A two-level trust-region method for optimal control problems with radiative transfer

M. Herty[1] and G. Thömmes[1]

Abstract

We consider optimal control problems for the radiative transfer equation with a distributed source as control variable and the radiation intensity as state variable minimising a quadratic functional of tracking type. The problem is solved using a trust-region method where the model in the trust-region subproblem uses the numerically more efficient well-known P_N approximations. This leads to a two-level method based on the radiative transfer equations on the fine level and the P_N equations on the coarse level. Numerical results show the feasability of the new approach and confirm that it can lead to significant benefits in terms of computational costs.

1 Introduction

Radiation transport [22, 23] is central to many technical processes – e.g. glass cooling [26, 27, 34], gas turbine combustion chambers [29, 30] or combustion car engines – and appears also in medical applications like radiation therapy [4, 5, 13]. These phenomena are modeled by the radiative transfer equations, which are challenging from the point of view of numerical solution because of the high dimensionality of the problem. This has led to the development of various approximate equations ranging from moment expansions like P_N or diffusion equations like simplified P_N (SP_N), among others [19, 16, 17, 28, 32]. In recent years optimal control problems have gained growing interest since efficient methods for solving the underlying radiative transfer equations have been developed such that the goal of controlling radiation in a desired way can be addressed using computer simulations, see e.g. [31, 1, 3, 21]. Owing to the high complexity of the full equations, however, many optimisation approaches are based on simpler approximate models whenever they are appropriate. They allow to solve the optimisation problems with acceptable computational costs. Our approach is based on the full radiative transfer equations and at the same time tries to make use of the numerically efficient approximations by employing them as models in the subproblems of a trust-region method.

We consider an optimal control problem in the realm of radiative transfer with a tracking-type cost functional for given functions $\bar{R}, \bar{Q} : D \to \mathbb{R}$,

$$\tilde{F}(R,Q) = \frac{\alpha_1}{2}\int_D (R-\bar{R})^2 dx + \frac{\alpha_2}{2}\int_D (\int_{S^2} Q d\omega - \bar{Q})^2 dx. \tag{1}$$

[1]TU Kaiserslautern, PO Box 3049, D-67653 Kaiserslautern, GERMANY, {herty,thoemmes}@mathematik.uni-kl.de. This work has been supported by the Kaiserslautern Excellence Cluster DASMOD and the DAAD PPP Program Canada 2006–2008.

Here, $R(x) = \int_{S^2} I(x,\omega)d\omega$ denotes the radiosity or total flux corresponding to the space and direction dependent intensity I.

The intensity $I(x,\omega) : \mathbb{R}^d \times S^2 \to \mathbb{R}$ is computed by solving the radiative transfer equation,

$$\epsilon\omega \cdot \nabla I + (\sigma_s + \sigma_a)I = \frac{\sigma_s}{4\pi} \int_{S^2} I \, d\omega' + Q(x,\omega), \tag{2}$$

where d is the space dimension of the underlying domain, S^2 is the sphere in $\mathbb{R}^3$ and σ_s and σ_a are problem dependent scattering and absorption parameters [23] and ϵ is a scaling factor of the equation, i.e., $\epsilon = x^{ref}/(\sigma_a^{ref} + \sigma_s^{ref})$. The equation contains the source term $Q(x,\omega)$ which can be interpreted as an exterior source or sink of radiation energy. The external source is the control variable of our problem, see also section 2.

Existence results and first-order optimality conditions for this problem have been derived in [14, 25, 27] using adjoint calculus. Moreover,the corresponding P_N and SP_N optimal control problems can be found in the literature [15]. Starting from these results, we set up a numerical multilevel trust-region method based on the full radiative transfer system and simplified problems. There is a variety of literature on trust-region methods and we refer to [6] and the references therein for further information on this numerical method. The trust-region method we consider here has been inspired by [11], see also [24, 12, 8] and the discussion in section 2.

The paper is organised as follows. In section 2 we give an overview of the optimal control problem for the radiative transfer equation (fine level problem) and its approximations based on the P_N and SP_N equations (coarse level problem) and present the details of the two-level trust-region method we propose. The results of two test cases are summarized in section 3 illustrating the feasibility of the new method.

2 Optimal control problem and two-level trust region method

We are interested in fast and efficient numerical methods for solving optimal control problems in radiative transfer. A general mathematical formulation of the problems under consideration is the following: Let $x \in D$, where D is a bounded, convex subset with Lipschitz boundary of $\mathbb{R}^d$ with $d = 1, 2, 3$, and let $\omega \in S^2$. We assume that the radiative intensity $I = I(x,\omega)$ at position x propagating along direction ω satisfies the scaled radiative transfer equation (2). The equation is accompanied by boundary conditions on the incoming directions

$$I(x,\omega) = A, \quad n(x) \cdot \omega < 0, \tag{3}$$

where n denotes the outer normal on D. For simplicity we assume constant boundary data A. Even though equation (2) is a simplification of the full

time- and frequency-dependent radiative transfer equation, it is still a valid model when dealing, for example, with grey media or when the mean free path of the radiation is small compared to the characteristic length. More details on modeling aspects can be found e.g. in [23].

The optimisation problem for determining a distributed control $Q(x, \omega)$ then reads

$$\min_{R,Q} \tilde{F}(R, Q) \text{ subject to (2).} \tag{4}$$

Problem (4) has been studied in [14] and an adjoint calculus has been derived. Furthermore, the corresponding optimality conditions have been investigated in [15] and [9]. Here, we contribute to this investigation a numerical algorithm for solving (4) based on a multilevel model hierarchy in numerical optimisation. This will be accomplished by using the multilevel trust-region method proposed by Gratton et al. [11] as well as by the fact, that there exists a well-known model hierarchy for (2), namely the P_N and SP_N approximations [16, 19, 17, 7].

Usually, for a numerical solution to (4) by a descent-type method for the reduced cost functional F defined by $F(Q) := \tilde{F}(R(Q), Q)$ where $R(Q)$ is the radiosity obtained by solving (2). As derived in [14], the gradient of the reduced cost functional $\nabla_Q F$ at Q is given by

$$\nabla_Q F(Q)[\delta Q] = \alpha_2(Q - \frac{1}{4\pi}\bar{Q}) \cdot \delta Q + J \cdot \delta Q, \tag{5}$$

where $J = J(x, \omega)$ is the adjoint variable solving the corresponding equation

$$-\epsilon\omega \cdot \nabla J + (\sigma_a + \sigma_s)J = \frac{\sigma_s}{4\pi} \int_{S^2} J \, d\omega' + \alpha_1(R - \bar{R}), \tag{6a}$$

$$J(x, \omega) = 0, \; n(x) \cdot \omega > 0, \tag{6b}$$

and, again, $R = \int_{S^2} I d\omega$, where I the solution to (2). A full discretisation of (2) and (6) in angle ω and space x variables leads to a high-dimensional discrete system which has to be solved at least once for each descent step. In the sequel, we first introduce various coarse models approximating the radiative transfer equation, the cost functional and its reduced gradient. Second, we include the coarse models in an optimisation for the fine model, i.e., (2), to reduce the numerical effort.

