Available online at www.sciencedirect.com
ScienceDirect
Journal of Computational Physics 223 (2007) 731–758
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

Solving Maxwell's equations using the ultra weak variational formulation

T. Huttunen [a,*], M. Malinen [a], P. Monk [b]

[a] Department of Physics, University of Kuopio, P.O. Box 1627, 70211 Kuopio, Finland
[b] Department of Mathematical Sciences, University of Delaware, Newark, DE 19711, USA

Received 13 January 2006; received in revised form 3 October 2006; accepted 3 October 2006
Available online 29 November 2006

Abstract

We investigate the ultra weak variational formulation for simulating time-harmonic Maxwell problems. This study has two main goals. First, we introduce a novel derivation of the UWVF method which shows that the UWVF is an unusual version of the standard upwind discontinuous Galerkin (DG) method with a special choice of basis functions. Second, we discuss the practical implementation of an electromagnetic UWVF solver. In particular, we propose a method to avoid the conditioning problems that are known to hamper the use of the UWVF for problems in general geometries and inhomogeneous media. In addition, we show how to implement the PML in the UWVF to accurately approximate physically unbounded problems and discuss the parallelization of the UWVF. Three-dimensional numerical simulations are used to examine the feasibility of the UWVF for simulating wave propagation in inhomogeneous media and scattering from complex structures.

Keywords: Plane wave; Variational method; Electromagnetism; Maxwell's equations; Numerical methods; Ultra weak variational formulation

1. Introduction

This paper is concerned with the development and validation the ultra weak variational formulation (UWVF) for Maxwell's equations proposed by Cessenat [4] and described in [5]. This method is based on a volume mesh and uses plane wave solutions of the Maxwell system on each element in the grid in the discretization procedure. Hence, the UWVF method resembles Trefftz type finite element techniques in which plane waves are also used to dicretize the Maxwell's equations [20,23] although the variational formulation is different.

We first give an alternative derivation of the UWVF that shows that the UWVF is a discontinuous Galerkin (DG) method but differs from standard methods in having a special choice of basis functions (coming from

* Corresponding author. Tel.: +358 17 162 050; fax: +358 17 162 585.
E-mail addresses: tomi.huttunen@uku.fi (T. Huttunen), mamaline@messi.uku.fi (M. Malinen), monk@math.udel.edu (P. Monk).

0021-9991/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2006.10.016

the plane waves mentioned above) and a special choice of degrees of freedom. Due to the connection with DG methods, it is no surprise that the UWVF shares several features of the finite element method. Convergence can be obtained by either refining the grid or increasing the number of basis functions per element (within certain limits which we shall describe). Due to the underlying DG scheme, having different numbers of basis functions per element is easy to implement and, it turns out, essential for obtaining a more robust UWVF.

In our implementation of the UWVF a tetrahedral mesh is used, and thus the UWVF can easily approximate complicated geometric structures in the same way as finite element methods. In addition the UWVF can handle variable coefficients in the Maxwell system (for example a variable conductivity), but unlike the finite element method, the coefficients must be piecewise constant.

In the original UWVF, the number of basis functions per element is fixed, and this can cause severe ill-conditioning of the resulting matrix problem if too many directions are used on a small element. We have investigated this for UWVF for the Helmholtz equation [16] where we proposed to use a variable number of directions per element and to control the global condition number of the UWVF matrix by controlling the local condition number of certain submatrices via a choice of the number of directions element by element through the grid (thus avoiding an expensive global condition number calculation). We shall show here that this scheme can also be used for the Maxwell UWVF. We also propose a direct method for estimating the number of plane waves on an element and show that this can improve performance.

The UWVF uses a different variational formulation to either traditional DG or finite element methods for Maxwell's equations. The resulting discrete matrix equation then has different properties (distribution of eigenvalues [4]) to the standard finite element matrices. We find that a simple biconjugate gradient scheme can be used to solve this system, at least for the matrices encountered in this study (and provided the conditioning problem mentioned previously is handled). Since the biconjugate gradient algorithm only requires matrix multiplication we can easily parallelize the UWVF on a message passing system parallel computer. We shall give details of this and report on the associated performance later in the paper.

The basic UWVF of [4] approximates unbounded domains via a low order absorbing boundary condition and this is a weakness. One goal of this paper is to show how to implement the perfectly matched layer (PML) of Bérenger [1] in the UWVF to give a better absorbing boundary.

In the numerical examples section we shall compare the UWVF and an edge element code from the COMSOL multi-physics commercial finite element package [10]. We find that we can obtain comparable accuracy for the model problem with sharply less memory use and with less computer time. This suggests that the UWVF may be advantageous in some situations. We have found this to be particularly evident as the wave number increases when COMSOL can run out of memory.

