Mathematical Models and Methods in Applied Sciences

Certified Reduced Order Methods for Optimal Treatment Planning

Bahodir Ahmedov
Aachen Institute for Advanced Study in Computational Engineering Science (AICES), RWTH Aachen University, Schinkelstr. 2, 52062 Aachen, Germany
ahmedov@aices.rwth-aachen.de

Martin A. Grepl
Institut für Geometrie und Praktische Mathematik, RWTH Aachen University, Templergraben 55, 52056 Aachen, Germany,
grepl@igpm.rwth-aachen.de

Michael Herty
Institut für Geometrie und Praktische Mathematik, RWTH Aachen University, Templergraben 55, 52056 Aachen, Germany,
herty@igpm.rwth-aachen.de

Received (Day Month Year)
Revised (Day Month Year)
Communicated by (xxxxxxxxxx)

We study numerical methods for inverse problems arising in cancer therapy treatment under uncertainty. The interest is on efficient and reliable numerical methods that allow to determine the influence of possible unknown parameters on the treatment plan for cancer therapy. A kinetic transport equation is used to model the evolution of charged particles in tissue. A mixed variational framework is presented and existence and uniqueness of a weak solution is established. The optimality system is approximated using a low–dimensional reduced basis formulation based on a P_N -FE discretization. We derive *a posteriori* bounds for the error in the reduced basis solution of the optimal control problem with respect to the solution of the P_N -FE discretization. Numerical results in slab geometry are presented to confirm the validity of our approach.

Keywords: optimal radiotherapy, reduced basis methods, *a posteriori* error estimation, P_N - methods, mixed variational formulation

AMS Subject Classification: 85A25, 49J20, 65K10, 65M15, 92C50

1. Introduction

We are interested in the numerical discretization of inverse problems in cancer therapy. In particular, we focus on treatment planing using ionizing radiation. Cancer therapy requires the understanding of many aspects ranging from tumor growth to applied problems in engineering. In particular, the understanding of cancer tis-

sue and its evolution has given rise to complex models and it is, at this point, far from being fully understood; we refer to Ref. 6 for a review. New models may come in different flavors as, for example, recently introduced phenomenological models[35,36,31]. Our focus is on inverse problems related to the transport process of radiation through tissue. The actual influence of the radiation on tissue and the evolution of cancer are simplified using only proportionality arguments[41]. Other models may be substituted here in the future. However, the propagation process of radiation through tissue is at least quite well understood from a physical point of view[2,20]. In this work we aim to exploit the structure of this transport process to address an inverse problem, which is one aspect of radiation treatment planing. The aim of radiation treatment is to deposit enough energy in cancer cells so that they are destroyed while healthy tissue around the cancer cells should be harmed as little as possible.

The design of the precise treatment plan and the corresponding dose computation based on the geometry of cancer and healthy cells is still a topic of active research. We follow here an approach requiring mathematical modeling and optimization techniques[41]; for a review on existing technologies we refer to Refs. 8, 19. The mathematical modeling of dose calculation using a kinetic transport equation has recently gained attention[20,29,42,44,17,21]. It has been argued that the transport based dose calculations have the same computational complexity and accuracy as Monte Carlo simulations[7]. However, as opposed to the Monte Carlo methods, they allow to exploit structural information for analytical and numerical purposes[43,17]. For optimal treatment planing problems governed by kinetic transport equation a variety of analytical and numerical results could therefore be established over the past years.

The focus of the present paper is on the numerical analysis of dose computation problems. In particular, we investigate the dependence of the optimal treatment plan with respect to uncertain parameters of the problem. The guiding example is the location of tissue cells with different radiative properties. In practice, those are obtained through imaging processes and may only be reliable up to a certain error. A quantification of the effect of this error on the optimal treatment plan and a fast recomputation of the dose for online adjustment of treatment plans is certainly desirable. A first attempt using feedback control has been studied in Ref. 15.

Here, we proceed differently and pursue the reduced basis (RB) method, a model order reduction technique which allows for a low–dimensional approximation of the parameterized problem. Various model order reduction techniques have been employed over the past years to speed up the solution of optimal control problems, e.g., proper orthogonal decomposition (POD)[45], reduction based on intertial manifolds[24], and RB methods[26,34,27]. However, since the solution of the reduced optimal control problem is generally suboptimal, rigorous and efficiently evaluable *a posteriori* error bounds are crucial to assess the quality of the solution[26,28,34,45]. To this end, we first introduce a variational formulation of the kinetic transport equation — previously studied in the context of forward simulation in Refs. 11, 12,

13 — and extend the discretization towards the treatment planning problem. We also note that an RB approach for a transport model has been previously considered in Ref. 37 and transport-dominated problems in Ref. 9. Based on the work in Ref. 28, we then develop efficient and reliable *a posteriori* error bounds for the RB solution to the radiation treatment planning problem, i.e., for the error in the optimal control and the associated cost functional. Finally, we employ the reduced solution to quantify L^∞ and L^2-deviations of the treatment plan in terms of the uncertainty in the geometry parameters.

In the following we present more details on the optimal treatment planning problem. The starting point is a kinetic transport equation for particle transport in a medium. This equation can be derived from underlying particle dynamics used in clinical radiotherapy[20,14]. The full model includes all physical effects and gives computationally similar results to established commerical Monte–Carlo codes[7]. For a full review on the underlying equations and properties we refer to Refs. 20, 16. Since our focus is on the framework used in inverse problems, we consider a simplified setting to illustrate the ideas. The simplifications are the following: (i) the energy dependence of the traveling particles are neglected; (ii) the geometry is strongly simplified leading to spatially one–dimensional objects — this corresponds to a one–dimensional cut through a CT scan, whereas the full model would clearly require to resolve a three–dimensional body; and (iii) the scattering and absorption of the tissue is only spatially dependent, but not dependent on the energy or traveling direction of particles. The last point is a simplification regarding the behavior of tissue under radiation. We refer to Ref. 16 for a model including all the previous effects.

Under the previous assumptions, the particle transport through tissue is described by

$$\mu\partial_x\psi(x,\mu) + \sigma_t(x)\psi(x,\mu) = \sigma_s(x)\int_{-1}^{1} s(x,\mu\cdot\mu')\psi(x,\mu')\mathrm{d}\mu' + q(x). \tag{1.1}$$

Here, $\psi : [0,1]\times[-1,1] \to \mathbb{R}$ describes the probability density of particles at a spatial point $x \in [0,1]$ and traveling with direction (speed) $\mu \in [-1,1]$. The value $\psi(x,\mu)\mathrm{d}x\mathrm{d}\mu$ can thus be thought of as the probability to have a particle in the element $\mathrm{d}x\mathrm{d}\mu$. The particles travel with speed μ in x-direction leading to the transport term $\mu\partial_x\psi$. Along their path the particles may undergo scattering and absorption. The rate of scattering and absorption is described by $\sigma_t(x)$ and may dependent on x due to different possible tissues at different positions x. Particles traveling with speed μ' may also be scattered towards the current direction μ. Therefore, a source term with rate $\sigma_s(x)$ appears on the right–hand side of the equation integrating all possible directions μ'. The scattering may also dependent on the angle of μ and μ'. This dependence is represented by the function $s : [0,1]\times[-1,1] \to \mathbb{R}$ in the integral kernel. Finally, a source term $q : [0,1] \to \mathbb{R}$ may introduce new particles according to our treatment plan. The function q is later determined by solving an optimal control problem.

We note that scattering is determined by the total cross-section σ_t and by the total scattering cross-section σ_s, which are are non-negativ functions from $[0,1]$ to $\mathbb{R}$. It is clear that the values of $\sigma_{t,s}$ *strongly dependent* on the underlying geometry. Scattering and adsorption is by magnitudes different in water and bone tissue. This indicated by the dependence on x of those functions. The function $q(\cdot)$ is the treatment plan (or control) and acts as an external source independent of μ. The function s is also called the scattering kernel and given in case of radiotherapy by the simplified Henvey-Greenstein formula

$$s(x,\eta) := \frac{1-g(x)^2}{4\pi(1+g(x)^2-2g(x)\eta)^{3/2}},$$

where g is the average cosine of the scattering angle. For high energy particles, small angle and energy changes are very likely, thus the scattering kernel is very forward–peaked. Boundary conditions for (1.1) are prescribed at the inflow part, i.e., for $\mu > 0$ at $x = 0$ and for $\mu < 0$ at $x = 1$. At those points we prescribe zero boundary conditions $\psi(x,\mu) = 0$ for simplicity. Note that Eq. (1.1) is also the basic model for transport of non–interacting particles in astrophysics, neutron transport and other applications. It is referred to as the monochromatic radiative transfer equation[10,25,16,12].

As in Ref. 41, we assume that the amount of destroyed cells in a small volume, be they cancer or healthy cells, is directly proportional to the dose

$$D(x) = \int_{-1}^{1} \psi(x,\mu)\mathrm{d}\mu, \tag{1.2}$$

deposited in that volume. As discussed previously, this is clearly a very simple model neglecting many aspects of cancer growth. However, it is clear that a model for dose distribution and cancer damage needs to dependent on $D(x)$. Any improved model on the dependence on D may be included in future inverse problems using a similar methodology as outlined below.

The computational domain is divided into tumor tissue, normal tissue and a region at risk: $\Omega = [0,1] = \Omega_T \cup \Omega_N \cup \Omega_R$. We prescribe a desired dose distribution $\bar{D}(x)$, which usually has a constant value in Ω_T and is zero elsewhere. The problem of optimal treatment planning is then to find an external beam distribution q such that

$$\bar{J}(\psi) = \frac{\alpha_T}{2}\int_{\Omega_T}(D-\bar{D})^2\mathrm{d}x + \frac{\alpha_N}{2}\int_{\Omega_N}(D-\bar{D})^2\mathrm{d}x + \frac{\alpha_R}{2}\int_{\Omega_R}(D-\bar{D})^2\mathrm{d}x$$

is minimal. Here, α_T, α_N, and α_R are (positive) weights determining the trade-off between the terms in the cost. Defining $\alpha(x) = \alpha_T\chi_T(x) + \alpha_N\chi_N(x) + \alpha_R\chi_R(x)$, where $\chi_{T,N,R}(x)$ are appropriate characteristic functions, we formulate the optimal treatment planning problem therefore as

$$\min J(\psi,q) = \int_0^1 \frac{\alpha(x)}{2}(D-\bar{D})^2\,\mathrm{d}x + \frac{1}{2}\int_0^1 q(x)^2\,\mathrm{d}x \text{ subject to (1.1).} \tag{1.3}$$

We are interested in the optimal treatment plan $q(x)$ obtained as the numerical solution to problem (1.3). Recall that the material properties σ_t and σ_s may differ in various regions of the domain and thus depend on the exact geometry, i.e., location of tumor, normal tissue, and region at risk. These locations, however, may change during treatment. Our goal is thus to investigate the dependence of the optimal plan $q(x)$ on the geometry parameters. Since re-planning (finding an optimal plan $q(x)$ for new values of the geometry parameters) is often expensive in terms of computational time, we apply the RB method to the optimal control problem (1.3) allowing to efficiently investigate the parametric dependence.

In Section 2 we present a variational discretization similar to the one proposed in Ref. 11 and establish existence and uniqueness of the optimal control $q(x)$. In this way we provide the theoretical background for the discretization of problem (1.3). In order to assess geometry changes in problem (1.3) we introduce the parameterized geometry as well as the parameterized optimal control problem in Section 3. We state the first order optimality conditions characterising the optimal control q. We subsequently derive the numerical scheme for the solution of the radiative transfer equation. The RB method for the geometry dependent optimal radiotherapy model (1.3) is presented in Section 4, where we first introduce the spaces and approximation. And subsequently we develop *a posteriori* error bounds for the optimal control and the associated cost functional value. The performance of the *a posteriori* error bounds is shown in Section 5. Here, we also present numerical examples where we explore uncertainty properties of the optimal control when input parameters are subject to random variations.

We try to meet the standards of reproducible research in the computation sciences, laid out e.g. by LeVeque[30]. The source code, which is designed in MATLAB®, along with files to generate all figures and results of the paper as well as additional functions and examples are available online[1].

2. Mixed Variational Formulation

2.1. *Kinetic Transport Equation*

We introduce a variational formulation of Eq. (1.1) and problem (1.3) in order to apply reduced order modeling techniques. The variational formulation of (1.1) has been studied among others in Ref. 12 and we follow here closely their presentation. We denote by $\mathcal{D} := [0,1] \times [-1,1]$ the bounded convex domain of the problem and also define the spatial domain $\mathcal{I} := [0,1]$. We next introduce the L^2-scalar products $(\cdot,\cdot)_{\mathcal{D}}$ and $(\cdot,\cdot)_{\mathcal{I}}$ and induced norms $\|\cdot\|_{\mathcal{D}}^2 = (\cdot,\cdot)_{\mathcal{D}}$ and $\|\cdot\|_{\mathcal{I}}^2 = (\cdot,\cdot)_{\mathcal{I}}$ on the sets $\mathcal{D}$ and $\mathcal{I}$, respectively. The inflow boundary is denoted by $\mathcal{D}_+ := \{(x,\mu) : \ x = 0, \mu > 0 \text{ or } x = 1, \mu < 0\}$, the remaining boundary by $\mathcal{D}_- = \partial\mathcal{D}\backslash\mathcal{D}_+$. In view of the arising operators we define the Lebesgue space of square integrable functions over $\mathcal{D}$ and $\mathcal{I}$ as $X := L^2(\mathcal{D})$ and $Y := L^2(\mathcal{I})$ with scalar products $(\psi,\phi)_X = (\psi,\phi)_{\mathcal{D}}$ and $(\psi,\phi)_Y = (\psi,\phi)_{\mathcal{I}}$; furthermore, we introduce $\mathbb{X} := \{\psi \in X : \mu\partial_x\psi \in X\}$ and $\mathbb{X}_{bc} := \{\psi \in \mathbb{X} : \psi = 0 \text{ on } \mathcal{D}_+\}$.