2.1 Model hierarchies for radiative transfer

The first hierarchy of coarse models are the simplified P_N approximations (SP_N). They are good approximate models if the medium can be assumed to be optically thick, i.e., $\epsilon \ll 1$, see [18, 35]. We shortly recall the SP_1, SP_2 and SP_3 approximations, which are commonly-used alternatives of (2).

Defining the total scattering cross section $\sigma_t = \sigma_a + \sigma_s$ and the total flux of the source $q(x) = \int_{S^2} Q(x,\omega)d\omega$, the SP_1 approximation reads

$$-\frac{\epsilon^2}{\sigma_t}\nabla^2\varphi + \sigma_a\varphi = q, \quad x \in D, \tag{7a}$$

$$\text{b.c.} \quad \varphi + \frac{2\epsilon}{3\sigma_t} n \cdot \nabla\varphi = \pi A, \quad x \in \partial D. \tag{7b}$$

The SP_2 approximation is expressed in terms of an auxiliary variable $\tilde{\varphi} = \varphi + \frac{4\sigma_a}{5\sigma_t}(\varphi - \frac{1}{\sigma_a}q)$:

$$-\epsilon^2 \frac{5\sigma_t + 4\sigma_a}{15\sigma_t}\nabla^2\tilde{\varphi} + \sigma_a\tilde{\varphi} = q, \tag{8a}$$

$$\text{b.c.} \quad \tilde{\varphi} + \frac{4\epsilon(5\sigma_t + 4\sigma_a)}{45\sigma_t^2} n \cdot \nabla\tilde{\varphi} = \frac{5\sigma_t - 8\sigma_a}{15\sigma_t} q + \frac{2(5\sigma_t + 4\sigma_a)}{15\sigma_t}\pi A. \tag{8b}$$

Finally, the SP_3 approximation is given by two coupled elliptic equations for φ and a second auxiliary variable denoted by $\hat{\varphi}$:

$$-\frac{\epsilon^2}{3\sigma_t}\nabla^2(\varphi + 2\hat{\varphi}) + \sigma_a\varphi = q, \tag{9a}$$

$$-\frac{9\epsilon^2}{35\sigma_t}\nabla^2\hat{\varphi} + \sigma_t\hat{\varphi} - \frac{2}{5}\sigma_a\varphi = -\frac{2}{5}q, \tag{9b}$$

$$\text{b.c.} \quad \varphi + \frac{5}{16}\hat{\varphi} + \frac{\epsilon}{6\sigma_t} n \cdot \nabla\varphi + \frac{2\epsilon}{9\sigma_t} n \cdot \nabla\hat{\varphi} = \pi A, \tag{9c}$$

$$\text{b.c.} \quad -\frac{1}{16}\varphi + \frac{5}{16}\hat{\varphi} + \frac{3\epsilon}{14\sigma_t} n \cdot \nabla\hat{\varphi} = -\frac{\pi}{4}A, \tag{9d}$$

$$\tag{9e}$$

For the subsequent optimisation we note that the previous equations are an approximation in the following way: given a source term $Q(x,\omega)$, the fine model yields the radiosity $R = \int_{S^2} I(x,\omega)d\omega$, where I is the solution of (2). The coarse SP_N model produces an approximation φ to this radiosity R. The SP_N is a coarse grid model since all quantities are independent of ω. The cost functional for the coarse SP_N model reads

$$\tilde{f}^{SP}(\varphi, q) = \frac{\alpha_1}{2}\int_D (\varphi - \bar{R})^2 dx + \frac{\alpha_2}{2}\int_D (q - \bar{Q})^2 dx, \tag{10}$$

and depending on the order $N = 1, 2, 3$ a single evaluation of (10) consists of a solution of one or two second order partial differential equations. Furthermore, using adjoint calculus an expression for the gradient of the reduced functional $f^{SP}(q) := \tilde{f}^{SP}(\varphi(q), q)$ where $\varphi(q)$ is the solution to (7), (8) or (9), respectively, has been derived in [15]:

$$\nabla_q f^{SP}(q)[\delta q] = \alpha_2(q - \bar{Q}) \cdot \delta q + \psi \cdot \delta q. \tag{11}$$

According to the order N of the approximation, the adjoint variable ψ is given for SP_1 by

$$-\frac{\epsilon^2}{3\sigma_t}\nabla^2\psi + \sigma_a\psi = \alpha_1(\varphi - \bar{R}),$$
$$\text{b.c.} \quad \psi + \frac{2\epsilon}{3\sigma_t} n \cdot \nabla\psi = 0,$$

and for SP_2 with auxiliary variable $\tilde{\psi} = \psi + \frac{4\sigma_s}{5\sigma_t}(\psi - \frac{\alpha_1}{\sigma_a}(\varphi - \bar{R}))$ we have

$$-\epsilon^2\frac{5\sigma_a + 9\sigma_s}{15\sigma_t}\nabla^2\tilde{\psi} + \sigma_a\tilde{\psi} = \alpha_1(\varphi - \bar{R}),$$
$$\text{b.c.} \quad \tilde{\psi} + \frac{4\epsilon(5\sigma_t + 4\sigma_a)}{45\sigma_t^2} n \cdot \nabla\tilde{\psi} = 0.$$

Furthermore, for SP_3 the adjoint system for the two unknowns ψ and $\hat{\psi}$ is

$$-\frac{\epsilon^2}{3\sigma_t}\nabla^2(\psi + 2\hat{\psi}) + \sigma_a\psi = \alpha_1(\varphi - \bar{R}),$$
$$-\frac{9\epsilon^2}{35\sigma_t}\nabla^2\hat{\psi}\sigma_t\hat{\psi} - \frac{2}{5}\sigma_a\psi = -\frac{2}{5}\alpha_1(\varphi - \bar{R})$$
$$\text{b.c.} \quad \psi + \frac{5}{16}\hat{\psi} + \frac{\epsilon}{6\sigma_t} n \cdot \nabla\psi + \frac{2\epsilon}{9\sigma_t} n \cdot \nabla\hat{\psi} = 0,$$
$$\text{b.c.} \quad -\frac{1}{16}\psi + \frac{5}{16}\hat{\psi} + \frac{3\epsilon}{14\sigma_t} n \cdot \nabla\hat{\psi} = 0,$$

In these adjoint equations, $\varphi = \varphi(q)$ is the solution to (7), (8) or (9), respectively, for a given control q.