Of course, besides FEM, there are many other commonly encountered alternatives to the UWVF. Integral equation techniques handle the unbounded domain without trouble, and can also handle complex geometry (see for example [8]). Their main disadvantages are that penetrable media need to be implemented either by systems of integral equations on each interface between piecewise constant regions or via volume integral equations (alternatively boundary element and finite element methods can be coupled – see for example [13]). In addition to solve the resulting linear system (which also requires to evaluate singular integrals in a careful way) it is necessary to use a fast operator evaluation strategy like the fast multipole method (FMM) [12]. Elsewhere, one of us (Monk) and E. Darrigrand have started to test coupling boundary integral equations using the FMM together with the UWVF to provide an alternative mesh truncation procedure [11].

We also mention the popular finite integration method (FIT) [24,25]. This is essentially a volume based finite difference scheme so the representation of curved or complicated boundaries requires special care. The implementation of impedance boundary conditions is then even more complex. Of course the simple structured mesh of a finite difference method implies a more efficient code and potentially special solution methods when compared to similar finite element methods.

Returning to a discussion of the UWVF, we note that the theoretical convergence properties of the UWVF are not as well understood as for the other methods mentioned above (a dispersion analysis is not yet available). In addition, the solution is not computed everywhere but only on the faces of the tetrahedral mesh (termed the "skeleton" of the mesh) and requires a post-processing step to obtain the solution away from the skeleton. However the far field pattern can easily be calculated directly from the UWVF solution [4]. A goal of this paper is to show practical ways to avoid or control the problems we have mentioned above

(we shall not address theoretical aspects here beyond presenting a novel and unifying derivation of the method). In summary, we shall address several practical issues related to using the UWVF as follows:

- We shall discuss the choice of the number of basis functions per element to balance the competing needs of accuracy and conditioning.
- We shall show how to implement a PML in the UWVF and demonstrate that it provides enhanced performance compared to the basic low order boundary condition in the standard UWVF.
- We shall show how to post-process the computational results to approximate the solution away from the mesh skeleton.
- We shall show that the method can solve scattering problems such as scattering in a layered medium and scattering from a sphere where exact solutions are available, as well as scattering from the NASA almond where we compare to published results and results from a finite element method. We shall also show one way to implement a point source in the UWVF.

As part of these investigations we shall also exhibit the performance of the linear system solver (BiCGStab), and investigate the frequency dependence of the solution. We shall try to show that the UWVF is a viable and useful Maxwell solver.

We finish this introduction by describing in detail the Maxwell system and boundary conditions approximated by the basic UWVF of Cessenat and Després [4,5]. Suppose Ω is a bounded polyhedral domain in $\mathbb{R}^3$. We want to approximate the electric field $\boldsymbol{E}$ and magnetic field $\boldsymbol{H}$ that satisfy the following time-harmonic Maxwell system

$$-\mathrm{i}\omega\epsilon\boldsymbol{E} - \nabla\times\boldsymbol{H} = 0 \quad \text{in } \Omega, \tag{1}$$

$$-\mathrm{i}\omega\mu\boldsymbol{H} + \nabla\times\boldsymbol{E} = 0 \quad \text{in } \Omega. \tag{2}$$

Here ω is the temporal frequency of the field, and ϵ and μ are, respectively, the permittivity (complex valued in general) and permeability (real valued) of the material in Ω. In particular, for ϵ, the real part $\Re(\epsilon)$ is bounded and strictly positive. The imaginary part $\Im(\epsilon)$ is bounded and non-negative, and μ is real, bounded and strictly positive. The field is supposed to satisfy the following generalized impedance boundary condition

$$-\boldsymbol{E}\times\boldsymbol{n} + \sigma(\boldsymbol{H}\times\boldsymbol{n})\times\boldsymbol{n} = Q(\boldsymbol{E}\times\boldsymbol{n} + \sigma(\boldsymbol{H}\times\boldsymbol{n})\times\boldsymbol{n}) + \boldsymbol{g} \quad \text{on } \Gamma = \partial\Omega, \tag{3}$$

where Q is a real scalar function of position on the boundary with $|Q| \leqslant 1$, $\boldsymbol{n}$ is the unit outward normal on Γ and $\boldsymbol{g}$ is a tangential vector field giving the boundary condition. The positive parameter σ defined on the boundary Γ (beware this is not the conductivity!) helps specify the boundary condition. Often we choose $\sigma = \sqrt{\mu}/\sqrt{|\epsilon|}$, but more generally we allow σ to be any bounded strictly positive real function on Γ.

Note that the boundary condition (3) is rather general and for special choices of Q can be used to implement several standard boundary conditions. Choosing $Q = 1$ we get $\boldsymbol{E}\times\boldsymbol{n} = -2\boldsymbol{g}$ which is the standard perfect electrically conducting boundary condition. If $Q = 0$ we obtain the impedance boundary data. $-\boldsymbol{E}\times\boldsymbol{n} + \sigma(\boldsymbol{H}\times\boldsymbol{n})\times\boldsymbol{n} = \boldsymbol{g}$ and with an appropriate choice of $\sigma = \sqrt{\mu_0}/\sqrt{\epsilon_0}$ where ϵ_0 and μ_0 are the electromagnetic parameters of free space we have the lowest order absorbing boundary condition that can be used to truncate the computational domain in a scattering calculation. Finally, $Q = -1$ gives a magnetic wall condition useful for approximating surfaces with very high permeability (or to implement a symmetry boundary condition).

