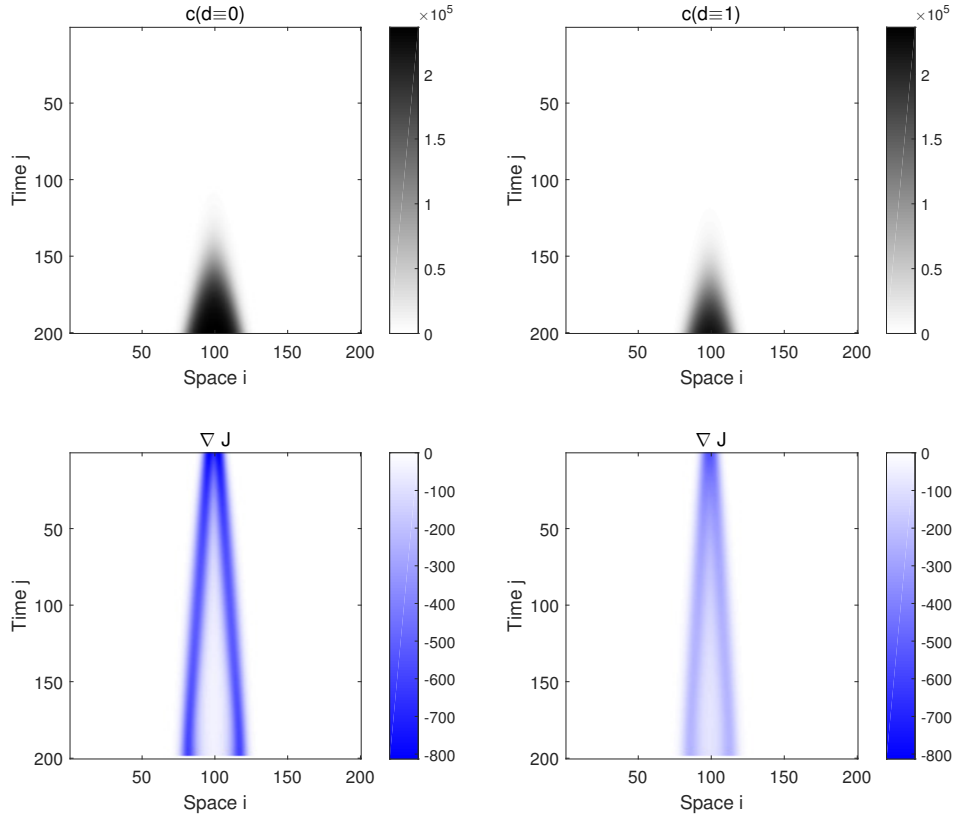
FIGURE 2. Simulation of the model (3), and calculated from adjoint sensitivity analysis the spatiotemporal gradient of the objective function (15) with respect to the radiation dose distribution d for different irradiation values. Left panels: radiation dose distribution d equal to zero, right panels: uniform irradiation $d \equiv 1$

6. **Spatiotemporal radiation optimization.** To optimize the radiation dose distribution d we used the gradient descent method with additional constraints:

- Maximum cumulative dose S_{max}:

$$S = \sum_{j=1}^{j_{max}} \sum_{i=1}^{i_{max}} d_i^j, \quad S \leq S_{max} \tag{17}$$

- Maximum point dose d_{max}:

$$d_i^j \leq d_{max}, \quad j = 1, 2, \ldots, j_{max}, \quad i = 1, 2, \ldots, i_{max} \tag{18}$$

To incorporate these constraints during optimization we modified the defined objective function J and the way in which the input signal of radiation dose distribution is generated.