Following Refs. 10, 17, we impose the following assumptions on the coefficients in Eq. (1.1)

$$s, \sigma_t, \sigma_s \geq 0, \quad \sigma_t(x), \sigma_s(x) \in L^\infty(\mathcal{I}), \quad \int_{-1}^{1} s(x,\eta)d\eta \leq c_0, \ \forall x \in \mathcal{I}, \tag{2.1a}$$

$$\sigma_t(x) - \sigma_s(x) \int_{-1}^{1} s(x,\eta)d\eta \geq \hat{\beta} > 0, \ \forall x \in \mathcal{I}, \tag{2.1b}$$

Clearly, the scattering and absorption rates and kernels need to be non–negative. The requirement to be bounded is a very weak assumption also natural for realistic rates. The kernel s needs to be integrable for the equation to be meaningful, and this is in particular true for the Henvey–Greenstein kernel introduced in the previous paragraph. Therefore, the first assumption is satisfied by any available physical scattering and absorption functions. The second assumptions links the scattering and absorption ratios. It essentially requires that there is absorption in a medium, or – formulated differently – that there are no void or vacuum regions present. This assumptions guarantees the inf-sup stability and uniqueness of the solution. However, in realistic CT data this assumption is also usually fulfilled. Therefore, the imposed assumptions do not present a serious restriction to realistic scenarios.

We next define the transport and source term operator as

$$\mathcal{A} : \mathbb{X} \to X, \ \ \mathcal{A}\psi = \mu\partial_x\psi, \tag{2.2}$$

$$\mathcal{C} : X \to X, \ \ C\psi = \sigma_t\psi - \sigma_s \int s(x, \mu\cdot\mu')\psi(x,\mu')\mathrm{d}\mu'. \tag{2.3}$$

We can then restate Eq. (1.1) as an operator equation on $\mathbb{X}$ and a strong solution $\psi \in \mathbb{X}_{bc}$ fulfills

$$\mathcal{A}\psi + \mathcal{C}\psi = q, \ \ a.e.\,(x,\mu) \in \mathcal{D}. \tag{2.4}$$

The main properties of the operators (2.2) and (2.3) have been analysed for example in Ref. 10 and are summarized as follows: Under assumption (2.1), $\mathcal{A}$ is a linear and bounded operator, and $\mathcal{C}$ is linear, self–adjoint with respect to X, bounded, coercive and therefore invertible. The estimations on C are as follows:

$$\begin{aligned}(C\phi,\phi)_{\mathcal{D}} &\leq \|\sigma_t\|_{L^\infty}\|\phi\|^2_{\mathcal{D}} \\ &\quad + \|\sigma_s\|_{L^\infty}\left\|\int \sqrt{s(x,\mu\mu')}\phi(x,\mu')\sqrt{s(x,\mu'\mu)}\phi(x,\mu)dx\mathrm{d}\mu\mathrm{d}\mu'\right\| \\ &\leq (\|\sigma_t\|_{L^\infty} + 2\|\sigma_s\|_{L^\infty}c_0)\|\phi\|^2_{\mathcal{D}}, \\ (C\phi,\phi)_{\mathcal{D}} &\geq \int_{\mathcal{D}}\left(\sigma_t - \sigma_s\int_{-1}^{1} s(x,\mu\mu')\mathrm{d}\mu'\right)\phi^2(x,\mu)\mathrm{d}x\mathrm{d}\mu \geq \hat{\beta}\|\phi\|^2_D.\end{aligned}$$

Furthermore, for any $q \in Y$, there exists a unique (strong) solution $\psi \in \mathbb{X}_{bc}$ to Eq. (1.1). Define the operator mapping q to ψ, i.e.,

$$\mathcal{E} : Y \to \mathbb{X}_{bc} \subset X, \ \ \mathcal{E}(q) = \psi, \tag{2.5}$$

where ψ is the solution to (1.1). According to Ref. 17 the operator $\mathcal{E}$ is a bounded linear operator on X and we have $\|\mathcal{E}(q)\|_X \leq \frac{1}{\beta}\|q\|_{\mathcal{I}}$.

A variational formulation is obtained using an even–odd splitting in the velocity space given by

$$\psi^{\pm}(x,\mu) = \frac{1}{2}\left(\psi(x,\mu) \pm \psi(x,-\mu)\right).$$

The splitting suggests to search for solutions $\psi = \psi^+ + \psi^-$ having the regularity $\psi^+ \in \mathbb{X}$ and $\psi^- \in X$. The space of such functions is denoted by $\mathcal{X}$, i.e., $\mathcal{X} := \{\psi = \psi^+ + \psi^- : \psi^+ \in \mathbb{X}, \psi \in X\}$. We introduce the following scalar product

$$(v,w)_{\mathcal{X}} = (\mathcal{A}v^+, \mathcal{A}w^+)_{\mathcal{D}} + (v,w)_{\mathcal{D}} + (v^+, w^+)_{\mathcal{T}}, \tag{2.6}$$

and associated norm $\|v\|^2_{\mathcal{X}} = \|\mathcal{A}v^+\|^2_{\mathcal{D}} + \|v\|^2_{\mathcal{D}} + \|v^+\|^2_{\mathcal{T}}$ on $\mathcal{X}$, where

$$(v^+, w^+)_{\mathcal{T}} := \sum_{x\in\{0,1\}} \int_{-1}^{1} |\mu|\, w^+(x,\mu)\, v^+(x,\mu)\,\mathrm{d}\mu.$$

We derive the variational formulation of (1.1) for zero boundary conditions as follows. Assume $\psi \in \mathbb{X}$ is a strong solution. Multiplying (1.1) by a test function $\phi \in \mathbb{X}$ and integrating over $\mathcal{D}$ yields

$$\begin{aligned} 0 &= \int_{\mathcal{D}} \mu\,(\psi^+_x + \psi^-_x)\,(\phi^+ + \phi^-)\,\mathrm{d}x\,\mathrm{d}\mu \\ &\quad + \int_{\mathcal{D}} \left(\sigma_t\,\psi\,\phi - \sigma_s \int \psi(x,\mu')\,s(x,\mu'\cdot\mu)\,\mathrm{d}\mu'\,\phi(x,\mu)\right)\mathrm{d}x\,\mathrm{d}\mu - \int_{\mathcal{D}} q\,\phi\,\mathrm{d}x\,\mathrm{d}\mu \\ &= \int_{\mathcal{D}} \mu\,\left(\phi^+\,\psi^-_x + \phi^-\,\psi^+_x\right)\mathrm{d}x\,\mathrm{d}\mu + (C\phi,\psi)_{\mathcal{D}} - (q,\phi)_{\mathcal{D}}. \end{aligned}$$

At $x = 1$ we have $\psi^+(\mu) = \frac{1}{2}\psi(\mu) = \psi^-(\mu)$ for $\mu > 0$ and $\psi^+(\mu) = \frac{1}{2}\psi(-\mu) = -\psi^-(\mu)$ for $\mu < 0$ and vice versa at $x = 0$. Therefore, $\mu\psi^-(\mu) = |\mu|\psi^+(\mu)$ and $\mu\psi^-(\mu) = -|\mu|\psi^+(\mu)$. Hence, integration by parts yields

$$\int_{\mathcal{D}} \mu\,\phi^+\,\psi^-_x\,\mathrm{d}x\,\mathrm{d}\mu = -\int_{\mathcal{D}} \mu\,\phi^+_x\,\psi^-\,\mathrm{d}x\,\mathrm{d}\mu + \sum_{x\in\{0,1\}} \int_{-1}^{1} |\mu|\,\phi^+(x,\mu)\,\psi^+(x,\mu)\,\mathrm{d}\mu.$$

We thus define bilinear forms $b(\cdot,\cdot) : Y \times \mathcal{X} \to \mathbb{R}$ and $a(\cdot,\cdot) : \mathcal{X}\times\mathcal{X} \to \mathbb{R}$ by

$$b(q,\phi) = (q,\phi)_{\mathcal{D}} \tag{2.7}$$

$$\begin{aligned} a(\psi,\phi) &= (\phi^-, \mathcal{A}\psi^+)_{\mathcal{D}} - (\psi^-, \mathcal{A}\phi^+)_D + (C\phi,\psi)_{\mathcal{D}} \\ &\quad + \sum_{x\in\{0,1\}} \int_{-1}^{1} |\mu|\phi^+(x,\mu)\psi^+(x,\mu)\mathrm{d}\mu, \end{aligned} \tag{2.8}$$

respectively; we refer to Ref. 12 for the case $\mathcal{D} = \mathbb{R}^2 \times \mathcal{S}^2$. Note that since C is self-adjoint, a is symmetric. Furthermore, b is bounded on $\mathcal{X}$ since

$$b(q,\phi) \leq 2\|q\|_{\mathcal{I}}\|\phi\|_D \leq 2\|q\|_{\mathcal{I}}\|\phi\|_{\mathcal{X}}.$$

Next, we prove that the bilinear form a is continuous and inf–sup stable. We first note that a is bounded since

$$\begin{aligned} a(\psi,\phi) &\leq \|\phi^-\|_D\|\mathcal{A}\psi^+\|_D + \|\psi^-\|_D\|\mathcal{A}\phi^+\|_D + \|C\phi\|_{\mathcal{D}}\|\psi\|_D + \|\phi^+\|_{\mathcal{T}}\|\psi^+\|_T \\ &\leq \gamma_a\|\phi\|_{\mathcal{X}}\|\psi\|_{\mathcal{X}}, \end{aligned}$$

where $\gamma_a = 4(\|\sigma_t\|_{L^\infty} + 2\|\sigma_s\|_{L^\infty}c_0)$ and we used the Cauchy-Schwarz inequality and $\|\phi^-\|_{\mathcal{D}} \leq \|\phi\|_{\mathcal{D}}$. It remains to show the inf–sup stability which requires to prove the following two estimates

$$\sup_{\|w\|_{\mathcal{X}}=1} |a(v,w)| \geq \beta\|v\|_{\mathcal{X}} \quad \text{and} \quad \sup_{\|v\|_{\mathcal{X}}=1} |a(v,w)| \geq \beta\|w\|_{\mathcal{X}}.$$

Since a is symmetric it suffices to prove one of the above. In the case $w = 0$ there is nothing to prove. We obtain for $v \in \mathbb{X}$, $v \neq 0$ For $w = v$ and $w = \mathcal{A}v^+$ we obtain

$$a(v,v) = (Cv,v) + \|v^+\|_{\mathcal{T}}^2 \geq \hat{\beta}\|v\|_{\mathcal{D}}^2 + \|v^+\|_{\mathcal{T}}^2 \geq \min\{\hat{\beta},1\}(\|v\|_{\mathcal{D}}^2 + \|v^+\|_{\mathcal{T}}^2).$$

A simple computation shows that we have $(\mathcal{A}v^+)^+ = 0$, $(\mathcal{A}v^+)^- = \mathcal{A}v^+$, and therefore

$$\begin{aligned} a(v,\mathcal{A}v^+) &= ((\mathcal{A}v^+)^-,\mathcal{A}v^+)_{\mathcal{D}} - (v^-,\mathcal{A}(\mathcal{A}v^+)^+)_{\mathcal{D}} + (\mathcal{C}\mathcal{A}v^+,v)_{\mathcal{D}} + ((\mathcal{A}v^+)^+,v^+)_{\mathcal{T}} \\ &= \|\mathcal{A}v^+\|_{\mathcal{D}}^2 + (\mathcal{A}v^+,Cv)_D \geq \|\mathcal{A}v^+\|_{\mathcal{D}}^2 - \|\mathcal{A}v^+\|_D\|Cv\|_D \\ &\geq \frac{1}{2}\|\mathcal{A}v^+\|_{\mathcal{D}}^2 - \frac{\gamma_a}{8}\|v\|_D^2. \end{aligned}$$

Hence, for $w = v - \mathcal{A}v^+$ we obtain

$$a(v,w) \geq \min\left\{\frac{1}{2}, \min\{\hat{\beta},1\} - \frac{\gamma_a}{8}\right\}\|v\|_{\mathcal{X}}^2.$$

Since $\|w\|_{\mathcal{X}} \leq 2\|v\|_{\mathcal{X}}$ we obtain – provided that $\min\{\hat{\beta},1\} - \frac{\gamma_a}{8} > 0$ – the inequality

$$a(v,w) \geq \beta_{\mathrm{LB}}\|v\|_{\mathcal{X}}\|w\|_{\mathcal{X}}$$

where

$$\beta_{\mathrm{LB}} := \frac{1}{2}\min\left\{\frac{1}{2}, \min\{\hat{\beta},1\} - \frac{\gamma_a}{8}\right\}. \tag{2.9}$$

We summarize the findings as follows using the generalized Lax–Milgram Lemma, resp. Babuzka–Aziz Lemma[3,4].