2. Derivation of the UWVF

In this section, we shall derive the UWVF for the basic Maxwell system (1)–(3). Our derivation, which differs from that of Cessenat and Després [4,5], highlights the connection between the UWVF and the classical flux splitting discontinuous Galerkin method for symmetric hyperbolic systems (see for example [18]). In fact we shall show that the UWVF can be viewed as a discontinuous Galerkin method with a special choice of test and trial space.

2.1. The continuous problem

Let $\tau_h = \{K\}$ denote a regular finite element mesh of elements K of maximum diameter h covering Ω. In principle, this mesh can be quite general allowing for mixing various elements (cubes, tetrahedra, prisms, etc.), but in our implementation we use a tetrahedral finite element mesh since it is easy to generate and fits general boundaries reasonably well. Hence we shall assume that each element K is a tetrahedron and hence has triangular faces (so simplifying some integrals that need to be performed during the calculation).

We now proceed along standard lines to derive a discontinuous Galerkin method for the Maxwell system recalling first the integration by parts identity that for any $\boldsymbol{a}$ and $\boldsymbol{b}$ sufficiently smooth on K

$$\int_K \nabla \times \boldsymbol{a} \cdot \bar{\boldsymbol{b}}\, \mathrm{d}V = \int_K \boldsymbol{a} \cdot \nabla \times \bar{\boldsymbol{b}}\, \mathrm{d}V + \int_{\partial K} \boldsymbol{n}^K \times \boldsymbol{a} \cdot \bar{\boldsymbol{b}}\, \mathrm{d}A,$$

where the over-line represents complex conjugate and $\boldsymbol{n}^K$ is the unit outward normal to the boundary ∂K of K. Now let $\boldsymbol{\xi}^K$ and $\boldsymbol{\psi}^K$ denote smooth test vector functions on an element in the mesh. Multiplying (1) and (2) by the complex conjugate of $\boldsymbol{\xi}^K$ and $\boldsymbol{\psi}^K$ and integrating over K using the above integration by parts identity to move the curl off the trial functions we obtain

$$\int_K (-\mathrm{i}\omega\epsilon \boldsymbol{E} \cdot \overline{\boldsymbol{\xi}^K} - \boldsymbol{H} \cdot \nabla \times \overline{\boldsymbol{\xi}^K})\, \mathrm{d}V = \int_{\partial K} \boldsymbol{n}^K \times \boldsymbol{H} \cdot \overline{\boldsymbol{\xi}^K}\, \mathrm{d}A,$$
$$\int_K \left(-\mathrm{i}\omega\mu \boldsymbol{H} \cdot \overline{\boldsymbol{\psi}^K} + \boldsymbol{E} \cdot \nabla \times \overline{\boldsymbol{\psi}^K}\right) \mathrm{d}V = -\int_{\partial K} \boldsymbol{n}^K \times \boldsymbol{E} \cdot \overline{\boldsymbol{\psi}^K}\, \mathrm{d}A.$$

Adding the two equations and reordering the left-hand side we obtain

$$\int_K (\boldsymbol{E} \cdot (\overline{\mathrm{i}\omega\bar{\epsilon}\boldsymbol{\xi}^K - \nabla \times \boldsymbol{\psi}^K}) + \boldsymbol{H} \cdot (\overline{\mathrm{i}\omega\mu\boldsymbol{\psi}^K - \nabla \times \boldsymbol{\xi}^K}))\, \mathrm{d}V = \int_{\partial K} (\boldsymbol{n}^K \times \boldsymbol{H} \cdot \overline{\boldsymbol{\xi}^K} - \boldsymbol{n}^K \times \boldsymbol{E} \cdot \overline{\boldsymbol{\psi}^K})\, \mathrm{d}A,$$

where we have used the fact that μ is assumed to be real valued. Usually, in the derivation of the discontinuous Galerkin method, we would now specify how to compute the "fluxes" or surface currents $\boldsymbol{n}^K \times \boldsymbol{E}$ and $\boldsymbol{n}^K \times \boldsymbol{H}$ from approximate discontinuous fields, but in this case we now make an important assumption that is the essential part of the UWVF. We assume that $\boldsymbol{\xi}^K$ and $\boldsymbol{\psi}^K$ satisfy the following adjoint Maxwell system on K:

$$\mathrm{i}\omega\bar{\epsilon}\boldsymbol{\xi}^K + \nabla \times \boldsymbol{\psi}^K = 0 \quad \text{in } K, \tag{4}$$
$$\mathrm{i}\omega\mu\boldsymbol{\psi}^K - \nabla \times \boldsymbol{\xi}^K = 0 \quad \text{in } K. \tag{5}$$

