Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 12–06

PHYSIOLOGIC PARAMETER ESTIMATION USING INVERSE PROBLEMS

ISABEL N. FIGUEIREDO AND CARLOS LEAL

Abstract: The estimation of *in vivo* physiologic parameters is an important, but difficult, issue in bio-medicine. Therefore the development of mathematical techniques predicting these parameter values is very relevant. In a previous work we have proposed a convection-diffusion-shape model, which correlates colonic crypt patterns with the cellular kinetics ocuring inside the crypts (this correlation is significant in the context of colorectal cancer). This model involves several physiologic parameters, for which only qualitative information is available in the literature, as for instance, the birth rate of proliferative cells. In this paper we present a framework for estimating this birth rate parameter, in a colonic crypt, assuming that the distribution of proliferative cells is known in that crypt. More precisely, we resolve an inverse problem, where the unknown coefficient field, the birth rate, is connected to the observed measurements, the proliferative cell density, through a partial differential equation. This inverse problem is a PDE-constrained optimization problem, highly non linear and time-dependent, which is solved by an inexact Newton method. Some test simulations illustrate the efficacy of the proposed parameter inversion, and forecast its application with real patient data.

Keywords: convection-diffusion equations, inverse problem, optimality system, finite elements.

AMS Subject Classification (2010): 35K99, 35J15, 65M32, 49K20, 65M60, 65M06.

1. Introduction

The epithelium of the colon is perforated by millions of small crypts, which play a crucial role in colon physiology. Each crypt has a cylindrical tube shape, that is closed at the bottom and with a round opening in the top, directed at the lumen's colon. Different types of cells fill the crypt. These are aligned along the crypt wall: stems cells are believed to reside in the bottom of the crypt, transit cells along the middle part of the crypt axis and differentiated cells at the top of the crypt. The colon epithelium undergoes a

Received February 20, 2012.

This work was partially supported by the research project UTAustin/MAT/0009/2008 of the UT Austin | Portugal Program (http://www.utaustinportugal.org/) and by CMUC and FCT (Portugal), through European program COMPETE/FEDER.

complete renewal, by means of a programmed mechanism driven by the cellular kinetics inside the crypts [14, 28, 29]. In normal human colonic crypts, the cells renew completely each 5-6 days (see [23]), through an harmonious and ordered procedure which includes the proliferation of cells, their migration along the crypt wall towards the top and their apoptosis, as they reach the orifice of the crypt and the cell cycle is finished. Under normal conditions the colonic cellular mechanism is regulated by biochemical and biomechanical signals. Any disfunction of this process might originate loss of homeostasis in colonic crypts, promoting an increase of the proliferative rate and a decrease in the death rate of cells, and neoplasia results (see [5, 9, 11, 16, 21, 25, 27]).

There are many works, in which convection-diffusion equations are used for describing tissue-growth, tumor-growth, cellular kinetics (see for instance [15, 32, 31]). In particular, in a previous work (see [6]) we have proposed and implemented a convection-diffusion-shape model for simulating and predicting the morphogenesis of a colonic crypt. This model couples a parabolic type equation, whose unknown is the proliferative cell density inside a single colonic crypt, with an elliptic equation whose unknown is the pressure, related to the convective velocity of the proliferative cells by Darcy's law (we assume the cells flow through the crypt, like fluid through a porous medium [10, 22]). The parameters involved in this convection-diffusion system are the rate of birth, the rate of death, and the diffusion, of the proliferative cell density. There are many articles in the literature reporting qualitative information about these parameters, but on the other hand there is very few references, regarding quantitative values for these parameters (see for example [1, 17, 18]). In reality these type of parameters are very difficult, or even impossible, to be measured directly and *in vivo*, using experimental or imaging techniques. Therefore, if other related quantities are available and can be measured (as for instance, the distribution of proliferative cells, which may be inferred using CEM images; CEM (confocal laser endomicroscopy) is a new diagnosis technique that enables the histological examination of suspicious tissues in real-time during an ongoing endoscopy), then solution methods for inverse problems are, undeniably, valuable tools for estimating these parameters. In an inverse problem, the unknown parameters are associated to the data or observable measurements, through a partial differential equation (see for instance [13, 20, 30] for examples and solution methods for inverse problems), and the goal is to minimize the misfit between the data and the prediction of the model (the latter is the solution of the partial differential

[24] I-M. Shih et al. Top-down morphogenesis of colorectal tumors. *Proceedings of the National Academy of Sciences of the United States*, 98(5):2640–2645, 2001.
[25] O. M Sieber, K. Heinimann, and I. P. Tomlinson. Genomic instabilitythe engine of tumorigenesis? *Nat Rev Cancer*, 3(9):701–708, 2003.
[26] The Mathworks, Inc. *http://www.matlab.com.*
[27] I. Tomlinson, P. Sasieni, and W. Bodmer. How many mutations in a cancer? *Am J Pathol*, 160(3):755–758, 2002.
[28] I. M. M. van Leeuwen, H. M. Byrne, O. E. Jensen, and J. R. King. Crypt dynamics and colorectal cancer: advances in mathematical modelling. *Cell Proliferation*, 39:157–181, 2006.
[29] I. M. M. van Leeuwen, C. M. Edwards, M. Ilyas, and H. M. Byrne. Towards a multiscale model of colorectal cancer. *World Journal of Gastroenterology*, 13(9):1399–1407, 2007.
[30] C. R. Vogel. *Computational methods for inverse problems*. Frontiers in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002.
[31] A. C. Walter. *A Comparison of Continuum and Cell-based Models of Colorectal Cancer*. PhD thesis, University of Nottingham, March 2009.
[32] J. P. Ward and J. R. King. Mathematical modelling of avascular-tumor growth. *IMA Journal of Mathematics Applied in Medicine and Biology*, 14:39–69, 1997.

Isabel N. Figueiredo
CMUC, Department of Mathematics, University of Coimbra, 3001-454 Coimbra, Portugal
E-mail address: isabelf@mat.uc.pt

Carlos Leal
CMUC, Department of Mathematics, University of Coimbra, 3001-454 Coimbra, Portugal
E-mail address: carlosl@mat.uc.pt

[4] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.
[5] E. R. Fearon and B. Vogelstein. A genetic model for colorectal tumorigenesis. *Cell*, 61(5):759–767, 1990.
[6] I.N. Figueiredo, C. Leal, G. Romanazzi, and B. Engquist P.N. Figueiredo. A convection-diffusion-shape model for aberrant colonic crypt morphogenesis. *Computing and Visualization in Science*, (in press), 2012.
[7] Isabel N. Figueiredo, Pedro N. Figueiredo, and Nuno Almeida. Image-driven parameter estimation in absorption-diffusion models of chromoscopy. *SIAM Journal on Imaging Sciences*, 4(3):884–904, 2011.
[8] P. Figueiredo, M. Donato, M. Urbano, H. Goulão, H. Gouveia, C. Sofia, M. Leitão, and Diniz Freitas. Aberrant crypt foci: endoscopic assessment and cell kinetics characterization. *International Journal of Colorectal Disease*, 24(4):441–450, 2009.
[9] J. Galle, M. Loeffler, and D. Drasdo. Modeling the effect of deregulated proliferation and apoptosis on the growth dynamics of epithelial cell populations in vitro. *Biophysical Journal*, 88:62–75, 2005.
[10] H. P. Greenspan. On the growth and stability of cell cultures and solid tumors. *Journal of Theoretical Biology*, 56(2):229–242, 1976.
[11] D. Hanahan and R. A. Weinberg. The hallmarks of cancer. *Cell*, 7(100):57–70, 2000.
[12] C. Hogea, C. Davatzikos, and G. Biros. An image-driven parameter estimation problem for a reaction-diffusion glioma growth model with mass effects. *J. Math. Biol.*, 56(6):793–825, 2008.
[13] K. Ito and K. Kunisch. *Lagrange multiplier approach to variational problems and applications*, volume 15 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.
[14] M. D. Johnston, C.M. Edwards, W. F. Bodmer, P. K. Maini, and S. J. Chapman. Mathematical modeling of cell population dynamics in the colonic crypt and in colorectal cancer. *Proceedings of the National Academy of Sciences of the United States*, 104(10):4008–4013, 2007.
[15] J. R. King and S. J. Franks. Mathematical analysis of some multi-dimensional tissue-growth models. *European Journal of Applied Mathematics*, 15:273–295, 2004.
[16] F. P. Michor, Y. Iwasa, H. Rajagopalan, C. Lenguauer, and M.A. Nowak. Linear model of colon cancer initiation. *Cell Cycle*, 3(3):358–362, 2004.
[17] P. J. Murray, C. E. Edwards, M. J. Tindall, and P. K. Maini. From a discrete to a continuum model of cell dynamics in one dimension. *Physical Review E*, 80(031912), 2009.
[18] P. J. Murray, A. Walter, A. G. Fletcher, C. E. Edwards, M. J. Tindall, and P. K. Maini. Comparing a discrete and continuum model of the intestinal crypt. *Physical Biology*, 8(026011), 2011.
[19] J. Nocedal and S. J. Wright. *Numerical optimization*. Springer Series in Operations Research and Financial Engineering. Springer, New York, second edition, 2006.
[20] N. Petra and G. Stadler. Model variational inverse problems governed by partial differential equations. *The Institute for Computational Engineering and Sciences, The University of Texas at Austin, USA*, ICES Report No. 11-05, 2011.
[21] A. G. Renehan, Haboubi N. J. ODwyer, S. T., and C. S. Potten. Early cellular events in colorectal carcinogenesis. *Colorectal Dis.*, 4(2):76–89, 2002.
[22] T. Roose, S. Jonathan Chapman, and P. K. Maini. Mathematical models of avascular tumor growth. *SIAM Review*, 49(2):179–208, 2007.
[23] M. H. Ross, G. I. Kaye, and W. Pawlina. *Histology : A text and atlas*. Lippincott Williams & Wilkins, Philadelphia, 2003.

equation, whose coefficients are the unknown parameters). In particular, we refer the papers [7, 12], where inverse problems have been formulated and solved for estimating physiologic parameters, by taking medical images as the observable measurements.

