Available online at www.sciencedirect.com
ScienceDirect
Journal of Computational Physics 227 (2008) 6071–6105
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

Sparse adaptive finite elements for radiative transfer

G. Widmer *, R. Hiptmair, Ch. Schwab

Seminar für Angewandte Mathematik, Eidgenössische Technische Hochschule, CH-8092 Zürich, Switzerland

Received 2 February 2007; received in revised form 19 November 2007; accepted 17 February 2008
Available online 7 March 2008

Abstract

The linear radiative transfer equation, a partial differential equation for the radiation intensity $u(\mathbf{x},\mathbf{s})$, with independent variables $\mathbf{x} \in D \subset \mathbb{R}^n$ in the physical domain D of dimension $n = 2, 3$, and angular variable $\mathbf{s} \in S^2 := \{\mathbf{y} \in \mathbb{R}^3 : |\mathbf{y}| = 1\}$, is solved in the $n+2$-dimensional computational domain $D \times S^2$. We propose an adaptive multilevel Galerkin finite element method (FEM) for its numerical solution. Our approach is based on (a) a stabilized variational formulation of the transport operator, (b) on so-called sparse tensor products of two hierarchic families of finite element spaces in $H^1(D)$ and in $L^2(S^2)$, respectively, and (c) on wavelet thresholding techniques to adapt the discretization to the underlying problem. An *a priori* error analysis shows, under strong regularity assumptions on the solution, that the sparse tensor product method is clearly superior to a discrete ordinates method, as it converges with essentially optimal asymptotic rates while its complexity grows essentially only as that for a linear transport problem in $\mathbb{R}^n$. Numerical experiments for $n = 2$ on a set of example problems agree with the convergence and complexity analysis of the method and show that introducing adaptivity can improve performance in terms of accuracy vs. number of degrees even further.

Keywords: Radiative transfer; Wavelet finite elements; Sparse grids; Adaptive finite elements

1. Introduction

We are concerned with the numerical solution of the stationary monochromatic radiative transfer equation [1] on a bounded Lipschitz domain $D \subset \mathbb{R}^n, n = 2$ or 3, with fully absorbing cold walls without scattering.

We identify a direction $\mathbf{s}$ with a point on the sphere S^2 and are looking for the intensity $u(\mathbf{x},\mathbf{s})$, satisfying

$$\mathbf{s} \cdot \nabla_{\mathbf{x}} u(\mathbf{x},\mathbf{s}) + \kappa(\mathbf{x}) u(\mathbf{x},\mathbf{s}) = \kappa(\mathbf{x}) f(\mathbf{x}), \quad (\mathbf{x},\mathbf{s}) \in D \times S^2, \tag{1}$$

$$u(\mathbf{x},\mathbf{s}) = 0, \quad \mathbf{x} \in \partial D, \quad \mathbf{s} \cdot \mathbf{n}(\mathbf{x}) < 0, \tag{2}$$

$\mathbf{n}(\mathbf{x})$ is the outer unit normal on the boundary, $\kappa \geqslant 0$ the absorption coefficient, $f \geqslant 0$ the blackbody intensity.

* Corresponding author. Tel.: +41 44 632 60 38.
E-mail addresses: gisela.widmer@math.ethz.ch (G. Widmer), ralf.hiptmair@math.ethz.ch (R. Hiptmair), christoph.schwab@math.ethz.ch (Ch. Schwab).

0021-9991/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2008.02.025

Regarding the direction $\mathbf{s}$ as a mere parameter, the equation can be solved by line integration for any given position $(\mathbf{x}, \mathbf{s})$. However, as the equation is stated in five (respectively four, for $n = 2$) dimensions, this strategy is too expensive in order to compute the intensity field $u(\mathbf{x}, \mathbf{s})$ with a fine resolution.

Popular methods to solve the radiative transfer problem are, apart from Monte Carlo schemes, the method of spherical harmonics (in particular the P_1 approximation) or the discrete ordinates method. Overviews of numerical methods for radiative transfer can e.g. be found in [1] or [2] and some recent developments in [3].

The method of spherical harmonics is based on a semi-discretization in the solid angle by expanding the intensity into a truncated series of spherical harmonics, which leads to a coupled system of equations in space only. For the P_1 approximation, the equations (1) and (2) boil down to a diffusion equation. The P_N approximation is only suitable when the intensity function is near-isotropic, as the approximation rate with respect to the number of spherical harmonics is very poor for highly non-isotropic functions (cf. [1, Sect. 15]).

In the discrete ordinates method (often referred to as S_N), the equation is solved for N fixed directions. The method is very popular due to its simplicity, but suffers from so-called *ray effects* which require a fine angular resolution if localized emissive areas are present.

In most applications, the systems of equations arising from a P_N- or an S_N approximation are solved with finite difference or finite element schemes. In [4], for example, a least squares formulation is discretized with spherical harmonics in the solid angle and finite elements in space. Kanschat [5] uses a Petrov–Galerkin FE discretization with piecewise constant functions in the solid angle and streamline diffusion stabilization in the physical domain D.

The boundary value problem (1) is posed on an $n + 2$-dimensional domain. Thus, uniform refinement of standard finite element spaces incurs a drastic growth of the number of unknowns by a factor of 2^{n+2} for a moderate gain in extra accuracy. We will present a method to overcome this"curse of dimensionality" – already observed in [6] – for radiative transfer problems with sufficiently smooth absorption coefficients $\kappa(\mathbf{x})$ and blackbody intensity $f(\mathbf{x})$. For problems that lack the required smoothness, we propose an adaptive algorithm, based on *a posteriori* wavelet thresholding techniques, to reduce the number of degrees of freedom. Unlike some other methods for radiative transfer, our method does not require κ to be strictly positive.

The paper is structured as follows: in Section 2, we describe the problem setting and the scaled least squares variational formulation used for deriving our method. Such stabilized formulations have been used e.g. in [7] and in particular for the linear Boltzmann equation arising from the neutron transport problem in [8,4] or [9]. While in [8,9] the absorption coefficient is assumed to be bounded away from zero and the authors in [4] optimize the scaling parameter to balance the absorption and scattering effects for constant coefficients, we tailor the scaling parameter to provide coercivity and continuity estimates for partially degenerate absorption coefficient functions.

Section 3 presents the Galerkin discretization and the construction of the approximation spaces. A tensor product of piecewise linears in D and piecewise constant functions in the solid angle plus a one-point quadrature rule gives a formulation equivalent to the "discrete ordinates method" known to be computationally expensive. Our new idea in constructing approximation spaces is to combine only *tensor products of selected "detail" spaces in two hierarchies of finite element spaces*, one in the physical domain D and one in the solid angle S^2, respectively. The exclusion of most of the possible combinations of detail spaces in our construction without compromising asymptotic convergence rates is, in fact, the key to breaking the curse of dimensionality in this kind of problems.

The core development of this paper is contained in Sections 3.1–3.3. There, we detail the sparse tensor product method, applied to the radiative transfer equation, including the two wavelet bases we use in D and in S^2, respectively. Our approach is inspired by the idea of *hyperbolic cross approximations* and by so-called *sparse grids*. The latter were introduced in the finite element context by Zenger in [10]. Sparse grids have been used for solving a wide range of high-dimensional problems, such as e.g. numerical integration [11], the N-electron Schrödinger-equation in quantum chemistry [12], elliptic [13,14] and parabolic [15] partial differential equations, high-order FE methods [16] or integral equations [17] among others.

Our sparse tensor product method reduces the total number of degrees of freedom to the number of degrees of freedom in physical space D only (up to logarithmic terms). We prove this for solutions of sufficient smoothness and for sparse tensor products of FE-spaces with uniform mesh refinement in D and in S^2. Even without adaptivity, therefore, our sparse tensor method avoids the complexity of $n + 2$-dimensional problems

without compromising convergence orders, at least for sufficiently smooth solutions. For problems like light beams, however, the sparse tensor product method is not optimal in the sense of a *best N-term approximation* [18]. In order to also cover this class of problems, where the smoothness requirement is not fulfilled, we describe a particular adaptive sparse tensor product method in Section 4.

Adaptive sparse grid methods have for example been applied to the Helmholtz equation [19], elliptic PDE's [13] or a singular perturbation model [20]. More general adaptive wavelet methods can e.g. be found in [21–24] and the references cited therein. Starting from the sparse tensor product space described in Section 3.1, we iteratively update the space in order to obtain a solution that is close to a *best N-term approximation*. We emphasize that straightforward adaptive approaches in radiative transfer, as e.g. [5], where in a FE method with a fixed discretization in the solid angle the mesh in the physical domain D is refined according to *a posteriori* error indicators, will improve the constant in complexity estimates but can not break the curse of dimensionality in the solution of the $n+2$-dimensional problem. The sparse, adaptive tensor finite element method introduced here achieves this by adaptively resolving $D \times S^2$ – *interactions in multiscale representations of the solution in D and S^2, respectively*.

Section 5 contains numerical results. We compare the full tensor product approximation, which, in this case, is equivalent to a discrete ordinates method, to the sparse tensor product method with uniform mesh refinement and with the adaptive sparse tensor product method for a set of model problems.

Finally, in Section 6, we comment on extending the present approach to higher order and also on applying sparse tensorization to the P_N-family of methods.

2. Variational formulation

The non-scattering stationary monochromatic radiative transfer equation on a bounded Lipschitz domain $D \subset \mathbb{R}^3$ with fully absorbing cold walls reads:

$$(\mathbf{s}\cdot\nabla_{\mathbf{x}} + \kappa(\mathbf{x}))u(\mathbf{x},\mathbf{s}) = \kappa(\mathbf{x})f(\mathbf{x}) \quad \text{in } D\times S^2, \tag{3}$$

$$u(\mathbf{x},\mathbf{s}) = 0 \quad \text{on } \Gamma_-(\mathbf{s}), \mathbf{s}\in S^2, \tag{4}$$

where $\Gamma_-(\mathbf{s})$ denotes the inflow boundary defined by

$$\Gamma_-(\mathbf{s}) := \{\mathbf{x}\in\partial D : \mathbf{s}\cdot\mathbf{n}(\mathbf{x}) < 0\} \subset \partial D, \quad \mathbf{s}\in S^2. \tag{5}$$

If we assume that $\frac{\partial u(\mathbf{x})}{\partial z} = 0$, where $\mathbf{x} = (x,y,z)' \in \widetilde{D}\times\mathbb{R}, \widetilde{D}\subset\mathbb{R}^2$, the equation reduces to a 2-dimensional problem ($n=2$) for $u(\mathbf{x},\mathbf{s}) = u(\tilde{\mathbf{x}},\mathbf{s}), \tilde{\mathbf{x}} = (x,y,0)'$, in physical space. In that case, we ignore the third component of the inner product $\mathbf{s}\cdot\nabla_{\mathbf{x}}u$.

When regarding $\mathbf{s}$ as a mere parameter, the radiative transfer equation (3)–(4) reduces to a linear convection equation for the directed intensity $u(\cdot,\mathbf{s})$. It is well known that its standard Galerkin discretization by means of continuous trial functions is unstable (e.g. [25]). As we want the variational formulation to be independent of the discretization used and applicable to transparent media (i.e. $\kappa = 0$), we opt for the stabilized variational formulation proposed in [4],

i.e. we seek $u : D\times S^2 \mapsto \mathbb{R}$ as the minimizer of the quadratic least squares functional

$$J(u) := (\epsilon(\mathbf{s}\cdot\nabla_{\mathbf{x}}u + \kappa u - \kappa f), \mathbf{s}\cdot\nabla_{\mathbf{x}}u + \kappa u - \kappa f)_{L^2}, \tag{6}$$

where

$$\epsilon(\mathbf{x}) = \begin{cases} 1, & \kappa(\mathbf{x}) < \kappa_0, \\ \frac{1}{\kappa(\mathbf{x})}, & \kappa(\mathbf{x}) \geqslant \kappa_0 \end{cases} \tag{7}$$

with $\kappa_0 \approx 0.134$ (for details see [26]). In (6), we adopted the notation

$$(u,v)_{L^2} := (u,v)_{L^2(D\times S^2)} = \int_D \int_{S^2} uv \,\mathrm{d}\mathbf{s}\,\mathrm{d}\mathbf{x} \tag{8}$$

and the associated L^2-norm will be denoted by $\|\circ\|$.

For the proper statement of this minimization problem as well as of the FEM below, we define the Hilbert spaces

$$\mathcal{V} := \{u \in L^2(D \times S^2) : \mathbf{s} \cdot \nabla_{\mathbf{x}} u \in L^2(D \times S^2)\}. \tag{9}$$

We equip $\mathcal{V}$ in (9) with the norm $\| \circ \|_S$, defined by

$$\|u\|_S^2 := \|\mathbf{s} \cdot \nabla_{\mathbf{x}} u\|^2 + \|u\|^2 \tag{10}$$

and introduce the subsets

$$\mathcal{V}_0 := \{u \in \mathcal{V} : u = 0 \text{ on } \Gamma_-(\mathbf{s}), \mathbf{s} \in S^2\}. \tag{11}$$

Here, the homogeneous essential boundary condition on $\Gamma_-(\mathbf{s})$ is well defined as $\mathcal{V}_0$ is a closed, linear subspace of $\mathcal{V}$ due to the continuity of the corresponding trace operator, cf. [27,28] or [29, Chapter XXI, Section 2].

Next we introduce the bilinear form

$$a(u,v) := (\epsilon \mathbf{s} \cdot \nabla_{\mathbf{x}} u, \mathbf{s} \cdot \nabla_{\mathbf{x}} v)_{L^2} + (\epsilon \mathbf{s} \cdot \nabla_{\mathbf{x}} u, \kappa v)_{L^2} + (\epsilon \kappa u, \mathbf{s} \cdot \nabla_{\mathbf{x}} v)_{L^2} + (\epsilon \kappa u, \kappa v)_{L^2} \tag{12}$$

and define the "source" functional

$$l(v) := (\epsilon \kappa^2 f, v)_{L^2} + (\epsilon \kappa f, \mathbf{s} \cdot \nabla_{\mathbf{x}} v)_{L^2}. \tag{13}$$

Then the resulting linear variational problem reads: Seek $\tilde{u} \in \mathcal{V}_0$ such that

$$a(\tilde{u}, v) = l(v) \quad \forall v \in \mathcal{V}_0. \tag{14}$$

For $n = 2$ we further require that there is a constant $C > 0$ such that

$$\|\mathbf{s} \cdot \nabla_{\mathbf{x}} u\| \geqslant C\|u\|. \tag{15}$$

Then the following theorem holds [4].

Theorem 2.1. *For every non-negative and bounded κ the bilinear form $a(u,v)$ is continuous on $\mathcal{V} \times \mathcal{V}$ and coercive on $\mathcal{V}_0 \times \mathcal{V}_0$ equipped with the norm $\| \circ \|_S$. In particular, for every $f \in L^2(D \times S^2)$,there exists a unique weak solution $\tilde{u} \in \mathcal{V}_0$ of the stabilized variational form (14) and (4).*

Although the proofs in [4] are restricted to piecewise constant absorption coefficients, the extension to non-constant coefficients is straightforward (see e.g. [26]).

As the bilinear form $a(.,.)$ is symmetric and positive definite on $\mathcal{V}_0$, the expression

$$\|u\|_A := \sqrt{a(u,u)} \tag{16}$$

defines a norm on $\mathcal{V}_0$, to which we will refer as "energy", or $A(D \times S^2)$ norm below.

Remark 1. In this paper, we restrict ourselves to non-scattering radiative transfer and the least squares formulation stated above. However, the (adaptive) sparse tensor product method could also be applied to other variational formulations and to problems that consider scattering effects (see e.g. [4] or [9]).