In order to implement the maximum cumulative dose constraint (17) we applied an external *penalty function* approach by adding an extra term to the objective
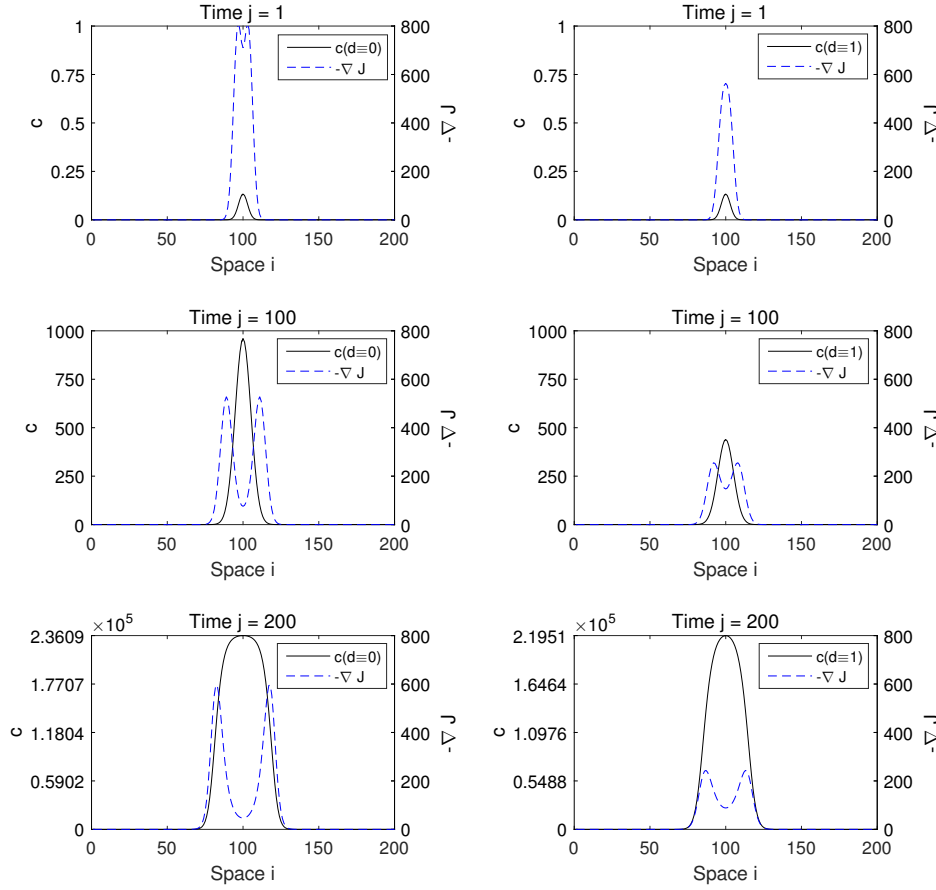
FIGURE 3. Simulation of the model (3) compared with the antigradient of objective function (15) with respect to the radiation dose distribution d for particular time moments $j = 1, 100, 200$. Left panels: radiation dose distribution d equal to zero, right panels: uniform irradiation $d \equiv 1$

function J in cases when the minimization algorithm exceeds S_{max}:

$$J' = \left\{ \begin{array}{ll} J & S \leq S_{max} \\ J + \frac{1}{2}k(S - S_{max})^2 & S > S_{max} \end{array} \right. \tag{19}$$

where k is a scalar positive coefficient increasing during the optimization.

The maximum point dose constraint (18) was implemented by adding a new signal v which is an argument of a hyperbolic tangent function (having an upper limit equal to one). The value of this function multiplied by d_{max} gives us the radiation dose distribution d at which all values do not exceed d_{max}:

$$d_i^j = d_{max} \tanh(v_i^j), \qquad j = 1, 2, \ldots, j_{max}, \quad i = 1, 2, \ldots, i_{max} \tag{20}$$