The aim of this paper is to show how the birth rate of proliferative cells (understood as a parameter field) could possible be estimated and validated from existing patient-data (for instance the density of proliferative cells), by resolving an inverse problem. More precisely, assuming that the proliferative cell mechanism, in the colonic crypt, is governed by the convection-diffusion system defined in [6], we show that, by resolving an inverse problem (that minimizes the misfit between two proliferative cell densities: one predicted by the model and the other actually been observed), then the birth rate of proliferative cells can be estimated, along all the colonic crypt. The location where the birth rate deviates from normal qualitative values can then be determined, and used for further clinical research and better understanding of the abnormal process leading to the increase or decrease of proliferative cells.

More specifically, the inverse problem, considered in this paper, is a PDE-constrained optimization problem, highly nonlinear, and time-dependent. We describe its formulation, the optimality conditions and explain the corresponding proposed solution algorithm, which is an inexact Newton-type algorithm: a Gauss-Newton-Conjugate-Gradient method (we refer to [19] for a comprehensive discussion about Newton methods and also to [20], where different examples of inverse problems are formulated and solved with Gauss-Newton-Conjugate-Gradient methods).

The layout of this paper is as follows: after this introduction, in section 2, the inverse problem is described and formulated as a PDE-constrained optimization problem. In section 3 a continuous formulation of the inverse problem is presented. In section 4 the numerical approximation and a methodology for its solution is explained and discussed. In section 5 some numerical simulations are reported, for demonstrating and evaluating, the formulation and solution methodology proposed in this paper. Finally some conclusions and future work are discussed in the last section.

2. Parameter field inversion in convection-diffusion transport

For defining the inversion problem proposed in this paper, and also for computational purposes, we use a two-dimensional (2D) version of a colonic crypt, obtained by unwrapping the crypt, which yields a rectangular domain. Thus we denote by $\Omega \subset \mathbb{R}^2$ a rectangular, open and bounded domain, symbolizing the (fixed) spatial domain occupied by the crypt, and $[0,T]$ a time interval (x and t denote hereafter arbitrary points in $\bar{\Omega}$ and $[0,T]$, respectively). The inverse problem is defined as

$$\min_{\alpha} J(\alpha) := \frac{1}{2}\int_0^T \int_\Omega |N(x,t) - N_d(x,t)|^2 \, dxdt + \frac{\gamma_1}{2}\int_\Omega |\alpha|^2 \, dx + \frac{\gamma_2}{2}\int_\Omega |\nabla\alpha|^2 \, dx \qquad (1)$$

where the unknown α is assumed to belong to the space $L^\infty(\Omega)$ (the set of measurable and essentially bounded scalar functions defined in Ω), and N is one component of the solution pair (N,p) of the time-dependent convection-diffusion system (see [6])

$$\begin{aligned}
N_t - \nabla\cdot(\nabla p N) - \nabla\cdot(D\nabla N) - (\alpha-\beta)N = 0 \quad &\text{in} \quad \Omega\times(0,T),\\
-\Delta p - \nabla\cdot(D\nabla N) - \alpha N = 0 \quad &\text{in} \quad \Omega\times(0,T),\\
p = 0 \quad &\text{in} \quad \Gamma_1\times(0,T),\\
\tfrac{\partial p}{\partial n} = 0 \quad &\text{in} \quad \big(\Gamma_2\cup\Gamma_3\cup\Gamma_4\big)\times(0,T),\\
N = 0 \quad &\text{in} \quad \Gamma_1\times(0,T),\\
\tfrac{\partial N}{\partial n} = 0 \quad &\text{in} \quad \big(\Gamma_2\cup\Gamma_3\cup\Gamma_4\big)\times(0,T),\\
N(.,0) = N_0(.) \quad &\text{in} \quad \Omega.
\end{aligned} \qquad (2)$$

In (2) the unknowns N and p represent, respectively, the density of proliferative cells and the pressure exerted by those cells in the crypt wall. The parameters α and β are, respectively, the rate of birth and death of proliferative cells, and D is the diffusion coefficient of these cells. The boundary of the crypt is split into four disjoint parts, Γ_1 (the top), Γ_2 and Γ_3 (the lateral boundaries), and Γ_4 (the bottom). We refer to [6] for a justification of the boundary conditions for N and p. We remark that in [6], $p = 1$ in $\Gamma_1\times(0,T)$, while here, in (2), $p = 0$, therefore for keeping exactly the same boundary conditions for p as in [6], we should consider for the pressure $p+1$ instead of p in (2) (this does not change the solution of (2), since only the gradient

TABLE 2. Perfomance of the GNCG method for Test 2

$GNit$	$CGit$	cost	$misfit$	reg	$\|\hat{\alpha}\|$	$\|grad\|$	$\|\alpha - true\,\alpha\|$
1	1	1.565938e-02	1.565533e-02	3.882688e-06	4.654272e+00	2.389825e-02	1.191270e+00
2	1	1.431946e-03	1.421627e-03	1.031402e-05	8.875193e-01	3.527189e-02	1.069731e+00
3	2	8.562983e-05	7.267626e-05	1.290504e-05	9.308110e-01	3.395806e-03	1.639861e-01
4	1	6.777264e-05	5.518686e-05	1.253286e-05	3.357337e-02	1.113041e-03	1.497473e-01
5	6	5.798690e-05	4.421412e-05	1.369337e-05	1.334136e-01	1.995245e-04	2.127643e-02
6	11	5.794314e-05	4.416346e-05	1.369754e-05	9.737750e-03	1.660610e-05	1.418488e-02
7	15	5.794311e-05	4.416332e-05	1.369767e-05	1.544920e-04	4.442516e-07	1.422052e-02
GN method converged after 7 iterations. Total number of CG iterations: 37 Total number of forward-adjoint solves: 44							

6. Conclusions

In this paper a PDE-constrained optimization problem has been formulated and solved by an inexact Newton method, for estimating the birth rate of proliferative cells in colonic crypts. We have assumed that the proliferative cell mechanism is driven by the convection-diffusion system defined in [6], and the fitting term of the inverse problem, fits the density of proliferative cells data (for instance obtained from medical images or by biopsy analysis) with the density of proliferative cells predicted by the model, by variations of the birth rate parameter. Details on the Newton method have been explained, with emphasis on the definition of the reduced Hessian, for solving the Newton step, by a conjugate-gradient method. The two test simulations described in this paper, with synthetic database, show a good performance of the methodology. Furthermore, a similar procedure could be used for estimating the other parameter fields involved in the convection-diffusion system (2) (as the rate of death or the diffusion of the proliferative cells). Though the simultaneous inversion of several parameter fields is computationally very expensive. In the future, we intend to apply this inverse problem with real medical databases (*i.e.* with densities of proliferative cells in colonic crypts, taken either from one individual or from different groups of individuals).

References

[1] E. B. Brown, Y. Boucher, S. Nasser, and R. K. Jain. Measurement of macromolecular diffusion coefficients in humans tumors. *Microvascular Research*, 67:231–236, 2004.
[2] COMSOL MULTIPHYSICS®. *http://www.comsol.com/*.
[3] D. Drasdo and M. Loeffer. Individual-based models to growth and folding in one-layered tissues: intestinal crypts and early development. *Nonlinear Analysis*, 47:245–256, 2001.

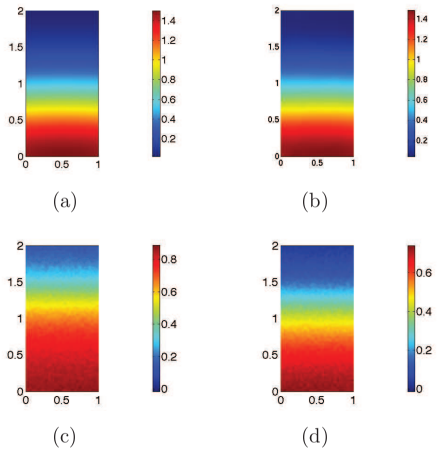


FIGURE 4. Test 2: True birth rate α in (a) and recovered α by the inverse problem in (b). Two instances of the noisy data cell density N_d at time $t = 3$ in (c) and $t = 20$ in (d).

term in the definition of $\mathbf{J}$; reg is the value of the sum of the regularizing terms (second and third terms in the definition of $\mathbf{J}$); $\|\hat{\alpha}\|$ is the L^2-norm of $\hat{\alpha}$ in the Gauss-Newton step; $\|grad\|$ is the L^2-norm of the control equation, *i.e* the derivative of $\mathbf{J}$ with respect to α; $\|\alpha - true\,\alpha\|$ is the L^2-norm of the difference between the true birth rate, $true\,\alpha$, and the birth rate predicted by the model α.