The second coarse model we investigate are the spherical harmonic approximations (P_N). In the case of a one-dimensional slab geometry and for a control $Q(x, \mu)$ the radiative transfer equation simplifies

$$\epsilon\mu\partial_x I + (\sigma_a + \sigma_s)I = \frac{\sigma_s}{2}\int_{-1}^{1} I \; d\mu' + Q, \tag{12}$$

where $\mu = cos(e_x, \omega) \in (-1, 1)$ is the cosine of the angle between direction and x-axis. We can further assume that x is normalised to $x \in [0, 1]$. The P_N approximations are obtained by assuming that $I(x, \mu)$ is approximated by a truncated expansion with respect to Legendre Polynomials P_l using $(N + 1)$ terms (see e.g. [19, 32])

$$I(x, \mu) \equiv \sum_{l=0}^{N} \frac{2l + 1}{2}\varphi_l(x)P_l(\mu).$$

The moments in this expansion are denoted φ_l, $l = 0, \ldots, N$. The P_N approximation is then a $N+1$ dimensional system of equations

$$\epsilon\partial_x \left(\frac{l+1}{2l+1}\varphi_{l+1} + \frac{l}{2l+1}\varphi_{l-1} \right) + (\sigma_a + \sigma_s)\varphi_l = \sigma_s\varphi_l + 2q_l, \quad l = 0, \ldots, N \tag{13}$$

where the intensity and source moments are given by

$$\varphi_l = \int_{-1}^{1} I(x,\mu)P_l(\mu)\, d\mu, \quad \text{and} \qquad q_l = \int_{-1}^{1} Q(x,\mu)P_l(\mu)\, d\mu,$$

(formally we set $\varphi_{-1} = \varphi_{N+1} = 0$). It may be noted that in this 1D situation the first moment φ_0 coincides with the integrated flux φ that appears in the SP_N approximations. Different boundary conditions can be imposed for (13) [20]. We impose Mark-type boundary conditions

$$\text{b.c.} \sum_{l=0}^{N} \varphi_l(0)\frac{2l+1}{2}P_l(\mu_k) = 0, \ \mu_k > 0$$
$$\text{and} \sum_{l=0}^{N} \varphi_l(1)\frac{2l+1}{2}P_l(\mu_k) = 0, \ \mu_k < 0,$$

where μ_k is the kth zero of the Legendre Polynom P_{N+1}.

The coarse cost functional is given by

$$\tilde{f}^P(\varphi_0, q_0) = \frac{\alpha_1}{2}\int_D (\varphi_0 - \bar{R})^2 dx + \frac{\alpha_2}{2}\int_D (q_0 - \bar{Q})^2 dx, \tag{14}$$

and each evaluation requires the solution of the $N+1$ transport equations (13). The gradient of the reduced cost functional f^P is given by [9]

$$\nabla_q f^p(q_0)[\delta q] = \alpha_2(q_0 - \bar{Q}) \cdot \delta q + \psi_0 \cdot \delta q, \tag{15}$$

where $\psi_l, l = 0, \ldots, N$ denote the adjoint variables obtained from the solution of the system

$$-\epsilon\partial_x \left(\frac{l+1}{2l+1}\psi_{l+1} + \frac{l}{2l+1}\psi_{l-1} \right) + (\sigma_a + \sigma_s)\psi_l = \sigma_s\psi_l + 2\alpha_1(\varphi_0 - \bar{R})\delta_{l0},$$

with $\psi_{-1} = \psi_{N+1} = 0$, and boundary conditions

$$\text{b.c.} \sum_{l=0}^{N} \psi_l(0)\frac{2l+1}{2}P_l(\mu_k) = 0, \ \mu_k < 0$$
$$\text{and} \sum_{l=0}^{N} \psi_l(1)\frac{2l+1}{2}P_l(\mu_k) = 0, \ \mu_k > 0.$$

Next, we incorporate the model hierarchies in a two-level trust-region method.

2.2 Two-level optimisation method

In [11] a general trust-region type method for multilevel models has been introduced. The primary examples are related to the multigrid method for partial differential equations, see [24]. However, we apply this concept in a two-level version to solve the optimal control problem for the radiative transfer equation (4). The independent variables are naturally divided into space and direction variables. This is exploited in the above mentioned models to construct approximations which contain only the space as independent variable. We use these reduced models as an alternative to the classical quadratic trust-region subproblem to compute a new iterate on the fine level. We explain the algorithm in the continuous setting and refer to section 3 for details of the implementation.

On the fine level we deal with direction dependent controls while at the coarse level controls are independent of directions. The transition from fine to coarse is accomplished by the restriction operator, which is realised by angular integration

$$q(x) \equiv r(Q)(x) = \int_{S^2} Q(x,\omega)\, d\omega,$$

and the prolongation from coarse to fine level is then

$$Q(x,\omega) \equiv p(q)(x,\omega) = \frac{1}{4\pi} q(x).$$

To correctly couple the coarse level functional in step k to the fine level, the coarse level iteration is started with a modified functional in each substep

$$\hat{f}_k(q_0 + \delta q) = f(q_0) + v \cdot \delta q + w, \tag{16}$$

where $q_0 = r(Q_k)$ is the restriction of the current fine level control, $v = r(\nabla_Q F(Q_k)) - \nabla_q f(q_0)$ is a gradient shift, and $w = F(Q_k) - f(q_0)$ shifts the functional values. (Here, f denotes either the function f^{SP_N} or f^P.) In this way we can make sure that the behaviour of coarse and fine level functionals match locally around q_0 and Q_k, in particular we enforce the gradient relation

$$\nabla_q \hat{f}(q_0) = r(\nabla_Q F(Q_k)).$$

The trust-region optimisation on the coarse level then proceeds in the standard way and returns a control q^*. The difference, $q^* - q_0$, is then projected to the fine grid, $\delta Q_k = p(q^* - q_0)$, where this step size proposal of the subproblem is processed as usual.

In the following trust-region algorithm we denote by $0 < \delta_1 \leq \delta_2 < 1$ the thresholds for the radius decrease and increase, the radius being decreased or increased by $0 < \gamma_1 < 1 < \gamma_2$, respectively. Furthermore, in the termination criterion we use upper bounds ϵ_a and ϵ_s for the difference in functional values and for size of a single step. Coarse level parameters are indicated with superscript 'c'.

I **Initialisation**. Set $k = 0$ and initialize the trust region radius Δ_0 and thresholds δ_1 and δ_2. Initialize the approximation to the reduced Hessian of the fine model $H_0 = Id$. Fix an initial guess for the control Q_0 and solve the fine model by computing $F_0 = F(Q_0)$ and $\nabla F_0 = \nabla F(Q_0)$.

II **Choice of the model**. Either we proceed using the fine model (Taylor step, III) or the coarse grid model (IV).

III **Fine-level step.** Solve the fine level minimisation problem at step k

$$\min_{\delta Q} F(Q_k + \delta Q) \text{ subject to } \|\delta Q\|_2 \leq \Delta_k \tag{17}$$

by a trust region method using a quadratic model approximation

$$M_k(\delta Q) = F_k + \nabla F_k \cdot \delta Q + \frac{1}{2}\delta Q^T \cdot H_k \cdot \delta Q. \tag{18}$$

Goto step V.