After adding these modifications we calculated the spatiotemporal gradient of the modified objective function J' (instead of J) with respect to the new signal v (instead of d) during the optimization procedure.
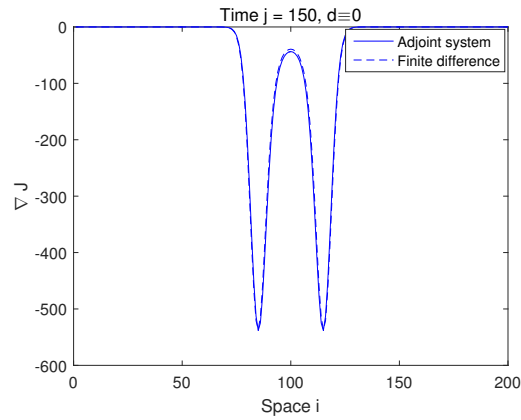
FIGURE 4. Comparison of the gradient calculated by the finite difference method and by adjoint sensitivity analysis
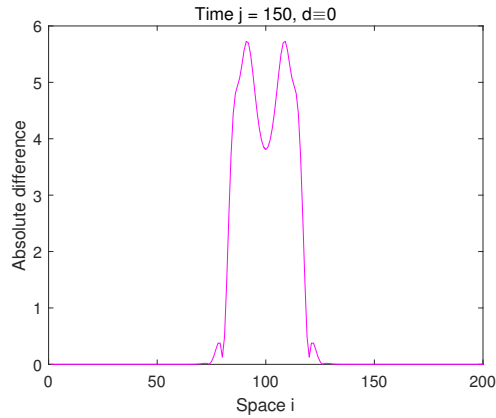

FIGURE 5. Absolute difference between gradients calculated by the finite difference method and by adjoint sensitivity analysis

The optimization was done for two sets of parameters acquired from [1]:

- $D = 27.7$, $\rho = 3.59$
- $D = 8.9$, $\rho = 50.29$

We chose these two cases with the minimal and the maximal value of ρ (the least and the most aggressive ones). Values of the constraints were $S_{max} = 3.8e4$ and $d_{max} = 4$ (radiation dose of 4 Gy is often used in radiotherapy, especially for palliative treatment [6]). The optimized radiation dose distribution and simulations of the model are shown in Figs. 6 and 7.