3. Galerkin discretization

From now on, the variational problem (14) will be considered on the space

$$V_0 := H^{1,0}(D \times S^2) \cap \mathcal{V}_0, \quad H^{1,0}(D \times S^2) = H^1(D) \otimes L^2(S^2). \tag{17}$$

Since V_0 is a proper, closed subspace of $\mathcal{V}_0$ (and, hence, of $\mathcal{V}$), the variational problem (14), *restricted to* V_0, admits a unique weak solution $\bar{u} \in V_0$. *In what follows, we shall assume that the two weak solutions, $\tilde{u} \in \mathcal{V}_0$ and $\bar{u} \in V_0$, of (14) coincide and denote this solution by u.* This is a regularity assumption stating that the weak solution $u \in \mathcal{V}$ of (3) and (4) belongs, in fact, to $H^1(D) \otimes L^2(S^2)$. Note that this assumption precludes line discontinuities of u in D which may arise due to transport along rays of discontinuous boundary data. The Galerkin discretization of (14) is obtained, as usual, by restricting $u = \tilde{u}$ and v in the weak formulation (14) to a

one-parameter family of finite dimensional subspaces $\{V_0^L\}_L$ of $\mathcal{V}_0$, where the superscript L will denote "level" of mesh refinement. This yields

$$u_L \in V_0^L : a(u_L, v) = l(v) \quad \forall v \in V_0^L. \tag{18}$$

Due to the coercivity and continuity of $a(\cdot,\cdot)$ on $\mathcal{V}_0 \times \mathcal{V}_0$, (18) admits a unique solution which satisfies the Galerkin orthogonality

$$\forall v \in V_0^L : a(u - u_L, v) = 0. \tag{19}$$

As a consequence of (19) and of Theorem 2.1, the error $e_L = u - u_L$ is quasioptimal in the $\|\circ\|_S$-norm (10), i.e. for every subspace V_0^L of V_0 which is closed in V_0 with respect to the $\|\circ\|$-norm we obtain

$$\|u - u_L\|_S \leqslant C(\kappa, D) \inf_{v_L \in V_0^L} \|u - v_L\|_S. \tag{20}$$

We exploit this stability to construct sequences V_0^L of subspaces of V_0 by sparse tensorization. Since the computational domain $D \times S^2$ is a cartesian product of domains, we build V_0^L out of tensor products of "component" finite element spaces in D and in S^2, respectively. Note, however, that due to the $\mathbf{s}$ dependence of the Dirichlet boundary $\Gamma_-(\mathbf{s}) \subset \partial D$, the subspaces V_0^L will generally *not* be of tensor product type, once the boundary condition (4) is imposed.

Let us start by giving the construction of the component spaces without boundary conditions. To this end, we equip the domain D with a triangular ($n = 2$) or tetrahedral ($n = 3$) mesh $\mathcal{T}_D^0$, and the sphere with a mesh $\mathcal{T}_{S^2}^0$ consisting of spherical triangles. The hierarchic mesh sequences $\mathcal{T}_D^l, \mathcal{T}_{S^2}^l, l = 1, \ldots, L$, are then obtained by uniform dyadic refinement of the coarse meshes (see Fig. 1).

On the hierarchic mesh sequences, we specify finite element (FE) spaces. In physical space D, the finite element space $V_D^L := S^{p,1}(D, \mathcal{T}_D^L) \subset H^1(D)$ consists of piecewise polynomial functions of degree $p \geqslant 1$ on the finest triangulation $\mathcal{T}_D^L$ which are continuous in the physical domain D. In the solid angle $\mathbf{s} \in S^2$, we use $V_{S^2}^L := S^{q,0}(S^2, \mathcal{T}_{S^2}^L) \subset L^2(S^2)$ of discontinuous, piecewise polynomials of degree $q \geqslant 0$ on the spherical triangles of $\mathcal{T}_{S^2}^L$. In the implementation ahead, we realized the simplest case $p = 1$ and $q = 0$.

Based on the FE spaces V_D^L and $V_{S^2}^L$ in the "component domains" D and S^2, we define the tensor product finite element space $V_0^L \subset V_0$ on the cartesian product mesh $\mathcal{T}_D^L \times \mathcal{T}_{S^2}^L$ at refinement level L by

$$V_0^L := V^L \cap \mathcal{V}_0 \tag{21}$$

with

$$V^L := V_D^L \otimes V_{S^2}^L = S^{p,1}(D, \mathcal{T}_D^L) \otimes S^{q,0}(S^2, \mathcal{T}_{S^2}^L). \tag{22}$$

The Galerkin discretized problem then reads: find $u_L(\mathbf{x}, \mathbf{s}) \in V_0^L$ such that

$$a(u_L, v_L) = l(v_L) \quad \forall v_L \in V_0^L. \tag{23}$$

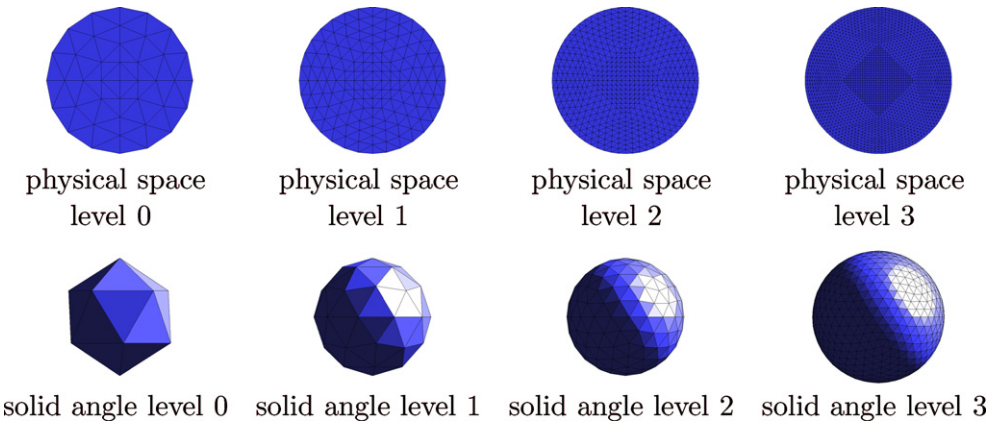

Fig. 1. Mesh hierarchies used in the sparse tensor product space.

Let $\alpha_i(\mathbf{x}),\ i = 1,\ldots,M_L = \dim V_D^L$, be a basis of V_D^L and $\beta_j(\mathbf{s}),\ j = 1,\ldots,N_L = \dim V_{S^2}^L$, a basis of $V_{S^2}^L$.

Remark 2. Note already at this stage the essential feature *that the basis function* $\alpha_i(\mathbf{x})$ *and* $\beta_j(\mathbf{s})$ *of the "component" spaces* V_D^L *and* $V_{S^2}^L$ *are assumed to be independent of* L. This means that *spaces and bases in* D *and* S^2 *are assumed to be hierarchical* which rules out, for example, the use of classical "hat" function finite element bases in D.

The approximate intensity $u_L \in V^L$ can be expressed in the tensor product form

$$u_L(\mathbf{x},\mathbf{s}) = \sum_{i=1}^{M_L}\sum_{j=1}^{N_L} u_{ij}\alpha_i(\mathbf{x})\beta_j(\mathbf{s}) \in V^L. \tag{24}$$

Then (23) leads to a linear system of equations for the $M_L \cdot N_L$ unknowns u_{ij}.

A natural choice of bases in the simplest case $p = 1$ (continuous, piecewise linear elements in D) and $q = 0$ (discontinuous, piecewise constant elements in S^2) is

- locally supported piecewise linear "hat functions" for V_D^L that is $\alpha_i(x_j) = \delta_{ij}$, where $\{x_1,\ldots,x_{M_L}\}$ is the set of vertices of $\mathcal{T}_D^L$,
- and the characteristic functions of the triangles of $\mathcal{T}_{S^2}^L$ as $\beta_j,\ j = 1,\ldots,N_L$.

In order to impose the boundary conditions in our implementation, which uses $p = 1$ and $q = 0$, we take the nodal basis in physical space and the characteristic functions of triangles $T \in \mathcal{T}_{S^2}^L$ as a basis in the solid angle S^2. We then set all degrees of freedom u_{ij} to zero, if i corresponds to a "inflow" boundary node $x_i \in \Gamma_-(\mathbf{s})$ for $\mathbf{s} \in \mathrm{supp}(\beta_j) \subset S^2$.

As mentioned previously, the Galerkin tensor product discretization with $q = 0$ is equivalent to a discrete ordinates method [1], if a one-point quadrature rule is applied in the solid angle. The discrete ordinates method is usually referred to as S_N, where $N = N_L$ is the number of discrete ordinates [1, Section 16]. As the number of degrees of freedom is roughly $M_L \cdot N_L$, where both M_L and N_L are the number of basis functions of a discretization in D and in S^2, respectively, this method is not able to overcome the curse of dimensionality even with adaptive mesh refinements in D and in S^2.

A first attempt is to try and keep both N_L and M_L small by picking *adapted bases* that offer a good representation of the solution with only a few degrees of freedom. The construction of such basis functions can be pursued via local adaptive mesh refinement starting from $\mathcal{T}_D^0$ and $\mathcal{T}_{S^2}^0$. Then, the ultimate trial and test functions are built according to (24). The limitations of this approach are evident: since the space $V_{S^2}^L$ is expected to provide good resolution of the radiation everywhere, $\mathcal{T}_{S^2}^L$ will usually have to be a fairly uniform mesh.

Yet, using all the product basis functions $\alpha_i(\mathbf{x})\beta_j(\mathbf{s}),\ i = 1,\ldots,M_L,\ j = 1,\ldots,N_L$, as in (24) may not be necessary at all, because only a few of them may really make a significant contribution to representing the final solution. Hence, a promising approach to obtaining efficient trial spaces is to select a few significant *product basis functions* of the form $\alpha_i(\mathbf{x})\beta_j(\mathbf{s})$ and let them span V^L. The *component basis functions* α_i and β_j may be chosen from large, even infinite, sets, which will not translate into prohibitively large discrete problems. This idea underlies the present approach to the Galerkin discretization of the radiative transfer problem which is based on *sparse tensor products* of the *hierarchic* component finite element spaces V_D^L and $V_{S^2}^L$. In the following, we shall discuss two choices of the sparse tensor product space – the *a priori* selection of combinations of *hierarchic* basis functions, and the *a posteriori*, adaptive selection of such combinations.

3.1. Sparse tensor product space

The selection of significant product basis functions needs to follow strict rules in order to overcome the curse of dimensionality without compromising accuracy. Such a set of rules for selecting basis functions *a priori* is offered by the framework of sparse grids [10,30]. In the following, we adapt these ideas to construct sparse tensor products of FE spaces in $H^{1,0}(D\times S^2) \simeq H^1(D)\otimes L^2(S^2)$. To do so, we have to exploit the *hierarchic, multilevel structure* of the sets $\{\alpha_i\}_i$ and $\{\beta_j\}_j$ of basis functions. They have to allow a partitioning into subsets associated with different *levels of resolution*. Crudely speaking, the level of a basis function corresponds

to its importance for approximating smooth functions: basis functions on level 0 make significant contributions, those on levels with a large index do not contribute much.

The hierarchic, multilevel structure of the finite element spaces will naturally arise from the nested triangulations $\mathcal{T}_D^l$ and $\mathcal{T}_{S^2}^l, l = 0, 1, 2, \ldots,$. On these, we define the corresponding nested sequences of finite element spaces

$$V_D^l := S^{p,1}(D, \mathcal{T}_D^l) \subset H^1(D), \quad V_{S^2}^l := S^{q,0}(S^2, \mathcal{T}_{S^2}^l) \subset L^2(S^2). \tag{25}$$

Here, $S^{p,1}(D, \mathcal{T}_D^l)$ denotes the continuous, piecewise polynomial functions of degree $p \geqslant 1$ on $\mathcal{T}_D^l$ and $S^{q,0}(S^2, \mathcal{T}_{S^2}^l)$ denotes the space of possibly discontinuous, piecewise polynomial functions of degree $q \geqslant 0$. These sequences of spaces are, in turn, nested, and there are so-called "detail spaces" $W_D^l, W_{S^2}^l$ such that

$$V_D^l = W_D^l \oplus V_D^{l-1}, \quad V_{S^2}^l = W_{S^2}^l \oplus V_{S^2}^{l-1}, \tag{26}$$

where $\oplus$ is the orthogonal direct sum with respect to the corresponding L^2 inner products.

Iterating (26), we see that for $l \geqslant 1$ the spaces V_D^l and $V_{S^2}^l$ possess an L^2- orthogonal decomposition into the detail subspaces W_D^l and $W_{S^2}^l$, respectively:

$$V_D^l = \bigoplus_{i=0}^{l} W_D^i, \quad V_{S^2}^l = \bigoplus_{i=0}^{l} W_{S^2}^i, \tag{27}$$

where we set, for notational convenience $W_D^0 := V_D^0$ and $W_{S^2}^0 := V_{S^2}^0$, respectively. With these definitions, the full tensor product space $V^L \subset H^1(D) \otimes L^2(S^2)$ at mesh refinement level L is easily seen to coincide with

$$V^L = V_D^L \otimes V_{S^2}^L = \bigoplus_{0 \leqslant l_1, l_2 \leqslant L} W_D^{l_1} \otimes W_{S^2}^{l_2}. \tag{28}$$

We shall consider the sparse tensor product space $\widehat{V}^L \subset V^L$ defined by

$$\widehat{V}^L := \bigoplus_{0 \leqslant l_1 + l_2 \leqslant L} W_D^{l_1} \otimes W_{S^2}^{l_2}. \tag{29}$$

This means that basis functions of level l in subspaces in D are tensorized only with basis functions in S^2 up to level $L - l$ and vice versa, see Fig. 2 for an illustration. We set

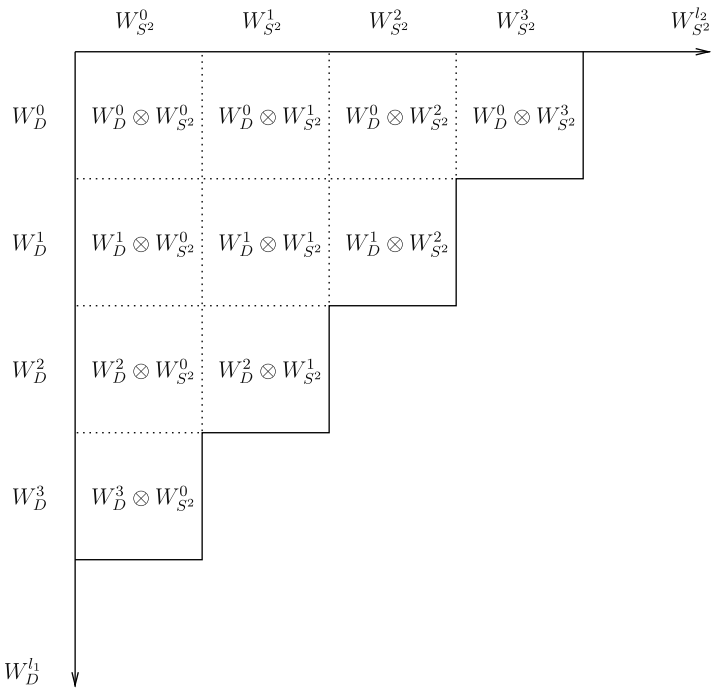

Fig. 2. Component spaces of the sparse tensor product space $\widehat{V}_L$ (for $L = 3$).

$$V_0^L := V^L \cap \mathcal{V}_0, \quad \widehat{V}_0^L \subset \widehat{V}^L \cap \mathcal{V}_0. \tag{30}$$

A simple counting argument also shows that

$$\dim V_0^L = O(N_L M_L), \quad \widehat{\mathcal{N}}_L := \dim \widehat{V}_0^L = O(N_L \log M_L + M_L \log N_L) \tag{31}$$

as $L \to \infty$.

3.2. Approximation properties

Since the sparse tensor product space $\widehat{V}^L$ is substantially smaller than the full tensor product space V^L, accuracy may be lost. We shall now estimate the rate of convergence of the sparse Galerkin approximations to u which show that, at least for smooth solutions, both spaces achieve the same asymptotic convergence rate.

To this end, we define the L^2-projection operators $P_D^l : L^2(D) \to V_D^l$ and $P_{S^2}^l : L^2(S^2) \to V_{S^2}^l$ with the convention that $P_D^{-1} = P_{S^2}^{-1} = 0$. Then the projector $\widehat{P}_L$ onto the sparse tensor product space of level L can be expressed as telescopic sum of projections onto products of the "detail" component spaces

$$\widehat{P}_L u(\mathbf{x}, \mathbf{s}) = \sum_{0 \leqslant l_1 + l_2 \leqslant L} (P_D^{l_1} - P_D^{l_1 - 1}) \otimes (P_{S^2}^{l_2} - P_{S^2}^{l_2 - 1}) u(\mathbf{x}, \mathbf{s}). \tag{32}$$

The projection $P_L = P_D^L \otimes P_{S^2}^L$ onto V^L can be represented analogously, if the summation extends over $0 \leqslant l_1, l_2 \leqslant L$ instead.