IV **Coarse-level step.**

a) Initialisation and restriction: Let $l = 0$ be the iteration index of the coarse level optimisation. Initialize the coarse level trust region radius $\Delta_0^c = \Delta_k$ and an approximation to the Hessian $H_0^c = Id$. Restrict to the coarse model $q_0 = r(Q_k)$. Evaluate the cost functional $f_0 = f(q_0)$ and the gradient $\nabla f_0 = \nabla f(q_0)$ by either (15) or (11), respectively.

b) Coarse functional: Compute the functional shift $w = F(Q_k) - f(q_0)$ and the gradient shift $v = r(\nabla_Q F(Q_k)) - \nabla_q f(q_0)$. Define the modified coarse functional by

$$\hat{f}(q) = f_0 + v \cdot (q - q_0) + w. \tag{19}$$

c) Solve the coarse level optimisation problem at step l

$$\min_{\delta q} \hat{f}(q_l + \delta q) \text{ subject to } \|\delta q\|_2 \leq \Delta_l^c \tag{20}$$

by the standard trust-region method using a quadratic model

$$m_l(\delta q) := \hat{f}_l + \nabla \hat{f}_l \cdot \delta q + \frac{1}{2}\delta q^T \cdot H_l^c \cdot \delta q.$$

d) Termination and prolongation: if $|\hat{f}_l - \hat{f}(q_l + \delta q)| < \epsilon_a$ or $||\delta q|| < \epsilon_s$ then stop coarse level iteration. Prolongate the control $\delta Q_k = p(q_l - q_0)$. Continue on fine level at step V.

e) Check acceptance: Determine the decrease within the coarse model

$$\rho = \frac{\hat{f}_l - \hat{f}(q_l + \delta q)}{m_l(0) - m_l(\delta q)}.$$

If $\rho \geq \delta_1$ then accept and update coarse level control $q_{l+1} = q_l + \delta q$. Evaluate new $\hat{f}_{l+1}$ and $\nabla \hat{f}_{l+1}$, which only involves the evaluation of the coarse level gradient (15) or (11), respectively. Update approximation H^c_{l+1} to the reduced Hessian by BFGS update formula. If $\rho > \delta_2$ then increase the trust-region radius $\Delta^c_{l+1} = \gamma^c_2 \Delta^c_l$. Set l:=l+1.

Else, when the step has not been accepted, decrease the trust-region radius $\Delta^c_l = \gamma^c_1 \Delta^c_l$.

Continue with step c).

V. **Termination.** if $|F_k - F(Q_k + \delta Q)| < \epsilon_a$ or $||\delta Q|| < \epsilon_s$ then Stop.

VI. **Check acceptance.** Determine decrease in the cost functional by comparing predicted and realized fine level descent

$$\rho = \frac{F_k - F(Q_k + \delta Q)}{M_k(0) - M_k(\delta Q)}.$$

If $\rho \geq \delta_1$ then accept and update the control $Q_{k+1} = Q_k + \delta Q$, and compute F_{k+1} and ∇F_{k+1} by (1) and (5). Furthermore, update the approximation H_{k+1} to the reduced Hessian by BFGS update formula. If $\rho > \delta_2$ then increase the trust-region radius $\Delta_{k+1} = \gamma_2 \Delta_k$. Set k:=k+1.

Else, when the step has not been accepted, decrease the trust-region radius $\Delta_k = \gamma_1 \Delta_k$.

Continue with step III.

Some remarks are in order. First, for simplicity we restricted ourselves to the most basic algorithm, but a variety of modifications can be applied, e.g., different update rules for the trust-region radius can be envisioned [11], other update formulas for the approximation to the reduced Hessian can be included and termination and update criteria can be level and iteration dependent. Second, the main computational advance stems from the fact that the coarse level steps IV.c-IV.e should be iterated several times before returning to the fine level. In these steps, we continue the optimisation without recomputing any fine level quantity. The performance of the complete algorithm then strongly depends on the quality of the approximative models and this will be investigated in section 3. Step IV.b guarantees convergence and first-order optimality for the combined algorithm since the following

relations hold true

$$\hat{f}(q_0) = F(Q_k), \quad \nabla_q \hat{f}(q_0) = \int_{S^2} \nabla_Q F(Q_k) d\omega. \tag{21}$$

Fourth, if we run the algorithm using only step III then we have a classical trust-region method applied to the fine model and no gain in numerical performance. The decisions in step II can be prescribed a priori or based on a comparison between fine and coarse grid gradients [11].

3 Numerical results

We implemented test cases in 1D using the DSA iterative scheme for the transport equations of the forward and adjoint equations on the *fine* level ([36], [37]). Equation (12) is discretised on an equidistant space grid using the diamond differencing scheme by evaluating intensity I and source q at the nodes $x_i + = i\Delta x,\ i = 0, \ldots, M$ and using averages $I_{i+\frac{1}{2}} = (I_{i+1} + I_i)/2$ and $q_{i+\frac{1}{2}} = (q_{i+1} + q_i)/2$. The iteration is started by choosing an initial iterate I^0_{ij} and computing the flux $\varphi^0_i = \sum_{j=1}^N I^0_{ij} w_j$. Then, for $k \geq 0$, the iteration proceeds in two substeps. First, the following transport equation with given right side is solved for the intermediate intensity $I^{k+\frac{1}{2}}_{ij}$

$$\epsilon\mu_j \frac{I^{k+\frac{1}{2}}_{i+1,j} - I^{k+\frac{1}{2}}_{ij}}{\Delta x} + \sigma_t I^{k+\frac{1}{2}}_{i+\frac{1}{2},j} = \frac{\sigma_s}{2}\varphi^k_{i+\frac{1}{2}} + q_{i+\frac{1}{2}},$$
$$\text{with b.c.} \quad I^{k+\frac{1}{2}}_{0,j} = A,\ \mu_j > 0, \quad I^{k+\frac{1}{2}}_{M,j} = A,\ \mu_j < 0.$$

This corresponds to the transport sweep in the source iteration method. Note that the sweep is done from left to right when $\mu_j > 0$, and from right to left when $\mu_j < 0$. Then the flux difference $\varphi^{k+\frac{1}{2}}_i = \sum_{j=1}^N (I^{k+\frac{1}{2}}_{ij} - I^k_{ij}) w_j$ is taken as source term for the computation of the correction $\delta\varphi^{k+\frac{1}{2}}$:

$$-\frac{\epsilon^2}{3\sigma_t} \frac{\delta\varphi^{k+\frac{1}{2}}_{i+1} - 2\delta\varphi^{k+\frac{1}{2}}_i + \delta\varphi^{k+\frac{1}{2}}_{i-1}}{\Delta x^2} + \sigma_a \frac{\delta\varphi^{k+\frac{1}{2}}_{i+1} + 2\delta\varphi^{k+\frac{1}{2}}_i + \delta\varphi^{k+\frac{1}{2}}_{i-1}}{4}$$
$$= \sigma_s \frac{\varphi^{k+\frac{1}{2}}_{i+1} - \varphi^k_{i+1}}{2} + \sigma_s \frac{\varphi^{k+\frac{1}{2}}_i - \varphi^k_i}{2},$$

with homogeneous boundary conditions on the left and right of the interval

$$\delta\varphi^{k+\frac{1}{2}}_0 - \frac{2\epsilon}{3\sigma_t} \frac{\delta\varphi^{k+\frac{1}{2}}_1 - \delta\varphi^{k+\frac{1}{2}}_0}{\Delta x} = 0,$$
$$\delta\varphi^{k+\frac{1}{2}}_M + \frac{2\epsilon}{3\sigma_t} \frac{\delta\varphi^{k+\frac{1}{2}}_M - \delta\varphi^{k+\frac{1}{2}}_{M-1}}{\Delta x} = 0.$$