Lemma 2.1. *Assume that the hypothesis* (2.1) *holds true. Assume further that*

$$\|\sigma_t\|_{L^\infty} + 2\|\sigma_s\|_{L^\infty}c_0 < 2\min\{\hat{\beta},1\}. \tag{2.10}$$

Then the bilinear form $a : \mathcal{X}\times\mathcal{X} \to \mathbb{R}$ defined in (2.8) *is linear, bounded and inf–sup stable with respect to the norm on $\mathcal{X}$. Furthermore, the bilinear form $b(\cdot,\cdot) : Y\times\mathcal{X} \to \mathbb{R}$ is bounded on $\mathcal{X}$. The variational problem*

$$a(\psi,\phi) = b(q,\phi), \quad \forall\phi \in \mathcal{X}, \tag{2.11}$$

thus has a unique solution $\psi \in \mathcal{X}$ which depends continuously on the data

$$\|\psi\|_{\mathcal{X}} \leq \frac{2}{\beta_{LB}} \|q\|_{\mathcal{I}}.$$

In the simplest case of constant s, σ_t and σ_s with s such that $c_0 = 1$ we have $\hat{\beta} = \sigma_t - \sigma_s =: \sigma_a$. We thus obtain the equivalent condition $\sigma_s < \frac{7}{3}\sigma_a$, indicating a ratio of absorption σ_a to scattering σ_s of roughly one–half.

2.2. *Optimal Control Problem*

From Lemma 2.1 we obtain the well–posedness of the control to state operator for the variational problem. We call this operator also $\mathcal{E}$ and it is defined by

$$\mathcal{E} : Y \to \mathcal{X},\ \mathcal{E}(q) = \psi, \tag{2.12}$$

where ψ solves (2.11). According to the previous lemma $\mathcal{E}$ is well–defined, $\mathcal{E}$ is a linear and bounded operator since we have $\|\mathcal{E}(q)\|_{\mathcal{X}} \leq \frac{2}{\beta_{\mathrm{LB}}}\|q\|_{\mathcal{I}}$. For notational convenience, we define the linear operator $\mathcal{L} : \mathcal{X} \to Y$ given by $\mathcal{L}\phi = \int_{-1}^{1} \phi(x,\mu)\mathrm{d}\mu$. The related optimization problem (1.3) using the variational formulation now reads

$$\begin{gathered}\min_{\psi\in\mathcal{X}, q\in Y} J(\psi, q) = \frac{1}{2}\int_{\mathcal{I}} \alpha(x)\left(\mathcal{L}\psi - \bar{D}\right)^2 \mathrm{d}x + \frac{1}{2}\int_{\mathcal{I}} q(x)^2 \mathrm{d}x \\ \text{subject to } (\psi, q) \in \mathcal{X} \times Y \text{ solves } \mathcal{E}(q) = \psi.\end{gathered} \tag{2.13}$$

It follows from our assumptions that there exists a unique optimal solution $(\psi^*, q^*) \in \mathcal{X} \times Y$ to (2.13)[32]; also see Ref. 17 for the specific case considered here. We summarize the result in the following proposition.

Proposition 2.1. *Assume that the hypothesis* (2.1) *and* (2.10) *hold true. Let* $\bar{D} = \int \bar{\psi} d\mu$ *with* $\bar{\psi} \in X$ *be given and assume* $\alpha > 0$ *with* $\alpha \in L^\infty(\mathcal{I})$.

The control problem (2.13) *then admits a unique solution* $q^* \in Y$. *The functions* $\psi^* \in \mathcal{X}$, $\lambda^* \in \mathcal{X}$, *and* $q^* \in Y$ *are a local minimum provided that the following first–order optimality conditions are fulfilled*

$$a(\psi^*, \phi) = b(q^*, \phi), \qquad \forall \phi \in \mathcal{X}, \tag{2.14}$$

$$a(\phi, \lambda^*) = \left(\alpha(\bar{D} - \mathcal{L}\psi^*), \mathcal{L}\phi\right)_Y, \qquad \forall \phi \in \mathcal{X}, \tag{2.15}$$

$$(q^*, \varphi)_Y - b(\varphi, \lambda^*) = 0, \qquad \forall \varphi \in Y. \tag{2.16}$$

3. Parameterized Optimal Control Problem

In radiotherapy treatment plan computations the spatial variation of the functions σ_t and σ_s depend on the geometry obtained from a CT scan of the patient's body. This geometry can only be determined up to a certain accuracy for reasons of movement of the patient during scanning or possible changes during the treatment. We are interested in a quantification of the optimal control q^* in terms of possible variations in σ_s and σ_t. We therefore consider a parameterized geometry and the

associated parameterized optimal control problem. For simplicity we assume that the scattering kernel $s = s(x, \mu \cdot \mu')$ is independent of x. However, the approach presented can also be extended to x-dependent scattering kernels.

3.1. *Geometric transformation and affine decomposition*

We consider the parameterized one-dimensional slab geometry sketched in Figure 1 on the top. We assume that the parameterized domain $\Omega_0(p)$ is divided into 5 subdomains $\Omega_0^i(p)$, $i = 1, \ldots, 5$, which are described by four parameters p_i, $i = 1, \ldots, 4$: the parameters p_1 and p_2 (resp. p_3 and p_4) denote the center and half-width of the domain $\Omega_0^2(p)$ (resp. $\Omega_0^4(p)$). More precisely, we have $\Omega_0(p) = \cup_{i=1}^5 \Omega_0^i(p)$, where $\Omega_0^1(p) = \chi_{[0,p_1-p_2]}(x)$, $\Omega_0^2(p) = \chi_{[p_1-p_2,p_1+p_2]}(x)$, $\Omega_0^3(p) = \chi_{[p_1+p_2,p_3-p_4]}(x)$, $\Omega_0^4(p) = \chi_{[p_3-p_4,p_3+p_4]}(x)$, and $\Omega_0^5(p) = \chi_{[p_3+p_4,1]}(x)$. We furthermore assume that the absorption and scattering coefficients σ_s and σ_t are known but possibly differ in the five subdomains so that we can write

$$\sigma_t(x) = \sum_{i=1}^{5} \sigma_t^i \, \chi^i(x) \quad \text{and} \quad \sigma_s(x) = \sum_{i=1}^{5} \sigma_s^i \, \chi^i(x), \tag{3.1}$$

where σ_t^i and σ_s^i, $i = 1, \ldots, 5$, are given and the $\chi^i(x)$ are the characteristic functions corresponding to subdomain i. We denote the admissible parameter domain for our 4-tuple (input) geometry parameter by $\mathcal{P}$, i.e., $p = (p_1, p_2, p_3, p_4) \in \mathcal{P} \subset \mathbb{R}^4$.

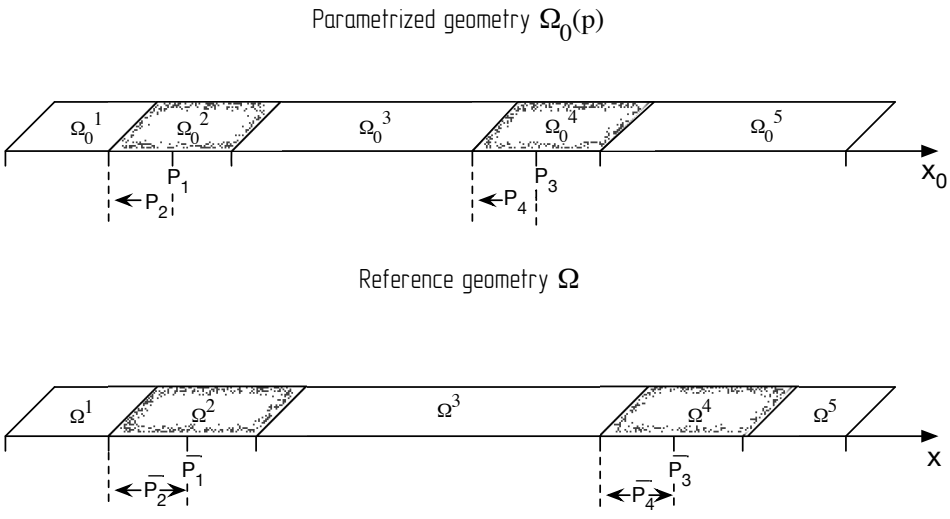


Fig. 1: Parameterized geometry $\Omega_0(p)$ and reference geometry Ω.

The efficiency of the RB method relies on an offline–online computational decomposition which requires that all bilinear and linear forms satisfy an affine parameter

dependence given by[39]

$$a(\psi,\phi;p) = \sum_{m=1}^{Q_a} \Theta_a^m(p)\, a^m(\psi,\phi), \quad \forall \psi,\phi \in \mathcal{X}, \tag{3.2}$$

$$b(q,\phi;p) = \sum_{m=1}^{Q_m} \Theta_b^m(p)\, b^m(q,\phi), \quad \forall q \in Y,\ \phi \in \mathcal{X}, \tag{3.3}$$

where the parameter-dependent coefficient functions $\Theta_{a,b}^m : \mathcal{P} \to \mathbb{R}$ are continuous and depend on p, but the bilinear forms $a^m(\cdot,\cdot) : \mathcal{X}\times\mathcal{X} \to \mathbb{R}$ and $b^m(\cdot,\cdot) : Y \times \mathcal{X} \to \mathbb{R}$ do *not* depend on p.

To obtain this form, we perform an affine mapping from the parameter dependent geometry $\Omega_0(p)$ to a parameter independent reference geometry $\Omega = \Omega_0(\bar{p})$ for some fixed reference parameter $\bar{p}$ sketched in Figure 1 on the bottom; also see Ref. 39 for a general discussion on geometry parameterizations and Ref. 1 for our specific case. We briefly outline the derivation and recall the bilinear forms

$$b(q,\phi;p) = (q,\phi)_{\mathcal{I}(p)} = \int_{\Omega_0(p)} \int_{-1}^{1} q\,\phi \,\mathrm{d}x_0\,\mathrm{d}\mu,$$

and

$$\begin{aligned} a(\psi,\phi;p) = &\int_{\Omega_0(p)} \int_{-1}^{1} \mu \left(\phi^- \partial_{x_0}\psi^+ - \psi^- \partial_{x_0}\phi^+\right) \mathrm{d}x_0\,\mathrm{d}\mu \\ &+ \int_{\Omega_0(p)} \int_{-1}^{1} \sigma_t\,\phi\,\psi\,\mathrm{d}x_0\,\mathrm{d}\mu + \sum_{x\in\{0,1\}} \int_{-1}^{1} |\mu|\,\phi^+(x,\mu)\,\psi^+(x,\mu)\mathrm{d}\mu \\ &- \int_{\Omega_0(p)} \int_{-1}^{1} \sigma_s \left\{ \int s(\mu,\mu')\phi(\mu')\mathrm{d}\mu' \right\} \psi\,\mathrm{d}x_0\,\mathrm{d}\mu, \end{aligned}$$

where the integrals in space are performed over $\Omega_0(p)$. The affine mapping for $x \in \Omega^k$ is given by $x_0 = G^k(p)x + c^k(\mu)$ and the inverse mapping is thus $x = (G^k(\mu))^{-1}(x_0 - c^k(\mu))$ for $k = 1,\ldots,5$. We have $\frac{\partial x}{\partial x_0} = (G^k(\mu))^{-1}$ and $d\Omega_0(\mu) = |\mathrm{det}G(\mu)|d\Omega = |G(\mu)d\Omega$. After mapping to the reference geometry we thus obtain (3.3) for $Q_b = 5$ with

$$\begin{aligned} &\theta_b^1(p) = \frac{p_1 - p_2}{\bar{p_1} - \bar{p_2}}, \qquad \theta_b^2(p) = \frac{p_2}{\bar{p_2}}, \qquad \theta_b^3(p) = \frac{p_3 - p_4 - p_1 - p_2}{\bar{p}_3 - \bar{p}_4 - \bar{p}_1 - \bar{p}_2}, \\ &\theta_b^4(p) = \frac{p_4}{\bar{p_4}}, \qquad \theta_b^5(p) = \frac{1 - p_3 - p_4}{1 - \bar{p_3} - \bar{p_4}}, \end{aligned}$$

and

$$b^m(\phi) = \int_{\Omega^m} \int_{-1}^{1} q\,\phi\,\mathrm{d}x\,\mathrm{d}\mu, \quad m = 1,\ldots,5.$$

Similarly, we obtain the affine decomposition (3.2) for $Q_a = 6$ with

$$\theta_a^1(p) = \frac{p_1 - p_2}{\bar{p_1} - \bar{p_2}}, \qquad \theta_a^2(p) = \frac{p_2}{\bar{p_2}}, \qquad \theta_a^3(p) = \frac{p_3 - p_4 - p_1 - p_2}{\bar{p_3} - \bar{p_4} - \bar{p_1} - \bar{p_2}},$$
$$\theta_a^4(p) = \frac{p_4}{\bar{p_4}}, \qquad \theta_a^5(p) = \frac{1 - p_3 - p_4}{1 - \bar{p_3} - \bar{p_4}}, \qquad \theta_a^6(p) = 1,$$

and

$$a^m(\psi, \phi) = \int_{\Omega^m} \int_{-1}^{1} \sigma_t \, \phi \, \psi \, \mathrm{d}x \, \mathrm{d}\mu - \int_{\Omega^m} \int_{-1}^{1} \sigma_s \left\{ \int s(\mu, \mu') \phi(\mu') \mathrm{d}\mu' \right\} \psi \, \mathrm{d}x \, \mathrm{d}\mu,$$

for $m = 1, \ldots, 5$; and

$$a^6(\psi, \phi) = \int_{-1}^{1} \left[\int_{\Omega} \mu \left(\phi^- \, \partial_x \psi^+ - \psi^- \, \partial_x \phi^+ \right) \mathrm{d}x + \sum_{x \in \{0,1\}} |\mu| \, \phi^+(x, \mu) \, \psi^+(x, \mu) \right] \mathrm{d}\mu.$$

The parameterized version of the weak formulation (2.11) can thus be stated as follows: given a parameter $\mu \in \mathcal{P}$ (and a control $q \in Y$), $\psi \in \mathcal{X}$ satisfies

$$a(\psi, \phi; p) = b(q, \phi; p), \quad \forall \phi \in \mathcal{X}.$$

We note that the parameterized bilinear form a satisfies the inf-sup condition

$$\beta(p) = \inf_{\psi \in \mathcal{X}} \sup_{\phi \in \mathcal{X}} \frac{a(\psi, \phi; p)}{\|\psi\|_{\mathcal{X}} \|\phi\|_{\mathcal{X}}} > 0, \tag{3.4}$$

where $\beta(p)$ is now parameter-dependent; similarly for the continuity of a and b.