For the first set of parameters $D = 27.7$, $\rho = 3.59$ the values of the objective function J' before (for radiation equals zero) and after optimization were 1.12e3 and 163.90, respectively. For the second set of parameters $D = 8.9$, $\rho = 50.29$ (the most aggressive tumor) the value of the objective function J' before optimization was 3.30e7 and after optimization decreased only to 3.27e7. These results are summarized in Table 1. One can see that in the second case the radiation therapy has little effect on tumor progression.
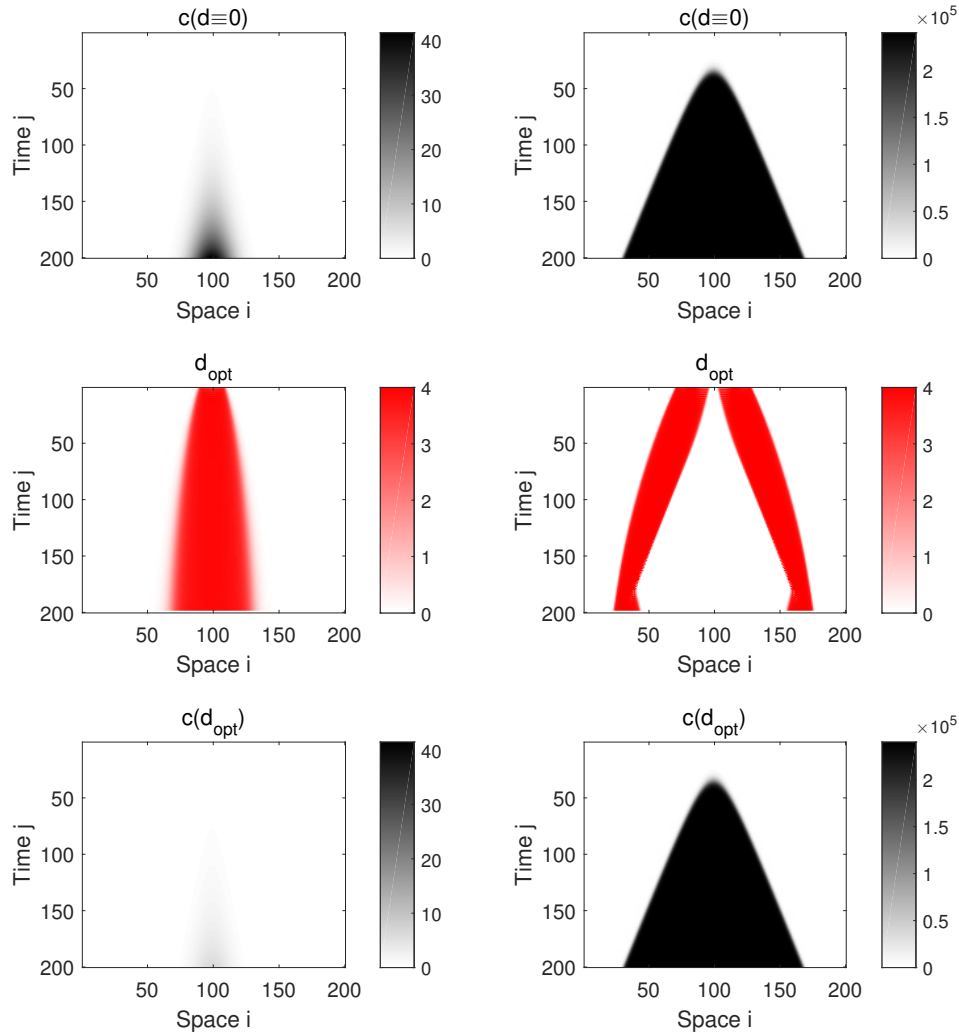
FIGURE 6. Simulations of the model (3) without irradiation, optimized radiation dose distribution d_{opt} and simulation of the model (3) with optimized radiation dose distribution $c(d_{opt})$ for two sets of parameters; left panels: $D = 27.7$, $\rho = 3.59$, right panels: $D = 8.9$, $\rho = 50.29$