With this assumption the above identity for $(\boldsymbol{E}, \boldsymbol{H})$ on K reduces to

$$\int_{\partial K} (\boldsymbol{n}^K \times \boldsymbol{H} \cdot \overline{\boldsymbol{\xi}^K} - \boldsymbol{n}^K \times \boldsymbol{E} \cdot \overline{\boldsymbol{\psi}^K})\, \mathrm{d}A = 0. \tag{6}$$

We now apply the usual discontinuous Galerkin upwind splitting method to this identity. Let

$$\boldsymbol{u}^K = \begin{pmatrix} \boldsymbol{E}|_K \\ \boldsymbol{H}|_K \end{pmatrix} \quad \text{and} \quad \boldsymbol{\phi}^K = \begin{pmatrix} \boldsymbol{\xi}^K \\ \boldsymbol{\psi}^K \end{pmatrix}$$

then (6) becomes

$$\int_{\partial K} D^K \boldsymbol{u}^K \cdot \boldsymbol{\phi}^K\, \mathrm{d}A = 0, \tag{7}$$

where the matrix D^K is given by

$$D^K = \left(\begin{array}{c|c} 0 & (Z^K)^T \\ \hline Z^K & 0 \end{array}\right) \quad \text{and} \quad Z^K = \begin{pmatrix} 0 & n_3^K & -n_2^K \\ -n_3^K & 0 & n_1^K \\ n_2^K & -n_1^K & 0 \end{pmatrix}.$$

Note that $Z^K \boldsymbol{a} = -\boldsymbol{n}^K \times \boldsymbol{a}$ for any vector $\boldsymbol{a}$ and $Z^K = -(Z^K)^{\mathrm{T}}$.

Flux splitting now amounts to a suitable factoring of D^K into positive and negative semi-definite parts corresponding to left and right going waves. To obtain the general UWVF we use a slightly more general factorization than usual. Let $\sigma > 0$ be defined on the faces of the mesh. On the boundary faces this is the function σ appearing in the boundary condition (3). For the remaining faces in the mesh our choice is given in Section 2.2 Eq. (12). To define the splitting of D^K let

$$L^{K,+} = \frac{1}{\sqrt{2\sigma}}(-Z^K, \sigma(Z^K)^2) \quad \text{and} \quad L^{K,-} = \frac{1}{\sqrt{2\sigma}}(-Z^K, -\sigma(Z^K)^2)$$

and define $D^{K,\pm} = \pm(L^{K,\pm})^{\mathrm{T}}(L^{K,\pm})$. Note that $L^{K,\pm}$ are 3×6 matrices so $D^{K,\pm}$ are 6×6 as required. Since $(Z^K)^{\mathrm{T}}(Z^K)^2\boldsymbol{a} = (\boldsymbol{n}^K \times (-\boldsymbol{n}^K \times (-\boldsymbol{n}^K \times \boldsymbol{a}))) = -\boldsymbol{n}^K \times \boldsymbol{a}$ for any vector $\boldsymbol{a}$, we have the identity $(Z^K)^{\mathrm{T}}(Z^K)^2 = (Z^K)$. A simple calculation then shows that $D^K = D^{K,+} + D^{K,-}$ with $D^{K,+}$ positive semidefinite and $D^{K,-}$ negative semidefinite. An important property of the splitting that we shall use later is that if elements K and K' share a common face then on that face we have (usingthefactthat $\boldsymbol{n}^K = -\boldsymbol{n}^{K'}$)

$$L^{K,-} = \frac{1}{\sqrt{2\sigma}}(-Z^K, -\sigma(Z^K)^2) = -\frac{1}{\sqrt{2\sigma}}(-Z^{K'}, \sigma(Z^{K'})^2) = -L^{K',+}.$$

Using the splitting of D^K and the factorization of each term in the splitting we may rewrite (7) as

$$\int_{\partial K} ((L^{K,+}\boldsymbol{u}^K)\cdot(L^{K,+}\boldsymbol{\phi}^K) - (L^{K,-}\boldsymbol{u}^K)\cdot(L^{K,-}\boldsymbol{\phi}^K))\,\mathrm{d}A = 0. \tag{8}$$

The discontinuous Galerkin flux splitting approach is then to couple the solution on adjacent elements using the second term in the above equation. Thus if K' is an element sharing a face with K we have (using the continuity properties of the solutions of Maxwell's equations across an interface in the absence of surface charges)

$$L^{K,-}\boldsymbol{u}^K = -L^{K',+}\boldsymbol{u}^{K'} \tag{9}$$

on the common face. For faces on the boundary Γ we use the boundary condition (3) written in the convenient form

$$L^{K,-}\boldsymbol{u}^K = -QL^{K,+}\boldsymbol{u}^K - \hat{\boldsymbol{g}} \quad \text{on } \partial K \cap \Gamma,$$

where $\hat{\boldsymbol{g}} = (1/\sqrt{2\sigma})\boldsymbol{g}$. Eq. (8) then becomes

$$\int_{\partial K} (L^{K,+}\boldsymbol{u}^K)\cdot(L^{K,+}\boldsymbol{\phi}^K)\,\mathrm{d}V + \sum_{K',\partial K'\cap\partial K = f\neq\emptyset}\int_f (L^{K',+}\boldsymbol{u}^{K'})\cdot(L^{K,-}\boldsymbol{\phi}^K)\,\mathrm{d}A$$
$$+ \sum_{\partial K\cap\Gamma = f\neq\emptyset}\int_f (QL^{K,+}\boldsymbol{u}^K + \hat{\boldsymbol{g}})\cdot(L^{K,-}\boldsymbol{\phi}^K)\,\mathrm{d}A = 0. \tag{10}$$