3.2. *Problem Statement*

We are now ready to state the parameterized optimal control problem as

$$\begin{aligned} \min_{\psi \in \mathcal{X}, q \in Y} J(\psi, q; p) = \frac{1}{2} \int_{\mathcal{I}(p)} \alpha \left(\mathcal{L}\psi - \bar{D} \right)^2 \mathrm{d}x + \frac{1}{2} \int_{\mathcal{I}(p)} q(x)^2 \mathrm{d}x \\ \text{s.t. } (\psi, q) \in \mathcal{X} \times Y \text{ solves } a(\psi, \phi; p) = b(q, \phi; p), \ \forall \phi \in \mathcal{X}. \end{aligned} \tag{3.5}$$

Note that the domain of integration, $\mathcal{I}(p)$, now also depends on the parameter p. It follows from Proposition 2.1 that there exists a unique optimal solution (ψ^*, q^*) to (3.5). Employing a Lagrangian approach, we again obtain the first-order optimality system consisting of the state equation, the adjoint equation, and the optimality equation: Given a parameter $p \in \mathcal{P}$, the optimal solution $(\psi^*, \lambda^*, q^*) \in \mathcal{X} \times \mathcal{X} \times Y$ satisfies

$$a(\psi^*, \phi; p) = b(q^*, \phi; p), \qquad \forall \phi \in \mathcal{X}, \tag{3.6a}$$
$$a(\phi, \lambda^*; p) = \left(\alpha(\bar{D} - \mathcal{L}\psi^*), \mathcal{L}\phi \right)_{Y(p)}, \qquad \forall \phi \in \mathcal{X}, \tag{3.6b}$$
$$(q^*, \varphi)_{Y(p)} - b(\varphi, \lambda^*; p) = 0, \qquad \forall \phi \in Y. \tag{3.6c}$$

3.3. P_N - *Expansion*

In general, we of course cannot expect to find an analytic solution to (3.6). We thus consider a tensor product approximation of (3.6): We employ a spectral $\mathrm{P_N}$ method for the angular domain and a finite element approach for the spatial domain. Since the state and the adjoint equations of the system (3.6) are structurally similar, we only present the numerical scheme for the state equation

$$a(\psi, \phi, p) = b(q, \phi; p), \quad \forall \phi \in \mathcal{X} \text{ and } q \in Y. \tag{3.7}$$

We express the angular dependence of the distribution function in terms of a Fourier series, i.e.,

$$\psi(x,\mu) = \sum_{\ell=0}^{\infty} \psi_\ell(x) \tfrac{2\ell+1}{2} P_\ell(\mu) \quad \text{and} \quad \phi(x,\mu) = \phi_k(x) P_k(\mu), \tag{3.8}$$

where P_ℓ are the Legendre polynomials of order ℓ and $\psi_\ell(x) \in Y$ for $\ell \in \mathbb{Z}$. Using properties of Legendre polynomials, we derive the even and odd components of ψ as

$$\psi^+ = \sum_{\ell} \psi_{2\ell}(x) \tfrac{4\ell+1}{2} P_{2\ell}(\mu) \quad \text{and} \quad \psi^- = \sum_{\ell} \psi_{2\ell+1}(x) \tfrac{4\ell+3}{2} P_{2\ell+1}(\mu) \tag{3.9}$$

By substituting the expansion into (3.7) we obtain (by suppressing the arguments and using properties of Legendre polynomials) the parameter dependent $\mathrm{P_N}$ equation

$$\sum_{m=1}^{Q_a} \Theta_a^m(p) a^m(\psi,\phi) = \sum_{m=1}^{Q_q} \Theta_q^m(p) b^m(q,\phi) \tag{3.10}$$

where $a^m(\psi,\phi)$ and $b^m(\psi,\phi)$ are defined as follows: for m = 1,...,6

$$b^m(q,\phi) = \int_{\Omega^m} q_k(x)\phi_k(x)\mathrm{d}x, \quad \text{for } k = 0, ...,$$

for m = 1,...,5

$$a^m(\psi,\phi) = \sigma_{tk}^m \int_{\Omega^m} \psi_k(x)\phi_k(x)\mathrm{d}x,$$

for m = 6 and k odd

$$a^6(\psi,\phi) = \int_\Omega \phi_k(x) \left[\tfrac{k+1}{2k+1} \partial_x \psi_{k+1}(x) + \tfrac{k}{2k+1} \partial_x \psi_{k-1}(x) \right] \mathrm{d}x,$$

for m = 6 and k even

$$\begin{aligned} a^6(\psi,\phi) = &- \int_\Omega \partial_x \phi_k(x) \left[\tfrac{k+1}{2k+1} \psi_{k+1}(x) + \tfrac{k}{2k+1} \psi_{k-1}(x) \right] \mathrm{d}x \\ &+ \sum_{\ell=0}^{\infty} \int_{-1}^{1} |\mu|\, \phi_k(0)\psi_\ell(0) \tfrac{2\ell+1}{2} P_\ell P_k \mathrm{d}\mu + \sum_{\ell=0}^{\infty} \int_{-1}^{1} |\mu|\, \phi_k(1)\psi_l(1) \tfrac{2\ell+1}{2} P_\ell P_k d\mu, \end{aligned}$$

where

$$\sigma_{tk}^m = \sigma_t^m + \sigma_{sk}^m, \quad \text{and} \quad \sigma_{sk}^m = 2\pi\sigma_s \int_{-1}^{1} P_k(\mu) s(\mu) \mathrm{d}\mu.$$

3.4. *Full discretization*

To approximate the solution of the $\mathrm{P_N}$ equations in the spatial domain we introduce a continuous and piecewise linear finite element approximation space of dimension N_h for the coefficient functions $\psi_\ell(x)$ and $q_\ell(x)$ in the expansion. These spaces are closed finite subspaces of $\mathcal{X}$ and Y, and defined as $\mathcal{X}_h := \mathcal{X}^+_{h,N} \oplus \mathcal{X}^-_{h,N}$ and $\mathcal{Y}_h \in Y$ respectively, where

$$\mathcal{X}^+_{h,N} := \{\psi_{\ell,i} : \psi(x,\mu) = \sum_{\ell=0}^{N}\sum_{i=0}^{N_h} \psi^i_{2\ell}\tfrac{4\ell+1}{2} P_{2\ell}(\mu)H_i(x)\},$$

$$\mathcal{X}^-_{h,N} := \{\psi_{\ell,i} : \psi(x,\mu) = \sum_{\ell=0}^{N}\sum_{i=0}^{N_h} \psi^i_{2\ell+1}\tfrac{4\ell+3}{2} P_{2\ell+1}(\mu)H_i(x)\}.$$

Here, $H_i(x)$ are the usual one-dimensional hat functions. Note that we employ the same finite element space for both even $\psi_{2\ell}(x)$ and odd $\psi_{2\ell+1}(x)$ components. The finite element approximation of the $\mathrm{P_N}$ equation serves as our "truth" approximation, i.e., we build the reduced basis approximation upon this $\mathrm{P_N}$ -FE approximation and shall measure the error with respect to the solution of the $\mathrm{P_N}$ -FE approximation.

Using this discretization, we obtain fully discrete optimality conditions. Given a parameter $p \in \mathcal{P}$, the truth optimal solution $(\psi^*_h, \lambda^*_h, q^*_h) \in \mathcal{X}_h \times \mathcal{X}_h \times \mathcal{Y}_h$ satisfies

$$a(\psi^*_h, \phi_h; p) = b(q^*_h, \phi_h; p), \qquad \forall \phi_h \in \mathcal{X}_h, \tag{3.11a}$$

$$a(\phi_h, \lambda^*_h; p) = \left(\alpha(\bar{D} - \mathcal{L}\psi^*_h), \mathcal{L}\phi_h\right)_{Y(p)}, \qquad \forall \phi_h \in \mathcal{X}_h, \tag{3.11b}$$

$$(q^*_h, \varphi_h)_{Y(p)} - b(\varphi_h, \lambda^*_h; p) = 0 \qquad \forall \varphi_h \in \mathcal{Y}_h, \tag{3.11c}$$

The optimality system is dimension of $(2N+1)N_h + N_h$, where N_h is the FE dimension, and N is number of terms in the $\mathrm{P_N}$ expansion.

Remark 3.1. We note that the fully discrete problem inherits the stability from the continuous setting since we have $\mathcal{A}\psi^+_h \in \mathcal{X}^-_h$ for every $\psi^+_h \in \mathcal{X}^+_h$. The discrete inf-sup condition

$$\beta_h(p) = \inf_{\psi_h \in \mathcal{X}_h} \sup_{\phi_h \in \mathcal{X}_h} \frac{a(\psi_h, \phi_h; p)}{\|\psi_h\|_{\mathcal{X}} \|\phi_h\|_{\mathcal{X}}} > 0 \tag{3.12}$$

thus holds for the FE-$\mathrm{P_1}$ approximation with the same constant as for the continuous problem; see Ref. 12 for details.

Remark 3.2. From the definition of the Legendre polynomials it directly follows that

$$\int_{-1}^{1} \psi(x,\mu)\mathrm{d}\mu = \int_{-1}^{1} \psi(x) P_0(\mu)\mathrm{d}\mu = \psi_0(x). \tag{3.13}$$

Therefore, $\mathcal{L}\psi(x,\mu) = \psi_0(x)$.

4. Reduced Basis Method

We now turn to the RB method. We first introduce the RB approximation of the truth optimality system (3.11), subsequently develop associated rigorous *a posteriori* error bounds, and finally comment on the offline-online decomposition.

4.1. *Approximation*

4.1.1. *Reduced Basis Spaces*

To begin, we assume that we are given the integrated reduced basis spaces

$$\mathcal{X}_{\mathcal{N}} = \text{span}\{\eta_n,\, 1 \le n \le \mathcal{N}\}, \quad 1 \le \mathcal{N} \le \mathcal{N}_{\max}, \tag{4.1}$$

where the η_n, $1 \le n \le \mathcal{N}$, are mutually $(\cdot,\cdot)_{\mathcal{X}}$-orthogonal basis functions. Furthermore, we assume that the reduced basis control spaces are given by

$$\mathcal{Y}_{\mathcal{M}} = \text{span}\{\xi_{\mathcal{M}},\, 1 \le m \le \mathcal{M}\}, \quad 1 \le \mathcal{M} \le \mathcal{M}_{\max}, \tag{4.2}$$

where the ξ_m, $1 \le m \le \mathcal{M}$, are mutually $(\cdot,\cdot)_Y$-orthogonal basis functions. In this work, we consider the proper orthogonal decomposition (POD) to generate these spaces: Let $\text{POD}_{\mathcal{X}}(\{v_k \in \mathcal{X} : 1 \le k \le n_{\text{train}}\}, \mathcal{N})$ return the $\mathcal{N}$ largest POD modes with respect to the $(\cdot,\cdot)_{\mathcal{X}}$ inner product (and normalized with respect to the $\mathcal{X}$-norm), and $\text{POD}_Y(\{v_k \in Y : 1 \le k \le n_{\text{train}}\}, \mathcal{M})$ return the $\mathcal{M}$ largest POD modes with respect to the $(\cdot,\cdot)_Y$ inner product (and normalized with respect to the Y-norm). We then introduce a finite but suitably large parameter train set $\Xi_{\text{train}} \subset \mathcal{P}$ of size n_{train} and define

$$\mathcal{X}_{\mathcal{N}_{\max}} = \text{POD}_{\mathcal{X}}(\{\psi^*(p) \in \mathcal{X}_h, \lambda^*(p) \in \mathcal{X}_h : p \in \Xi_{\text{train}}\}, \mathcal{N}_{\max}) \tag{4.3}$$

and

$$\mathcal{Y}_{\mathcal{M}_{\max}} = \text{POD}_Y(\{q^*(p) \in \mathcal{Y}_h : p \in \Xi_{\text{train}}\}, \mathcal{M}_{\max}). \tag{4.4}$$

Here, $\mathcal{N}_{\max}$ and $\mathcal{M}_{\max}$ are chosen based on the decay of the singular values in order to guarantee a desired approximation quality of the reduced spaces $\mathcal{X}_{\mathcal{N}}$ and $\mathcal{Y}_{\mathcal{M}}$, i.e., we require the sum of the neglected squared singular values to be less than some tolerance[22,45]. We note that, although the dimensions $\mathcal{N}_{\max}$ of the state/adjoint space and $\mathcal{M}_{\max}$ of the control space are, in principle, independent of each other, they both depend on the parametric complexity of the problem. For our numerical test case discussed in Section 5, for example, we obtain the same approximation error – again, based on the decay of the singular values – for $\mathcal{N} \approx 1.6\mathcal{M}$.