TABLE 1. Perfomance of the GNCG method for Test 1

$GNit$	$CGit$	cost	$misfit$	reg	$\|\hat{\alpha}\|$	$\|grad\|$	$\|\alpha - true\,\alpha\|$
1	1	2.865992e-02	2.865557e-02	4.207597e-06	3.774526e+00	2.022944e-02	1.394284e+00
2	1	5.659012e-04	5.497131e-04	1.620605e-05	1.254839e+00	4.575001e-02	7.864178e-01
3	4	1.539155e-04	1.067717e-04	4.715264e-05	5.498807e-01	1.745195e-03	4.604079e-01
4	5	1.361310e-04	7.783774e-05	5.829999e-05	1.350029e-01	4.893038e-04	3.982325e-01
5	9	1.350829e-04	7.483539e-05	6.025444e-05	2.939835e-02	7.785556e-05	3.854918e-01
6	12	1.350808e-04	7.489755e-05	6.019017e-05	1.577682e-03	3.553042e-06	3.855713e-01
GN method converged after 6 iterations. Total number of CG iterations: 32 Total number of forward-adjoint solves: 38							

and the laplacian of p are involved). The function N_0 is the initial data, *i.e.*, the density of proliferative cells at time $t = 0$.

The objective functional J in (1) involves a data misfit (the first term, with N_d being the data, possibly noisy) and two regularizing terms (γ_1 and γ_2 are regularizing constants). This inverse problem (1)-(2) can be interpreted in the following sense : to find the birth rate of proliferative cells α (considered as a parameter field $\alpha(x)$), from measurements of the cell density N_d (estimated, or validated, for an individual or a group of individuals, over the entire spatio-time domain), by matching the spatio-temporal evolution of the cell density predicted by the model, $N(x,t)$, with the corresponding cell density measurements $N_d(x,t)$.

3. Optimality system and Newton method

In this section we define the optimality system and Newton method in a continuous functional framework. Though this continuous formalism will not be used, fully, in the paper, it is necessary to describe it in order to derive the finite element matrices for the discrete problem (see sections 4 and 5: in this latter section we employ the software COMSOL MULTIPHYSICS® [2] for extracting the finite element matrices, using precisely the weak formulations described in the current section).

We consider the Sobolev space $H^1(\Omega) := \{u :\ u \in L^2(\Omega),\ \partial_i u \in L^2(\Omega),\ i = 1,2\}$ (with $\partial_i u$ the two spatial partial derivatives of u in the weak sense), and the space $W := \{u :\ u \in H^1(\Omega),\ u_{|\Gamma_1} = 0\}$. In the sequel, we denote the L^2- inner product by $(.,)$, *i.e.* for any scalar functions u, v defined on Ω

$$(u,v) := \int_\Omega u(x)v(x)dx.$$

Moreover, we introduce the functional space $U := L^2\big(0,T;W\big)$ involving time, which consists of all strongly measurable functions $z : [0,T] \to W$ with

$$\|z\|_{L^2(0,T;W)} := \Big(\int_0^T \|z(.,t)\|^2_{H^1(\Omega)}dt\Big)^{\frac{1}{2}} < +\infty \tag{3}$$

(see [4] for a detailed definition of Sobolev spaces and functional spaces involving time).

We form the Lagrangian functional $\mathcal{L}$ associated to problem (1), by adding to the objective functional J a duality pairing of the convection-diffusion system (2) with a Lagrange multiplier (M,q) (also known as the adjoint pair

variable).

$$\begin{aligned}\mathcal{L}(\alpha,N,p,M,q) \;:=\; J(\alpha)+\int_0^T \Big((N_t,M)+(\nabla pN,\nabla M)+(D\nabla N,\nabla M)-\big((\alpha-\beta)N,M\big)\\ +(\nabla p,\nabla q)+(D\nabla N,\nabla q)-(\alpha N,q)\Big)dt,\end{aligned} \tag{4}$$

where $\alpha \in L^\infty(\Omega)$, $(N,p) \in U \times U$ and $(M,q) \in U' \times U'$ (with U' the dual of U) . Then, the first-order optimality system is

$$\mathcal{L}'(\alpha,N,p,M,q)(\tilde\alpha,\tilde N,\tilde p,\tilde M,\tilde q)=0, \tag{5}$$

with $\mathcal{L}'$ the first derivative of $\mathcal{L}$, and for arbitrary functions $(\tilde\alpha,\tilde N,\tilde p,\tilde M,\tilde q)$ in $L^\infty(\Omega)\times U\times U\times U'\times U'$, with $\tilde N(.,0)=0$.

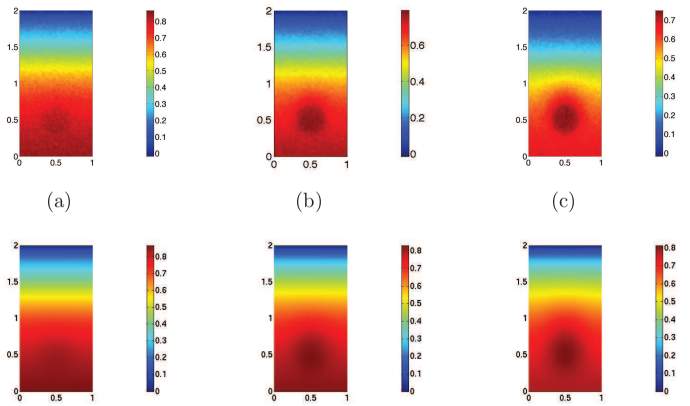
The variation of the Lagrangian functional with respect to (M,q) leads to the weak formulation of the state system (2) (also called forward system). The variation with respect to α, in the direction of $\tilde\alpha$, is the control equation

$$\mathcal{L}_\alpha(\alpha,N,p,M,q)(\tilde\alpha) := \gamma_1(\alpha,\tilde\alpha)+\gamma_2(\nabla\alpha,\nabla\tilde\alpha)-\int_0^T\Big((N,M\tilde\alpha)+(N,q\tilde\alpha)\Big)dt=0, \tag{6}$$

and the variation with respect to N and p, in the direction of $\tilde N$ and $\tilde p$, respectively, yields the adjoint system

$$\begin{aligned}\mathcal{L}_N(\alpha,N,p,M,q)(\tilde N) \;:=\; &\int_0^T\Big((N-N_d,\tilde N)+(\tilde N_t,M)+(\nabla p\tilde N,\nabla M)+(D\nabla\tilde N,\nabla M)\\ &-\big((\alpha-\beta)\tilde N,M\big)+(D\nabla\tilde N,\nabla q)-(\alpha\tilde N,q)\Big)dt=0,\\ \mathcal{L}_p(\alpha,N,p,M,q)(\tilde p) := &\int_0^T\Big((\nabla\tilde pN,\nabla M)+(\nabla\tilde p,\nabla q)\Big)dt=0.\end{aligned} \tag{7}$$

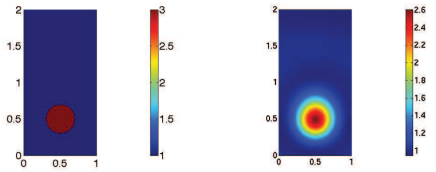
The strong form of this adjoint system is



(a) (b) (c)

(d) (e) (f)

FIGURE 2. Test 1: Noisy data cell density N_d (top row) and cell density N predicted by the model (bottom row), for time $t=4$ in (a)-(d), $t=10$ in (b)-(e) , and at $t=19$ in (c)-(f).

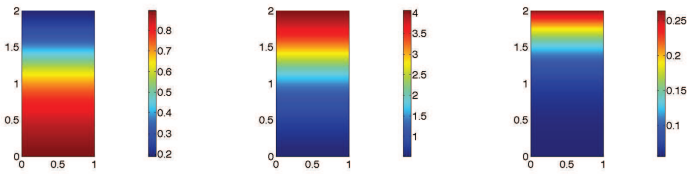


(a) (b)

FIGURE 3. Test 1: True birth rate α in (a) and recovered α by the inverse problem in (b).

For both tests we have checked the convergence and accuracy of the algorithm, by considering finer meshes, with more finite elements, and decreasing the tolerance.

The Tables 1 and 2 illustrate the performance of the GNGC algorithm for the two tests: *GNit* stands for the Gauss-Netwon iteration; *CGit* represents the number of iterations in the Conjugate-Gradient algorithm; *cost* is the value of the objective functional $\mathbf{J}$; $misfit$ represents the value of the first



(a) (b) (c)

FIGURE 1. Common data for both tests. (a) Initial cell density $\tilde{N}_0$. (b) Rate of death cells $\tilde{\beta}$. (c) Non constant diffusion $\tilde{D}_2$.

The graphics of the data parameters $\tilde{N}_0$, $\tilde{\beta}$ and $\tilde{D}_2$ are shown in Figure 1. Hereafter we omit the symbol tilde for the spatio-temporal synthetic domain, variables, unknowns and parameters, in order to simplify the notations.

For starting the GNCG method, the initial proliferative rate α_0 is 3.0, for all the two test problems. The stopping criterion for the Gauss-Newton algorithm is either the L^2-norm of the control equation (*i.e*, the L^2-norm of the derivative of J with respect to α), or the L^2-norm of $\hat{\alpha}$ in the Gauss-Newton step, less than a prescribed termination tolerance, set to 5×10^{-6} in all the tests. The implementation is done in MATLAB® [26]; we use COMSOL MULTIPHYSICS® [2] only for extracting the finite element matrices (for that we follow the same type of implementation described in [20]).

5.3. Tests 1 and 2. For the first test we suppose the diffusion coefficient is defined by D_2, and we generate a data cell density N_d, by solving the forward problem (2) for a birth rate α exhibiting high values in a circular region at the bottom of the crypt, as illustrated in Figure 3 (a). We add noise to this synthetic data N_d. The Figure 2 shows three time instances of the cell density N predicted by the inverse problem, as well as the noisy data cell density N_d, for the same time instances. The solution of the inverse problem is depicted in Figure 3 (b).