This is essentially the UWVF of Cessenat and Després before discretization but to make the connection more obvious we define

$$\mathscr{X}^K = L^{K,+}\boldsymbol{u}^K|_{\partial K}, \mathscr{Y}^K = L^{K,+}\boldsymbol{\phi}^K|_{\partial K} \quad \text{and} \quad F_K(\mathscr{Y}^K) = -L^{K,-}\boldsymbol{\phi}^K|_{\partial K}.$$

Then (10) becomes the problem of finding $\mathscr{X}^K$ on the faces ∂K of each element K such that

$$\int_{\partial K} \mathscr{X}^K \cdot \overline{\mathscr{Y}^K}\,\mathrm{d}A - \sum_{K',\partial K'\cap\partial K = f\neq\emptyset}\int_f \mathscr{X}^{K'}\cdot\overline{F_K(\mathscr{Y}^K)}\,\mathrm{d}A - \sum_{\partial K\cap\Gamma = f\neq\emptyset}\int_f Q\mathscr{X}^K\cdot\overline{F_K(\mathscr{Y}^K)}\,\mathrm{d}A$$
$$= \int_{\partial K}\hat{\boldsymbol{g}}\cdot\overline{F_K(\mathscr{Y}^K)}\,\mathrm{d}A \tag{11}$$

for all appropriate $\mathscr{Y}^K$. This is the UWVF for Maxwell's equations before discretization (exactly that of Cessenat and Després using their choice of σ, see (12)). A natural question is what is the correct space for the trial and test functions. It turns out that if $L_t^2(\partial K)$ denotes the space of square integrable fields on ∂K that are tangent to ∂K then we can seek $\mathscr{X}^K \in L_t^2(\partial K)$ on each K such that (11) holds for all $\mathscr{Y}^K \in L_t^2(\partial K)$ and all elements K. The only difficulty is to see that in this case, given any $\mathscr{Y}^K \in L_t^2(\partial K)$, we can find a solution $\boldsymbol{\phi}^K$ (in

$H(\mathrm{curl};K)^2$) of (4), (5) that satisfies the generalized impedance boundary condition $L^{K,+}\boldsymbol{\phi}^K = \mathscr{Y}^K$ on ∂K and furthermore that $F_K(\mathscr{Y}^K) \in L_t^2(\partial K)$. This can be proved as usual by studying the impedance problem on a bounded domain (see [4,19]). In particular, it can be shown [4] that (11) has exactly one solution for any $\hat{\boldsymbol{g}} \in L_t^2(\Gamma)$ and that if $\boldsymbol{u}^K$ is computed element-wise using the Maxwell system (1), (2) together with the boundary condition $L^{K,+}\boldsymbol{u}^K = \mathscr{X}^K$, then the resulting piecewise defined function is exactly the solution $\boldsymbol{u}$ that satisfies (1)–(3).

We remark that the above derivation of the UWVF extends in a simple way to the equations of elasticity and to the Helmholtz equation written as a first order system (indeed to a general class of symmetric hyperbolic equations).

2.2. The discrete UWVF

It remains only to discretize the UWVF and here we follow exactly [4]. Hence we only give enough details to make this paper self contained and refer the reader to the original source for a detailed discussion of the discrete problem. For any element K let $\epsilon^K = \epsilon|_K$ (i.e. the value of ϵ on K which we have assumed constant) and $\mu^K = \mu|_K$.

In order to obtain exactly the UWVF of Cessenat and Després we choose the auxiliary function σ on a face f in the mesh as follows

$$\sigma|_f = \begin{cases} \text{the function appearing in (3)} & \text{if } f \text{ is a boundary face,} \\ \sqrt{\mu_f/\epsilon_f} & \text{otherwise,} \end{cases} \tag{12}$$

where if f is a face common to elements K and K', we choose

$$\epsilon_f = \sqrt{|\epsilon^K \epsilon^{K'}|} \quad \text{and} \quad \mu_f = \sqrt{\mu^K \mu^{K'}}$$

and we recall that μ is assumed to be real and positive.