The generation of the POD basis requires n_{train} solutions of the truth optimality system (3.11) and may thus be expensive for large n_{train}. Another option to generate the reduced basis spaces is the greedy sampling procedure discussed in Ref. 28, which is generally more efficient thus preferred for large train samples.

Finally, we note that the primal and dual solutions are integrated in the space $\mathcal{X}_{\mathcal{N}}$. There are two reasons to do so: the first one is related to the stability of the reduced basis optimality system and discussed below in Remark 4.1; the second one

is due to the derivation of the *a posteriori* error bound for the cost functional in Section 4.2.2. To this end, we use the dual weighted residual approach from Ref. 5 which requires that the state ansatz space coincides with the adjoint test space (and vice versa). In order to fulfill this requirement on the spaces in combination with the stability considerations discussed below, we have to use an integrated state and adjoint space $\mathcal{X}_\mathcal{N}$.

4.1.2. *Galerkin Projection*

We next replace the truth approximation of the PDE constraint in (3.5) by its reduced basis approximation. The reduced basis optimal control problem is thus given by

$$\begin{aligned} \min_{\psi_\mathcal{N}\in\mathcal{X}_\mathcal{N}, q_\mathcal{M}\in Y_\mathcal{M}} J(\psi_\mathcal{N}, q_\mathcal{M}; p) &= \frac{1}{2}\int_{\mathcal{I}(p)} \alpha\left(\mathcal{L}\psi_\mathcal{N} - \bar{D}\right)^2 \mathrm{d}x + \frac{1}{2}\int_{\mathcal{I}(p)} q_\mathcal{M}^2 \mathrm{d}x \\ \text{s.t. } (\psi_\mathcal{N}, q_\mathcal{M}) \in \mathcal{X}_\mathcal{N} \times Y_\mathcal{M} &\text{ solves } a(\psi_\mathcal{N}, \phi; p) = b(q_\mathcal{M}, \phi; p), \forall \phi \in \mathcal{X}_\mathcal{N}. \end{aligned} \tag{4.5}$$

We can also directly state the associated first-order optimality system: For a given parameter $p \in \mathcal{P}$, the optimal solution $(\psi_\mathcal{N}^*, \lambda_\mathcal{N}^*, q_\mathcal{M}^*) \in \mathcal{X}_\mathcal{N} \times \mathcal{X}_\mathcal{N} \times Y_\mathcal{M}$ satisfies the system of equations

$$a(\psi_\mathcal{N}^*, \phi; p) = b(q_\mathcal{M}^*, \phi; p), \qquad \forall \phi \in \mathcal{X}_\mathcal{N}, \tag{4.6a}$$

$$a(\phi, \lambda_\mathcal{N}^*; p) = \left(\alpha(\bar{D} - \mathcal{L}\psi_\mathcal{N}^*), \mathcal{L}\phi\right)_{Y(p)}, \qquad \forall \phi \in \mathcal{X}_\mathcal{N}, \tag{4.6b}$$

$$(q_\mathcal{M}^*, \varphi)_{Y(p)} - b(\varphi, \lambda_\mathcal{N}^*; p) = 0, \qquad \forall \varphi \in Y_\mathcal{M}. \tag{4.6c}$$

The reduced basis optimality system is only of dimension $2\mathcal{N} + \mathcal{M}$ and can be evaluated efficiently using an offline–online computational decomposition, see Section 4.3.

Remark 4.1. There are two necessary ingredients to guarantee the stability of the reduced basis optimality system: the stability of the reduced state and adjoint equation, and the stability of the overall optimality system (4.6). Concerning the latter, we note that using integrated spaces for the state and adjoint preserves the stability of the reduced optimality system[34].

Concerning the former, we first recall the difference between coercive and noncoercive (inf-sup stable) problems. For coercive problems, the stability of the reduced basis approximation is directly inherited from the high-dimensional problem and thus guaranteed if a standard Galerkin projection is used. For noncoercive problems such as the state and adjoint equations in (4.6), however, a standard Galerkin projection does not guarantee the stability of the reduced basis approximation, i.e., $\beta_h(p) > 0$ in (3.12) does not generally imply that $\beta_\mathcal{N}(p) > 0$, where

$$\beta_\mathcal{N}(p) = \inf_{\psi\in\mathcal{X}_\mathcal{N}} \sup_{\phi\in\mathcal{X}_\mathcal{N}} \frac{a(\psi_\mathcal{N}, \phi_\mathcal{N}; p)}{\|\psi_\mathcal{N}\|_\mathcal{X}\|\phi_\mathcal{N}\|_\mathcal{X}} \tag{4.7}$$

is the inf-sup constant associated with the reduced basis approximation. More sophisticated approaches, e.g. enriching the test space with supremizer functions and

employing a Petrov-Galerkin approximation, do restore guaranteed stability at some additional complexity and cost[38,40,18].

In this paper, however, we will consider only the standard Galerkin approach and verify the stability numerically. The reason is that although the stability of a Galerkin reduced basis approximation cannot be theoretically proven, many problems turn out to be numerically stable in practice; see Ref. 33 for a detailed discussion. The addition of supremizers would thus only increase the computational cost, but have not benefit on the approximation property or the stability.

To provide numerical evidence we compute the inf-sup constant (4.7) for a regular parameter grid of size $5^4 = 625$ and different RB dimensions $\mathcal{N}$. The minimum and maximum value of the inf-sup constant $\beta_h(p)$ for the P_1 - FE approximation observed over the parameter sample are 3.3323 and 4.7027, respectively. We present the minimum value of the reduced basis inf-sup constant $\beta_{\mathcal{N}}(p)$ for different $\mathcal{N}$ in Table 1. We observe that $\beta_{\mathcal{N}}(p) > 0$ is indeed satisfied and that $\beta_{\mathcal{N}}(p)$ is very close to the minimum of $\beta_h(p)$ for $\mathcal{N} \geq 25$.

$\mathcal{N}$	5	10	15	25	35
$\min_{p\in\mathcal{P}_{\mathrm{ex}}} \beta_{\mathcal{N}}(p)$	3.0183	3.258	3.321	3.3323	3.3323

Table 1: Minimum inf-sup constants $\beta_{\mathcal{N}}(p)$ over Ξ_{test} for various values of $\mathcal{N}$.

4.2. *A Posteriori Error Estimation*

We turn to the *a posteriori* error estimation procedure. We follow the ideas presented in Ref. 28 and extend the bound introduced there to the radiation treatment planning problem.

To begin, we assume that we are given a positive lower bound $\beta_{\mathrm{LB}}(p) : \mathcal{P} \to \mathbb{R}$ for the inf-sup constant $\beta_h(p) > 0$ defined in (3.12). Furthermore, we assume that we have upper bounds available for the (parameter-dependent) constant

$$\gamma_{\mathcal{L}}^{\mathrm{UB}}(p) \geq \|\mathcal{L}\|_{\mathcal{X}\to Y(p)} := \sup_{\phi\in\mathcal{X}} \frac{\|\mathcal{L}\phi\|_{Y(p)}}{\|\phi\|_{\mathcal{X}}}, \tag{4.8}$$

and the continuity constant of the bilinear form $b(\cdot,\cdot;p)$

$$\gamma_b^{\mathrm{UB}}(p) \geq \gamma_b(p) := \sup_{\psi\in\mathcal{Y}_h} \sup_{\phi\in\mathcal{X}_h} \frac{b(\psi,\phi;p)}{\|\psi\|_{Y(p)}\|\phi\|_{\mathcal{X}}} > 0, \quad \forall p \in P. \tag{4.9}$$

We also define the state, adjoint, and control optimality errors as $e^{\psi}(p) = \psi_h^*(p) - \psi_{\mathcal{N}}^*(p)$, $e^{\lambda}(p) = \lambda_h^*(p) - \lambda_{\mathcal{N}}^*(p)$, and $e^{q}(p) = q_h^*(p) - q_{\mathcal{M}}^*(p)$, respectively. In the following we often drop the dependence on p to simplify notation.

4.2.1. *Error Bound for Optimal Control*

We obtain the following result for the error in the optimal control.

Proposition 4.1. *Let q_h^* and $q_{\mathcal{M}}^*$ be the optimal solutions to the truth and reduced basis optimal control problem, respectively. The error in the optimal control satisfies for all parameters $p \in \mathcal{P}$*

$$\|q_h^* - q_{\mathcal{M}}^*\|_{Y(p)} \leq \Delta_{\mathcal{M}}^q(p) \equiv \frac{1}{2}\left(\|r_q(\cdot;p)\|_{Y(p)'} + \frac{\gamma_b^{\mathrm{UB}}(p)}{\beta_{LB}(p)}\|r_\lambda(\cdot;p)\|_{\mathcal{X}'}\right) + \frac{1}{2}\left[\left(\frac{8}{\beta_{LB}(p)}\|r_\psi(\cdot;p)\|_{\mathcal{X}'}\|r_\lambda(\cdot;p)\|_{\mathcal{X}'} + \frac{\|\sqrt{\alpha}\|_{L^\infty(\mathcal{I})}^2\gamma_{\mathcal{L}}^{\mathrm{UB}}(p)^2}{\beta_{LB}(p)^2}\|r_\psi(\cdot;p)\|_{\mathcal{X}'}^2\right) + \left(\|r_q(\cdot;p)\|_{Y(p)'} + \frac{\gamma_b^{\mathrm{UB}}(p)}{\beta_{LB}(p)}\|r_\lambda(\cdot;p)\|_{\mathcal{X}'}\right)^2\right]^{\frac{1}{2}}. \quad (4.10)$$

Proof. We first note that the state error satisfies the error residual equation

$$a(e^\psi, \phi; p) = r_\psi(\phi; p) + b(e^q, \phi; p), \quad \forall \phi \in \mathcal{X}_h, \quad (4.11)$$

where the residual of the state equation is given by $r_\psi(\phi, p) = b(q_{\mathcal{M}}^*, \phi, p) - a(\psi_{\mathcal{N}}^*, \phi, p),\ \forall \phi \in \mathcal{X}_h$. Choosing $\phi = T_p e^\psi$ as test function in (4.11), where the supremizer $T_p : \mathcal{X}_h \to \mathcal{X}_h$ is given by $(T_p w, \phi)_{\mathcal{X}} = a(w, \phi; p),\ \forall \phi \in \mathcal{X}_h$, we obtain

$$\beta_{\mathrm{LB}}(p)\|T_p e^\psi\|_{\mathcal{X}}\|e^\psi\|_{\mathcal{X}} \leq a(e^\psi, T_p e^\psi; p) = r_\psi(T_p e^\psi; p) + b(e^q, T_p e^\psi; p).$$

Invoking (4.9) it follows that

$$\|e^\psi\|_{\mathcal{X}} \leq \frac{1}{\beta_{\mathrm{LB}}(p)}(\|r_\psi(\cdot;p)\|_{\mathcal{X}'} + \gamma_b^{\mathrm{UB}}(p)\|e^q\|_{Y(p)}). \quad (4.12)$$

Similarly, we note that the adjoint error satisfies the error residual equation

$$a(\phi, e^\lambda; p) = r_\lambda(\phi; p) - \left(\alpha\mathcal{L}e^\psi, \mathcal{L}\phi\right)_{Y(p)}, \quad \forall \phi \in \mathcal{X}_h, \quad (4.13)$$

where the adjoint residual is given by $r_\lambda(\phi, p) = \left(\alpha(\bar{D} - \mathcal{L}\psi_{\mathcal{N}}^*), \mathcal{L}\phi\right)_{Y(p)} - a(\phi, \lambda_{\mathcal{N}}^*; p),\ \forall \phi \in \mathcal{X}_h$. Choosing as test function $\phi = T_p e^\lambda$ in (4.13) and noting from (4.8) and the Cauchy-Schwarz Inequality that

$$\left(\alpha\mathcal{L}e^\psi, \mathcal{L}T_p e^\lambda\right)_{Y(p)} \leq \|\sqrt{\alpha}\|_{L^\infty(\mathcal{I})}\,\gamma_{\mathcal{L}}^{\mathrm{UB}}(p)\,\|\sqrt{\alpha}\mathcal{L}e^\psi\|_{Y(p)}\,\|T_p e^\lambda\|_{\mathcal{X}} \quad (4.14)$$

we arrive at

$$\|e^\lambda\|_{\mathcal{X}} \leq \frac{1}{\beta_{\mathrm{LB}}(p)}(\|r_\lambda(\cdot;p)\|_{\mathcal{X}'} + \|\sqrt{\alpha}\|_{L^\infty(\mathcal{I})}\gamma_{\mathcal{L}}^{\mathrm{UB}}(p)\|\sqrt{\alpha}\mathcal{L}e^\psi\|_{Y(p)}). \quad (4.15)$$