The new iterate for the flux is eventually updated

$$\varphi_i^{k+1} = \varphi_i^{k+\frac{1}{2}} + \delta\varphi_i^{k+\frac{1}{2}}.$$

After the iteration has stopped, we obtain a numerical solution for the intensity by performing an additional sweep with the final flux. The coarse level P_N and SP_N approximations corresponding to the one-dimensional transfer equation were discretised using standard finite differences.

The basis of our two-level optimisation algorithm is a BFGS trust-region method to solve the optimisation problem on both levels. The solution of the classical subproblem with quadratic functional in a ball with given radius, Δ, can be reduced to the problem of solving linear systems and a nonlinear scalar equation [38]. Directly using the symmetric matrix, H, given by the BFGS updates leads to a descent direction, d, by solving $Hd = g$, where g denotes the gradient of the functional, using Cholesky decomposition. It is accepted if $\|d\| \leq \Delta$. Otherwise, we compute modified directions based on a parameter, λ, by solving $(H + \lambda Id)d = g$ and try to find a zero of the nonlinear function $\Psi(\lambda) = \|d(\lambda)\| - \Delta$. Since $\Psi(0) > 0$ and since it can be shown that for $\lambda \geq \bar{\lambda} \equiv \frac{\|g\|}{\Delta}$ we have $\Psi(\lambda) \leq 0$, a numerical method can be used to find an approximate root, λ^*, with an appropriate descent direction $d = d(\lambda^*)$. We used the bisection method starting with the interval $[0, \bar{\lambda}]$.

3.1 Dependence of the performance of the two-level trust region method on parameters

In the first example we considered the optimisation problem (4) with the functional given by (1) and numerically study the dependence of the two-level trust region method on different parameters, e.g. coarse model used, grid width, ratio α_1/α_2, etc. The setting is as follows: We considered a given box-shaped source by applying a convolution to the indicator function χ on the interval $[0.25, 0.75]$ with a Gaussian $G_{0,\sigma}$ of variance $\sigma = \Delta x$ to get a smooth source term $\bar{Q} = \chi * G$. We then solved the radiative transfer equation (2) to obtain $\bar{R}$. This source together with the corresponding intensity were the desired states $(\bar{R}, \bar{Q})$ used in functional (1). Boundary values $I_l = 0$ and $I_r = 0$ for the intensity were prescribed for ingoing directions at the left and the right. In the transport equation the scattering parameters were $\sigma_a = 1$ and $\sigma_s = 1$. An equidistant grid $x_i = i\Delta x$ in the unit interval $[0, 1]$ was used for space discretisation, and N_g points and weights (μ_k, w_k) resulting from double Gaussian quadrature were used for the angular discretisation. We started the optimisation with a zero source term $Q_{init} = 0$. Iterations were started with a trust region radius of $\Delta_0 = 20$. The radius was decreased when the trust region ratio

$$\rho = \frac{F(Q) - F(Q + \delta Q)}{M(0) - M(\delta Q)}$$

was smaller than the lower threshold $\delta_1 = 0.25$ by multiplying with the factor $\gamma_1 = 0.25$. Conversely, the radius was increased by $\gamma_2 = 4$ when the ratio was larger than $\delta_2 = 0.75$. The same thresholds and factors were used on the coarse level. We used $\epsilon_a = 10^{-6}$ and $\epsilon_s = 10^{-6}$ as tolerances for the fine level stopping criterion based on the functional value $|F - M| < \epsilon_a$ and the step size $\|d\| < \epsilon_s$. On the coarse level, where only approximate solutions are needed, the tolerances were less strict: $\epsilon_a^c = 10^{-2}$ and $\epsilon_s^c = 10^{-2}$. If not stated otherwise, we use the SP_1 equations as coarse grid model since they are considered to be a good approximation of the radiative transfer equation. We fixed the number of coarse level iterations per fine level iteration at a value of 4, i.e. a fine step was followed by four coarse steps, which produced good results in our experiments. For each run, the number of fine and coarse functional evaluations and the run-times (CPU time) are recorded.

The first numerical test consisted in a comparison of the full trust region method with the two-level trust region method using SP_1. The weights of the two terms in the functional were $\alpha_1 = 1$ and $\alpha_2 = 100$ and $N_x = 50$ and $N_g = 8$ in this case. Figures 1 and 2 display the iteration history and the final results of the original trust-region method and the new two-level algorithm, respectively. We clearly observe faster convergence in the two-level algorithm and recover the boxed-shape control.

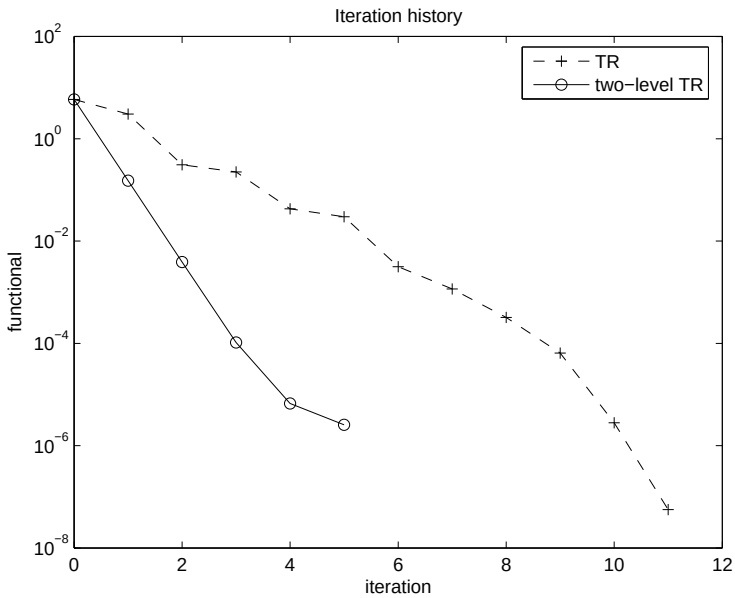

Figure 1: Comparison of the iteration histories for the reconstruction of the box source.