In the second test we suppose the diffusion coefficient is constant and defined by D_1. As before, we generate a data cell density N_d, by solving the forward problem (2) for a birth rate α as illustrated in Figure 4 (a). We add noise to this synthetic data N_d. The recovered α, which is the solution of the inverse problem, is depicted in Figure 4 (b). The Figures 4 (c) and (d) display two time instances of the noisy data cell density N_d.

$$
\begin{aligned}
-M_t + \nabla p \cdot \nabla M - \nabla \cdot (D \nabla M) - (\alpha - \beta) M &= \nabla \cdot (D \nabla q) \\
&\quad -(N - N_d) + \alpha q \quad \text{in} \quad \Omega \times (0,T), \\
-\Delta q - \nabla \cdot (N \nabla M) &= 0 \quad \text{in} \quad \Omega \times (0,T), \\
M &= 0 \quad \text{in} \quad \Gamma_1 \times (0,T), \\
\big(D \nabla (M+q)\big) \cdot n &= 0 \quad \text{in} \quad \big(\Gamma_2 \cup \Gamma_3 \cup \Gamma_4\big) \times (0,T), \\
q &= 0 \quad \text{in} \quad \Gamma_1 \times (0,T), \\
(N \nabla M + \nabla q) \cdot n &= 0 \quad \text{in} \quad \big(\Gamma_2 \cup \Gamma_3 \cup \Gamma_4\big) \times (0,T), \\
M(.,T) &= 0 \quad \text{in} \quad \Omega.
\end{aligned}
\tag{8}
$$

For effectively solving the optimality system (5), the Newton's method can be used. It computes an update direction $(\hat{\alpha}, \hat{N}, \hat{p}, \hat{M}, \hat{q})$ from the following Newton step, for the Lagrangian functional

$$
\mathcal{L}''(\alpha, N, p, M, q)[(\hat{\alpha}, \hat{N}, \hat{p}, \hat{M}, \hat{q}), (\tilde{\alpha}, \tilde{N}, \tilde{p}, \tilde{M}, \tilde{q})] = -\mathcal{L}'(\alpha, N, p, M, q)(\tilde{\alpha}, \tilde{N}, \tilde{p}, \tilde{M}, \tilde{q}), \tag{9}
$$

for all variations $(\tilde{\alpha}, \tilde{N}, \tilde{p}, \tilde{M}, \tilde{q})$, with $\mathcal{L}''$ the second derivative of $\mathcal{L}$. We remark that (9) is equivalent to the following system involving the second partial derivatives of $\mathcal{L}$

$$
\begin{aligned}
\mathcal{L}_{NN}(.)(\hat{N},\tilde{N}) + \mathcal{L}_{Np}(.)(\hat{p},\tilde{N}) + \mathcal{L}_{N\alpha}(.)(\hat{\alpha},\tilde{N}) + \mathcal{L}_{NM}(.)(\hat{M},\tilde{N}) + \mathcal{L}_{Nq}(.)(\hat{q},\tilde{N}) &= -\mathcal{L}_N(.)(\tilde{N}) \\
\mathcal{L}_{pN}(.)(\hat{N},\tilde{p}) + \mathcal{L}_{pp}(.)(\hat{p},\tilde{p}) + \mathcal{L}_{p\alpha}(.)(\hat{\alpha},\tilde{p}) + \mathcal{L}_{pM}(.)(\hat{M},\tilde{p}) + \mathcal{L}_{pq}(.)(\hat{q},\tilde{p}) &= -\mathcal{L}_p(.)(\tilde{p}) \\
\mathcal{L}_{\alpha N}(.)(\hat{N},\tilde{\alpha}) + \mathcal{L}_{\alpha p}(.)(\hat{p},\tilde{\alpha}) + \mathcal{L}_{\alpha\alpha}(.)(\hat{\alpha},\tilde{\alpha}) + \mathcal{L}_{\alpha M}(.)(\hat{M},\tilde{\alpha}) + \mathcal{L}_{\alpha q}(.)(\hat{q},\tilde{\alpha}) &= -\mathcal{L}_\alpha(.)(\tilde{\alpha}) \\
\mathcal{L}_{MN}(.)(\hat{N},\tilde{M}) + \mathcal{L}_{Mp}(.)(\hat{p},\tilde{M}) + \mathcal{L}_{M\alpha}(.)(\hat{\alpha},\tilde{M}) + \mathcal{L}_{MM}(.)(\hat{M},\tilde{M}) + \mathcal{L}_{Mq}(.)(\hat{q},\tilde{M}) &= -\mathcal{L}_M(.)(\tilde{M}) \\
\mathcal{L}_{qN}(.)(\hat{N},\tilde{q}) + \mathcal{L}_{qp}(.)(\hat{p},\tilde{q}) + \mathcal{L}_{q\alpha}(.)(\hat{\alpha},\tilde{q}) + \mathcal{L}_{qM}(.)(\hat{M},\tilde{q}) + \mathcal{L}_{qq}(.)(\hat{q},\tilde{q}) &= -\mathcal{L}_q(.)(\tilde{q})
\end{aligned}
\tag{10}
$$

where we have omitted the argument (α, N, p, M, q) (and replaced it by the notation $(.)$), for simplifying the notations. The explicit formulas for the left-hand sides in (10) are

$$\int_0^T \Big[(\hat{N},\tilde{N}) + (\nabla\hat{p},\nabla M\tilde{N}) - (\hat{\alpha},M\tilde{N}) - (\hat{\alpha},q\tilde{N}) - (\hat{M}_t + (\alpha-\beta)\hat{M},\tilde{N}) + (\nabla p,\nabla\hat{M}\tilde{N}) + (D\nabla\hat{M},\nabla\tilde{N}) - (\alpha\hat{q},\tilde{N}) + (D\nabla\hat{q},\nabla\tilde{N})\Big]dt + (\tilde{N}(.,T),\tilde{M}(.,T)) = -\mathcal{L}_N(.)(\tilde{N})$$

$$\int_0^T \Big[(\hat{N}\nabla M,\nabla\tilde{p}) + (N\nabla\hat{M},\nabla\tilde{p}) + (\nabla\hat{q},\nabla\tilde{p})\Big]dt = -\mathcal{L}_p(.)(\tilde{p})$$

$$\int_0^T \Big[-(\hat{N},M\tilde{\alpha}) - (\hat{N},q\tilde{\alpha}) - (\hat{M},N\tilde{\alpha}) - (\hat{q},N\tilde{\alpha})\Big]dt + \gamma_1(\hat{\alpha},\tilde{\alpha}) + \gamma_2(\nabla\hat{\alpha},\nabla\tilde{\alpha}) = -\mathcal{L}_\alpha(.)(\tilde{\alpha})$$

$$\int_0^T \Big[(\hat{N}_t,\tilde{M}) - \big((\alpha-\beta)\hat{N},\tilde{M}\big) + (\nabla p,\hat{N}\nabla\tilde{M}) + (D\nabla\hat{N},\nabla\tilde{M}) + (\nabla\hat{p},N\nabla\tilde{M}) - (\hat{\alpha},N\tilde{M})\Big]dt = -\mathcal{L}_M(.)(\tilde{M})$$

$$\int_0^T \Big[(D\nabla\hat{N},\nabla\tilde{q}) - (\alpha\hat{N},\tilde{q}) + (\nabla\hat{p},\nabla\tilde{q}) - (\hat{\alpha},N\tilde{q})\Big]dt = -\mathcal{L}_q(.)(\tilde{q}) \tag{11}$$

4. The discrete problem

We discretize now the inverse problem and compute the corresponding optimality conditions. The space-time discretization is based on continuous finite elements, for the space variable, and finite differences (using an implicit backward Euler scheme), for the time variable.

4.1. The discrete cost functional. Let $\mathbf{J}$ be the discretized cost functional in (1)

$$\min_\alpha \mathbf{J}(\alpha) := \frac{1}{2}(\bar{\mathbf{N}} - \bar{\mathbf{N}}_d)^T\bar{\mathbf{Q}}(\bar{\mathbf{N}} - \bar{\mathbf{N}}_d) + \frac{\gamma_1}{2}\alpha^T\mathbf{K}\alpha + \frac{\gamma_2}{2}\alpha^T\mathbf{R}\alpha \tag{12}$$

where α represents the finite element discretization of the birth rate, $\bar{\mathbf{N}} = [\mathbf{N}^1\,\mathbf{N}^2\ldots\mathbf{N}^s]$ is the space-time discretization of N (s is the total number of time steps, and $\mathbf{N}^i$ stands for the spatial degrees of freedom of the $i-$th time step). Furthermore, $\bar{\mathbf{N}}_d$ is the discretized space-time measurement data, $\mathbf{K}$ and $\mathbf{R}$ are the mass and stiffness (finite element) matrices, respectively, and $\bar{\mathbf{Q}}$ is the discretized space-time operator, which is a block diagonal matrix,

$(\tilde{N},\tilde{p},\tilde{D},\tilde{\alpha},\tilde{\beta})$ and (N,p,D,α,β), is given by

$$\tilde{N} := N, \quad \tilde{p} := \frac{120}{360^2}p, \quad \tilde{D} := \frac{120}{360^2}D \quad \tilde{\alpha} := 120\alpha, \quad \tilde{\beta} := 120\beta. \tag{39}$$

5.2. Data for the test cases. For the two test cases we use the following: 20×40 finite elements for discretizing the spatial domain $\tilde{\Omega} = (0,1)\times(0,2)$; the time step size $\Delta\tilde{t} = \frac{1}{20} = 0.05$; the regularizing parameters γ_1 and γ_2 are set equal to 10^{-5}; the termination tolerance tol is equal 5×10^{-6}.