We next note that the error in the optimal control satisfies

$$(e^q, \varphi)_{Y(p)} - b(\varphi, e^\lambda; p) = r_q(\varphi; p), \quad \forall \varphi \in \mathcal{Y}_h, \quad (4.16)$$

where the residual is given by $r_q(\varphi, p) = b(\varphi, \lambda^*_{\mathcal{N}}; p) - (q^*_{\mathcal{M}}, \varphi)_{Y(p)},\ \forall \varphi \in \mathcal{Y}_h$. We choose as test functions $\phi = e^\lambda$ in (4.11), $\phi = e^\psi$ in (4.13), and $\varphi = e^q$ in (4.16) and obtain

$$a(e^\psi, e^\lambda; p) - b(e^q, e^\lambda; p) = r_\psi(e^\lambda; p), \tag{4.17}$$

$$a(e^\psi, e^\lambda; p) + \left(\alpha \mathcal{L} e^\psi, \mathcal{L} e^\psi\right)_{Y(p)} = r_\lambda(e^\psi; p), \tag{4.18}$$

$$(e^q, e^q)_{Y(p)} - b(e^q, e^\lambda; p) = r_q(e^q; p). \tag{4.19}$$

Adding (4.19) and (4.18) and subtracting (4.17) yields

$$(e^q, e^q)_{Y(p)} + \left(\alpha \mathcal{L} e^\psi, \mathcal{L} e^\psi\right)_{Y(p)} = -r_\psi(e^\lambda; p) + r_\lambda(e^\psi; p) + r_q(e^q; p) \tag{4.20}$$

and hence

$$\begin{aligned}\|e^q\|^2_{Y(p)} + \|\sqrt{\alpha}\mathcal{L} e^\psi\|^2_{Y(p)} \le \|r_\psi(\cdot; p)\|_{\mathcal{X}'} \|e^\lambda\|_{\mathcal{X}} \\ + \|r_\lambda(\cdot; p)\|_{\mathcal{X}'} \|e^\psi\|_{\mathcal{X}} + \|r_q(\cdot; p)\|_{Y(p)'} \|e^q\|_{Y(p)}.\end{aligned} \tag{4.21}$$

By substituting (4.12) and (4.15) into(4.21) we obtain

$$\begin{aligned}\|e^q\|^2_{Y(p)} + \|\sqrt{\alpha}\mathcal{L} e^\psi\|^2_{Y(p)} \le\ & \|r_q(\cdot; p)\|_{Y(p)'} \|e^q\|_{Y(p)} \\ & + \|r_\psi(\cdot; p)\|_{\mathcal{X}'} \frac{1}{\beta_{\mathrm{LB}}(p)} (\|r_\lambda\|_{\mathcal{X}'} + \|\sqrt{\alpha}\|_{L^\infty(\mathcal{I})} \gamma^{\mathrm{UB}}_{\mathcal{L}}(p) \|\sqrt{\alpha}\mathcal{L} e^\psi\|_{Y(p)}) \\ & + \|r_\lambda(\cdot; p)\|_{\mathcal{X}'} \frac{1}{\beta_{\mathrm{LB}}(p)} (\|r_\psi\|_{\mathcal{X}'} + \gamma^{\mathrm{UB}}_b(p) \|e^q\|_{Y(p)}).\end{aligned}$$

Furthermore, it follows from Young's inequality that

$$\begin{aligned}\frac{\|\sqrt{\alpha}\|_{L^\infty(\mathcal{I})} \gamma^{\mathrm{UB}}_{\mathcal{L}}(p)}{\beta_{\mathrm{LB}}(p)} \|r_\psi(\cdot; p)\|_{\mathcal{X}'} \|\sqrt{\alpha}\mathcal{L} e^\psi\|_{Y(p)} \\ \le \frac{\|\sqrt{\alpha}\|^2_{L^\infty(\mathcal{I})} \gamma^{\mathrm{UB}}_{\mathcal{L}}(p)^2}{4\beta_{\mathrm{LB}}(p)^2} \|r_\psi(\cdot; p)\|^2_{\mathcal{X}'} + \|\sqrt{\alpha}\mathcal{L} e^\psi\|^2_{Y(p)}.\end{aligned}$$

By combining the last two equations we obtain

$$\begin{aligned}\|e^q\|^2_{Y(p)} \le & \|r_q(\cdot; p)\|_{Y(p)'} \|e^q\|_{Y(p)} + \frac{2}{\beta_{\mathrm{LB}}(p)} \|r_\psi(\cdot; p)\|_{\mathcal{X}'} \|r_\lambda(\cdot; p)\|_{\mathcal{X}'} \\ & + \frac{\gamma^{\mathrm{UB}}_b(p)}{\beta_{\mathrm{LB}}(p)} \|r_\lambda(\cdot; p)\|_{\mathcal{X}'} \|e^q\|_{Y(p)} + \frac{\|\sqrt{\alpha}\|^2_{L^\infty(\mathcal{I})} \gamma^{\mathrm{UB}}_{\mathcal{L}}(p)^2}{4\beta_{\mathrm{LB}}(p)^2} \|r_\psi(\cdot; p)\|^2_{\mathcal{X}'}\end{aligned}$$

This can be written in the form of a quadratic inequality for $\|e^q\|_{Y(p)}$ by

$$A\|e^q\|^2_{Y(p)} + B\|e^q\|_{Y(p)} + C \le 0 \tag{4.22}$$

where

$$\begin{aligned}A &= 1 \\ B &= -\left(\|r_q(\cdot; p)\|_{Y(p)'} + \frac{\gamma^{\mathrm{UB}}_b(p)}{\beta_{\mathrm{LB}}} \|r_\lambda(\cdot; p)\|_{\mathcal{X}'}\right) \\ C &= -\left(\frac{2}{\beta_{\mathrm{LB}}(p)} \|r_\psi(\cdot; p)\|_{\mathcal{X}'} \|r_\lambda(\cdot; p)\|_{\mathcal{X}'} + \frac{\|\sqrt{\alpha}\|^2_{L^\infty(\mathcal{I})} \gamma^{\mathrm{UB}}_{\mathcal{L}}(p)^2}{4\beta_{\mathrm{LB}}(p)^2} \|r_\psi(\cdot; p)\|^2_{\mathcal{X}'}\right)\end{aligned}$$

which is satisfied iff

$$\Delta_{\mathcal{N}}^{-} \leq \|e^q\|_{Y(p)} \leq \Delta^{+} \quad \text{with} \quad \Delta^{\pm} = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}.$$

The result follows by setting $\Delta_{\mathcal{N}}^{q}(p) = \Delta_{\mathcal{N}}^{+}$. □

4.2.2. *Error Bound for Cost Functional*

Given the error bound for the optimal control we may readily derive a bound for the error in the cost functional. To this end, we follow the approach described in Ref. 28 and first derive a bound for the optimality error in the state and adjoint solution.

Lemma 4.1. *The state optimality error, $e^{\psi} = \psi^* - \psi_{\mathcal{N}}^*$, is bounded by*

$$\|e^{\psi}\|_{\mathcal{X}} \leq \Delta_{\mathcal{N}}^{\psi}(p) = \frac{1}{\beta_{LB}(p)}(\|r_{\psi}(\cdot;p)\|_{\mathcal{X}'} + \gamma_b^{\mathrm{UB}}(p)\Delta_{\mathcal{M}}^{q}(p)), \quad \forall p \in \mathcal{P} \tag{4.23}$$

Proof. The proof follows from the inequality (4.12) and Proposition 4.1. □

Lemma 4.2. *The adjoint optimality error, $e^{\lambda} = \lambda^* - \lambda_{\mathcal{N}}^*$, is bounded by*

$$\|e^{\lambda}\|_{\mathcal{X}} \leq \Delta_{\mathcal{N}}^{\lambda}(p) = \frac{1}{\beta_{LB}(p)}(\|r_{\lambda}(\cdot;p)\|_{\mathcal{X}'} + \|\alpha\|_{L^{\infty}(\mathcal{I})}\,\gamma_{\mathcal{L}}^{\mathrm{UB}}(p)^2\,\Delta_N^{\psi}(p)), \quad \forall p \in \mathcal{P} \tag{4.24}$$

Proof. We follow the same steps leading to inequality (4.15), but note that

$$\left(\alpha\mathcal{L}e^{\psi}, \mathcal{L}T_p e^{\lambda}\right)_{Y(p)} \leq \|\alpha\|_{L^{\infty}(\mathcal{I})}\,\gamma_{\mathcal{L}}^{\mathrm{UB}}(p)^2\,\|e^{\psi}\|_{\mathcal{X}}\,\|T_p e^{\lambda}\|_{\mathcal{X}}. \tag{4.25}$$

The result then follows by invoking Lemma 4.1 to bound $\|e^{\psi}\|_{\mathcal{X}}$. □

The error in the cost functional can thus be estimated as follows.

Proposition 4.2. *Let $J^* = J(\psi^*, q^*; p)$ and $J_{\mathcal{N}}^* = J(\psi_{\mathcal{N}}^*, q_{\mathcal{N}}^*; p)$ be the optimal cost functional value for the truth and RB solutions. The error then satisfies*

$$\begin{aligned} |J^* - J_{\mathcal{N}}^*| \leq \Delta_{\mathcal{N}}^{J}(p) = \frac{1}{2}\Big(&\|r_{\psi}(\cdot;p)\|_{\mathcal{X}'}\Delta_{\mathcal{N}}^{\lambda}(p) + \|r_{\lambda}(\cdot;p)\|_{\mathcal{X}'}\Delta_{\mathcal{N}}^{\psi}(p) \\ &+ \|r_q(\cdot;p)\|_{Y(p)'}\Delta_{\mathcal{M}}^{q}(p)\Big), \quad \forall p \in \mathcal{P} \end{aligned} \tag{4.26}$$

Proof. We use the standard result from Ref. 5 to bound the cost functional error by

$$\begin{aligned} |J^* - J_{\mathcal{N}}^*| &= \frac{1}{2}\left(r_{\psi}(e^{\lambda};p) + r_{\lambda}(e^{\psi};p) + r_q(e^q;p)\right) \\ &\leq \frac{1}{2}\big(\|r_{\psi}(\cdot;p)\|_{\mathcal{X}'}\|e^{\lambda}\|_{\mathcal{X}} + \|r_{\lambda}(\cdot;p)\|_{\mathcal{X}'}\|e^{\psi}\|_{\mathcal{X}} \\ &\quad + \|r_q(\cdot;p)\|_{Y(p)'}\|e^q\|_{Y(p)}\big), \qquad \forall p \in \mathcal{P}. \end{aligned}$$

The result follow from Lemma 4.1 and 4.2 and Proposition 4.1. □

4.3. *Computational Procedure*

Due to the affine parameter dependence of all linear and bilinear forms, the reduced basis optimality system (4.6) can be efficiently evaluated using the standard offline-online decomposition. We thus skip the details here and refer to Ref. 28, where the computational procedure including computational costs are discussed.

The evaluation of the control and cost functional error bound requires computation of

(1) the dual norms of the state, adjoint, and optimality equation residuals $\|r_\psi(\cdot;p)\|_{\mathcal{X}'}$, $\|r_\lambda(\cdot;p)\|_{\mathcal{X}'}$, and $\|r_q(\cdot;p)\|_{Y(p)'}$; respectively; and
(2) the constants $\beta_{\mathrm{LB}}(p)$, $\gamma_{\mathcal{L}}^{\mathrm{UB}}(p)$, and $\gamma_b^{\mathrm{UB}}(p)$.

All of these quantities can be computed using an offline-online decomposition. Since the approach is also standard, we refer to Ref. 39 for details concerning the dual norms of the state and adjoint residuals and to Ref. 28 for the remaining quantities. We note, however, that the online computational complexity to evaluate all involved quantities depends only on the dimension of the reduced spaces and is *independent* of the dimension of the underlying truth approximation.

5. Numerical Results

We present a numerical example to confirm the validity of the proposed approach. The numerical schemes and methods have been implemented in MATLAB® and are available online[1]. The source package is designed in a such way that different example files do not require changes in solver files(`solver.m`, `RB_solver.m`). For each example we created a separate file (e.g. `ex_forw_constMat.m`) which defines the problem parameters and executes corresponding solver file. The data between the example and solver files (and also various functions) interchanges using a MATLAB® struct `par` which stores all the problem parameters, e.g., the problem name (`par.name`), definition of functions for material parameters (`par.sigma_a`, `par.sigma_sm`), profile of a dose distribution (`par.dose`) and so on; we refer to Ref. 1 for more details.

5.1. *Model Problem*

For our model problem we consider the parameterized geometry introduced in Section 3.2 and sketched in Figure 1. The admissible range $\mathcal{P}_1$ to $\mathcal{P}_4$ of the four geometry parameters p_1 to p_4 are summarized in Table 2, with the full parameter domain given by $\mathcal{P} = \mathcal{P}_1 \times \mathcal{P}_2 \times \mathcal{P}_3 \times \mathcal{P}_4$. We choose as reference parameter $\bar{p} = [0.2; 0.07; 0.8; 0.07]$. The absorption coefficient σ_a, scattering cross-section σ_s, and total cross-section $\sigma_t = \sigma_a + \sigma_s$ are assumed piece-wise constant in the five subdomains introduced in Section 3.2 and are also given in Table 2.

We assume that the tumor is located in the subdomain $\Omega_T = [0.38, 0.62]$ of the reference domain, i.e., we also allow the tumor to be changing in size with the

paramters. The desired dose should be maximal in the tumor region and negligibly small in the other areas. We therefore set the desired dose as sketched in Fig. 2. We set the regularization parameter $\alpha = 10$ over the whole domain.

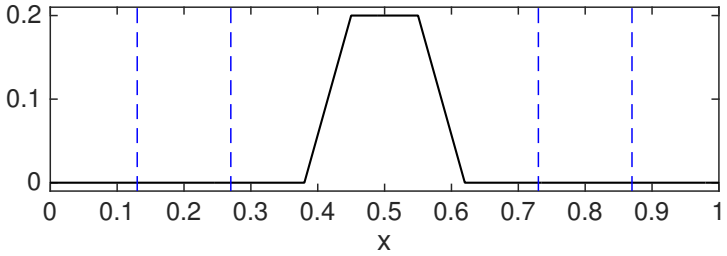


Fig. 2: Profile of parameterized desired dose $\bar{\psi}(x)$ in reference geometry Ω.