In the second numerical test the dependence of the CPU times on the

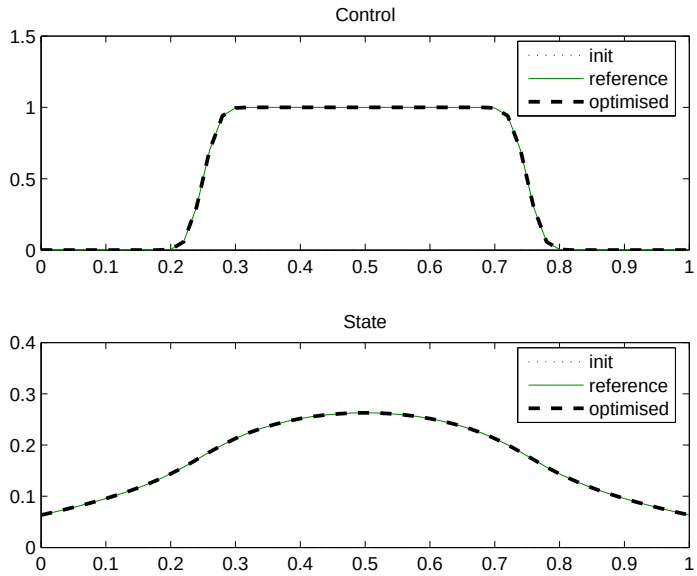

Figure 2: Optimisation results for control and state coincide with reference values.

discretization was studied. We varied the number of discretization points in space N_x (see Table 1) as well as the number of discretization points for the directions N_g (see Table 2). In all tables f refers to the number of iterations on the fine grid and c on the coarse grid. We observe faster convergence for the two-level algorithm as well as independence of the number of iterations from the discretisation level.

	two-level TR		TR	
N_x	iter (f/c)	time [sec]	iter (f)	time [sec]
25	3(3/3)	0.23	10(10)	0.64
50	5(5/8)	0.61	11(11)	1.61
75	3(3/6)	1.21	12(12)	4.95
100	3(3/6)	2.11	13(13)	11.06

Table 1: Iterations, evaluations and run-time depending on space grid.

Third, we studied the influence of the weighting parameters α_1 and α_2. Here, we used $N_x = 50$ and $N_g = 8$ discretisation points in space and angular variables, respectively. The results are given in Table 4 and 3 and show again the faster convergence of the two-level algorithm in terms of CPU time and number of iterations.

	two-level TR		TR	
N_g	iter (f/c)	time [sec]	iter (f)	time [sec]
4	3(3/6)	0.15	6(6)	0.23
8	5(5/8)	0.61	11(11)	1.61
16	3(5/8)	1.86	25(25)	20.63
32	5(5/8)	19.34	46(46)	233.91

Table 2: Iterations, evaluations and run-time depending on angle discretisation.

	two-level TR		TR	
α_2	iter (f/c)	time [sec]	iter (f)	time [sec]
10	2(2/4)	0.30	17(17)	2.51
100	5(5/8)	0.61	11(11)	1.61
1000	5(5/18)	0.68	10(20)	2.26
10000	5(5/18)	0.68	16(31)	6.17

Table 3: Iterations, evaluations and run-time depending on the relative weight of the source term in the functional.

		two-level TR		TR	
α_1	α_2	iter (f/c)	time [sec]	iter (f)	time [sec]
0.1	10	2(2/4)	0.27	17(17)	2.51
1	100	5(5/8)	0.61	11(11)	1.61
10	1000	5(5/18)	0.68	11(20)	2.34
100	10000	5(11/18)	1.24	16(33)	6.51

Table 4: Iterations, evaluations and run-time depending on the scaling of the functional.

Fourth, we studied the influence of the a priori choice on the number of coarse level steps in the two-level algorithm (Table 5). When no coarse level steps were used we had the standard trust-region algorithm. In this example $N_x = 50, N_g = 8$ and $\alpha_1/\alpha_2 = 100$. As expected, the computational time decreased when we used coarse level steps. However, if we used too many coarse level steps we no longer obtained appropriate iterates and needed more fine level steps until convergence.

Finally, we investigated the influence of different coarse models on the two-level trust region algorithm. We used the SP_N and the P_N hierarchy and recorded iteration numbers and computing times in Table 6. With larger N the computational times for the P_N model increased due to the increase in the number of equations. Therefore, the higher order P_N models are not

coarse	iter (f/c)	time [sec]
0	11(11/0)	1.61
1	2(2/2)	0.30
2	3(3/4)	0.40
3	4(4/6)	0.50
4	5(5/8)	0.61
5	4(4/8)	0.52

Table 5: Iterations, evaluations and run-time depending on the number of coarse to fine grid steps.

recommended in the two-level algorithm, since the gain in a reduced number of optimization steps is not as large as the additional computational cost for the P_N model.

	SP_N			P_N	
N	iter (f)	time [sec]	N	iter (f)	time [sec]
1	5(8)	0.61	1	7(7)	0.61
2	5(8)	0.61	3	12(12)	3.11
3	5(8)	0.65	5	12(12)	7.30

Table 6: Iterations, evaluations and run-time depending on the order N of the SP_N and P_N approximations, respectively.

3.2 Source inversion problem with angular dependend source

In a second example we considered a reference intensity $\bar{R}$ corresponding to an angular and space dependent source term $\bar{Q}$. The angular dependence is given by superposition of Legendre polynomials (see Figure 3)

$$Q_{ref}(x,\mu) = \frac{1}{2}G_{0.5,0.1}(x)(2 * P_0(\mu) + P_1(\mu) + P_5(\mu))$$

This source radiates a total flux given by a Gaussian $G_{0.5,0.1}(x)$ centered at $x = 0.5$ with variance $\sigma = 0.1$ in the domain. Note that any of our coarse models only have space dependent sources. Even with the higher order P_3 approximations we cannot resolve this source, due to the fifth order polynomial in Q_{ref}.

The source inversion problem can be stated as minimizing (1) without knowing Q_{ref}, which is a typical case in practice. We chose the weights $\alpha_1 = 100$ and $\alpha_2 = 1$ in the functional and set $N_x = 50$ and $N_g = 8$. We used SP_1 as coarse model. In this example, two coarse steps were performed per fine step. The iteration history and the final results of the original

and the two-level trust-region method for the optimal control problem are shown in figures 4 and 5. We observe the smaller number of iterations for the two-level trust region and nearly coinciding optimal controls q for both the full trust-region and the two-level algorithm. Furthermore, it is interesting to note that in the two-level case the control source has little angular dependence while the overall source flux nevertheless coincides well with the standard trust-region result (figures 3, 5).