The key to Cessenat and Després' discrete UWVF is to choose the basis functions for $\mathscr{X}^K$ and $\mathscr{Y}^K$ in a way that allows the action of the operator F_K to be computed easily. Various choices are possible, but for them all it is necessary to assume that ϵ and μ are piecewise constant on the mesh. We shall assume this from now on in the paper. Under this restriction we use plane wave solutions to approximate $\boldsymbol{\phi}^K$. In practice, following [4], a suitable family of plane waves are generated on K by choosing p_K directions $\boldsymbol{d}_\ell^K$, $|\boldsymbol{d}_\ell^K| = 1$, $1 \leqslant \ell \leqslant p_K$ (we use the optimal spherical codes from the website [22]), and then defining a unit real polarization vector $\boldsymbol{\psi}_{0,\ell}^K$ orthogonal to $\boldsymbol{d}_\ell$. From this we compute the complex polarizations (sometimes called the "right and left polarizations")

$$\boldsymbol{F}_\ell^K = \boldsymbol{\psi}_{0,\ell}^K + \mathrm{i}\boldsymbol{\psi}_{0,\ell}^K \times \boldsymbol{d}_\ell^K \quad \text{and} \quad \boldsymbol{G}_\ell^K = \boldsymbol{\psi}_{0,\ell}^K - \mathrm{i}\boldsymbol{\psi}_{0,\ell}^K \times \boldsymbol{d}_\ell^K, \quad 1 \leqslant \ell \leqslant p_K.$$

It is then easy to verify that the functions $(\boldsymbol{\xi}_\ell^{F,K}, \boldsymbol{\psi}_\ell^{F,K})$ given by

$$\boldsymbol{\xi}_\ell^{F,K} = \sqrt{\mu^K}\boldsymbol{F}_\ell^K \exp(\mathrm{i}k\boldsymbol{d}_\ell^K \cdot \boldsymbol{x}) \quad \text{and} \quad \boldsymbol{\psi}_\ell^{F,K} = \mathrm{i}\sqrt{\overline{\epsilon^K}}\boldsymbol{F}_\ell^K \exp(\mathrm{i}k\boldsymbol{d}_\ell^K \cdot \boldsymbol{x}),$$

where $k = \omega\sqrt{\mu^K\overline{\epsilon^K}}$ satisfy the adjoint Maxwell system 4, 5 on K. Similarly, the pair $(\boldsymbol{\xi}_\ell^{G,K}, \boldsymbol{\psi}_\ell^{G,K})$ given by

$$\boldsymbol{\xi}_\ell^{G,K} = \sqrt{\mu^K}\boldsymbol{G}_\ell^K \exp(\mathrm{i}k\boldsymbol{d}_\ell^K \cdot \boldsymbol{x}) \quad \text{and} \quad \boldsymbol{\psi}_\ell^{G,K} = -\mathrm{i}\sqrt{\overline{\epsilon^K}}\boldsymbol{G}_\ell^K \exp(\mathrm{i}k\boldsymbol{d}_\ell^K \cdot \boldsymbol{x})$$

are an independent set of solutions of the adjoint Maxwell system. These functions in turn generate plane waves

$$\boldsymbol{\phi}_\ell^{F,K} = \begin{pmatrix} \boldsymbol{\xi}_\ell^{F,K} \\ \boldsymbol{\psi}_\ell^{F,K} \end{pmatrix} \quad \text{and} \quad \boldsymbol{\phi}_\ell^{G,K} = \begin{pmatrix} \boldsymbol{\xi}_\ell^{G,K} \\ \boldsymbol{\psi}_\ell^{G,K} \end{pmatrix}$$

for $1 \leqslant \ell \leqslant p_K$. The reason for this somewhat complex choice of plane wave is that $\boldsymbol{\xi}_\ell^{G,K} \cdot \overline{\boldsymbol{\xi}_{\ell'}^{F,K}} = 0$ for any $1 \leqslant \ell, \ell' \leqslant p_K$ which provides some extra sparseness in certain matrices to be defined shortly. We can now define the approximation to $\mathscr{X}^K$ denoted $\mathscr{X}_h^K$ and given by

$$\mathscr{X}_h^K = \sum_{\ell=1}^{p_K} x_\ell^K L^{K,+} \boldsymbol{\phi}_\ell^{F,K} + x_{p_K+\ell}^K L^{K,+} \boldsymbol{\phi}_\ell^{G,K}.$$

To compute the unknown expansion coefficients $\{x_\ell^K\}_{\ell=1}^{p_K}$ for each element K in the mesh we follow the usual Galerkin approach of substituting $\mathscr{X}_h^K$ in place of $\mathscr{X}^K$ in (11) and choosing the test functions $\mathscr{Y}^K$ to be successively the basis function $\mathscr{Y}^K = L^{K,+}\boldsymbol{\phi}_\ell^{F,K}$, $1 \leqslant \ell \leqslant p_K$ and $\mathscr{Y}^K = L^{K,+}\boldsymbol{\phi}_\ell^{G,K}$, $1 \leqslant \ell \leqslant p_K$. We note that F_K is easy to compute for these basis functions since

$$F_K(L^{K,+}\phi_\ell^{F,K}) = -L^{K,-}\boldsymbol{\phi}_\ell^{F,K}$$

and similarly for the "G" basis functions. By enumerating the tetrahedra we can form a vector of unknowns $\vec{x}$ of length $M = \sum_{K\in\tau_h} 2p_K$ containing x_ℓ^K, $1 \leqslant \ell \leqslant 2p_K$ in the same order as the elements are numbered. This vector satisfies that matrix equation

$$(D - C)\vec{x} = \vec{f}, \tag{13}$$

where D is the block diagonal Hermitian $M \times M$ matrix resulting from the first term on the left-hand side of (11), C is the sparse $M \times M$ matrix resulting from the remaining two terms on the left-hand side and $\vec{f}$ is the load vector given by the right-hand side after choosing $\mathscr{Y}^K$ to be each of the discrete basis functions.