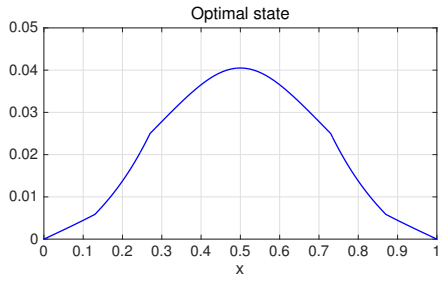
For the truth discretization we employ a piece-wise linear finite element approximation space for the state, adjoint, and control variables. The number of degrees of freedom is $\dim(\mathcal{X}_h) = \dim(\mathcal{Y}_h) = N_h = 200$. The total dimension of the truth optimality system is thus $5N_h = 1000$. In Fig. 3 we present sample solutions for the optimal state and control for two parameter values.

We construct the reduced basis spaces $\mathcal{Y}_\mathcal{M}$ and $\mathcal{X}_\mathcal{N}$ according to the POD procedure described in Section 4. For this purpose, we employ the train sample $\Xi_{\text{train}} = \mathcal{P}_1^{\text{train}} \times \mathcal{P}_2^{\text{train}} \times \mathcal{P}_3^{\text{train}} \times \mathcal{P}_4^{\text{train}}$ consisting of $\text{n}_{\text{train}} = 324$ equidistant parameter points over $\mathcal{P}$. We also introduce a parameter test sample Ξ_{test} of size $\text{n}_{\text{test}} = 100$ with a uniform-random distribution in $\mathcal{P}$. The model problem for these parameters is implemented in the example file `ex_optim_pDose_constMat.m`.

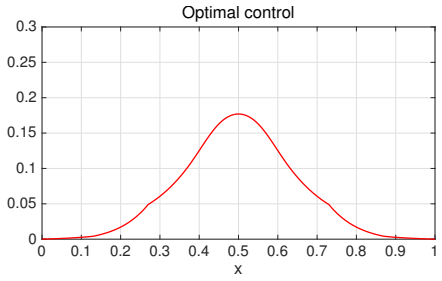
N_h	200
Ω_T	[0.38, 0.62]
Ω_N	(0.1, 0.38)×(0.62, 0.9)
Ω_R	[0, 0.1]× [0.9, 1]
α	10
$\sigma_a^{1,\dots,5}$	[4; 7; 4; 7; 4]
$\sigma_s^{1,\dots,5}$	[1; 2; 1; 2; 1]
$\bar{p}$	[0.2; 0.07; 0.8; 0.07]
Ξ_{test}	100 random parameters

$\mathcal{P}_1$	[0.15 ; 0.25]
$\mathcal{P}_2$	[0.04 ; 0.1]
$\mathcal{P}_3$	[0.15 ; 0.25]
$\mathcal{P}_4$	[0.04 ; 0.1]
$\mathcal{P}_1^{\text{train}}$	[0.15; 0.17; 0.19; 0.21; 0.23; 0.25]
$\mathcal{P}_2^{\text{train}}$	[0.04; 0.07; 0.10]
$\mathcal{P}_3^{\text{train}}$	[0.75; 0.77; 0.79; 0.81; 0.83; 0.85]
$\mathcal{P}_4^{\text{train}}$	[0.04; 0.07; 0.10]
Ξ_{train}	$\mathcal{P}_1^{\text{train}} \times \mathcal{P}_2^{\text{train}} \times \mathcal{P}_3^{\text{train}} \times \mathcal{P}_4^{\text{train}}$

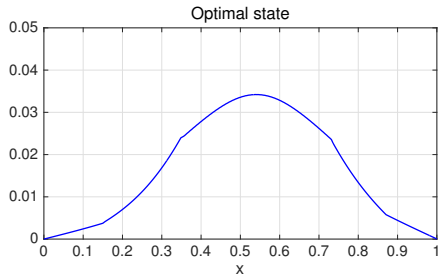
Table 2: Parameters for the numerical example.



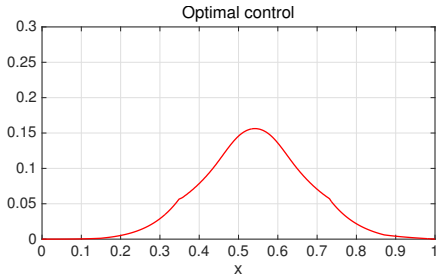
(a) for parameter values $\{0.2; 0.07; 0.8; 0.07\}$



(b) for parameter values $\{0.2; 0.07; 0.8; 0.07\}$



(c) for parameter values $\{0.25; 0.1; 0.8; 0.07\}$



(d) for parameter values $\{0.25; 0.1; 0.8; 0.07\}$

Fig. 3: Zero-th moments of optimal state $\psi_0^*(x)$ and optimal control $q_0^*(x)$ for various parameter values with constant regularization parameter $\alpha = 10$ in the original geometry $\Omega_0(p)$.

5.2. *Error Estimators*

We first consider the performance of the *a posteriori* error bounds. We use a global upper bound for the (parameter-dependent) constants $\gamma_{\mathcal{L}}^{\mathrm{UB}}(p)$ defined in (4.8) and $\gamma_b^{\mathrm{UB}}(p)$ defined in (4.9). To this end, we compute the values for $\|\mathcal{L}\|_{\mathcal{X}\to Y(p)}$ and $\gamma_b(p)$ over a regular parameter grid of size $5^4 = 625$ to find that $\|\mathcal{L}\|_{\mathcal{X}\to Y(p)}$ varies in the range 1.00 to 1.271 and $\gamma_b(p)$ varies from 1.00 to 1.615; we therefore set $\gamma_{\mathcal{L}}^{\mathrm{UB}}(p) = 1.271$ and $\gamma_b^{\mathrm{UB}}(p) = 1.615$. For the inf-sup lower bound $\beta_{\mathrm{LB}}(p)$ we also use the global lower bound $\beta_{\mathrm{LB}}(p) = 3.3323$ given by the minimum value of $\beta_h(p)$ as discussed in Remark 4.1. We are aware that these upper and lower bounds are not guaranteed to hold for all $p \in \mathcal{P}$ and that tighter bounds can most likely be obtain using, for example, the successive constraint method (SCM)[23]. However, we observed that the minimum and maximum stated above always occurred at a "corner" of the parameter domain. Since these parameter values are included in our parameter grid, we are confident — although we do not have a proof — that the global bounds are reliable. More importantly, the variation of these (parameter-dependent) constants is fairly small over the parameter domain and the gain of using the SCM thus negligible. Note that tighter bounds would simply result in slightly sharper *a posteriori* error bounds.

In Figure 4 we present, as a function of $\mathcal{M}$, the maximum absolute errors and error bounds for the optimal state and adjoint variable, the optimal control, and the associated cost functional. The errors and bounds are averaged over the test sample Ξ_{test} and we set $\mathcal{N} = 2\mathcal{M}$. We note that our *a posteriori* error bounds considerably overestimate the actual errors. We also note, however, the bound for the control and the cost become sharper as $\mathcal{M}$ increases, with the effectivity for the control bound reaching $\mathcal{O}(10)$ for $\mathcal{M} = 50$. Given the exponential convergence of the reduced basis solution and the fact that for larger values of $\mathcal{M}$ the error bounds are well below the acceptable tolerance, we consider the proposed bounds to be a useful tool to certify the reduced basis solution.

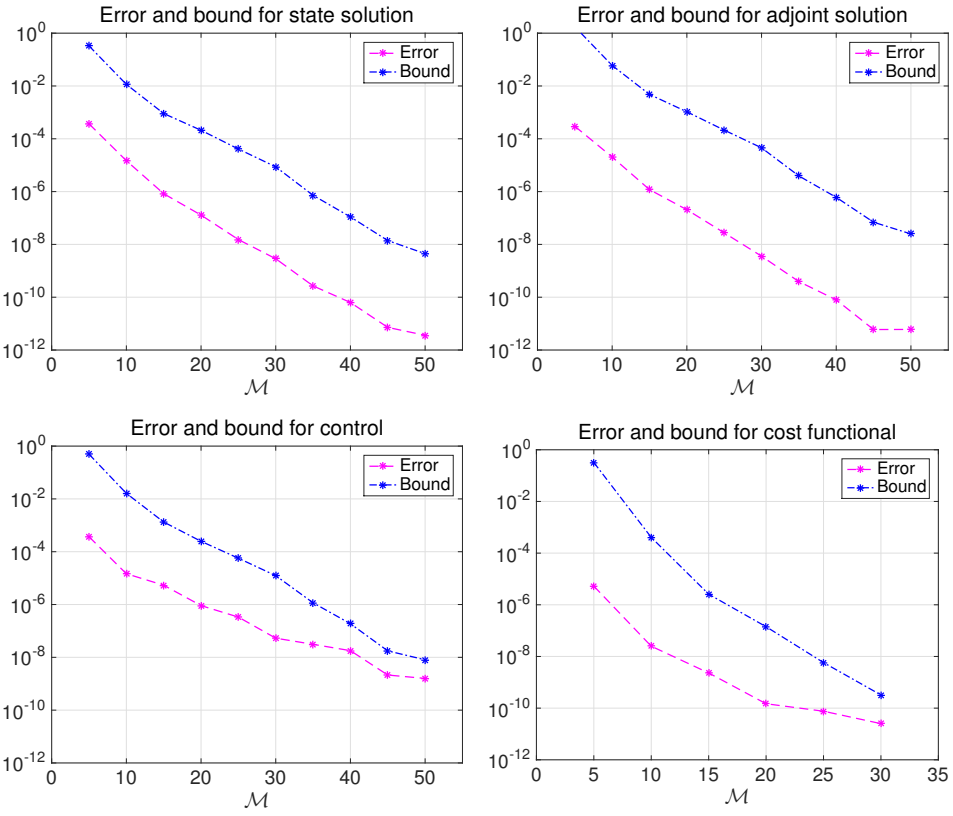


Fig. 4: Maximum errors and error bounds for state, adjoint, control and cost functional as a function of $\mathcal{M}$.

One reason for the overestimation of the state and adjoint error presumably lies in the transport character of the kinetic transport equation, where the pure Galerkin approximation does not yield a sufficiently tight relation between the error and the residual. In this context we also refer to Ref. 37, where similarly pessimistic reduced

basis *a posteriori* bounds have been observed for a kinetic transport model problem. An attempt to improve the efficiency of the bounds based on ideas presented in Ref. 9 is a topic of future research.

The example file for the computation of the errors and error bounds is implemented in `ex_RBoptim_constMat_mRBspace.m`. The results and figures can be reproduced by running this file.

5.3. *Uncertainty Quantification of Treatment Plans*

We now turn to the efficient computation of the uncertainty in the treatment plan under geometric variations obtained from the patient's CT scan. We assume that the geometry parameters are not known precisely, but instead are random variables distributed by a normal distribution with a fixed mean and variance, i.e., $p_i \sim \mathcal{N}(\bar{p}_i, \sigma_i^2)$, $i = 1, \ldots, 4$, where $\bar{p}_i$ denotes the mean and σ_i^2 the variance.

We employ the reduced basis approximation in combination with a Monte Carlo simulation to quantify the uncertainty of the optimal treatment plan under parametric variations. More precisely, we compute the expected value of the optimal control, $\mathrm{E}(q^*(p))$, where p is drawn from $\mathcal{N}(\bar{p}_i, \sigma_i^2)$ using the MC simulation and compare the result with the optimal control for the mean parameter value, i.e., $q^*(\bar{p})$. We also compute the difference $|\mathrm{E}(q^*(p)) - q^*(\bar{p})|$ for different variances and plot the $L^\infty-$ and L^2-norm of the difference as a function of the variance. This allows to answer the question: up to which noise level in the geometry data the optimal treatment plan is still reliable?

We first keep $p_2 = p_4 = 0.07$ fixed and only consider variations in p_1 and p_3. In Figure 5 on the left we plot the mean optimal control, i.e., the optimal control for the mean parameter values $\bar{p}_1 = 0.2$ and $\bar{p}_3 = 0.8$, and the expected value of the optimal control for the variances $\sigma = 0.04$ and 0.05. In Figure 5 on the right we plot the difference between the mean and expected value in the $L^\infty-$ and L^2-norm over the variance. In Figure 6 we show analogous results for fixed $p_1 = 0.2$ and $p_3 = 0.8$ and varying in p_2 and p_4 and in Figure 7 for the case where all parameters are varying. In all cases we observe a close to linear dependence of the difference between the mean and expected value on the variance. We also observe that the uncertainty in p_2 and p_4 has a slightly larger effect on the optimal treatment plan than the uncertainty in p_1 and p_3. The difference from the mean control (resp. treatment plan) of course occurs if all four parameter are uncertain.

In this study, we employed the reduced basis approximation only within the MC simulation to efficiently propagate and assess the uncertainty in the parameters. Future work will address the use of the *a posteriori* error bounds in the uncertainty quantification, i.e., to also provide certified error bounds for the expected value of the optimal treatment plan.