Further, especially for the second (the worse) case, the shape of the optimal spatiotemporal radiation dose distribution is visibly different than the tumor density profile. This raises an interesting question: what is the reason for such an effect? In general, the radiation is applied at points and times where the tumor density reaches average values (non-zero and not saturated — less than k_t). When the density of the tumor reaches the maximal environment capacity k_t then the optimal irradiation level decreases. This is because of the form of the mathematical model of the tumor growth; the last term in (1), responsible for the killing action, disappears at the maximal density of the tumor and the optimal spatiotemporal radiation protocol
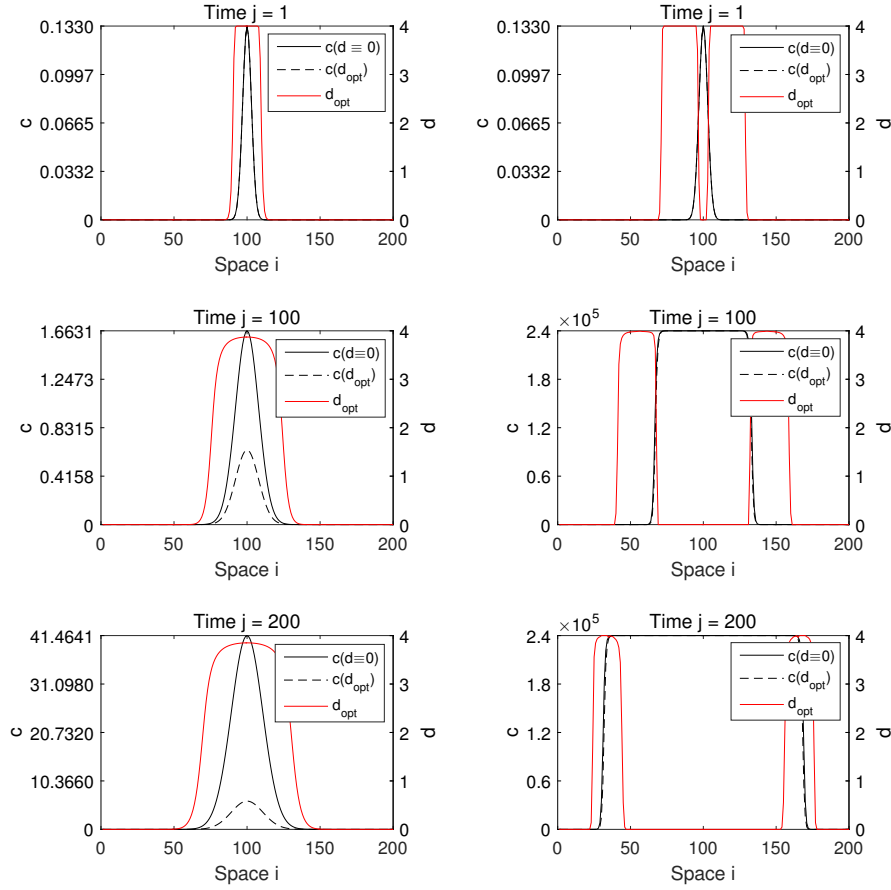
FIGURE 7. Simulations of the model (3) without irradiation $c(d \equiv 0)$, optimized radiation dose distribution d_{opt} and simulation of the model (3) with optimized radiation dose distribution $c(d_{opt})$ for particular time moments $j = 1$, 100, 200. Left panels: $D = 27.7$, $\rho = 3.59$, right panels: $D = 8.9$, $\rho = 50.29$.

is not applied to these places and times. Of course, this raises the question of the need to further improve the model used here, and in our future work we plan to apply our approach to different models and to assess the results.

Parameters	$D{=}27.7$, $\rho{=}3.59$	$D{=}8.9$, $\rho{=}50.29$
$J'(d \equiv 0)$	1.12e3	3.30e7
$J'(d_{opt})$	163.90	3.27e7

TABLE 1. Values of the objective function J' without irradiation and with optimized radiation dose distribution for two sets of model parameters

7. Conclusions. In this work we propose a novel approach to optimization of spatiotemporal radiotherapy using a mathematical model of tumor growth described by partial differential equations. This allows to control the radiation dose which can be different for different points of the tumor and surrounding tissue and for different times. The optimization is based upon adjoint sensitivity analysis which gives information about the spatiotemporal sensitivity (being a function of time and space) of a given scalar objective function characterizing the model's solution. The original partial differential model is transformed into a discrete difference scheme and then the adjoint sensitivity analysis and the radiation protocol is optimized. In future work, we plan to apply our approach to optimization of fractioned radiotherapy which is more realistic and applicable in clinical reality. We will also try to minimize other forms of the objective function by using the same adjoint sensitivity analysis.

REFERENCES

[1] D. Corwin, C. Holdsworth, R. C. Rockne, A. D. Trister, M. M. Mrugala, J. K. Rockhill, R. D. Stewart, M. Phillips and K. R. Swanson, *Toward Patient-Specific, Biologically Optimized Radiation Therapy Plans for the Treatment of Glioblastoma,* PLoS ONE 8, 2013.

[2] K. Fujarewicz and A. Galuszka, Generalized backpropagation through time for continuous time neural networks and discrete time measurements Artificial Intelligence and Soft Computing - ICAISC 2004 (eds. L. Rutkowski, J. Siekmann, R. Tadeusiewicz and L. A. Zadeh), *Lecture Notes in Computer Science*, **3070** (2004), 190–196.