In order to compute these matrices, integrals need to be performed over the faces of the tetrahedra in the mesh. For example, in order to compute the matrix D we must evaluate integrals of the form

$$\int_{\partial K} (L^{K,+}\boldsymbol{\phi}_\ell^{F,K}) \cdot \overline{(L^{K,+}\boldsymbol{\phi}_{\ell'}^{F,K})}\, \mathrm{d}A$$

for $1 \leqslant \ell, \ell' \leqslant p_K$ and similar integrals involving the "G" basis functions (as well as integrals involving both "F" and "G" basis functions for other terms in (11)). Using the definition of the operator $L^{K,+}$ and the basis function $\boldsymbol{\phi}_\ell^{F,K}$ we see that

$$L^{K,+}\boldsymbol{\phi}_\ell^{F,K} = \frac{1}{\sqrt{2\sigma}}\left(-\sqrt{\mu^K} Z^K \boldsymbol{F}_\ell^K + \mathrm{i}\sigma\sqrt{\epsilon^K}(Z^K)^2 \boldsymbol{F}_\ell^K\right)\exp(\mathrm{i}k\boldsymbol{d}_\ell^K \cdot \boldsymbol{x})$$

with a similar expression for $L^{K,+}\boldsymbol{\phi}_{\ell'}^{F,K}$. On each planar face f of K the matrix Z^K is constant, and we also assume that σ is also constant on f thus

$$\int_{\partial K} (L^{K,+}\boldsymbol{\phi}_\ell^{F,K}) \cdot \overline{(L^{K,+}\boldsymbol{\phi}_\ell^{F,K})}\, \mathrm{d}A = \sum_{f\subset\partial K} \frac{1}{2\sigma|_f}\mathscr{Z}_\ell^K|_f \cdot \overline{\mathscr{Z}_{\ell'}^K}|_f \int_f \exp(\mathrm{i}k(\boldsymbol{d}_\ell^K - \boldsymbol{d}_{\ell'}^K)\cdot\boldsymbol{x})\, \mathrm{d}A,$$

where $\mathscr{Z}_\ell^K = -\sqrt{\mu^K} Z^K \boldsymbol{F}_\ell^K + \mathrm{i}\sigma\sqrt{\epsilon^K}(Z^K)^2\boldsymbol{F}_\ell^K$ and similarly $\mathscr{Z}_{\ell'}^K$. The vector part of this expression is easy to compute and relatively cheap since the dot product must be done once per face. The main difficulty in computing this term is the complex exponential integral. Fortunately, Cessenat [4] shows that the integral over the face can be computed in closed form as a difference of sinc functions. Although some care needs to be taken to avoid cancellation errors this integral is then easy to implement.

The remaining integrals in (11) can be computed in the same way. The result is that we can compute a block diagonal inner product matrix D corresponding to the first term on the left-hand side of (11). This matrix is Hermitian and positive definite (as long as the wave directions are distinct), but may become severely ill-conditioned if too many directions are used. We shall return to this point when we discuss the numerical implementation of the method in Section 3. We also construct a matrix C corresponding to the remaining terms on the left-hand side of (11). This is a more general matrix that couples the expansion coefficients on a tetrahedron K to the expansion coefficients on tetrahedra K' that share a face with K.

The right-hand side in the matrix equation (13) can also be computed in the same way provided the source vector $\vec{\boldsymbol{g}}$ involves only complex exponentials (such as is the case when solving scattering problems when a plane incident wave strikes a given scatterer and it is desired to compute the scattered field). Otherwise quadrature is needed to compute the right-hand side. This is rather expensive due to the oscillatory nature of the basis functions and our current implementation does not allow general boundary data for this reason.

In other applications of the UWVF in two dimensions we have found that when the boundary of the domain is smooth, the accuracy of the UWVF can be improved by using curved edges to the elements adjacent

to the boundary. This requires using quadrature to compute the integrals on such curved edges. So far we have not implemented this scheme in 3D but a future improvement to our UWVF code would be to improve the accuracy of the boundary representation in this way (at the cost of an increase in computer time to compute the matrices D and C). Currently, we use a refined grid near to curved boundaries and allow the elements to grow rapidly away from the boundary. We have also not tested the UWVF with non-conforming meshes.

3. Implementation

In the previous section we provided a novel derivation of the UWVF and summarized how, after choosing the number of plane wave directions on each element, we can compute a matrix system (13) that must be solved in order to obtain the coefficients of the surface unknown $\mathscr{X}_h^K$. In this section, we discuss several practical choices needed to improve the basic UWVF.