In order to describe the approximation properties of the sparse tensor product space $\widehat{V}_L$, we follow [17,15] and introduce anisotropic Sobolev spaces with fractional derivatives. We start by defining, for $m, n \in \mathbb{N}_0$, the anisotropic Sobolev spaces

$$H^{s,t}(D \times S^2) := H^s(D) \otimes H^t(S^2), \tag{33}$$

which can, for integer values of s and of t, equivalently be defined by

$$\{u \in L^2(D \times S^2) | D_{\mathbf{x}}^{\alpha} D_{\mathbf{s}}^{\beta} u \in L^2(D \times S^2), 0 \leqslant |\alpha| \leqslant s, 0 \leqslant |\beta| \leqslant t\}, \tag{34}$$

where for $\alpha \in \mathbb{N}_0^n$, $D_{\mathbf{x}}^{\alpha}$ denotes the αth weak derivative with respect to $\mathbf{x} \in D$; we denote its order by $|\alpha| = \alpha_1 + \cdots + \alpha_n$. Analogously, for $\beta \in \mathbb{N}_0^2$, $D_{\mathbf{s}}^{\beta}$ denotes the weak derivative with respect to $\mathbf{s} \in S^2$ and we denote its order $|\beta| = \beta_1 + \beta_2$.

We equip the anisotropic space with the norm

$$\|u\|_{H^{s,t}}^2 := \sum_{\substack{0 \leqslant |\alpha| \leqslant s \\ 0 \leqslant |\beta| \leqslant t}} \|D_{\mathbf{x}}^{\alpha} D_{\mathbf{s}}^{\beta} u\|_{L^2(D \times S^2)}^2. \tag{35}$$

For arbitrary $s, t \geqslant 0$, we define $H^{s,t}(D \times S^2)$ by tensorization and interpolation.

For functions $u(\mathbf{x}, \mathbf{s}) \in H^{1,0}(D \times S^2)$ that are sufficiently smooth, the following theorem gives crucial approximation properties of the sparse grid space, see [26] for the proof.

Theorem 3.1. *Under the assumption that $p = q + 1 \geqslant 1$ and that $u \in H^{p+1,p}(D \times S^2)$, the best-approximation in the full tensor product space V^L satisfies the asymptotic error estimate*

$$\inf_{v \in V} \|u - v\|_A \lesssim \|u - P_L u\|_S \lesssim \|u - P_L u\|_{H^{1,0}(D \times S^2)} \tag{36}$$

$$\lesssim h_L^p \|u\|_{(H^{p+1,0} \cap H^{1,q+1})(D \times S^2)} \tag{37}$$

and best-approximation in the sparse tensor product space $\widehat{V}^L$ satisfies the error estimate

$$\inf_{v \in \widehat{V}} \|u - v\|_A \lesssim \|u - \widehat{P}_L u\|_S \lesssim \|u - \widehat{P}_L u\|_{H^{1,0}(D \times S^2)} \tag{38}$$

$$\lesssim h_L^p |\log h_L| \|u\|_{H^{p+1,p}(D \times S^2)}. \tag{39}$$

The number of degrees of freedom required for these solutions behaves, as $L \to \infty$, in the full tensor product spaces and in the sparse tensor product spaces, respectively, as

$$\dim(V_0^L) \sim N_L M_L \quad \text{and} \quad \dim(\widehat{V}_0^L) \sim N_L \log M_L + M_L \log N_L. \tag{40}$$

Here, $\sim$ and $\lesssim$ indicates equivalence up to constants which may depend on p, s, etc. but which are independent of h_L and of N_L, M_L as $L \to \infty$.

We see that, up to a logarithmic factor, the convergence rates attainable with the full and the sparse tensor product discretizations are identical, while the number of degrees of freedom in the sparse tensor product space is, again up to logarithmic factors, the same as that of the component spaces. In effect, the use of the sparse tensor product space $\widehat{V}_0^L$ in (18) reduces the FEM complexity from $3+2$-dimensional domain to essentially that of a FE computation in a 3-dimensional domain while retaining (up to logarithmic terms) the asymptotic rate of convergence *provided that $u \in \mathcal{V}_0$ is sufficiently smooth.*

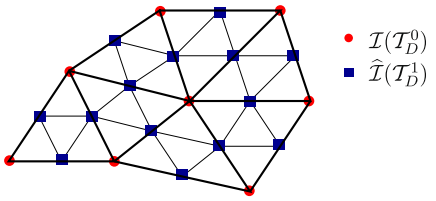

Fig. 3. Index sets $\mathcal{I}(\mathcal{T}_D^0)$ and $\widehat{\mathcal{I}}(\mathcal{T}_D^1)$.

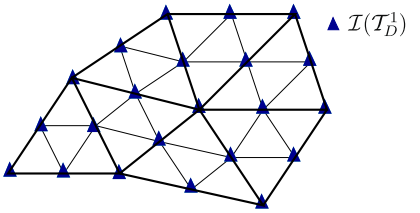

Fig. 4. Index set $\mathcal{I}(\mathcal{T}_D^1)$.

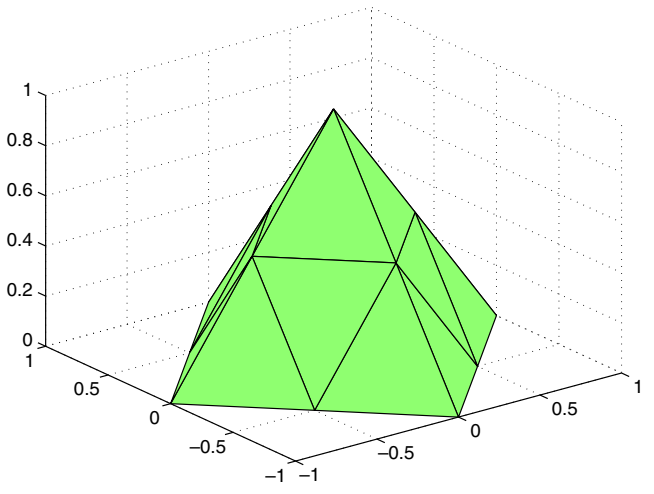

Fig. 5. $\phi_{l-1}^k(\mathbf{x})$.

This cannot be guaranteed for general absorption and emission data. However, even if the solution does not satisfy the regularity requirements needed for the approximation rates stated above, in most regions of the product domain, the solution will be sufficiently smooth. So, even in these cases the sparse tensor product space is a suitable *initial guess* based on which a space with better approximation properties can be constructed by adaptive wavelet thresholding techniques (see Section 4).

3.3. Wavelet finite element bases

Any implementation of the sparse tensor product method does require bases of the multilevel "detail" subspaces $W_D^l, W_{S^2}^l$. As we already indicated in Remark 2, the usual "hat" function finite element shape functions are *not* hierarchical and therefore unsuitable to realize a sparse tensor product method. Best suited for our purposes are either hierarchical or wavelet finite element bases. The latter offer the following important advantages over the former:

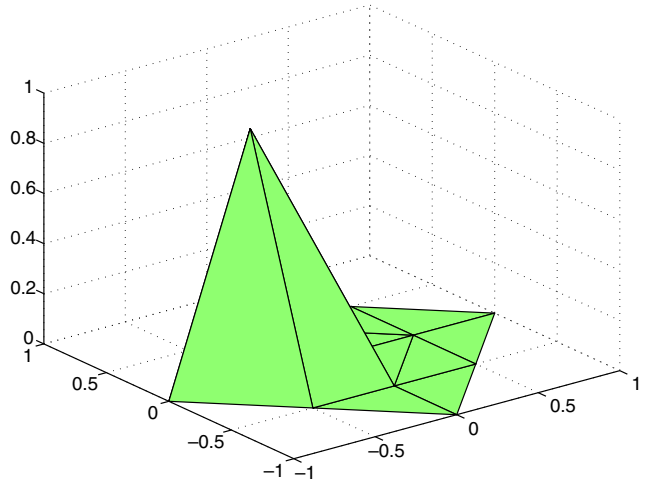

Fig. 6. $\phi_l^j(\mathbf{x})$.

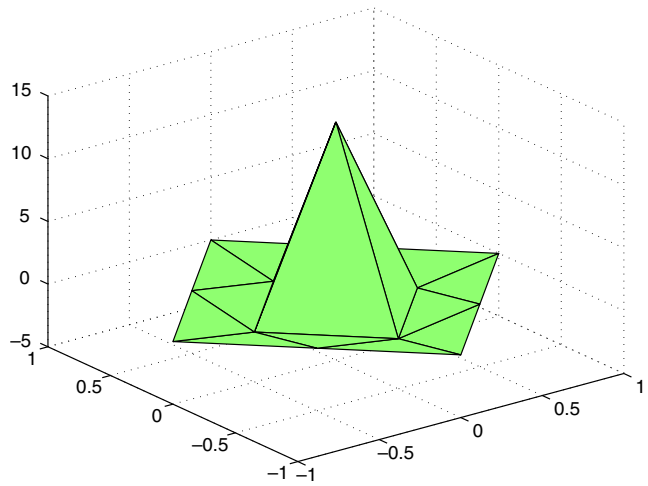

Fig. 7. $\theta_l^k(\mathbf{x})$.

(a) wavelets exhibit vanishing moments which allows to characterize the regularity of the intensity u in terms of the coefficient decay in its wavelet expansion. This is the basis for using thresholding of wavelet coefficients to steer adaptive refinements in both, D and S^2, simultaneously,
(b) (not exploited or addressed yet in the present paper) for scattering operators that are non-local on S^2, wavelet type basis functions allow for compression of the corresponding stiffness matrices without sacrificing convergence orders which means that even in this case the computational costs are of order $\mathcal{O}(N \log M + M \log N)$, where M is the number of degrees of freedom in physical space and N is the number of degrees of freedom in solid angle.

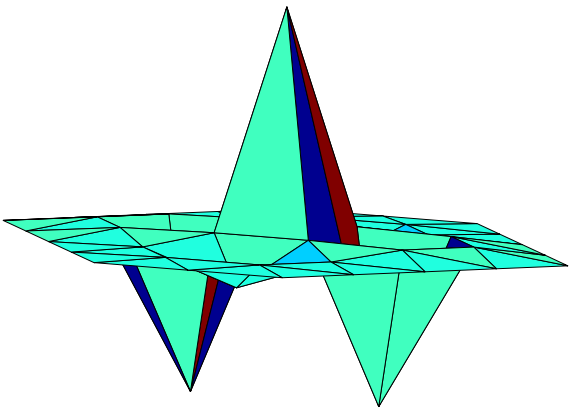

Fig. 8. Piecewise linear FE wavelet.

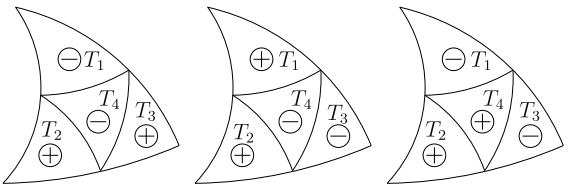

Fig. 9. Haar wavelet basis functions on spherical triangle subdivided into four triangles $T_1, \ldots, T_4$.

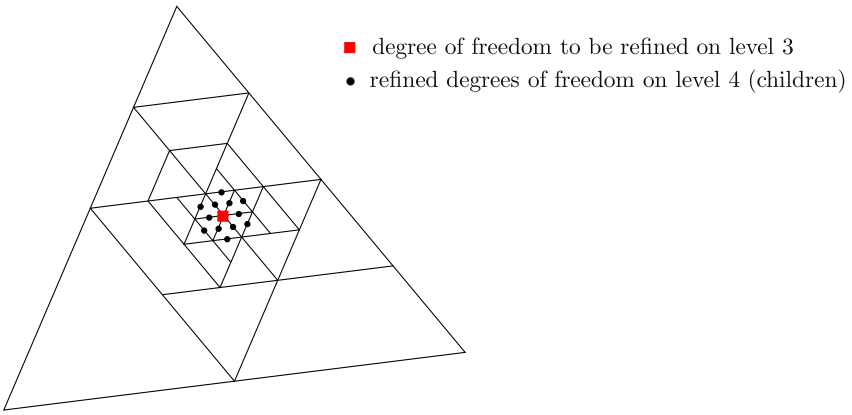

Fig. 10. Refinement of a degree of freedom in 2D space.

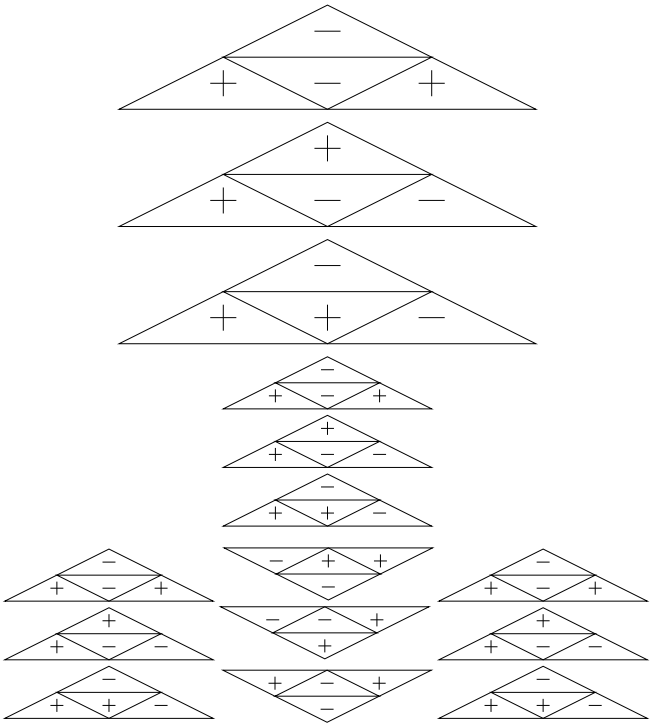

Fig. 11. Parents–children relationship of Haar wavelet functions on nested triangulations on the sphere.

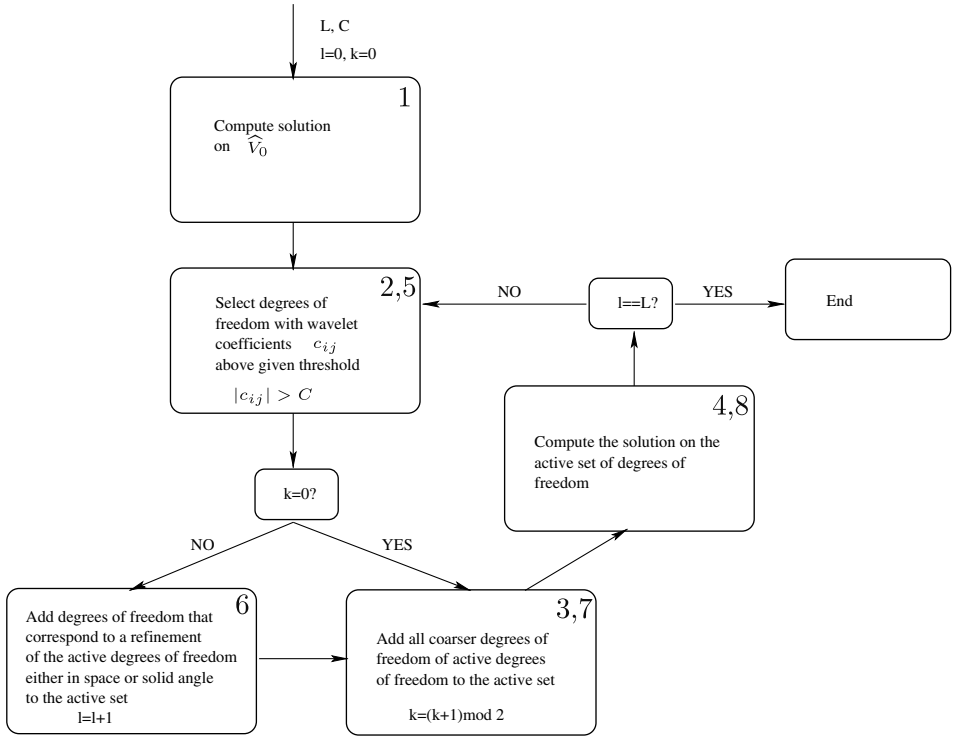

Fig. 12. Flowchart of the algorithm.

To date, several constructions of wavelet FE in general domains D are available. In particular, we use in D the isotropic, piecewise linear and continuous finite element wavelet basis described in [31]. They are constructed level-wise. On level 0, the basis functions $\psi_0^m(\mathbf{x}),\ m \in \mathcal{I}(\mathcal{T}_D^0)$, are the standard hat functions on the coarsest mesh. With $\mathcal{I}(\mathcal{T}_D^l)$ we denote the index set of vertices of the mesh $\mathcal{T}_D^l$ and with $\widehat{\mathcal{I}}(\mathcal{T}_D^l)$ the index set of vertices of $\mathcal{T}_D^l$ that do not belong to $\mathcal{T}_D^{l-1}$ (see Figs. 3 and 4).