[3] K. Fujarewicz, M. Kimmel and A. Swierniak, On fitting of mathematical models of cell signaling pathways using adjoint systems, *Math. Biosci. Eng.,* **2** (2005), 527–534.

[4] K. Fujarewicz, M. Kimmel, T. Lipniacki and A. Swierniak, Adjoint systems for models of cell signaling pathways and their application to parameter fitting, *IEEE/ACM Transacations On Computational Biology And Bioinformatics*, **4** (2007), 322–335.

[5] K. Fujarewicz and K. Łakomiec, Parameter estimation of systems with delays via structural sensitivity analysis, *Discrete and Continuous Dynamical Systems-series B*, **19** (2014), 2521–2533.

[6] P. Hoskin, A. Kirkwood, B. Popova, P. Smith, M. Robinson, E. Gallop-Evans, S. Coltart, T. Illidge, K. Madhavan, C. Brammer, P. Diez, A1. Jack and I. Syndikus, 4 Gy versus 24 Gy radiotherapy for patients with indolent lymphoma (FORT): a randomised phase 3 non-inferiority trial, *Lancet Oncology*, **15** (2014), 457–463.

[7] M. Jakubczak and K. Fujarewicz, Application of adjoint sensitivity analysis to parameter estimation of age-structured model of cell cycle, in *Information Technologies in Medicine*, (eds. E. Pietka, P. Badura, J. Kawa and W. Wieclawek), Advances in Intelligent Systems and Computing, **472**, Springer International Publishing, (2016), 123–131.

[8] K. Łakomiec and K. Fujarewicz, Parameter estimation of non-linear models using adjoint sensitivity analysis, *Advanced Approaches to Intelligent Information and Database Systems*, Springer, 2014, 59–68.

[9] K. Łakomiec, S. Kumala, R. Hancock, J. Rzeszowska-Wolny and K. Fujarewicz, Modeling the repair of DNA strand breaks caused by γ-radiation in a minichromosome, *Physical Biology*, **11** (2014), 045003.

[10] R. Rockne, E. C. Alvord Jr., J. K. Rockhill and K. R. Swanson, A mathematical model for brain tumor response to radiation therapy, *J. Math. Biol.*, **58** (2009), 561–578.

[11] R. Rockne, J. K. Rockhill, M. Mrugala, A. M. Spence, I. Kalet, K. Hendrickson, A. Lai, T. Cloughesy, E. C. Alvord Jr and K. R. Swanson, Predicting the efficacy of radiotherapy in individual glioblastoma patients in vivo: A mathematical modeling approach *Phys. Med. Biol.*, **55** (2010), 3271–3285.

[12] R. C. Rockne, A. D. Trister, J. Jacobs, A. J. Hawkins-Daarud, M. L. Neal, K. Hendrickson, M. M. Mrugala, J. K. Rockhill, P. Kinahan, K. A. Krohn and K. R. Swanson, Addendum to "A patient-specific computational model of hypoxia-modulated radiation resistance in glioblastoma using ^{18}F-FMISO-PET", *Journal of the Royal Society Interface*, **12** (2015), 20150927.

[13] R. Rockne, J. K. Rockhill, M. Mrugala, A. M. Spence, I. Kalet, K. Hendrickson, A. Lai, T. Cloughesy, E. C. Alvord and K. R. Swanson, Reply to comment on: "Predicting the efficacy of radiotherapy in individual glioblastoma patients in vivo: A mathematical modeling approach", *Physics in medicine and biology*, **61** (2016), 2968–2969.

[14] A. Swierniak, M. Kimmel, J. Smieja, K. Puszynski and K. Psiuk-Maksymowicz, *System Engineering Approach to Planning Anticancer Therapies*, Springer, 2016.

E-mail address: Krzysztof.Fujarewicz@polsl.pl
E-mail address: Krzysztof.Lakomiec@polsl.pl