It is known, that in normal colonic crypts, the proliferative cells are essentially located in the lower two-thirds of the crypt, with a strong activity at the bottom of the crypt, while the fully differentiated and apoptotic cells are located in the upper third part of the crypt (see [8, **?**, 24]). Accordingly, we choose for the initial condition $\tilde{N}_0$ a decreasing function of the height of the crypt, that must be zero in the upper third top part

$$\tilde{N}_0(\tilde{x},\tilde{y}) := 0.5(1+\frac{2}{\pi})\arctan\big(\frac{4/3-\tilde{y}}{0.45}\big), \qquad \forall(\tilde{x},\tilde{y})\in(0,1)\times(0,2). \tag{40}$$

Also because the proliferative cells are essentially located in the lower two-thirds of the crypt, then in normal colonic crypts, the birth rate of proliferative cells should be essentially a decreasing function of the height of the crypt, that must be zero in the upper third top part. In test 2 we adopt this definition for the birth rate (for generating the synthetic data), see Figure 4 (a), as opposed to test 1, where we assume an abnormal birth rate, see Figure 3(a). For the rate of death of the proliferative cell density, we just adopt, in both tests, the reverse definition of a normal birth rate. The rate of death $\tilde{\beta}$ is then prescribed as

$$\beta(\tilde{x},\tilde{y}) := 2.5(1-\frac{2}{\pi})\arctan\big(\frac{4/3-\tilde{y}}{0.45}\big), \qquad \forall(\tilde{x},\tilde{y})\in(0,1)\times(0,2) \tag{41}$$

For the diffusion coefficient $\tilde{D}$ we consider two possible choices: either a constant $\tilde{D}_1$ or a function $\tilde{D}_2$, with

$$\tilde{D}_1 := 10^{-2}, \qquad \tilde{D}_2(\tilde{x},\tilde{y}) := \frac{0.05}{(1+\frac{2}{\pi})\arctan\big(\frac{4/3-\tilde{y}}{0.45}\big)}, \quad \forall(\tilde{x},\tilde{y})\in(0,1)\times(0,2). \tag{42}$$

The second choice $\tilde{D}_2$ relies on [17, 18], where a one-dimensional model is proposed with a diffusion coefficient that has an inverse quadratic dependence on cell number density, so it should be bigger at the top of the crypt.

Remark 4.2. *For avoiding the possibility of having α negative (the birth rate of proliferative cells verifies $\alpha \geq 0$), we can extend the method to incorporate the bound $\alpha \geq 0$, by adding a logarithmic barrier term with barrier parameter μ to the cost functional (i.e. $\mu \ln \alpha$). Then, in the control equation* (26) *the term $\frac{-\mu}{\alpha}$ (interpreted as a finite element vector) should be added, and in the Hessian matrix* (28) *a diagonal matrix $\mathbf{Z}$ with components $\frac{\mu}{\alpha^2}$ should be added to $\gamma_1 \mathbf{K} + \gamma_2 \mathbf{R}$.*

5. Numerical tests

Here we report the results obtained for two test-cases, we have simulated, for analyzing the performance of our formulation and methodology. But firstly we reformulate the inverse problem in a synthetic domain, and define the common values for the data in both tests (which are solved in the synthetic domain).

5.1. Nondimensionlization. The average number of cells in a normal human colonic crypt is 120 cells in height (from bottom to top) and 60 cells in perimeter. Since the size of each cell is about 6 to 10 μm [3] (where $\mu m = 10^{-6} m$) we define $\Omega = (0, 360) \times (0, 720)$, and consider the unit length to be $1 \mu m$. Regarding the time interval $[0, T]$, it is measured in hours, and $T = 120$, by assuming that 5 days is the average time for a cell to reach the top of the crypt, given an initial starting position.

In order to perform the numerical tests we nondimensionalize problem (2) as follows. By doing the change of coordinates

$$\begin{array}{rcl} [0,360]\times[0,720] & \to & [0,1]\times[0,2] \\ (x,y) & \to & (\tilde{x}=\frac{x}{360}, \tilde{y}=\frac{y}{360}) \end{array} \quad \text{and} \quad \begin{array}{rcl} [0,120] & \to & [0,1] \\ t & \to & \tilde{t}=\frac{t}{120} \end{array} \tag{37}$$

the two first equations in (2) become

$$\begin{aligned} N_{\tilde{t}} - \tfrac{120}{360^2}\nabla\cdot(\nabla p N) - \tfrac{120}{360^2}\nabla\cdot(D\nabla N) - 120(\alpha-\beta)N = 0, \\ -\tfrac{120}{360^2}\Delta p - \tfrac{120}{360^2}\nabla\cdot(D\nabla N) - 120\alpha N = 0. \end{aligned} \tag{38}$$

Thus, by solving the inverse problem (1)-(2) in the nondimensionalized domain $\tilde{\Omega} \times (0, \tilde{T})$, with $\tilde{\Omega} = (0,1) \times (0,2)$ and $(0, \tilde{T}) = (0,1)$, the correspondence between the new and original unknowns and parameters, respectively

i.e.

$$\bar{\mathbf{Q}} = \begin{bmatrix} \frac{\Delta t}{2}\mathbf{K} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Delta t\mathbf{K} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \Delta t\mathbf{K} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \frac{\Delta t}{2}\mathbf{K} \end{bmatrix} \tag{13}$$

with Δt the time step size.

4.2. The discrete forward problem. Denoting by $\mathbf{N}_0$ the discretized initial data N_0 and by $\bar{\mathbf{p}} = [\mathbf{p}^1\, \mathbf{p}^2 \dots \mathbf{p}^s]$ the space-time discretization of the pressure p, then the discretized forward system (2) is defined by

$$\begin{aligned} \mathbf{N}^1 &= \mathbf{N}_0 \\ \mathbf{R}\mathbf{p}^i + (\mathbf{R}_D - \mathbf{K}_\alpha)\mathbf{N}^i &= \mathbf{0}, \quad i = 1, \dots, s \\ \mathbf{K}\frac{\mathbf{N}^{i+1} - \mathbf{N}^i}{\Delta t} + \big(\mathbf{C}(\mathbf{p}^i) + \mathbf{R}_D - \mathbf{K}_{\alpha-\beta}\big)\mathbf{N}^{i+1} &= \mathbf{0}, \quad i = 1, \dots, s-1 \end{aligned} \tag{14}$$

where $\mathbf{K}_\alpha$, $\mathbf{K}_{\alpha-\beta}$ and $\mathbf{R}_D$ are modified mass and stiffness matrices, depending on α, $\alpha - \beta$ and D, respectively, and $\mathbf{C}(\mathbf{p}^i)$ is also a finite element matrix, depending on the pressure $\mathbf{p}^i$. For convenience we introduce the following matrices,

$$\mathbf{A} = \mathbf{R}_D - \mathbf{K}_\alpha, \quad \mathbf{L} = \mathbf{K} + \Delta t(\mathbf{R}_D - \mathbf{K}_{\alpha-\beta}), \quad \mathbf{C}^i = \mathbf{C}(\mathbf{p}^i) \tag{15}$$

where we remark that $\mathbf{A}$ depends on D and α, and $\mathbf{L}$ depends on D, α and β. Then the discrete state system (14) can be written in matrix form as

$$\bar{\mathbf{S}}\bar{\mathbf{X}} = \bar{\mathbf{F}}, \tag{16}$$

where the discrete forward operator $\bar{\mathbf{S}}$ is defined by

$$\bar{\mathbf{S}} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{A} & \mathbf{R} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -\mathbf{K} & \mathbf{0} & \mathbf{L}+\Delta t\mathbf{C}^1 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A} & \mathbf{R} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -\mathbf{K} & \mathbf{0} & \mathbf{L}+\Delta t\mathbf{C}^2 & \dots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{L}+\Delta t\mathbf{C}^{s-2} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{A} & \mathbf{R} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & -\mathbf{K} & \mathbf{0} & \mathbf{L}+\Delta t\mathbf{C}^{s-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \mathbf{A} & \mathbf{R} \end{bmatrix} \tag{17}$$

and

$$\bar{\mathbf{X}} = \begin{bmatrix} \mathbf{X}^1 \\ \mathbf{X}^2 \\ \vdots \\ \mathbf{X}^s \end{bmatrix} = \begin{bmatrix} \begin{pmatrix} \mathbf{N}^1 \\ \mathbf{p}^1 \end{pmatrix} \\ \begin{pmatrix} \mathbf{N}^2 \\ \mathbf{p}^2 \end{pmatrix} \\ \vdots \\ \begin{pmatrix} \mathbf{N}^s \\ \mathbf{p}^s \end{pmatrix} \end{bmatrix}, \qquad \bar{\mathbf{F}} = \begin{bmatrix} \mathbf{N}_0 \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}. \tag{18}$$