On a higher level $l > 0$, the construction of the wavelet functions $\psi_l^m(\mathbf{x}), m \in \widehat{\mathcal{I}}(\mathcal{T}_D^l)$, is based on the meshes $\mathcal{T}_D^l$ and $\mathcal{T}_D^{l-1}$. For $j \in \widehat{\mathcal{I}}(\mathcal{T}_D^l)$, we define $\phi_l^j(\mathbf{x})$ to be the hat function of vertex j on mesh $\mathcal{T}_D^l$.

We now construct a family of functions $\theta_l^i(\mathbf{x}) \in S^{1,1}(\mathcal{T}_D^l), i \in \mathcal{I}(\mathcal{T}_D^{l-1})$, that satisfy $(\theta_l^i, \phi_{l-1}^k)_{L^2(D)} \simeq \delta_{ik}$.

As it can easily be verified, the piecewise linear functions $\theta_l^i(\mathbf{x})$ with

$$\theta_l^i(\mathbf{v}) = \begin{cases} 14, & \mathbf{v} \text{ is vertex } x_i \text{ of mesh } \mathcal{T}_D^{l-1} \\ -1, & \mathbf{v} \text{ is neighboring vertex of } x_i \text{ on mesh } \mathcal{T}_D^l \\ 0, & \mathbf{v} \text{ is any other vertex of mesh } \mathcal{T}_D^l \end{cases} \tag{41}$$

fulfill this condition. Examples of the functions $\phi_{l-1}^k(\mathbf{x}), \phi_l^j(\mathbf{x})$ and $\theta_l^k(\mathbf{x})$ are displayed in Figs. 5–7.

The wavelets on level l are then obtained by

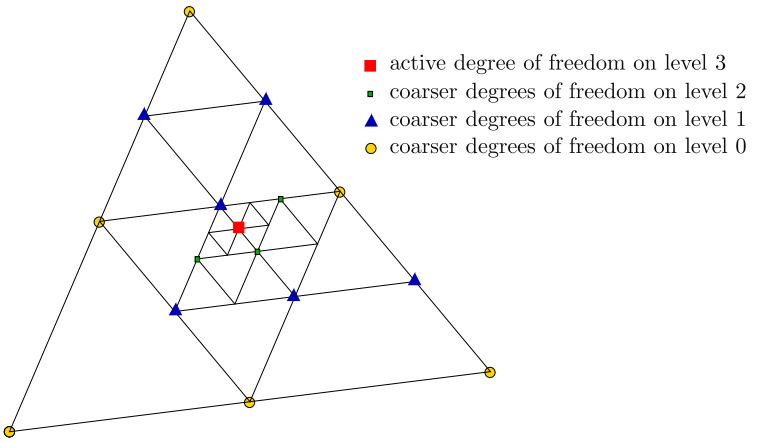

Fig. 13. Coarser degrees of freedom in 2D space that have to be contained in the mesh when a degree of freedom on the finest level is active.

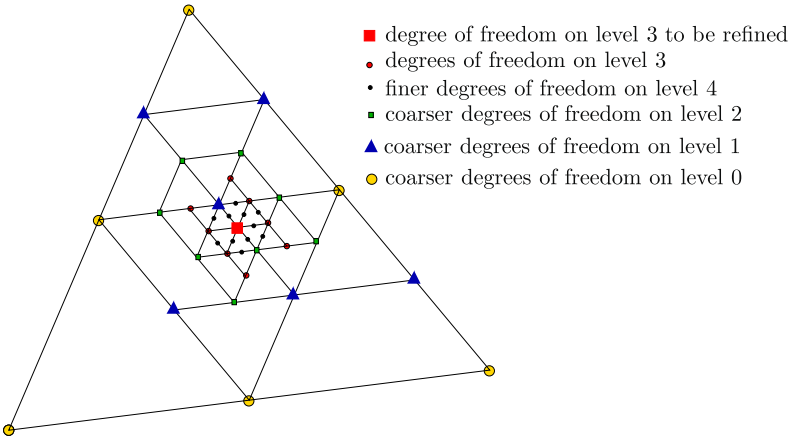

Fig. 14. Active degrees of freedom in 2D space after refining a degree of freedom on level 3.

$$\psi_l^m(\mathbf{x}) = \phi_l^m(\mathbf{x}) - \sum_{k \in \mathcal{I}(\mathcal{T}_D^{l-1})} \frac{(\phi_l^m, \phi_{l-1}^k)_{L^2(D)}}{(\theta_l^k, \phi_{l-1}^k)_{L^2(D)}} \theta_l^k(\mathbf{x}), \quad m \in \widehat{\mathcal{I}}(\mathcal{T}_D^l). \tag{42}$$

An example is shown is Fig. 8. With $W_D^j = \operatorname{span}\{\psi_j^m, m \in \widehat{\mathcal{I}}(\mathcal{T}_D^j)\}$, $V_D^j = \operatorname{span}\{\psi_i^m,\ 0 \leqslant i \leqslant j,\ m \in \widehat{\mathcal{I}}(\mathcal{T}_D^i)\}$, they fit into the framework of Section 3.1.

The functions $\frac{\psi_l^m(\mathbf{x})}{\|\psi_l^m\|}, m \in \widehat{\mathcal{I}}(\mathcal{T}_D^l), l \in \mathbb{N}$, form a Riesz basis for $L^2(D)$, whereas the scaled functions $\tilde{\psi}_l^m(\mathbf{x}) = 2^{-l}\frac{\psi_l^m(\mathbf{x})}{\|\psi_l^m\|}, l \in \mathbb{N}, m \in \widehat{\mathcal{I}}(\mathcal{T}_D^l)$ form a Riesz basis for $H^1(D)$ (see e.g. [24] and the references therein). As a consequence, in a wavelet decomposition of

$$u = \sum_{l,m} u_{lm}\tilde{\psi}_m^l \quad \in H^1(D) \tag{43}$$

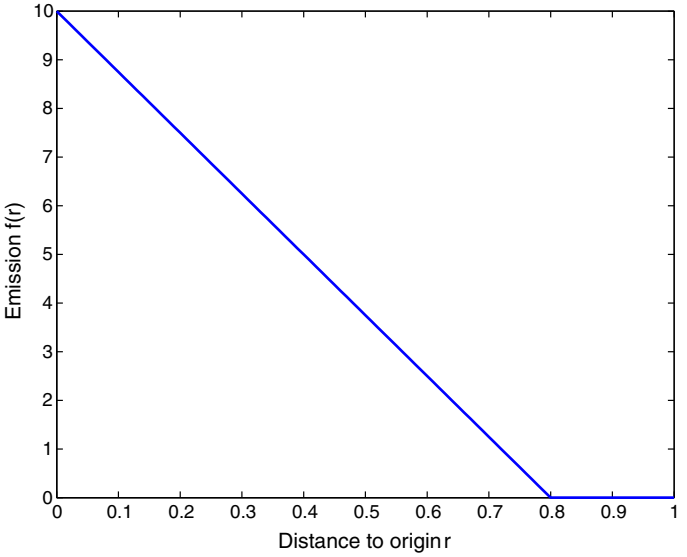

Fig. 15. Blackbody intensity $f(\mathbf{x})$ of Example 1.

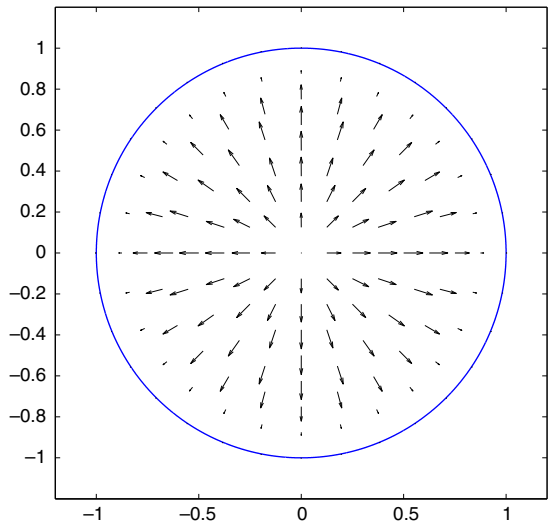

Fig. 16. Heat flux of Example 1 (reference solution).

the size of a coefficient u_{lm} provides direct information about the contribution of the corresponding basis function $\tilde{\psi}_m^l$ with respect to the $H^1(D)$-norm.

On unstructured, triangular meshes on S^2, we use agglomerated Haar wavelets (see e.g. [32]) that are slightly adapted for the sphere. As we use piecewise constant functions, the number of degrees of freedom corresponds to the number of spherical triangles.

On refinement level 0 in S^2, χ_0^n are the characteristic functions on the triangles T_n of the coarsest triangulation. On higher levels $l > 0$, the basis functions χ_l^n are based on the meshes $\mathcal{T}_{S^2}^l$ and $\mathcal{T}_{S^2}^{l-1}$, where $n \in$ index set on level l. The support of χ_l^n is a triangle on $\mathcal{T}_{S^2}^{l-1}$. On each sub-triangle $T_i, i = 1, \ldots, 4$,

$$\chi_l^n = \pm \frac{\frac{1}{|T_i|}}{\sqrt{\sum_{k=1}^4 \frac{1}{|T_k|}}}, \tag{44}$$

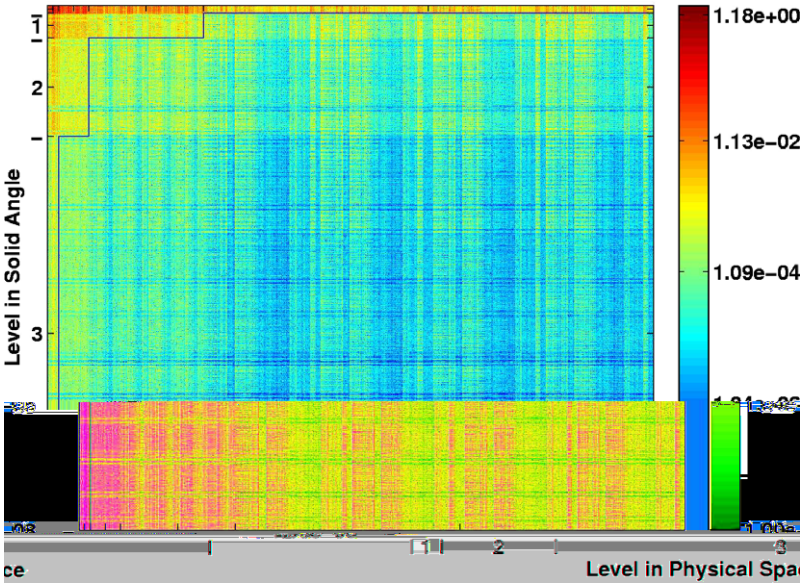

Fig. 17. Size of wavelet coefficients of the full tensor product solution for Example 1.

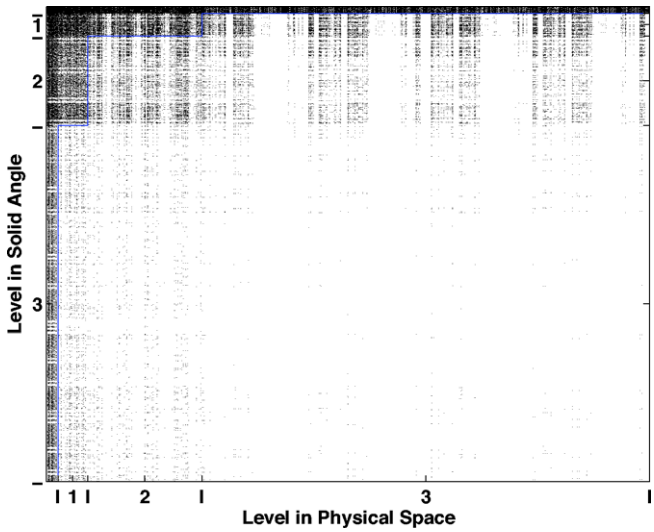

Fig. 18. $\widehat{\mathcal{N}}_L = 149120$ largest wavelet coefficients of the full tensor product solution for Example 1 at level $L = 3$.

where $|T_i|$ denotes the area of the spherical triangle T_i, see Fig. 9. This ensures L^2-orthogonality between the different levels.

4. Adaptivity

4.1. Theory

The sparse tensor method described in the previous section is a powerful tool to reduce the number of degrees of freedom in the discretization in the case of an intensity function that is smooth with respect to physical space as well as solid angle.

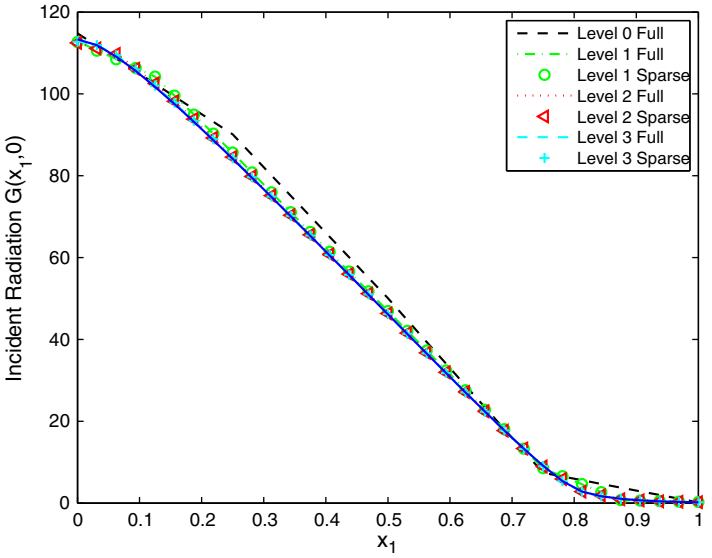

Fig. 19. Incident radiation for Example 1.

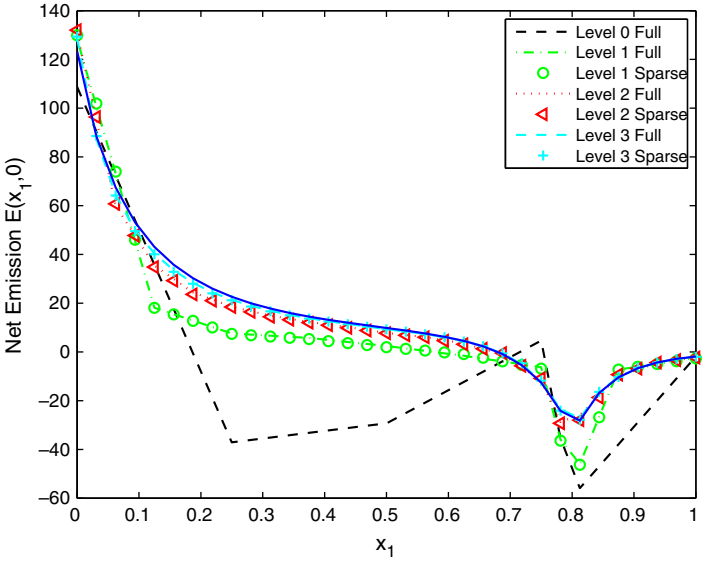

Fig. 20. Net emission for Example 1.

However, for a large number of problems, this *a priori* choice of degrees of freedom is not asymptotically optimal in the sense of a *best N-term approximation*: assuming that we have at hand a wavelet basis $\{\psi_\lambda(\mathbf{x},\mathbf{s})\}_{\lambda\in\nabla}$ indexed by a multi-index λ in the index set ∇ which spans the space $\mathcal{V}_0$, such as the product wavelet basis constructed in the previous section, the best N-term approximation u^*_Λ of the intensity $u(\mathbf{x},\mathbf{s})$ with respect to a norm $\|\circ\|_*$ on $D\times S^2$ is an approximation of u from the space $V_0^\Lambda\subset\mathcal{V}_0$ given by

$$V_0^\Lambda := \operatorname{span}\{\psi_\lambda : \lambda\in\Lambda\} \tag{45}$$

for some index set $\Lambda\subset\nabla$ with $\widehat{\mathcal{N}}_L = \#\Lambda = \dim V_0^\Lambda$ many indices chosen such that

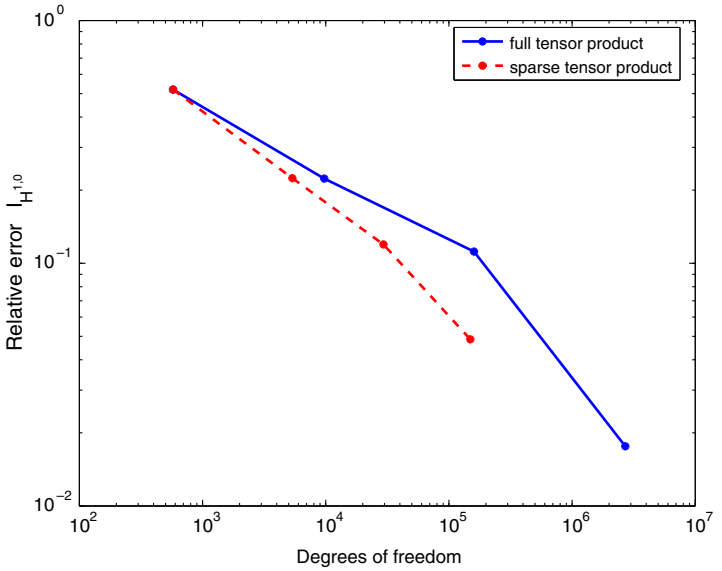

Fig. 21. Ex. 1: Relative intensity error in the $H^{1,0}(D\times S^2)$-norm.