Table 7 compares the two algorithms for different scalings of the functional. By increasing the value of α_1 for fixed value of α_2, we can give the source term a lower weight compared with the flux-tracking term. In this way the control looses regularity and the functional has less convexity, which makes the problem harder to solve, while at the same time a closer approximation of $\bar{R}$ is enforced. It revealed that the two-level algorithm gives comparable results when there is strong convexity of the functional, i.e., α_1 small. However, the performance deteriorated when α_1 was large.

	two-level			TR		
α_1	iter (f/c)	time [sec]	$\|R-\bar{R}\|$	iter (f)	time [sec]	$\|R-\bar{R}\|$
10^0	7(7/7)	0.16	1.2252	20(20)	2.90	1.1928
10^1	7(11/7)	1.12	0.3942	34(34)	4.88	0.3152
10^2	19(37/35)	3.22	0.0953	59(59)	8.67	0.0450
10^3	11(31/16)	2.06	0.0726	42(42)	6.14	0.0054

Table 7: Iterations, evaluations and run-time depending on the scaling of the functional. The second weight was $\alpha_2 = 1$.

4 Conclusions

Following the approach of Gratton et al. [11], we developed a two-level trust-region method for optimal control problems with radiative transfer which uses the P_N approximations as the trust-region subproblem instead of the usual quadratic model. The P_N approximations, in particular the SP_N approximations, have been shown to be computationally more efficient for radiative transfer problems in many cases. Since the trust-region subproblem based on these approximations is better suited than the standard quadratic function as a model of the orginal functional, the two-level method significantly reduces the number of iterations. This more than compensates the slightly higher cost of the subproblem and leads to an overall reduction of total run-time. In this way the two-level approach is computationally more efficient for these particular optimal control problems.

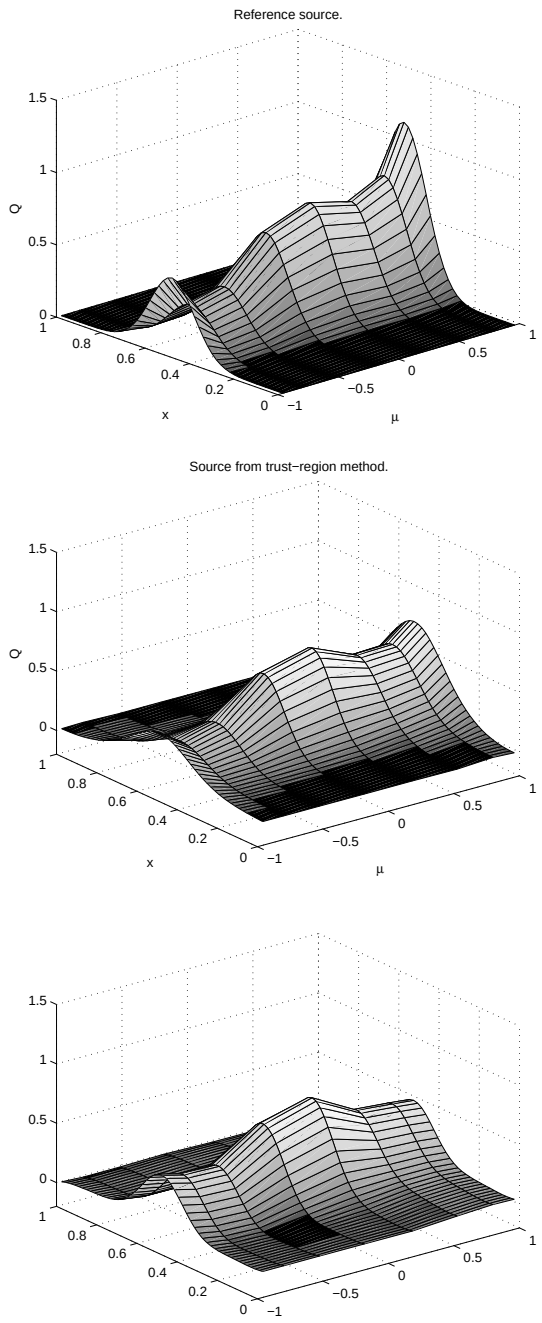

Figure 3: Comparison of the reference source (top) and source given by the control resulting from the standard trust-region method (middle) and the two-level algorithm (bottom).

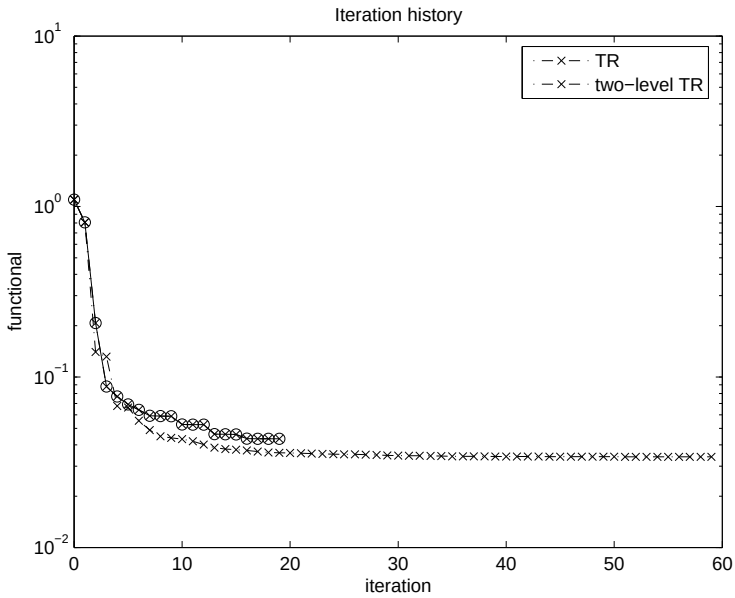

Figure 4: Comparison of the iteration histories for angular dependent source.

References

[1] V. I. Agoshkov, C. Bardos, *Optimal Control Approach in Inverse Radiative Heat Transfer Problems: The Problem of the Boundary Function*, **5**, ESAIM: Control Optimisation and Calculus of Variations, pp. 259-278, (2000)

[2] E. D. Aydin, C. R. E. Oliveira, and A. J. H. Goddard, *A comparison between transport and diffusion calculations using finite element-spherical harmonics radiation transport method*, Med. Phys. **29** (2002), 2013–2023.

[3] J. M. Banoczi, C. T. Kelley, *A Fast Multilevel Algorithm for the Solution of Nonlinear Systems of Conductive-Radiative Heat Transfer Equations in Two Space Dimensions*, **20**, SIAM J. Sci. Comp., pp. 1214–1228, (1999)

[4] C. Börgers, *The radiation therapy planning problem*, IMA Volumes in Mathematics and its applications, vol. 110, Springer-Verlag, 1999.

[5] K. K. Bucci, A. Bevan, and M. Roach III, *Advances in radiation therapy: conventional to 3d, to IMRT, to 4d, and beyond*, CA Cancer J. Clin. **55** (2005), 117–134.