4.3. The discrete Lagrangian. By setting

$$\bar{\mathbf{Y}} = \begin{bmatrix} \mathbf{Y}^1 \\ \mathbf{Y}^2 \\ \vdots \\ \mathbf{Y}^s \end{bmatrix} = \begin{bmatrix} \begin{pmatrix} \mathbf{M}^1 \\ \mathbf{q}^1 \end{pmatrix} \\ \begin{pmatrix} \mathbf{M}^2 \\ \mathbf{q}^2 \end{pmatrix} \\ \vdots \\ \begin{pmatrix} \mathbf{M}^s \\ \mathbf{q}^s \end{pmatrix} \end{bmatrix} \tag{19}$$

then the discrete Lagrangian is defined by (compare with (4))

$$\mathcal{L}(\alpha, \bar{\mathbf{X}}, \bar{\mathbf{Y}}) := \mathbf{J}(\alpha) + \mathbf{Y}^T(\bar{\mathbf{S}}\bar{\mathbf{X}} - \bar{\mathbf{F}}). \tag{20}$$

Hereafter we also use the notations $\bar{\mathbf{M}} = [\mathbf{M}^1\,\mathbf{M}^2 \ldots \mathbf{M}^s]$, $\bar{\mathbf{q}} = [\mathbf{q}^1\,\mathbf{q}^2 \ldots \mathbf{q}^s]$, when referring to the adjoint variables $(\bar{\mathbf{M}}, \bar{\mathbf{q}})$. Introducing the following block-matrices (each of which with $s \times s$ blocks)

$$\bar{\mathbf{A}} = \begin{bmatrix} \mathbf{A} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{A} & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A} \end{bmatrix} \qquad \bar{\mathbf{R}} = \begin{bmatrix} \mathbf{R} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{R} & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{R} \end{bmatrix},$$

$$\bar{\mathbf{B}} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} \\ -\mathbf{K} & \mathbf{L} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{K} & \mathbf{L} & \ldots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & -\mathbf{K} & \mathbf{L} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -\mathbf{K} & \mathbf{L} \end{bmatrix}, \quad \bar{\mathbf{C}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \Delta t\mathbf{C}^1 & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \Delta t\mathbf{C}^{s-1} \end{bmatrix}. \tag{21}$$

Therefore, the computation of the Newton direction $\hat{\alpha}$, in (31), using the approximate reduced Hessian, and by the conjugate gradient method reduces to computing the vector

$$\begin{aligned} \mathbf{H}\hat{\alpha} &:= (\gamma_1\mathbf{K} + \gamma_2\mathbf{R})\hat{\alpha} + \Big(\frac{\partial \bar{\mathbf{B}}}{\partial \alpha}\bar{\mathbf{N}}\Big)^T \hat{\bar{\mathbf{M}}} + \Big(\frac{\partial \bar{\mathbf{A}}}{\partial \alpha}\bar{\mathbf{N}}\Big)^T \hat{\bar{\mathbf{q}}} \\ &:= (\gamma_1\mathbf{K} + \gamma_2\mathbf{R})\hat{\alpha} - \Delta t \sum_{i=2}^{s} \Big(\frac{\partial \mathbf{K}_{\alpha-\beta}}{\partial \alpha}\mathbf{N}^i\Big)^T \hat{\mathbf{M}}^i - \sum_{i=1}^{s} \Big(\frac{\partial \mathbf{K}_{\alpha}}{\partial \alpha}\mathbf{N}^i\Big)^T \hat{\mathbf{q}}^i. \end{aligned} \tag{36}$$

The terms involving the derivatives, with respect to the pressure and the birth rate, on (34)-(35)-(36), are nonlinear terms, which are obtained directly from the derivatives of the continuous Lagrangian, as explained in (24).

We conclude this section by summarizing the main steps of the Gauss-Newton-Conjugate-Gradient method (GNCG method), we propose, for solving the inverse problem:

Algorithm

Step 1. Set $k = 0$ and initialize with α_0.

Step 2. For $k \geq 0$, and with α_k, solve the state and adjoint systems (14) and (25), respectively, and get $(\bar{\mathbf{N}}_k, \bar{\mathbf{p}}_k)$, $(\bar{\mathbf{M}}_k, \bar{\mathbf{q}}_k)$. Then solve the system

$$\mathbf{H}\hat{\alpha}_k = -\mathcal{L}_\alpha(\alpha_k, \bar{\mathbf{N}}_k, \bar{\mathbf{p}}_k, \bar{\mathbf{M}}_k, \bar{\mathbf{q}}_k)$$

by the conjugate gradient method, where the vector $\mathbf{H}\hat{\alpha}_k$ is defined by the formula (36) (the reduced Hessian $\mathbf{H}$ is defined in (33)). Update the birth rate of proliferative cells by

$$\alpha_{k+1} = \alpha_k + \theta_k\hat{\alpha}_k$$

where θ_k is a suitable chosen step length, such that the cost functional is sufficiently decrease at α_{k+1} (this can be achieved by choosing a θ_k that satisfies the Armijo condition or the Wolfe condition [19]).

Step 3. Terminate when either the norm of the gradient of $\mathbf{J}$ with respect to α at α_k, or, the norm of $\hat{\alpha}_k$ in the Gauss-Newton step is sufficiently small (*i.e*, smaller than a prescribed tolerance *tol*).

Remark 4.1. *We observe that in Step 2, the procedure for choosing the step length in the line search, involves the computation of the cost functional, which means a solution for the state system (which is nonlinear) must be computed for each trial* α_{k+1}. *This is computationally costly, since new finite element matrices should be built, see* (14).

where $\mathbf{H}$ is the reduced Hessian, defined by

$$\mathbf{H} := \mathbf{H}_{22} + \mathbf{H}_{32}^T\,\mathbf{H}_{31}^{-T}\Big(\mathbf{H}_{11}\,\mathbf{H}_{31}^{-1}\,\mathbf{H}_{32} - \mathbf{H}_{12}\Big) - \mathbf{H}_{21}\,\mathbf{H}_{31}^{-1}\,\mathbf{H}_{32}. \tag{32}$$

Remark that this reduced Hessian involves the inverse of the state operator and the inverse of its adjoint ($\mathbf{H}_{31}$ and $\mathbf{H}_{13} = \mathbf{H}_{31}^T$). The reduced Hessian system (31) can be solved by an iterative method, as for instance the conjugate gradient method. We observe as well that the Newton direction $\hat{\alpha}$ (solution of (31)) is a descent direction for the inverse problem (12), if the reduced Hessian (see (32)), or some appropriate approximation of it, is positive definite. Therefore, in order to guarantee the positive definiteness of the reduced Hessian we neglect the block matrices $\mathbf{H}_{12}$, $\mathbf{H}_{21}$, and also the sub-block matrices $\mathcal{L}_{\bar{\mathbf{N}}\bar{\mathbf{p}}}$ and $\mathcal{L}_{\bar{\mathbf{p}}\bar{\mathbf{N}}}$ in $\mathbf{H}_{11}$. This leads to the following Gauss-Newton approximation of the Hessian in (32), which is always positive definite

$$\mathbf{H} := \mathbf{H}_{22} + \mathbf{H}_{32}^T\,\mathbf{H}_{31}^{-T}\tilde{\mathbf{H}}_{11}\,\mathbf{H}_{31}^{-1}\,\mathbf{H}_{32}, \qquad \text{with} \quad \tilde{\mathbf{H}}_{11} := \begin{pmatrix} \bar{\mathbf{Q}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}. \tag{33}$$

Then, returning to (29) and (30) (with $\mathbf{H}_{21} = \mathbf{H}_{12} = 0$ and $\tilde{\mathbf{H}}_{11}$ instead of $\mathbf{H}_{11}$, in the definition of the Hessian of
$\mathcal{L}$), the solution $(\hat{\bar{\mathbf{N}}}, \hat{\bar{\mathbf{p}}})$ of system (29) is defined by

$$\begin{aligned}
\hat{\mathbf{N}}^1 &= \mathbf{0}, \\
-\mathbf{K}\hat{\mathbf{N}}^{i-1} + \Big(\mathbf{L} + \Delta t\mathbf{C}^{i-1}\Big)\hat{\mathbf{N}}^i + \Delta t\Big(\frac{\partial \mathbf{C}^{i-1}}{\partial \mathbf{p}^{i-1}}\mathbf{N}^i\Big)\hat{\mathbf{p}}^i - \Delta t\Big(\frac{\partial \mathbf{K}_{\alpha-\beta}}{\partial \alpha}\mathbf{N}^i\Big)\hat{\alpha} &= \mathbf{0}, \quad \forall i = 2, \ldots, s \\
\mathbf{R}\hat{\mathbf{p}}^i + \mathbf{A}\hat{\mathbf{N}}^i - \Big(\frac{\partial \mathbf{K}_\alpha}{\partial \alpha}\mathbf{N}^i\Big)\hat{\alpha} &= \mathbf{0}, \quad \forall i = 1, \ldots, s
\end{aligned} \tag{34}$$

and the solution $(\hat{\bar{\mathbf{M}}}, \hat{\bar{\mathbf{q}}})$ of system (30) is defined by

$$\begin{aligned}
\mathbf{R}^T\hat{\mathbf{q}}^s &= \mathbf{0}, \\
\mathbf{R}^T\hat{\mathbf{q}}^i + \Delta t\Big(\frac{\partial \mathbf{C}^i}{\partial \mathbf{p}^i}\mathbf{N}^{i+1}\Big)^T\hat{\mathbf{M}}^{i+1} &= \mathbf{0}, \quad \forall i = s-1, \ldots, 1 \\
\tfrac{\Delta t}{2}\mathbf{K}\hat{\mathbf{N}}^s + (\mathbf{L} + \Delta t\mathbf{C}^{s-1})^T\hat{\mathbf{M}}^s + \mathbf{A}^T\hat{\mathbf{q}}^s &= \mathbf{0}, \\
\Delta t\mathbf{K}\hat{\mathbf{N}}^i + (\mathbf{L} + \Delta t\mathbf{C}^{i-1})^T\hat{\mathbf{M}}^i - \mathbf{K}\hat{\mathbf{M}}^{i+1} + \mathbf{A}^T\hat{\mathbf{q}}^i &= \mathbf{0}, \quad \forall i = s-1, \ldots, 2 \\
\tfrac{\Delta t}{2}\mathbf{K}\hat{\mathbf{N}}^1 + \hat{\mathbf{M}}^1 - \mathbf{K}\hat{\mathbf{M}}^2 + \mathbf{A}^T\hat{\mathbf{q}}^1 &= \mathbf{0}.
\end{aligned} \tag{35}$$