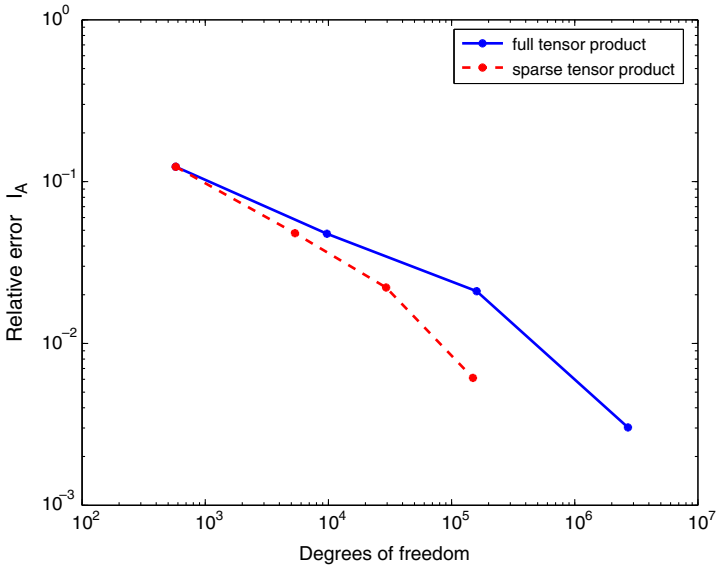

Fig. 22. Ex. 1: Relative intensity error in the $A(D\times S^2)$-norm.

$$\|u - u_\Lambda^*\|_* \leqslant \inf_{\substack{\Lambda \subset \nabla \\ \#\Lambda = N}} \min_{v_\Lambda \in V_0^\Lambda} \|u - v_\Lambda\|_*. \tag{46}$$

In other words, the best N-term approximation u_Λ^* is the best approximation to u if we are willing to spend N degrees of freedom in $D \times S^2$. It is obvious that the best N-term approximation u_Λ^* of u converges at least as fast as the sparse tensor product approximation $\hat{u}^L$ with $N = \dim \widehat{V}_0^L$ degrees freedom. Therefore, by Theorem 3.1, also the best N-term approximation u_Λ^Λ of the intensity will be able to break the curse of dimension (see [33] for theoretical background on this). What is more, however, is that the set of solutions u for which u_Λ^Λ attains the convergence rate of the sparse tensor product approximation $\hat{u}_L$ with $N = \dim \widehat{V}_0^L$ degrees of

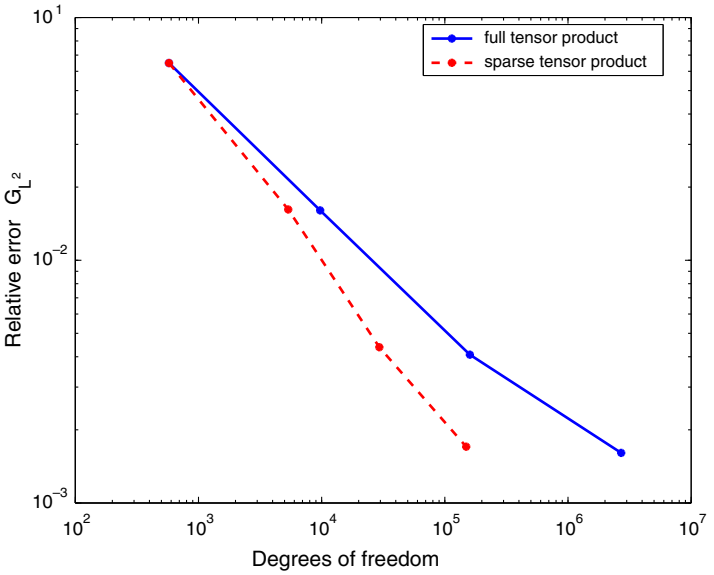

Fig. 23. Ex. 1: Relative error of the incident radiation in the $L^2(D)$-norm.

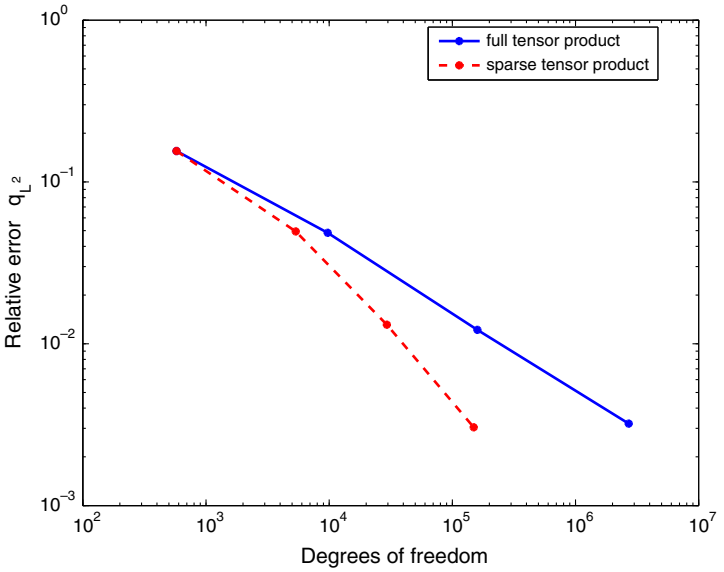

Fig. 24. Ex. 1: Relative error of the heat flux in the $L^2(D)$-norm.

freedom is much larger than $H^{p+1,p}(D \times S^2)$ – it is typically some Besov space (see [34] and the references there for details).

Consider a light beam, for example. On the one hand a very fine resolution in the direction of the ray as well as a fine spatial resolution, where the ray crosses the domain, is required. On the other hand, most degrees of freedom in physical space as well as in solid angle can be neglected as the intensity there is zero. In some real-world applications, the intensity will neither be smooth everywhere in the domain nor only consist of a light beam in vacuum. We therefore opt for an adaptive algorithm that selects the degrees of freedom that are relevant for the given problem. In our implementation, we exploit an important feature of best N-term approximation, namely that a *near best N-term approximation* in $H^{1,0}(D \times S^2)$ can be computed by *wavelet thresholding*, i.e. by simply keeping the N largest contributions to the solution, measured in the $\|\circ\|_{H^{1,0}(D\times S^2)}$-norm, of the wavelet expansion of u, see e.g. [35, Theorem 4.3.1].

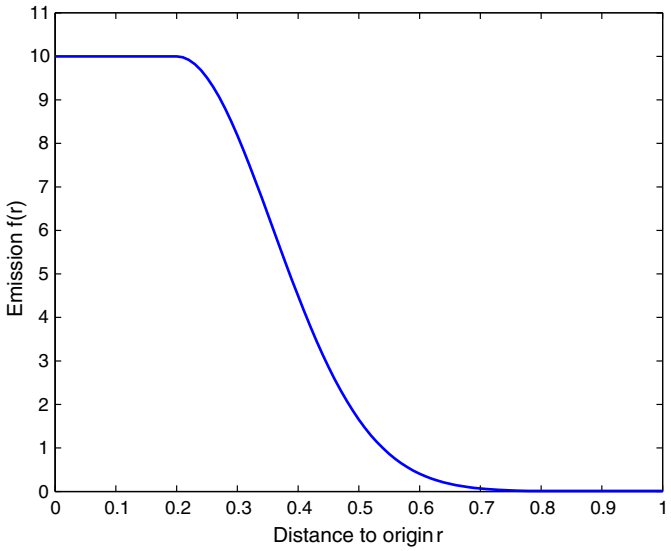

Fig. 25. Blackbody intensity for Example 2.

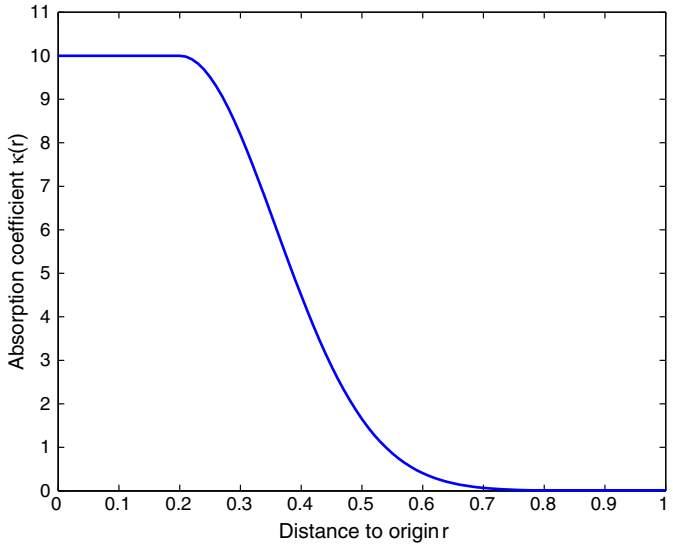

Fig. 26. Absorption coefficient for Example 2.

Adaptive sparse tensor product methods based on this idea have been applied to various problems, see e.g. [20–23]. A survey of adaptive wavelet techniques can be found, for example, in [24] and the references there.

Note that while the norm equivalence of the wavelet coefficients to the norm $\|\circ\|_{H^{1,0}(D\times S^2)}$ is rather straightforward to establish, the norm equivalence of the wavelet coefficients to the energy norm $\|\circ\|_A$ is to date open. Nevertheless, we propose to use thresholding of wavelet coefficients also in the present context and describe next the implementation of an adaptive discretization based on the *heuristic* ideas above.

4.2. Implementation

In order to describe the algorithm, we introduce a partial order (“parent–child relationship”) of the basis functions as follows:

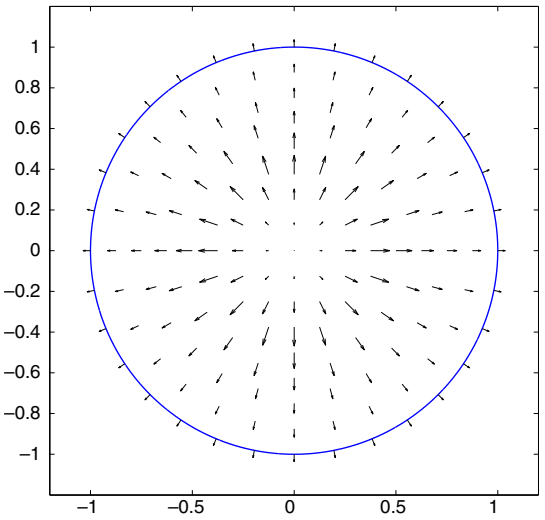

Fig. 27. Heat flux of Example 2 (reference solution).

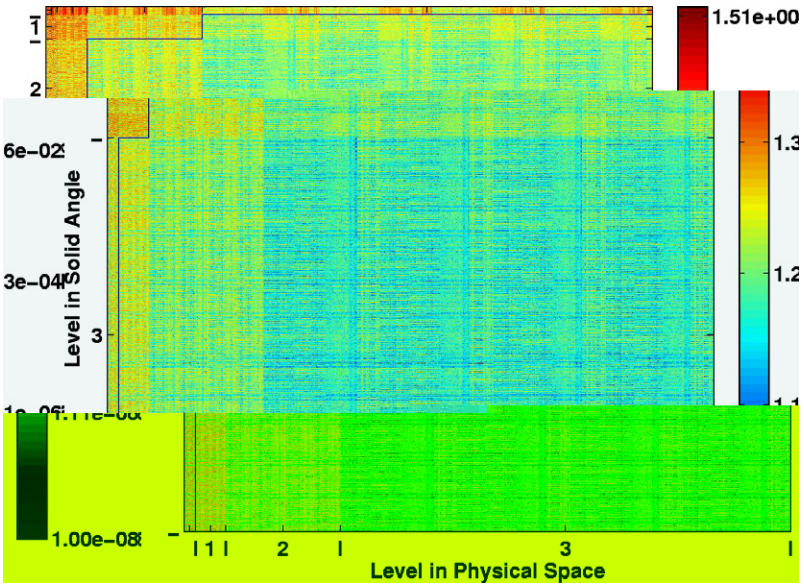

Fig. 28. Size of wavelet coefficients of the full tensor product solution for Example 2.

$$\psi(\mathbf{x}) <_D \psi'(\mathbf{x}) \text{ if } \psi(\mathbf{x}) \text{ is a child of } \psi'(\mathbf{x}), \tag{47}$$

$$\chi(\mathbf{s}) <_{S^2} \chi'(\mathbf{s}) \text{ if } \chi(\mathbf{s}) \text{ is a child of } \chi'(\mathbf{s}), \tag{48}$$

$\psi(\mathbf{x}), \psi(\mathbf{x})'$ being wavelets in physical space and $\chi(\mathbf{s}), \chi(\mathbf{s})'$ being wavelets on the sphere.

We restrict ourselves to the 2-dimensional case, as the numerical experiments are carried out for $n = 2$. Generalizations of all concepts below to $n = 3$ are straightforward.

In physical space, the children of a basis function corresponding to a given vertex correspond to the edge midpoints of the neighboring triangles of the vertex (see Fig. 10).

In the solid angle, the children are the twelve basis functions on the next level that overlap with the basis function to be refined. Fig. 11 shows three parent wavelet functions with their twelve children. Each parent has twelve children and each child three parents (except on the coarsest level).

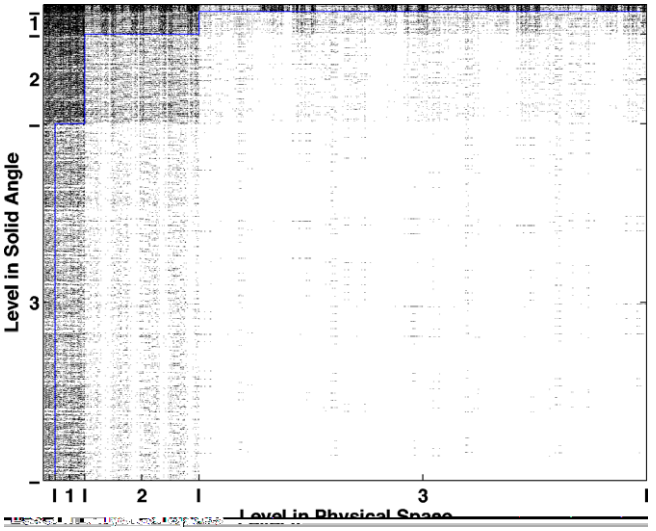

Fig. 29. $\widehat{\mathcal{N}}_L = 149120$ largest wavelet coefficients of the full tensor product solution for Example 2 at level $L = 3$.

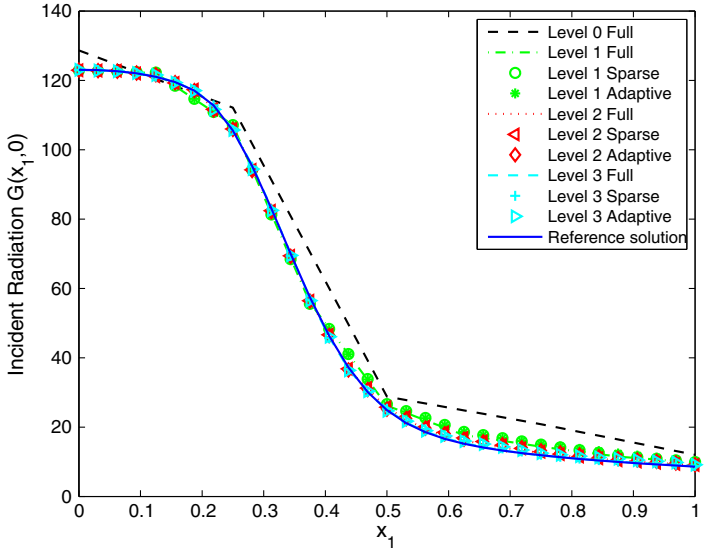

Fig. 30. Incident radiation for Example 2.

The adaptive algorithm proceeds level-wise (see also Fig. 12).