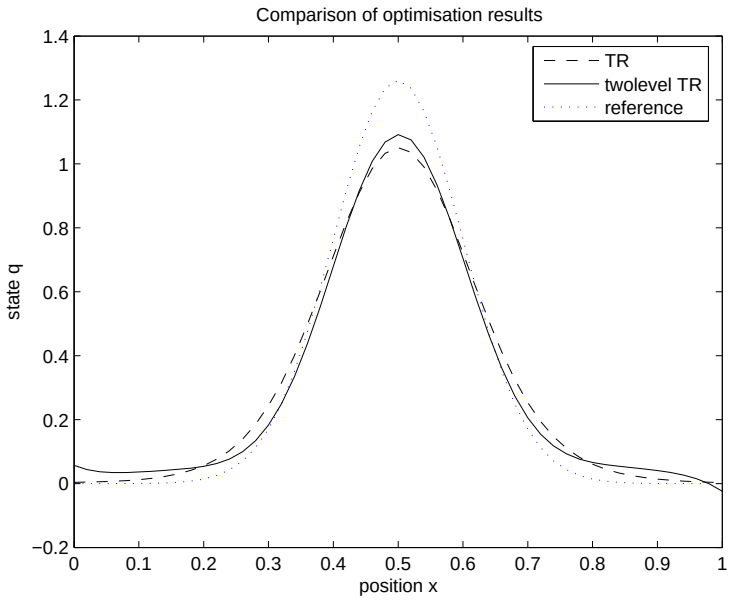

Figure 5: Optimisation results for control for angular dependent source. Plotted is the integrated quantity $q(x) = \int_{S^2} Q d\omega$.

[6] A. R. Conn, N. I. M. Gould, and P. L. Toint, *Trust-Region Methods*, MPS/SIAM Series on Optimization, Philadelphia, 2000.

[7] B. Dubroca and A. Klar, *Half moment closure for radiative transfer equations*, J. Comput. Phys. **180** (2002), 584–596.

[8] M. Fisher, *Minimization algorithms for variational data assimilation*, In Recent Developments in Numerical Methods for Atmospheric Modeling, ECMWF, (1982).

[9] M. Frank, M. Herty, M. Schäfer, *Optimal Treatment Planning in Radiotherapy based on Boltzmann Transport Calculations*, submitted 2006.

[10] M. Frank, M. Schäfer and R. Pinnau. *A Hierarchy of Approximations to the Radiative Heat Transfer Equations: Modelling, Analysis and Simulation.* M3AS, 15:643-665, 2006.

[11] S. Gratton, A. Sartenaer, and P. L. Toint, *Recursive trust-region methods for multiscale nonlinear optimization (part i): Global convergence and complexity*, Technical Report 04/06, (2005).

[12] A. Griewank and P. L. Toint, *Local convergence analysis for partititoned quasi-newton updates*, Numerische Mathematik, 39 (1982), pp. 429–448.

[13] H. Hensel, R. Iza-Teran, and N. Siedow, *Deterministic model for dose calculation in photon radiotherapy*, Phys. Med. Biol. **51** (2006), 675–693.

[14] M. Herty, R. Pinnau, M. Seaid, *On Optimal Control Problems in Radiative Transfer*, preprint, 2005

[15] M. Herty, R. Pinnau, G. Thömmes, *Asymptotic and Discrete Concepts For Optimal Control Problems in Radiative Transfer*, preprint 2006.

[16] E. W. Larsen, *Tutorial: The nature of transport calculations used in radiation oncology*, Transp. theory Stat. Phys. **26** (1997), 739.

[17] E. W. Larsen and J. B. Keller, *Asymptotic solution of neutron transport problems for small mean free path*, J. Math. Phys. **15** (1974), 75.

[18] E. W. Larsen, G: Thömmes, A. Klar, M. Seaïd, and T. Götz, *Simplified P_n approximations to the equations of radiative heat transfer in glass*, J. Comput. Phys. **183** (2002), 652–675.

[19] D. Levermore. *Moment closure hierarchies for kinetic theories.* J. Stat. Phys., 83, 1996.

[20] R. E. Marshak, *Note on the spherical harmonic method as applied to the milne problem on a sphere*, Phys. Rev., 71 (1947), pp. 443–447.

[21] S. Manservisi, K. Heusermann, *On some optimal control problems for the heat radiative transfer equation*, **5**, ESAIM: Control, Optimisation and Calculus of Variations, pp 425-444, (2000)

[22] D. Mihalas and B. S. Mihalas, *Foundations of Radiation Hydrodynamics.* Oxford University Press, New York, 1983.

[23] M. F. Modest, *Radiative Heat Transfer.* McGraw-Hill, 1993.

[24] S. G. Nash, *A multigrid approach to discretized optimization problems*, Optimization Methods and Software, 14 (2000), pp. 99–116.

[25] R. Pinnau, *Analysis of Optimal Boundary Control For Radiative Heat Transfer Modelled by the SP_1 System*, preprint 2005.

[26] R. Pinnau, A. Schulze. *Newton's Method for Optimal Temperature-Tracking of Glass Cooling Processes.* Submitted for publication, 2004.

[27] R. Pinnau and G. Thömmes. *Optimal boundary control of glass cooling processes.* M2AS, 120:1261, 2004.

[28] M. Schäfer, M. Frank, and R. Pinnau, *A hierarchy of approximations to the radiative heat transfer equations: Modelling, analysis and simulation*, Math. Meth. Mod. Appl. Sci. **15** (2005), 643–665.

[29] M. Seaïd, M. Frank, A. Klar, R. Pinnau and G. Thömmes, *Efficient Numerical Methods for Radiation in Gas Turbines*, J. Comp. Applied Math., **170**, pp 217-239, 2004

[30] M. Seaïd, A. Klar and R. Pinnau, *Numerical Solvers for Radiation and Conduction in High Temperature Gas Flows*, Flow, Turbulence and Combustion, **3**, pp 413-432, 2005

[31] 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), 721–744.

[32] H. Struchtrup, *On the number of moments in radiative transfer problems*, Ann. Phys. (N.Y.) **266** (1998), 1–26.

[33] J. Tervo and P. Kolmonen, *Inverse radiotherapy treatment planning model applying boltzmann-transport equation*, Math. Models. Methods. Appl. Sci. **12** (2002), 109–141.

[34] G. Thömmes, R. Pinnau, M. Seaid, T. Götz, and A. Klar. *Numerical methods and optimal control for glass cooling processes.* TTSP, 31(4-6):513–529, 2002.

[35] G. Thömmes, *Radiative Heat Transfer Equations for Glass Cooling Problems: Analysis and Numerics*, PhD Thesis, 2002.

[36] R. E. Alcouffe, *Diffusion synthetic acceleration methods for the Diamond-Difference Discrete-Ordinates Equations*, Nucl. Sci. Eng., 64:344-355, 1977.

[37] E. E. Lewis and W. F. Miller Jr., *Computational Methods of Neutron Transport*, John Wiley and Sons New York, 1984, reprint 1993.

[38] P. Spellucci, *Numerisch Verfahren der nichtlinearen Optimierung*, Birkhäuser Verlag, 1993.