then, the discrete Lagrangian (20) is equivalently defined by

$$\begin{aligned}
\mathcal{L}(\alpha, \bar{\mathbf{N}}, \bar{\mathbf{p}}, \bar{\mathbf{M}}, \bar{\mathbf{q}}) := \frac{1}{2}(\bar{\mathbf{N}} - \bar{\mathbf{N}}_d)^T\bar{\mathbf{Q}}(\bar{\mathbf{N}} - \bar{\mathbf{N}}_d) + \frac{\gamma_1}{2}\alpha^T\mathbf{K}\alpha + \frac{\gamma_2}{2}\alpha^T\mathbf{R}\alpha \\
+\bar{\mathbf{q}}^T(\bar{\mathbf{A}}\bar{\mathbf{N}} + \bar{\mathbf{R}}\bar{\mathbf{p}}) + \bar{\mathbf{M}}^T(\bar{\mathbf{B}} + \bar{\mathbf{C}})\bar{\mathbf{N}} - (\mathbf{M}^1)^T\mathbf{N_0}.
\end{aligned} \tag{22}$$

4.4. The discrete adjoint problem. The gradient of the discrete Lagrangian (22) is

$$\begin{aligned}
\mathcal{L}_{\bar{\mathbf{N}}}(\alpha, \bar{\mathbf{N}}, \bar{\mathbf{p}}, \bar{\mathbf{M}}, \bar{\mathbf{q}}) &:= \bar{\mathbf{Q}}(\bar{\mathbf{N}} - \bar{\mathbf{N}}_d) + \bar{\mathbf{q}}^T\bar{\mathbf{A}} + \bar{\mathbf{M}}^T(\bar{\mathbf{B}} + \bar{\mathbf{C}}) \\
\mathcal{L}_{\bar{\mathbf{p}}}(\alpha, \bar{\mathbf{N}}, \bar{\mathbf{p}}, \bar{\mathbf{M}}, \bar{\mathbf{q}}) &:= \bar{\mathbf{q}}^T\bar{\mathbf{R}} + \bar{\mathbf{M}}^T\frac{\partial \bar{\mathbf{C}}}{\partial \bar{\mathbf{p}}}\bar{\mathbf{N}} \\
\mathcal{L}_{\alpha}(\alpha, \bar{\mathbf{N}}, \bar{\mathbf{p}}, \bar{\mathbf{M}}, \bar{\mathbf{q}}) &:= \big(\gamma_1\mathbf{K} + \gamma_2\mathbf{R}\big)\alpha + \bar{\mathbf{q}}^T\frac{\partial \bar{\mathbf{A}}}{\partial \alpha}\bar{\mathbf{N}} + \bar{\mathbf{M}}^T\frac{\partial \bar{\mathbf{B}}}{\partial \alpha}\bar{\mathbf{N}} \\
\mathcal{L}_{\bar{\mathbf{M}}}(\alpha, \bar{\mathbf{N}}, \bar{\mathbf{p}}, \bar{\mathbf{M}}, \bar{\mathbf{q}}) &:= \bar{\mathbf{N}}^T(\bar{\mathbf{B}} + \bar{\mathbf{C}})^T - \bar{\mathbf{N}}_0^T\bar{\mathbf{W}} \\
\mathcal{L}_{\bar{\mathbf{q}}}(\alpha, \bar{\mathbf{N}}, \bar{\mathbf{p}}, \bar{\mathbf{M}}, \bar{\mathbf{q}}) &:= \bar{\mathbf{N}}^T\bar{\mathbf{A}}^T + \bar{\mathbf{p}}^T\bar{\mathbf{R}}^T
\end{aligned} \tag{23}$$

where $\bar{\mathbf{W}}$ is a block diagonal matrix, with diagonal blocks equal to zero, except the first which is the identity. In particular, we observe that the nonlinear terms $\bar{\mathbf{M}}^T\frac{\partial \bar{\mathbf{C}}}{\partial \bar{\mathbf{p}}}\bar{\mathbf{N}}$, $\bar{\mathbf{q}}^T\frac{\partial \bar{\mathbf{A}}}{\partial \alpha}\bar{\mathbf{N}}$, $\bar{\mathbf{M}}^T\frac{\partial \bar{\mathbf{B}}}{\partial \alpha}\bar{\mathbf{N}}$, appearing on (23) and involving the derivatives with respect to the pressure and the birth rate, have a direct relation with the corresponding terms in the continuous Lagrangian gradient (see (5)-(7)). In effect, the discretization of

$$\begin{aligned}
&\int_0^T (\nabla\tilde{p}N, \nabla M)dt \quad \text{yields} \quad \Big(\bar{\mathbf{M}}^T\frac{\partial \bar{\mathbf{C}}}{\partial \bar{\mathbf{p}}}\bar{\mathbf{N}}\Big)(\tilde{\mathbf{p}}) \\
&-\int_0^T (N, q\tilde{\alpha})dt \qquad \text{yields} \quad \Big(\bar{\mathbf{q}}^T\frac{\partial \bar{\mathbf{A}}}{\partial \alpha}\bar{\mathbf{N}}\Big)(\tilde{\alpha}) \\
&-\int_0^T (N, M\tilde{\alpha})dt \qquad \text{yields} \quad \Big(\bar{\mathbf{M}}^T\frac{\partial \bar{\mathbf{B}}}{\partial \alpha}\bar{\mathbf{N}}\Big)(\tilde{\alpha})
\end{aligned} \tag{24}$$

where $\tilde{\mathbf{p}}$ is the space-time discretization of $\tilde{p}$, and for $\tilde{\alpha}$, representing a variation of the birth rate parameter field, we have kept the same notation for its finite element discretization.

The discrete adjoint system is (compare with the continuous adjoint system (7))

$$
\begin{aligned}
&\mathcal{L}_{\mathbf{N}^s}(\alpha,\bar{\mathbf{N}},\bar{\mathbf{p}},\bar{\mathbf{M}},\bar{\mathbf{q}}) := \tfrac{\Delta t}{2}\mathbf{K}^T(\mathbf{N}^s-\mathbf{N}_d{}^s)+\mathbf{A}^T\mathbf{q}^s+(\mathbf{L}+\Delta t\mathbf{C}^{s-1})^T\mathbf{M}^s=\mathbf{0},\\
&\mathcal{L}_{\mathbf{N}^i}(\alpha,\bar{\mathbf{N}},\bar{\mathbf{p}},\bar{\mathbf{M}},\bar{\mathbf{q}}) := \Delta t\mathbf{K}^T(\mathbf{N}^i-\mathbf{N}_d{}^i)+\mathbf{A}^T\mathbf{q}^i+(\mathbf{L}+\Delta t\mathbf{C}^{i-1})^T\mathbf{M}^i-\mathbf{K}^T\mathbf{M}^{i+1}=\mathbf{0},\\
&\mathcal{L}_{\mathbf{N}^1}(\alpha,\bar{\mathbf{N}},\bar{\mathbf{p}},\bar{\mathbf{M}},\bar{\mathbf{q}}) := \tfrac{\Delta t}{2}\mathbf{K}^T(\mathbf{N}^1-\mathbf{N}_d{}^1)+\mathbf{A}^T\mathbf{q}^1+\mathbf{M}^1-\mathbf{K}^T\mathbf{M}^2=\mathbf{0},\\
&\mathcal{L}_{\mathbf{p}^s}(\alpha,\bar{\mathbf{N}},\bar{\mathbf{p}},\bar{\mathbf{M}},\bar{\mathbf{q}}) := \mathbf{R}^T\mathbf{q}^s=\mathbf{0},\\
&\mathcal{L}_{\mathbf{p}^i}(\alpha,\bar{\mathbf{N}},\bar{\mathbf{p}},\bar{\mathbf{M}},\bar{\mathbf{q}}) := \mathbf{R}^T\mathbf{q}^i+(\Delta t\mathbf{M}^{i+1})^T\frac{\partial\mathbf{C}^i}{\partial\mathbf{p}^i}\mathbf{N}^{i+1}=\mathbf{0},
\end{aligned}
\tag{25}
$$

where $i=s-1,\dots,1$, and $\frac{\partial(.)}{\partial\mathbf{p}^i}$ denotes the derivative with respect to $\mathbf{p}^i$. From the penultimate equation in (25) we obtain $\mathbf{q}^s=0$, because $\mathbf{R}^T$ (the transpose of the stiffness matrix) is positive definite. Therefore, the discrete adjoint system is a backwards-in-time Euler discretization.