(1) We start with the computation of the intensity on the coarsest tensor product space.
(2) Select the degrees of freedom where the wavelet coefficients are above a given threshold.
(3) We impose the constraint that complete trees have to be maintained. We therefore ensure that all ancestors of the selected degrees of freedom are also included in the set, i.e. if the tensor product of basis functions $\psi(\mathbf{x})\chi(\mathbf{s})$ is in the active set, we add recursively all degrees of freedom $\psi'(\mathbf{x})\chi'(\mathbf{s})$ with $(\psi = \psi', \chi <_{S^2} \chi')$ or $(\psi <_D \psi', \chi = \chi')$.
(4) We then recompute the solution with the active set of degrees of freedom before.
(5) We threshold the coefficients again.

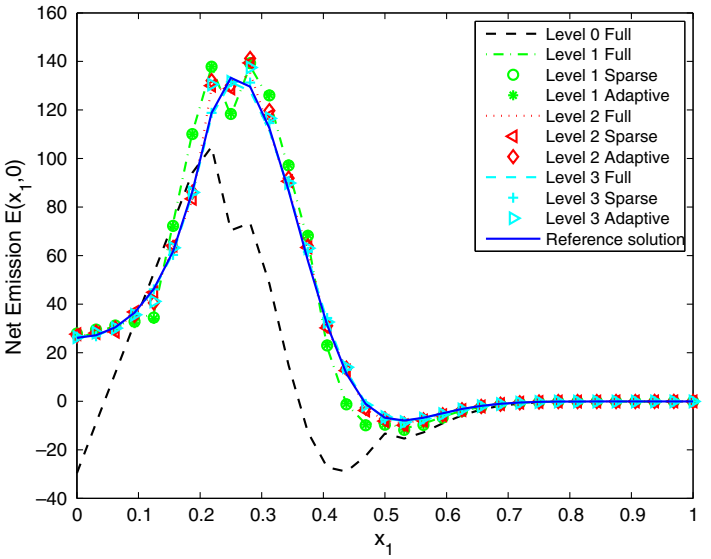

Fig. 31. Net emission for Example 2.

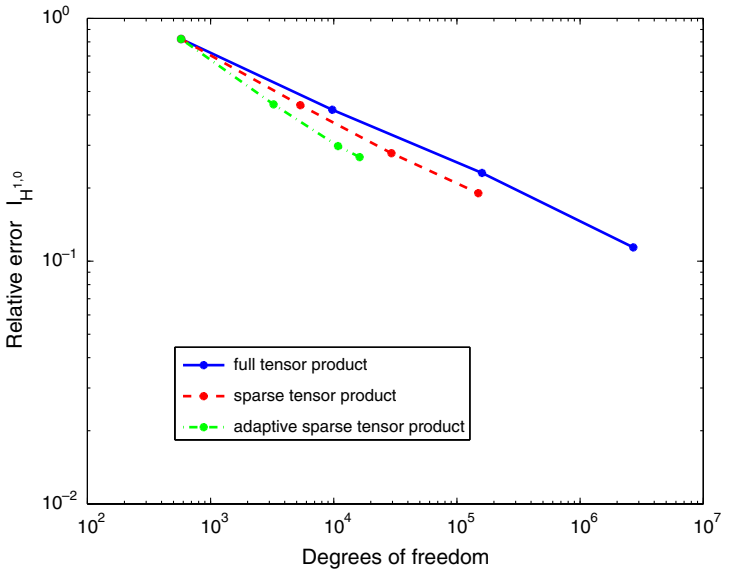

Fig. 32. Ex. 2: Relative intensity error in the $H^{1,0}(D \times S^2)$-norm.

(6) We now add the degrees of freedom that correspond to refinements in physical space or solid angle of the active degrees of freedom, i.e. in order to refine a tensor product degree of freedom $\psi(\mathbf{x})\chi(\mathbf{s})$, we add all degrees of freedom $\psi'(\mathbf{x})\chi'(\mathbf{s})$ with either $(\psi' = \psi, \chi' <_{S^2} \chi)$ or $(\psi' <_D \psi, \chi' = \chi)$.
(7) Again, we add all coarser degrees of freedom.
(8) After recomputing the solution, we repeat this procedure on higher levels starting from (2), until the finest level L is reached.

Although the solution has to be computed twice in each iteration step, the method is not too expensive, as a good initial guess for an iterative solver is available from the previous step.

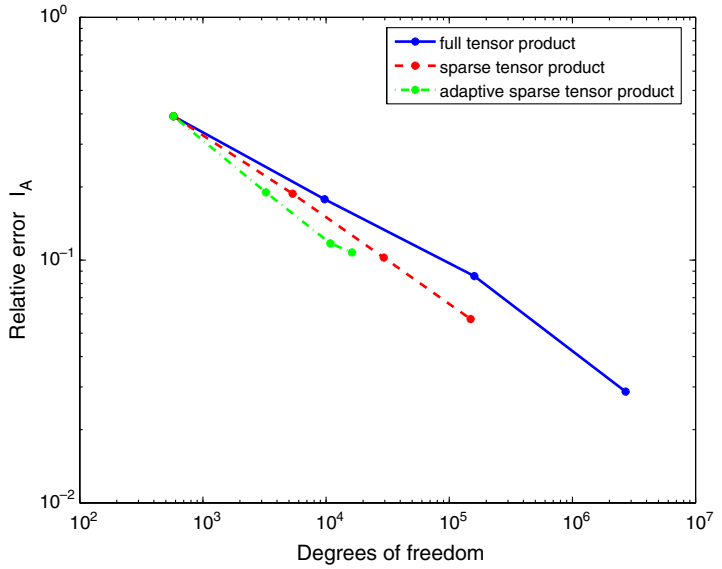

Fig. 33. Ex. 2: Relative intensity error in the $A(D \times S^2)$-norm.

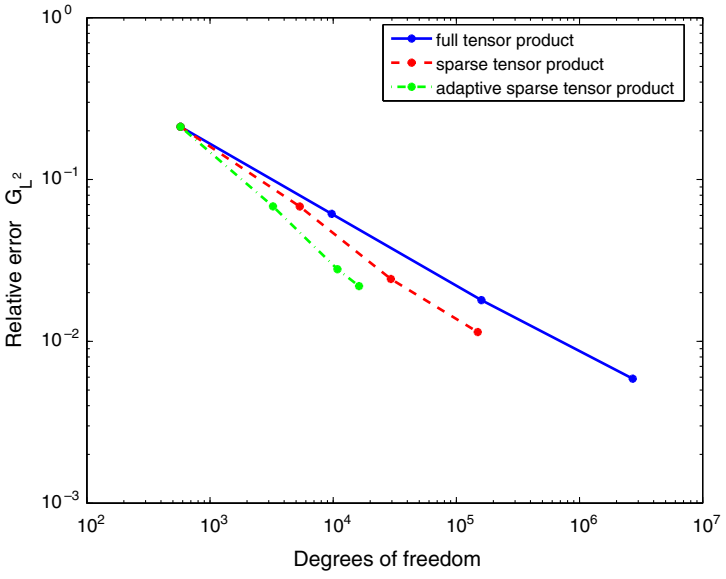

Fig. 34. Ex. 2: Relative error of the incident radiation in the $L^2(D)$-norm.

In Fig. 13, one can see which coarser degrees of freedom have to be active, if the set contains a certain vertex on the finest level. A consistent refinement of a degree of freedom is shown in Fig. 14.

As the set of active degrees of freedom is a subset of the degrees of freedom of a sparse tensor product space, the adaptive algorithm can be seen as an additional sparsification. However, in extreme examples as in the one of a narrow light beam, this algorithm allows to go up to much higher levels and the selected degrees of freedom will have little in common with the original sparse tensor product space.

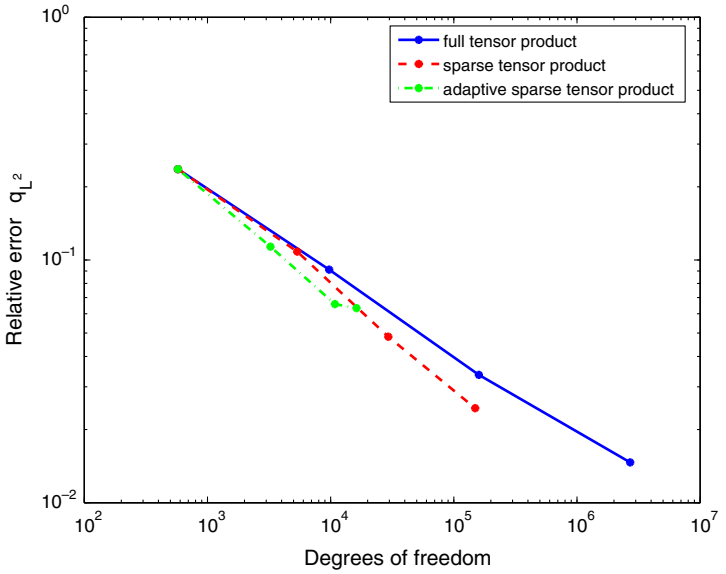

Fig. 35. Ex. 2: Relative error of the heat flux in the $L^2(D)$-norm.

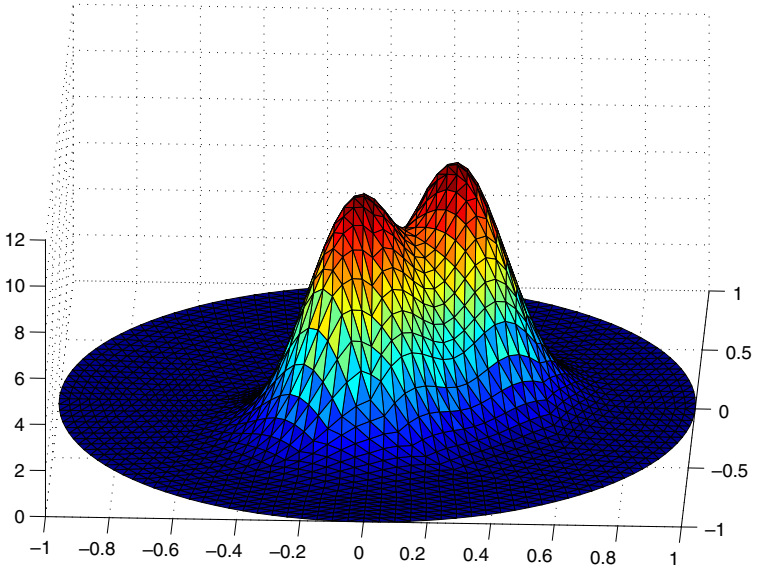

Fig. 36. Absorption coefficient for Example 3.

5. Numerical experiments

We test our method on the model circular 2D spatial domain $D = \{\mathbf{x} \in \mathbb{R}^2, |\mathbf{x}| \leqslant 1\}$, with different sets of emission and absorption data. The nested meshes we use for the definition of the multilevel hierarchies are shown in Fig. 1. In physical space, they consist of 41, 145, 545 and 2113 degrees of freedom, in solid angle of 20, 80, 320 and 1280 degrees of freedom. In order to compute the solution for a given set of degrees of freedom, we compute the stiffness matrices with respect to physical space and solid angle and use the conjugate gradient method with diagonal preconditioning to solve the linear system. We stop the iteration when the relative error in the energy norm is smaller than 10^{-6}. We impose the boundary conditions by projecting the solution onto the subspace described in Section 3 in each iteration step.

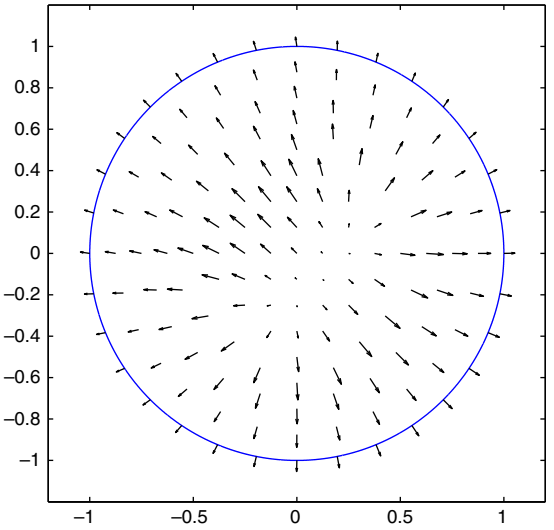

Fig. 37. Heat flux of Example 3 (reference solution).

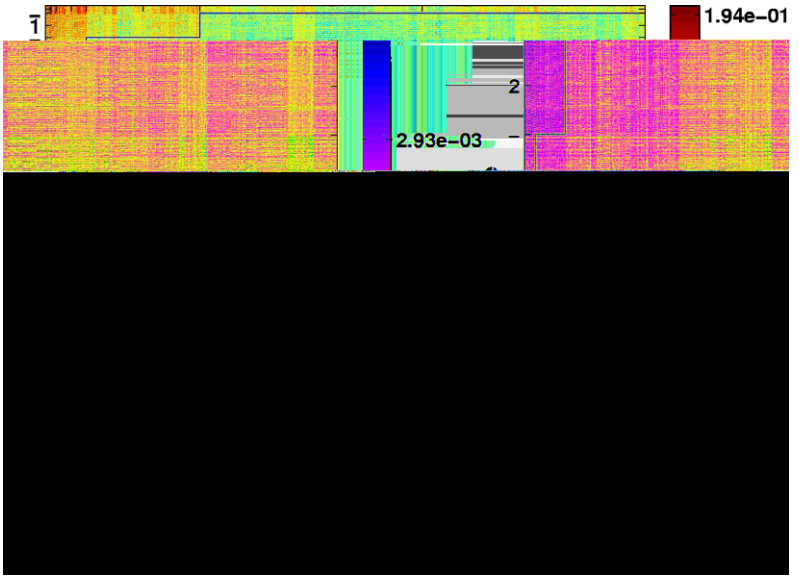

Fig. 38. Size of wavelet coefficients of the full tensor product solution for Example 3.

$$\text{Incident radiation } G(\mathbf{x}) = \int_{S^2} u(\mathbf{x}, \mathbf{s}) \mathrm{d}\mathbf{s}, \tag{49}$$

$$\text{and net emission } E(\mathbf{x}) = \kappa(\mathbf{x})(4\pi f(\mathbf{x}) - G(\mathbf{x})) \tag{50}$$

are important quantities in radiative transfer simulations. We plot profiles along the (positive) x_1-axis of those quantities for the different methods and levels.

In order to obtain error estimates, we compute reference solutions by line integration of the transport-reaction equation on level 3 in physical space and solid angle. We measure the error of the radiation intensity, the incident radiation and the heat flux in suitable norms.

We therefore define the following errors:

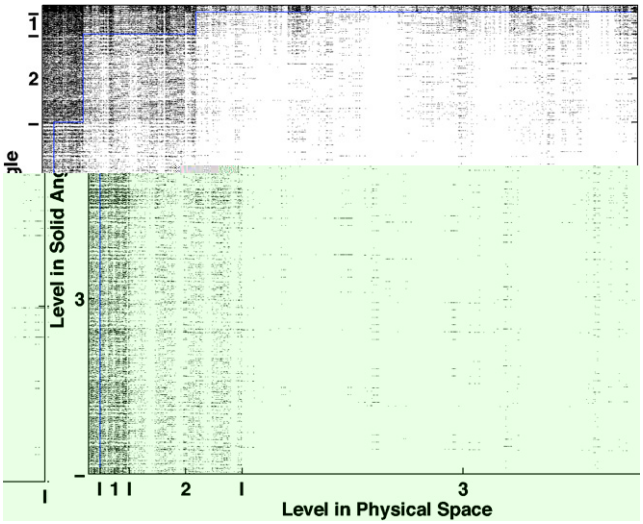

Fig. 39. $\widehat{\mathcal{N}}_L = 149,120$ largest wavelet coefficients of the full tensor product solution for Example 3 at level $L = 3$.