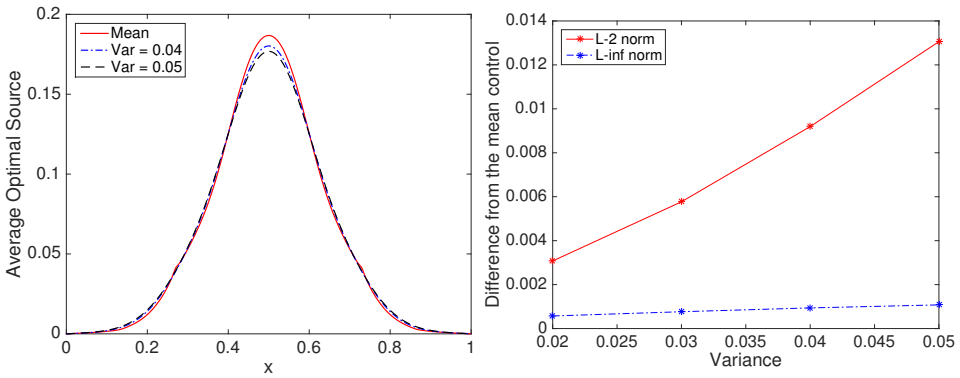


Fig. 5: Mean and expected value of optimal control over Ω (left) and the difference as a function of the variance (right) given uncertainty in p_1 and p_3; mean values $\bar{p}_1 = 0.2$ and $\bar{p}_3 = 0.8$; number of MC samples $n_{\mathrm{MC}} = 10^5$.

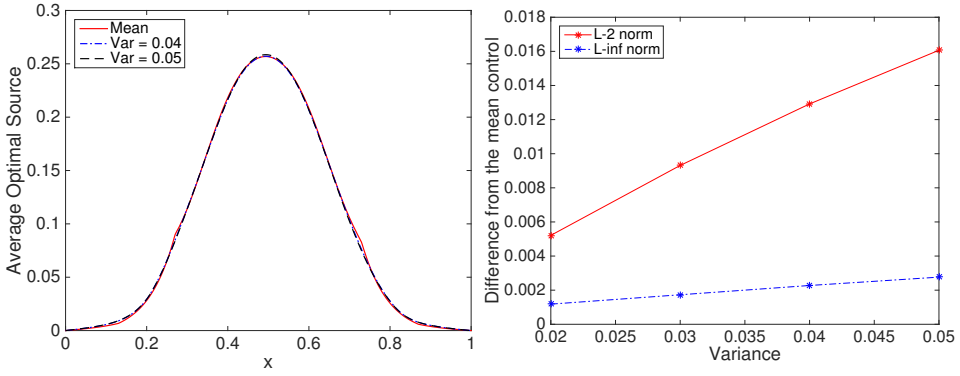


Fig. 6: Mean and expected value of optimal control over Ω (left) and the difference as a function of the variance (right) given uncertainty in p_2 and p_4; mean values $\bar{p}_2 = 0.07$ and $\bar{p}_4 = 0.07$; number of MC samples $n_{\mathrm{MC}} = 10^5$.

Acknowledgment

This work was supported by the Excellence Initiative of the German federal and state governments, the German Research Foundation through Grant GSC 111 and the Cluster of Excellence "Production technologies for high–wage countries," the BMBF KinOpt Project, and by the DFG STE2063/1-1. We would also like to thank Eduard Bader and Mark Kärcher of RWTH Aachen for helpful discussions and the reviewer for helpful commments which improved the readability of the paper.

References

1. B. Ahmedov, *Approximation and Reduced Models for Optimal Radiotherapy*, PhD

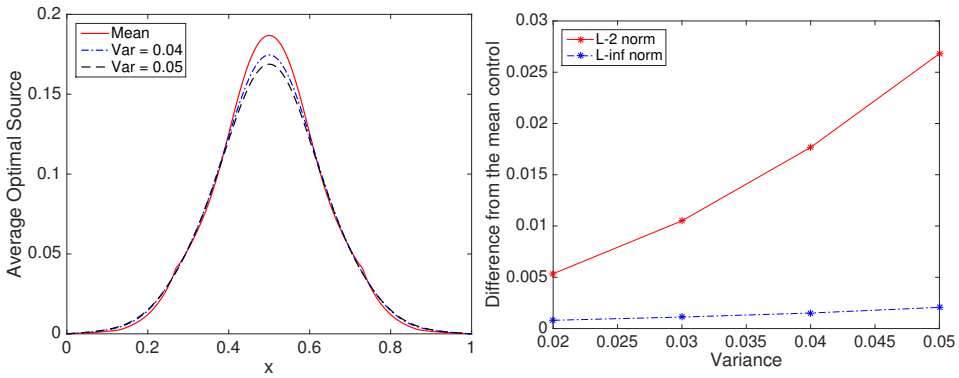


Fig. 7: Mean and expected value of optimal control over Ω (left) and the difference as a function of the variance (right) given uncertainty in p_1, p_2, p_3, and p_4; mean values $\bar{p}_1 = 0.2$, $\bar{p}_2 = 0.07$, $\bar{p}_3 = 0.8$, and $\bar{p}_4 = 0.07$; number of MC samples $n_{\mathrm{MC}} = 10^5$.

thesis, RWTH Aachen University, 2015 (in prep.).

2. P. Andreo, *Monte Carlo techniques in medical radiation physics*, Phys. Med. Biol., 36 (1991), pp. 861–920.
3. I. Babuška, *Error-bounds for finite element method*, Numer. Math., 16 (1970/1971), pp. 322–333.
4. I. Babuška and A. K. Aziz, *Survey lectures on the mathematical foundations of the finite element method*, in The mathematical foundations of the finite element method with applications to partial differential equations (Proc. Sympos., Univ. Maryland, Baltimore, Md., 1972), Academic Press, New York, 1972, pp. 1–359. With the collaboration of G. Fix and R. B. Kellogg.
5. R. Becker, H. Kapp, and R. Rannacher, *Adaptive finite element methods for optimal control of partial different equations: Basic concept*, SIAM J. Control Optim., 39 (2000), pp. 113–132.
6. N. Bellomo, N. K. Li, and P. K. Maini, *On the foundations of cancer modelling: selected topics, speculations, and perspectives*, Math. Models Methods Appl. Sci., 18 (2008), pp. 593–646.
7. C. Börgers, *The radiation therapy planning problem*, vol. 110 of IMA Volumes in Mathematics and its applications, Springer-Verlag, 1999.
8. K. K. Bucci, A. Bevan, and M. R. III, *Advances in radiation therapy: conventional to 3d, to IMRT, to 4d, and beyond*, CA Cancer J. Clin., 55 (2005), pp. 117–134.
9. W. Dahmen, C. Plesken, and G. Welper, *Double greedy algorithms: reduced basis methods for transport dominated problems*, ESAIM Math. Model. Numer. Anal., 48 (2014), pp. 623–663.
10. R. Dautray and J. L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology (v.6)*, Springer, Paris, 1993.
11. H. Egger and M. Schlottbom, *Analysis and regularization of problems in diffuse optical tomography*, SIAM J. Math. Anal., 42 (2010), pp. 1934–1948.
12. ———, *A mixed variational framework for the radiative transfer equation*, Math. Models Methods Appl. Sci., 22 (2012), pp. 1150014, 30.
13. ———, *An L^p theory for stationary radiative transfer*, Appl. Anal., 93 (2014), pp. 1283–1296.

14. M. Frank, H. Hensel, and A. Klar, *A fast and accurate moment method for dose calculation in electron radiotherapy*, SIAM J. Appl. Math., 67 (2007), pp. 582–603.
15. M. Frank, M. Herty, and M. Hinze, *Instantaneous closed loop control of the radiative transfer equations with applications in radiotherapy*, ZAMM Z. Angew. Math. Mech., 92 (2012), pp. 8–24.
16. M. Frank, M. Herty, and A. N. Sandjo, *Optimal radiotherapy treatment planning governed by kinetic equations*, Math. Models Methods Appl. Sci., 20 (2010), pp. 661–678.
17. M. Frank, M. Herty, and M. Schäfer, *Optimal treatment planning in radiotherapy based on Boltzmann transport calculations*, Math. Models Methods Appl. Sci., 18 (2008), pp. 573–592.
18. A.-L. Gerner and K. Veroy, *Certified reduced basis methods for parametrized saddle point problems*, SIAM Journal on Scientific Computing, 34 (2012), pp. A2812–A2836.
19. K. A. Gifford, J. L. H. Jr., T. A. Wareing, G. Failla, and F. Mourtada, *Comparioson of a finite-element multigroup discrete-ordinates code with Monte Carlo for radiotherapy calculations*, Phys. Med. Biol., 51 (2006), pp. 2253–2265.
20. H. Hensel, R. Iza-Teran, and N. Siedow, *Deterministic model for dose calculation in photon radiotherapy*, Phys. Med. Biol., 51 (2006), pp. 675–693.
21. M. Herty and A. N. Sandjo, *On optimal treatment planning in radiotherapy governed by transport equations*, Math. Models Methods Appl. Sci., 21 (2011), pp. 345–359.
22. P. Holmes, J. L. Lumley, and G. Berkooz, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Cambridge University Press, 1996.
23. D. B. P. Huynh, G. Rozza, S. Sen, and A. T. Patera, *A successive constraint linear optimization method for lower bounds of parametric coercivity and inf-sup stability constants*, C. R. Math., 345 (2007), pp. 473–478.
24. K. Ito and K. Kunisch, *Reduced order control based on approximate inertial manifolds*, Linear Algebra Appl., 415 (2006), pp. 531–541.
25. ———, *Reduced-order optimal control based on approximate inertial manifolds for nonlinear dynamical systems*, SIAM J. Numer. Anal., 46 (2008), pp. 2867–2891.
26. M. Kärcher and M. A. Grepl, *A posteriori error estimation for reduced order solutions of parametrized parabolic optimal control problems*, ESAIM Math. Model. Numer. Anal., 48 (2014), pp. 1615–1638.
27. M. Kärcher and M. A. Grepl, *A certified reduced basis method for parametrized elliptic optimal control problems*, ESAIM Control Optim. Calc. Var., 20 (2014).
28. M. Kärcher, M. A. Grepl, and K. Veroy, *Certified reduced basis methods for parametrized distributed optimal control problems*, SIAM J. Control Optim., (2015, submitted).
29. E. W. Larsen, *Tutorial: The nature of transport calculations used in radiation oncology*, Transp. theory Stat. Phys., 26 (1997), p. 739.
30. R. J. Le Veque, *Python tools for reproducible research on hyperbolic problems*, Comput. Sci. Eng., 11 (2009), p. 19.
31. E. A. B. F. Lima, R. C. Almeida, and J. T. Oden, *Analysis and numerical solution of stochastic phase-field models of tumor growth*, Numer. Methods Partial Differential Equations, 31 (2015), pp. 552–574.
32. J.-L. Lions, *Optimal control of systems governed by partial differential equations.*, Translated from the French by S. K. Mitter. Die Grundlehren der mathematischen Wissenschaften, Band 170, Springer-Verlag, New York-Berlin, 1971.
33. Y. Maday, A. T. Patera, and D. V. Rovas, *A blackbox reduced-basis output bound method for noncoercive linear problems*, Studies in Mathematics and its Applications,

D. Cioranescu and J. L. Lions, eds., Elsevier Science B.V., 31 (2002), pp. 533–569.
34. F. Negri, G. Rozza, A. Manzoni, and A. Quarteroni, *Reduced basis method for parametrized elliptic optimal control problems*, SIAM J. Sci. Comput., 35 (2013), pp. A2316 – A2340.
35. J. T. Oden, A. Hawkins, and S. Prudhomme, *General diffuse-interface theories and an approach to predictive tumor growth modeling*, Math. Models Methods Appl. Sci., 20 (2010), pp. 477–517.
36. J. T. Oden, E. E. Prudencio, and A. Hawkins-Daarud, *Selection and assessment of phenomenological models of tumor growth*, Math. Models Methods Appl. Sci., 23 (2013), pp. 1309–1338.
37. A. T. Patera and E. M. Rønquist, *Reduced basis approximation and a posteriori error estimation for a Boltzmann model*, Comput. Methods Appl. Mech. Engrg., 196 (2007), pp. 2925–2942.
38. C. Prud'homme, D. V. Rovas, K. Veroy, and A. T. Patera, *A mathematical and computational framework for reliable real-time solution of parametrized partial differential equations*, M2AN Math. Model. Numer. Anal., 36 (2002), pp. 747–771. Programming.
39. G. Rozza, D. B. P. Huynh, and A. T. Patera, *Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations*, Arch. Comput. Method. E., 15 (2008), pp. 229–275.
40. G. Rozza and K. Veroy, *On the stability of the reduced basis method for stokes equations in parametrized domains*, Comput. Methods Appl. Mech. Engrg., (2007), pp. 1244–1260.
41. D. M. Shepard, M. C. Ferris, G. H. Olivera, and T. R. Mackie, *Optimizing the delivery of radiation therapy to cancer patients*, SIAM Rev., 41 (1999), pp. 721–744.
42. J. Tervo and P. Kolmonen, *Inverse radiotherapy treatment planning model applying boltzmann-transport equation*, Math. Models. Methods. Appl. Sci., 12 (2002), pp. 109–141.
43. J. Tervo, P. Kolmonen, M. Vauhkonen, L. M. Heikkinen, and J. P. Kaipio, *A finite-element model of electron transport in radiation therapy and related inverse problem*, Inv. Probl., 15 (1999), pp. 1345–1361.
44. J. Tervo, M. Vauhkonen, and E. Boman, *Optimal control model for radiation therapy inverse planning applying the Boltzmann transport equation*, Lin. Alg. Appl., 428 (2008), pp. 1230–1249.
45. F. Tröltsch and S. Volkwein, *Pod a - posteriori error estimates for linear-quadratic optimal control problems*, Comput. Optim. Appl., 44 (2009), pp. 83–115.