Finally, the discrete control equation is (compare with the continuous adjoint equation (6))

$$
\begin{aligned}
\mathcal{L}_\alpha(\alpha,\bar{\mathbf{N}},\bar{\mathbf{p}},\bar{\mathbf{M}},\bar{\mathbf{q}}) &:= (\gamma_1\mathbf{K}+\gamma_2\mathbf{R})\alpha+\bar{\mathbf{q}}^T\frac{\partial\bar{\mathbf{A}}}{\partial\alpha}\bar{\mathbf{N}}+\bar{\mathbf{M}}^T\frac{\partial\bar{\mathbf{B}}}{\partial\alpha}\bar{\mathbf{N}}\\
&:= (\gamma_1\mathbf{K}+\gamma_2\mathbf{R})\alpha-\sum_{i=1}^{s}\mathbf{q}^{i\,T}\frac{\partial\mathbf{K}_\alpha}{\partial\alpha}\mathbf{N}^i-\Delta t\sum_{i=2}^{s}\mathbf{M}^i\frac{\partial\mathbf{K}_{\alpha-\beta}}{\partial\alpha}\mathbf{N}^i=\mathbf{0}
\end{aligned}
\tag{26}
$$

where $\frac{\partial(.)}{\partial\alpha}$ denotes the derivative with respect to α.

4.5. The Gauss-Newton-Conjugate-Gradient method. In this discrete setting, the Newton step on the optimality conditions is defined by (compare with (9))

$$
\begin{bmatrix}
\begin{pmatrix}\mathcal{L}_{\bar{\mathbf{N}}\bar{\mathbf{N}}} & \mathcal{L}_{\bar{\mathbf{N}}\bar{\mathbf{p}}}\\ \mathcal{L}_{\bar{\mathbf{p}}\bar{\mathbf{N}}} & \mathcal{L}_{\bar{\mathbf{p}}\bar{\mathbf{p}}}\end{pmatrix} & \begin{pmatrix}\mathcal{L}_{\bar{\mathbf{N}}\alpha}\\ \mathcal{L}_{\bar{\mathbf{p}}\alpha}\end{pmatrix} & \begin{pmatrix}\mathcal{L}_{\bar{\mathbf{N}}\bar{\mathbf{M}}} & \mathcal{L}_{\bar{\mathbf{N}}\bar{\mathbf{q}}}\\ \mathcal{L}_{\bar{\mathbf{p}}\bar{\mathbf{M}}} & \mathcal{L}_{\bar{\mathbf{p}}\bar{\mathbf{q}}}\end{pmatrix}\\
\begin{pmatrix}\mathcal{L}_{\alpha\bar{\mathbf{N}}} & \mathcal{L}_{\alpha\bar{\mathbf{p}}}\end{pmatrix} & \mathcal{L}_{\alpha\alpha} & \begin{pmatrix}\mathcal{L}_{\alpha\bar{\mathbf{M}}} & \mathcal{L}_{\alpha\bar{\mathbf{q}}}\end{pmatrix}\\
\begin{pmatrix}\mathcal{L}_{\bar{\mathbf{M}}\bar{\mathbf{N}}} & \mathcal{L}_{\bar{\mathbf{M}}\bar{\mathbf{p}}}\\ \mathcal{L}_{\bar{\mathbf{q}}\bar{\mathbf{N}}} & \mathcal{L}_{\bar{\mathbf{q}}\bar{\mathbf{p}}}\end{pmatrix} & \begin{pmatrix}\mathcal{L}_{\bar{\mathbf{M}}\alpha}\\ \mathcal{L}_{\bar{\mathbf{q}}\alpha}\end{pmatrix} & \begin{pmatrix}\mathcal{L}_{\bar{\mathbf{M}}\bar{\mathbf{M}}} & \mathcal{L}_{\bar{\mathbf{M}}\bar{\mathbf{q}}}\\ \mathcal{L}_{\bar{\mathbf{q}}\bar{\mathbf{M}}} & \mathcal{L}_{\bar{\mathbf{q}}\bar{\mathbf{q}}}\end{pmatrix}
\end{bmatrix}
\begin{bmatrix}\hat{\bar{\mathbf{N}}}\\ \hat{\bar{\mathbf{p}}}\\ \hat{\alpha}\\ \hat{\bar{\mathbf{M}}}\\ \hat{\bar{\mathbf{q}}}\end{bmatrix}
=-\begin{bmatrix}\mathcal{L}_{\bar{\mathbf{N}}}\\ \mathcal{L}_{\bar{\mathbf{p}}}\\ \mathcal{L}_{\alpha}\\ \mathcal{L}_{\bar{\mathbf{M}}}\\ \mathcal{L}_{\bar{\mathbf{q}}}\end{bmatrix}
\tag{27}
$$

where $(\hat{\bar{\mathbf{N}}},\hat{\bar{\mathbf{p}}},\hat{\alpha},\hat{\bar{\mathbf{M}}},\hat{\bar{\mathbf{q}}})$ is the update direction and both the first and second derivatives of the Lagrangian are evaluated at $(\alpha,\bar{\mathbf{N}},\bar{\mathbf{p}},\bar{\mathbf{M}},\bar{\mathbf{q}})$ (we have omitted it for simplifying the notations). The explicit definition for the Hessian matrix of the discrete Lagrangian functional $\mathcal{L}$, appearing on the left-hand side in (27), is easily derived from (23) yielding

$$
\begin{bmatrix}
\begin{pmatrix}\bar{\mathbf{Q}} & \bar{\mathbf{M}}^T\frac{\partial\bar{\mathbf{C}}}{\partial\bar{\mathbf{p}}}\\ \bar{\mathbf{M}}^T\frac{\partial\bar{\mathbf{C}}}{\partial\bar{\mathbf{p}}} & \mathbf{0}\end{pmatrix} & \begin{pmatrix}\bar{\mathbf{q}}^T\frac{\partial\bar{\mathbf{A}}}{\partial\alpha}+\bar{\mathbf{M}}^T\frac{\partial\bar{\mathbf{B}}}{\partial\alpha}\\ \mathbf{0}\end{pmatrix} & \begin{pmatrix}\bar{\mathbf{B}}^T+\bar{\mathbf{C}}^T & \bar{\mathbf{A}}^T\\ (\frac{\partial\bar{\mathbf{C}}}{\partial\bar{\mathbf{p}}}\bar{\mathbf{N}})^T & \bar{\mathbf{R}}^T\end{pmatrix}\\
\begin{pmatrix}\bar{\mathbf{q}}^T\frac{\partial\bar{\mathbf{A}}}{\partial\alpha}+\bar{\mathbf{M}}^T\frac{\partial\bar{\mathbf{B}}}{\partial\alpha} & \mathbf{0}\end{pmatrix} & \gamma_1\mathbf{K}+\gamma_2\mathbf{R} & \begin{pmatrix}(\frac{\partial\bar{\mathbf{B}}}{\partial\alpha}\bar{\mathbf{N}})^T & (\frac{\partial\bar{\mathbf{A}}}{\partial\alpha}\bar{\mathbf{N}})^T\end{pmatrix}\\
\begin{pmatrix}\bar{\mathbf{B}}+\bar{\mathbf{C}} & \frac{\partial\bar{\mathbf{C}}}{\partial\bar{\mathbf{p}}}\bar{\mathbf{N}}\\ \bar{\mathbf{A}} & \bar{\mathbf{R}}\end{pmatrix} & \begin{pmatrix}\frac{\partial\bar{\mathbf{B}}}{\partial\alpha}\bar{\mathbf{N}}\\ \frac{\partial\bar{\mathbf{A}}}{\partial\alpha}\bar{\mathbf{N}}\end{pmatrix} & \begin{pmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}
\end{bmatrix}
\tag{28}
$$

We now assume that $(\bar{\mathbf{N}},\bar{\mathbf{p}})$, and $(\bar{\mathbf{M}},\bar{\mathbf{q}})$ satisfy the state and adjoint equations, such that $\mathcal{L}_{\bar{\mathbf{N}}}=\mathcal{L}_{\bar{\mathbf{p}}}=\mathcal{L}_{\bar{\mathbf{M}}}=\mathcal{L}_{\bar{\mathbf{q}}}=\mathbf{0}$ in (27), and use a block elimination in (27) for removing the incremental state and adjoint variables $(\hat{\bar{\mathbf{N}}},\hat{\bar{\mathbf{p}}})$ and $(\hat{\bar{\mathbf{M}}},\hat{\bar{\mathbf{q}}})$. In fact, denoting by $\mathbf{H}_{ij}$, with $i,j=1,2,3$ (remark that $\mathbf{H}_{ji}=\mathbf{H}_{ij}^T$), the nine block matrices displayed in the Hessian of $\mathcal{L}$ (see (27) and (28)), then from the last block-line the incremental state variables $(\hat{\bar{\mathbf{N}}},\hat{\bar{\mathbf{p}}})$ can be written as a function of the incremental control variable $\hat{\alpha}$

$$
\begin{bmatrix}\hat{\bar{\mathbf{N}}}\\ \hat{\bar{\mathbf{p}}}\end{bmatrix}=-\mathbf{H}_{31}^{-1}\,\mathbf{H}_{32}\,\hat{\alpha},
\tag{29}
$$

and from the first block-line it results the following system for the incremental variables $(\hat{\bar{\mathbf{M}}},\hat{\bar{\mathbf{q}}})$

$$
\begin{bmatrix}\hat{\bar{\mathbf{M}}}\\ \hat{\bar{\mathbf{q}}}\end{bmatrix}=-\mathbf{H}_{13}^{-1}\Big(\mathbf{H}_{11}\begin{bmatrix}\hat{\bar{\mathbf{N}}}\\ \hat{\bar{\mathbf{p}}}\end{bmatrix}+\mathbf{H}_{12}\,\hat{\alpha}\Big).
\tag{30}
$$

Consequently, the Newton step (27) reduces to the system

$$
\mathbf{H}\hat{\alpha}=-\mathcal{L}_\alpha,
\tag{31}
$$