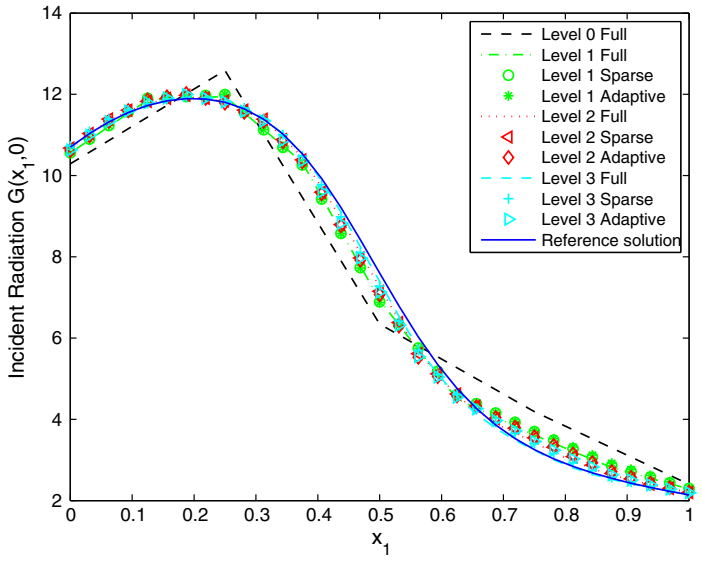

Fig. 40. Incident radiation for Example 3.

$$I_A := \frac{\|u - I_{\text{ref}}\|_{A(D\times S^2)}}{\|u\|_{A(D\times S^2)}}, \tag{51}$$

$$I_{H^{1,0}} := \frac{\|u - I_{\text{ref}}\|_{H^{1,0}(D\times S^2)}}{\|u\|_{H^{1,0}(D\times S^2)}}, \tag{52}$$

$$G_{L^2} := \frac{\|G - G_{\text{ref}}\|_{L^2(D)}}{\|G\|_{L^2(D)}} \tag{53}$$

and

$$q_{L^2} := \frac{\|\mathbf{q} - \mathbf{q}_{\text{ref}}\|_{L^2(D)}}{\|\mathbf{q}\|_{L^2(D)}}, \tag{54}$$

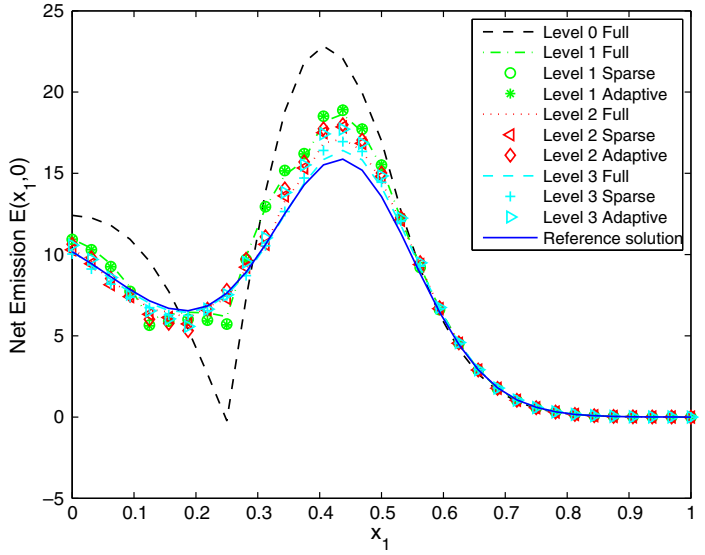

Fig. 41. Net emission for Example 3.

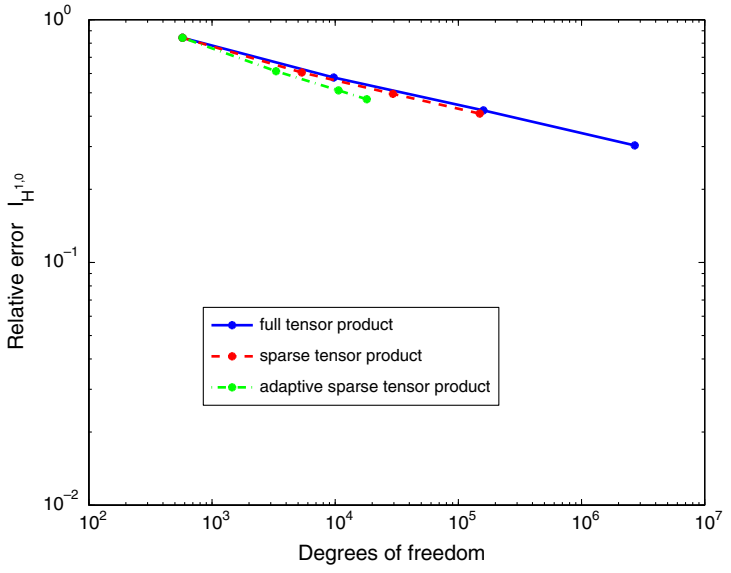

Fig. 42. Ex. 3: Relative intensity error in the $H^{1,0}(D \times S^2)$-norm.

where

$$\mathbf{q}(\mathbf{x}) := \int_{S^2} u(\mathbf{x}, \mathbf{s})\mathbf{s}\,d\mathbf{s} \tag{55}$$

is the heat flux and $I_{\text{ref}}, G_{\text{ref}}$ and $\mathbf{q}_{\text{ref}}$ are the reference solutions of the intensity, incident radiation and heat flux.

We visualize the efficiency of the (adaptive) sparse tensor product approximation by plotting the error in the intensity (resp. in the incident radiation or heat flux) versus the number of active degrees of freedom on levels 0 to 3.

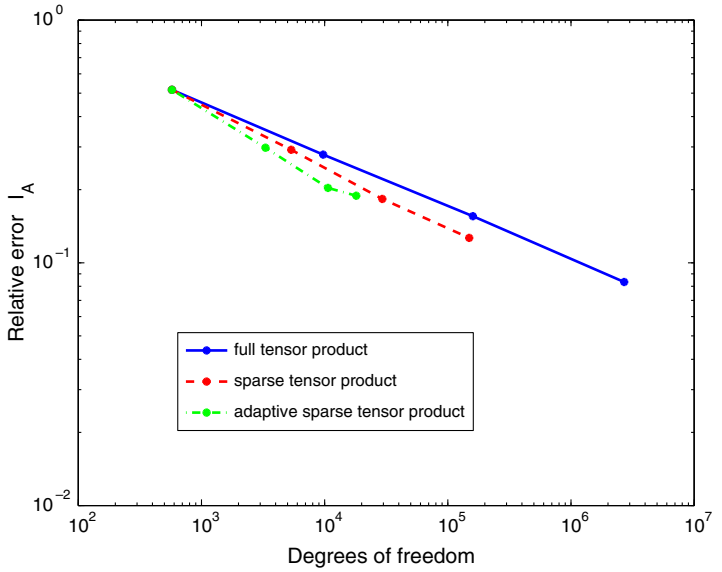

Fig. 43. Ex. 3: Relative intensity error in the $A(D \times S^2)$-norm.

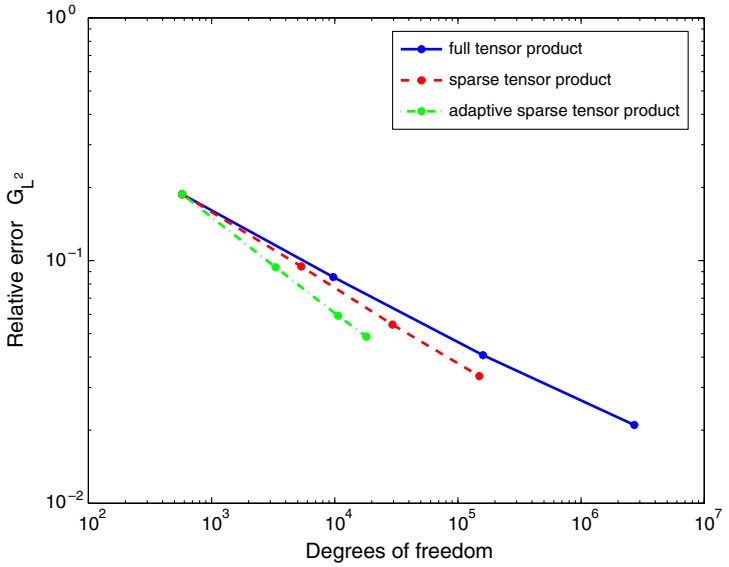

Fig. 44. Ex. 3: Relative error of the incident radiation in the $L^2(D)$-norm.

Example 1. Illustrates the performance of the sparse tensor product method, when the intensity is a smooth function with respect to physical space as well as solid angle due to a large absorption coefficient and zero emission near the boundary. The blackbody intensity is given in Fig. 15, while the absorption coefficient is 10 everywhere in the domain. The resulting heat flux of the reference solution is given in Fig. 16.

In Fig. 17 the absolute values of the coefficients with respect to the tensor product wavelet basis of the full tensor product solution are displayed, while Fig. 18 shows the $\widehat{\mathcal{N}}_L$ largest wavelet coefficients. Here, $\widehat{\mathcal{N}}_L$ is the number of degrees of freedom in the sparse tensor product space. In both figures, the sparse tensor product space corresponds to the area to the left and above the blue line. As expected, most of the $\widehat{\mathcal{N}}_L$ largest coefficients are contained in the sparse tensor product space.

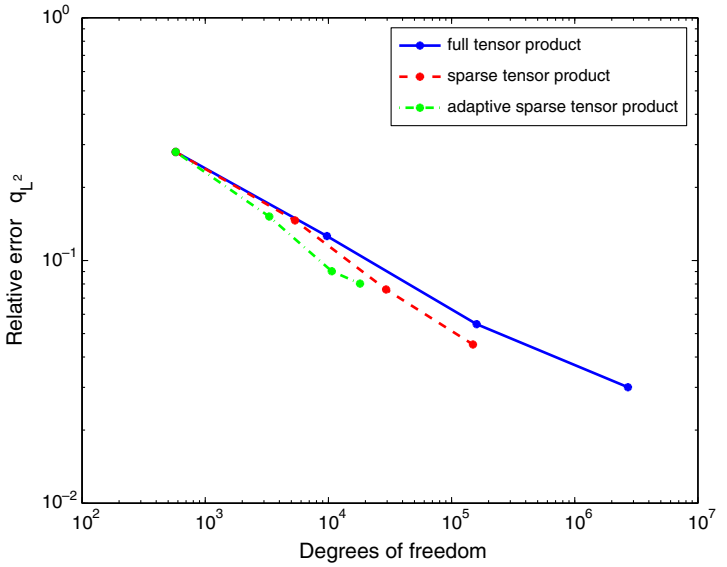

Fig. 45. Ex. 3: Relative error of the heat flux in the $L^2(D)$-norm.

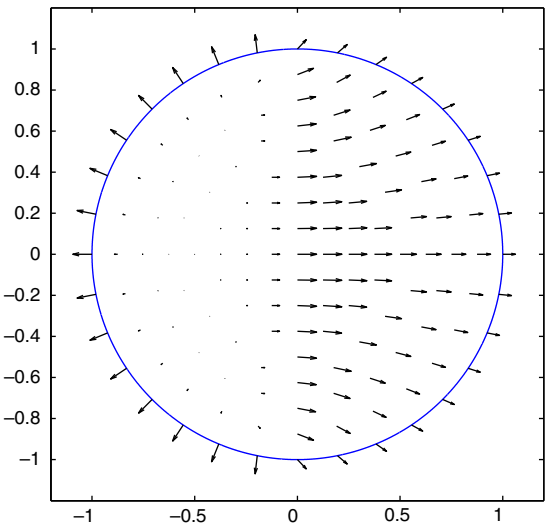

Fig. 46. Heat flux of Example 4 (reference solution).

The excellent approximation properties of the sparse tensor product space are also confirmed by the profiles of the incident radiation and net emission that are shown in Figs. 19 and 20 and the convergence results in Figs. 21–24: The sparse tensor product approximation is (almost) as accurate as the full tensor product approximation, while the number of degrees of freedom is reduced from $N \cdot M$ to $\mathcal{O}(N \log M + M \log N)$, where M is the number of degrees of freedom in physical space D and N the number of degrees of freedom in solid angle S^2.

Example 2 (*see Figs. 25 and 26*). There is a radiating zone in the center of the domain with an exponential decay of the blackbody intensity and the absorption coefficient to zero between $|\mathbf{x}| = 0.2$ and roughly $|\mathbf{x}| = 0.7$. In contrast to Example 1, the absorption coefficient varies between 0 and 10 and energy is emitted in the center and transported to the boundary of the domain, leading to the heat flux displayed in Fig. 27.

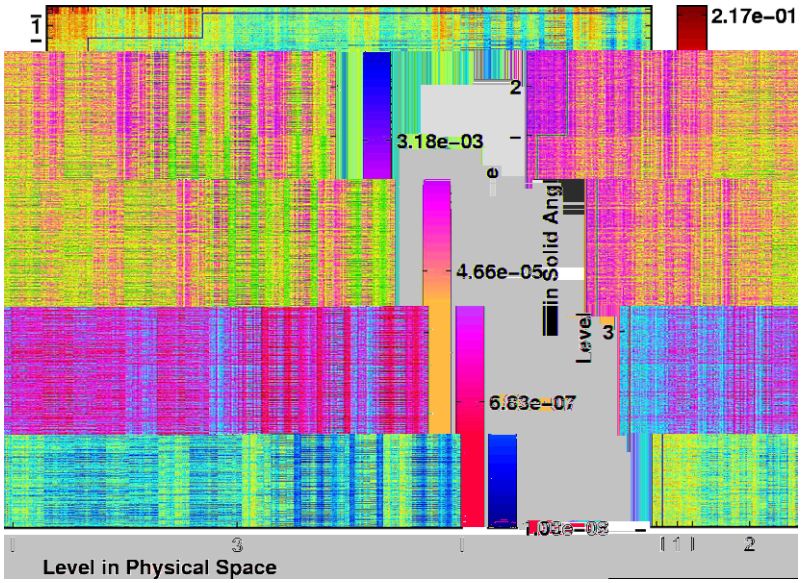

Fig. 47. Size of wavelet coefficients of the full tensor product solution for Example 4.

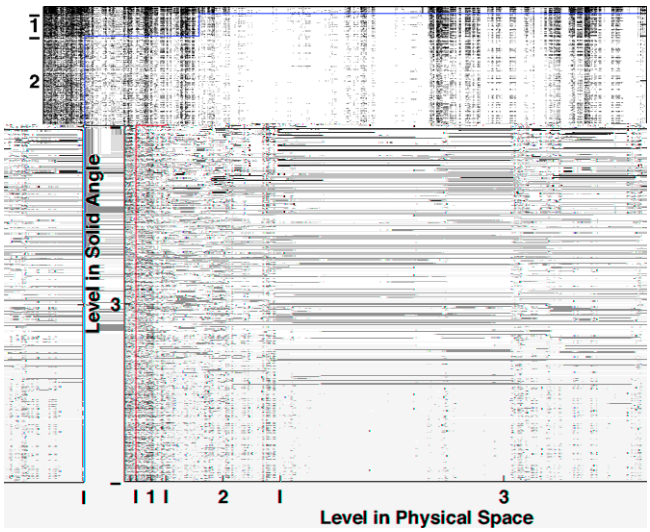

Fig. 48. $\widehat{\mathcal{N}}_L = 149120$ largest wavelet coefficients of the full tensor product solution for Example 4 at level $L = 3$.

Again, the structure of the wavelet coefficients indicates that the solution can be well approximated in the (adaptive) sparse tensor product space (Figs. 28–30).

The results of the incident radiation and net emission confirm that. With the exception of the adaptive sparse tensor approximation on level 3, the error increases only slightly when the (adaptive) sparse tensor product method is applied. Figs. 32–35 show that the sparse tensor product method is clearly superior to the full tensor product approximation. Introducing adaptivity, we further improve the ratio between the error and the number of degrees of freedom of the discretization. However, as in particular the results in the net emission and the heat flux in Figs. 31 and 35 indicate, the selection of the degrees of freedom in the adaptive algorithm could most likely be improved.

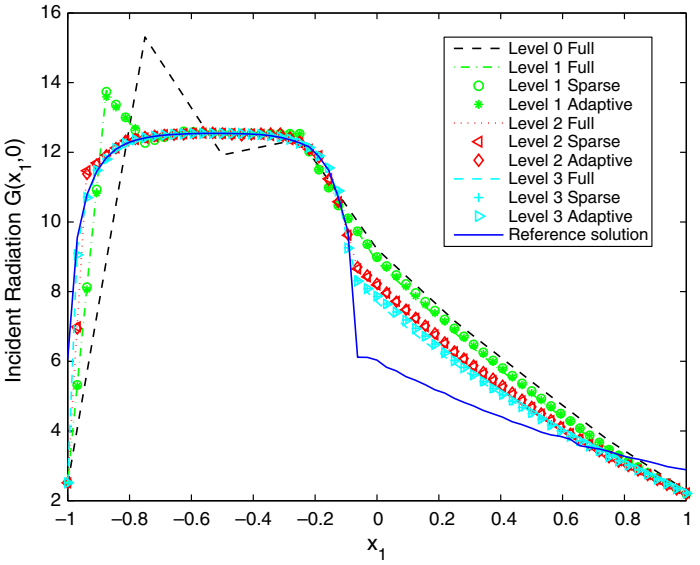

Fig. 49. Incident radiation for Example 4.

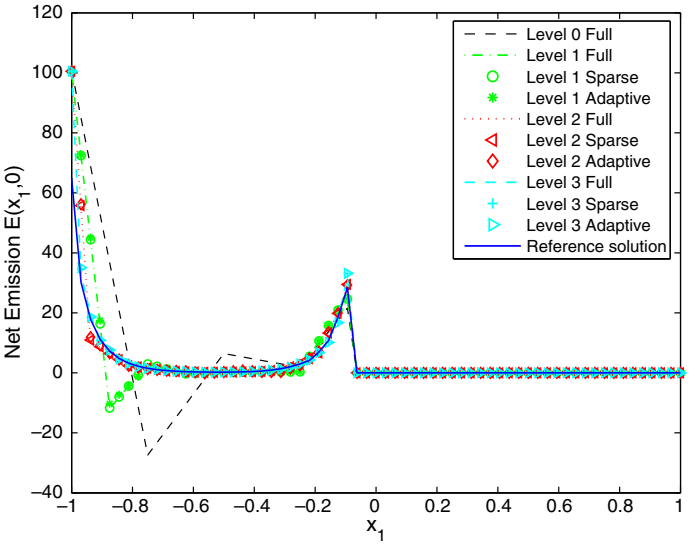

Fig. 50. Net emission for Example 4.

Example 3. The radiating source is not radially symmetric due to the absorption coefficient, which is given in Fig. 36. The blackbody intensity $f(\mathbf{x})$ is equal to 1 everywhere in the domain. This results in the non-symmetric heat flux given in Fig. 37. The results are displayed in Figs. 38–45 and confirm that radial symmetry is not required for the approximation properties.

Example 4. We test the performance of the methods for a problem, where the domain is divided into a strongly absorbing and a non-absorbing area. We choose the discontinuity not to match with the mesh in physical space and set the absorption coefficient as follows:

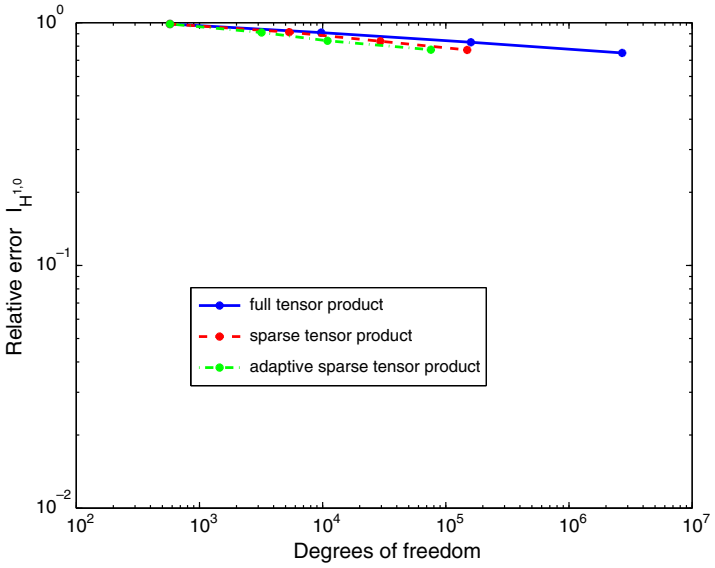

Fig. 51. Ex. 4: Relative intensity error in the $H^{1,0}(D \times S^2)$-norm.

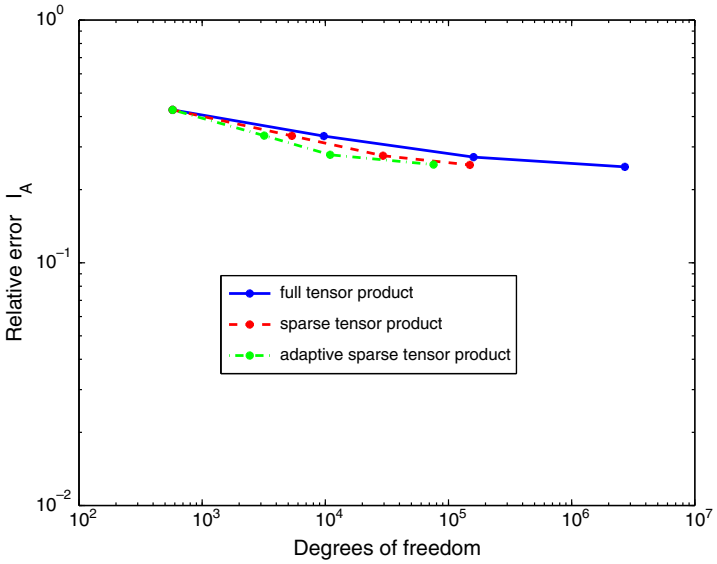

Fig. 52. Ex. 4: Relative intensity error in the $A(D \times S^2)$-norm.

$$\kappa(\mathbf{x}) = \begin{cases} 10, & x_1 < -\frac{1}{15}, \\ 0, & \text{otherwise.} \end{cases} \tag{56}$$

The blackbody intensity is equal to 1 everywhere in the domain. The resulting heat flux is shown in Fig. 46.

This problem is very hard to solve, as the intensity into directions with a negative s_1-component undergoes a sudden change around $x_1 = -\frac{1}{15}$ and for $x_1 < -\frac{1}{15}$ the large emission κf leads to steep gradients at the boundary. Therefore, we expect that a high resolution in physical space as well as solid angle is required in these regions. This is confirmed by Figs. 49–54 as all methods fail to produce an accurate solution, with the error mainly originating from an insufficient resolution at the boundary and the discontinuity.

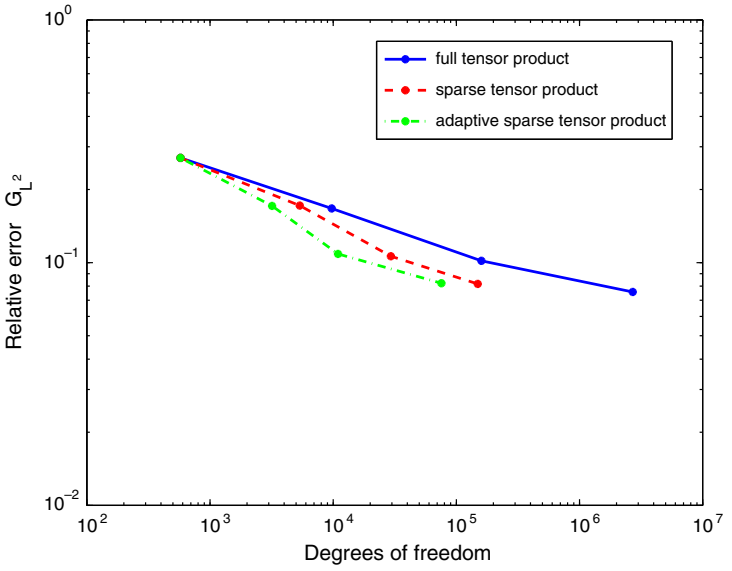

Fig. 53. Ex. 4: Relative error of the incident radiation in the $L^2(D)$-norm.

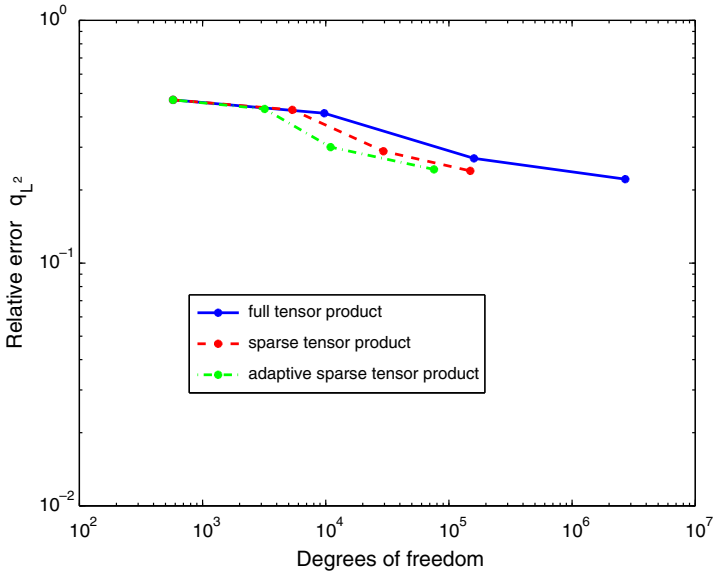

Fig. 54. Ex. 4: Relative error of the heat flux in the $L^2(D)$-norm.

However, reducing the number of degrees of freedom to the (adaptive) sparse tensor product space hardly influences the accuracy (Figs. 47–54). Therefore, the adaptive sparse tensor product method has the *potential* to include some degrees of freedom from higher levels to *improve the accuracy at affordable computational costs*. Due to limitations of the current implementation, we could not carry out such experiments.

Remark 3. The simple diagonally preconditioned CG-solver combined with projection methods to impose the boundary conditions is not satisfactory, as the number of iterations increases with the number of levels and the algorithm even does not converge to a relative error of 10^{-6} on level 3 for some problems. However, as we focus on the approximation properties of the different spaces, we content ourselves here with that straightforward method for the moment.

6. Conclusions

We have presented an efficient method to discretize the radiative transfer equation for arbitrary absorption coefficients without scattering. For solutions of sufficient regularity, the sparse tensor product approximation is (almost) as accurate as the full tensor product approximation, while the number of degrees of freedom is reduced from $N_L \cdot M_L$ to $\mathcal{O}(N_L \log M_L + M_L \log N_L)$, where N_L is the number of degrees of freedom in physical space D and M_L the number of degrees of freedom in solid angle S^2. Here, we used only the lowest degree finite elements, namely $p = 1$ in D and $q = 0$ in S^2. Even with this lowest order method, in numerical experiments our sparse tensor product method could achieve an accuracy comparable to that of the full tensor product method with only a fraction (1–20%) of the degrees of freedom. We obtained with a simple adaptive refinement strategy based on thresholding the solution's wavelet coefficients in various examples an additional reduction in the number of the degrees of freedom by a factor of 10 while still retaining the accuracy of the scheme. If the radiation intensity u is piecewise smooth, increasing the approximation order to $p > 1$ in D and to $q > 0$ in S^2 and applying wavelet coefficient thresholding allows a further reduction of degrees of freedom by selecting the most relevant contributions. The $p = 1, q = 0$ sparse tensor method realized here could be viewed as sparse tensor version of the discrete ordinate method S_N. Our construction is considerably more general, however: apart from raising p and q, other hierarchies of spaces in S^2, such as the span of spherical harmonics of order at most N, in place of the hierarchic sequence $V^L_{S^2}$ in (22) could be used.

Acknowledgment

Partial support by the Swiss Commission for Technology and Innovation (KTI) under Grant No. 7399.1 EPRPIW and partial support by ABB Corporate Research, Department of Electro-Technologies, Dättwil (Baden, CH) is acknowledged.

References

[1] M.F. Modest, Radiative Heat Transfer, second ed., Academic Press, Amsterdam, 2003.
[2] R. Siegel, J.R. Howell, Thermal Radiation Heat Transfer, 4th ed., Taylor and Francis, New York, 2002.
[3] R. Wehrse, W. Kalkofen, Advances in radiative transfer, Astron. Astrophys. Rev. 12 (1/2) (2006) 3–29.
[4] T.A. Manteuffel, K.J. Ressel, G. Starke, A boundary functional for the least-squares finite-element solution of neutron transport problems, SIAM J. Numer. Anal. 37 (2) (2000) 556–586.
[5] G. Kanschat, Parallel and adaptive Galerkin methods for radiative transfer problems, Ph.D. thesis, University of Heidelberg, 1996.
[6] R. Bellmann, Adaptive Control Processes: A Guided Tour, Princeton University Press, Princeton, 1961.
[7] B. Jiang, Non-oscillatory and non-diffusive solution of convection problems by the iteratively reweighted least squares finite element method, J. Comp. Phys. 105 (1993) 108–121.
[8] T.A. Manteuffel, K.J. Ressel, Least-squares finite-element solution of the neutron transport equation in diffusive regimes, SIAM J. Numer. Anal. 35 (2) (1998) 806–835.
[9] L. Bourhrara, New variational formulations for the neutron transport equation, Trans. Theory Stat. Phys. 33 (2) (2004) 93–124.
[10] C. Zenger, Sparse grids, in: Parallel Algorithms for Partial Differential Equations, Proceedings of the Sixth GAMM-Seminar, Kiel, Vieweg-Verlag, 1990, pp. 241–251.
[11] T. Gerstner, M. Griebel, Numerical integration using sparse grids, Numer. Algorithms 18 (3–4) (1998) 209–232.

[12] H. Yserentant, Sparse grid spaces for the numerical solution of the electronic Schrödinger equation, Numer. Math. 101 (2) (2005) 381–389.
[13] M. Griebel, Adaptive sparse grid multilevel methods for elliptic pdes based on finite differences, Computing 61 (2) (1998) 151–179.
[14] V.H. Hoang, C. Schwab, High-dimensional finite elements for elliptic problems with multiple scales, Multiscale Model. Simul. 3 (1) (2005) 168–194.
[15] T. Von Petersdorff, C. Schwab, Numerical solution of parabolic equations in high dimensions, M2AN Math. Model. Numer. Anal. 38 (1) (2004) 93–127.
[16] H.-J. Bungartz, A multigrid algorithm for higher order finite elements on sparse grids, Electron. Trans. Numer. Anal. 6 (1997) 63–77.
[17] M. Griebel, P. Oswald, T. Schiekofer, Sparse grids for boundary integral equations, Numer. Math. 83 (2) (1999) 279–312.
[18] R. DeVore, Nonlinear approximation, Acta Numer. 7 (1998) 51–150.
[19] R. Balder, C. Zenger, The solution of multidimensional real Helmholtz equations on sparse grids, SIAM J. Sci. Comput. 17 (3) (1996) 631–646.
[20] J. Nordmans, P.W. Hemker, Application of an adaptive sparse-grid technique to a model singular perturbation problem, Computing 65 (2000) 357–378.
[21] A. Cohen, W. Dahmen, R. DeVore, Adaptive wavelet methods for elliptic operator equations: convergence rates, Math. Comp. 70 (233) (2000) 27–75.
[22] A. Cohen, W. Dahmen, R. DeVore, Adaptive wavelet methods ii – beyond the elliptic case, Found. Comput. Math. 2 (3) (2002) 203–245.
[23] T. Gantumur, H. Harbrecht, R. Stevenson, An optimal adaptive wavelet method without coarsening of the iterands, Math. Comp. 76 (258) (2007) 615–629.
[24] A. Cohen, W. Dahmen, R. DeVore, Adaptive Wavelet Techniques in Numerical Simulation, John Wiley and Sons Ltd., New York, 2004.
[25] C. Johnson, Numerical Solution of Partial Differential Equations by the Finite Element Method, Cambridge University Press, Cambridge, UK, 1987.
[26] G. Widmer, R. Hiptmair, C. Schwab, Sparse adaptive finite elements for radiative transfer, Tech. Rep. 2007-01, SAM, ETH Zürich, Zürich, Switzerland, 2007.
[27] M. Cessenat, Théorèmes de trace pour des espaces de fonctions de la neutronique, C.R. Acad. Sci. Paris Sér. I Math. 299 (16) (1984) 831–834.
[28] M. Cessenat, Théorèmes de trace pour des espaces de fonctions de la neutronique, C.R. Acad. Sci. Paris Sér. I Math. 300 (3) (1985) 831–834.
[29] R. Dautray, J.L. Lions, Mathematical Analysis and Numerical Methods for Science and Technology, vol. 6, Springer-Verlag, Berlin, 1993.
[30] H.-J. Bungartz, M. Griebel, Sparse grids, Acta Numer. 13 (2004) 123–147.
[31] H. Nguyen, Finite element wavelets for solving partial differential equations, Ph.D. thesis, University of Utrecht, 2005.
[32] G. Schmidlin, Ch. Lage, Ch. Schwab, Rapid solution of first kind boundary integral equations in R^3, Eng. Anal. Bound. Elem. 27 (5) (2003) 469–490.
[33] P.-A. Nitsche, Best *N*-term approximation spaces for tensor product wavelet bases, Constr. Approx. 24 (1) (2006) 49–70.
[34] Ch. Schwab, R. Stevenson, Adaptive wavelet algorithms for elliptic pdes on product domains, Math. Comp. 77 (2008) 71–92.
[35] A. Cohen, Numerical Analysis of Wavelet Methods, Elsevier Amsterdam, Amsterdam, 2003.