3.1. Choice of the plane wave directions

The choice of the number and directions of the plane waves on each element has a critical influence on the accuracy of the discrete UWVF. Choosing too many directions on a given element can result in a very poorly conditioned matrix D (introduced in the previous section). Since our inversion scheme requires to compute D^{-1} this can cause the iterative method to fail to converge. Thus the choice of the number of directions on a given element requires a balance between accuracy and conditioning.

Cessenat and Després suggest the use of a fixed number of directions on all the elements in the mesh. In this case we recall the basic convergence result due to Cessenat [4].

Theorem 3.1. *Suppose $\epsilon = \mu = 1$ and that p directions are used on each element in a quasi-uniform and regular mesh. Then there is a set of directions $\{\boldsymbol{d}_n\}_{n=1}^p$, where $p = (N+1)(N+3)$ such that if $|Q| < 1$ on $\partial\Omega$ then*

$$\|\mathscr{X} - \mathscr{X}_h\|_{L^2(\partial\Omega)} \leqslant Ch^{N+1/2}$$

as the mesh size h decreases, provided the solution is sufficiently smooth.

This theorem tells us that, at least for smooth solutions to the Maxwell system, we can increase the order of convergence of the method by increasing the number of directions p per element (similar to an p-finite element method where the degree of the polynomial basis governs the order of convergence for smooth solutions). In addition it was shown in [9] that for any fixed mesh, $\|\mathscr{X} - \mathscr{X}_h\|_V \to 0$ as $p \to \infty$ for any electromagnetic field in $H(\mathrm{curl};\Omega)$ (again assuming $\epsilon = \mu = 1$) provided the directions are chosen to be suitable quadrature points on the unit sphere. These results suggest that choosing p large may be advantageous for accuracy.

Using the UWVF for acoustic problems, we have found that a uniform choice of p across all elements may lead to poor accuracy on large elements, or poor conditioning on small elements [16]. In [16] we advocated the practical choice of setting a maximum allowable condition number and choosing the number of directions p_K for each element K to be the largest number such that the block of D associated with K has condition number at or below the cutoff. Thus we emphasized the practical need for convergence of the iterative scheme at the expense of indirect control over accuracy. This choice works well for a serial program, but when we come to implement a parallel code we would like to predict approximately the number of directions per element quickly to help load balancing. This choice can later be refined element by element after the parallel job has been allocated to the processors as long as the number of directions per element does not change greatly.

We have adopted a heuristic for choosing the number of directions per element based on the analysis of the error in using plane waves to approximate an independent plane wave not in the basis. This analysis does not include polarization effects and is therefore incomplete, but it does suggest why plane waves are essentially equivalent to Bessel functions for building the basis. The analysis also does not apply to evanescent waves.

Suppose we wish to approximate a plane wave $\exp(\mathrm{i}k\boldsymbol{x}\cdot\boldsymbol{d})$, $|\boldsymbol{d}| = 1$, using a sum of plane waves in the directions $\boldsymbol{d}_1, \ldots, \boldsymbol{d}_{p_K}$ on an element K. We choose the origin of the coordinate system to be at the center of the inscribed sphere (having radius ρ_K) and denote by h_K the maximum distance of points on ∂K from this inscribed center. In fact h_K is roughly the radius of the element if the elements are regular. The assumption

behind the analysis we shall now give is that $k\rho_K$ is large so that the inscribed sphere is many wavelengths across (and of course kh_K is still larger). For any point $\boldsymbol{x}$ on the surface of the element we have $\rho_K \leqslant |\boldsymbol{x}| \leqslant h_K$. Thus we can seek to approximate $\exp(\mathrm{i}k\boldsymbol{d}\cdot\boldsymbol{x})$ to error ϵ in the maximum norm for large $k|\boldsymbol{x}| \geqslant k\rho_K$. Carayol and Collino [3] show that if

$$L_K \approx kh_K + \frac{1}{2}\left(\frac{2}{3}\right)^{3/2} W^{2/3}\left(\frac{2kh_K}{3\epsilon^2}\right)(kh_K)^{1/3} - \frac{1}{2} + \text{terms vanishing in } kh_K, \tag{14}$$

where W is the Lambert W function defined on $[1/\mathrm{e},\infty)$ by $W(t)\exp(W(t)) = t$, then for large kh_K we can truncate the Jacobi–Anger expansion to error ϵ using $(L_K+1)^2$ terms in the sum. Using the fact that L_K is an increasing function of kh_K and $k|\boldsymbol{x}|$ is also large on the surface of the element we have

$$\exp(\mathrm{i}k\boldsymbol{d}\cdot\boldsymbol{x}) \approx 4\pi\sum_{\ell=0}^{L_K}\sum_{m=-\ell}^{\ell}\mathrm{i}^\ell j_\ell(k|\boldsymbol{x}|)Y_\ell^m(\hat{\boldsymbol{x}})\overline{Y_\ell^m}(\boldsymbol{d})$$

to error ϵ where $\hat{\boldsymbol{x}} = \boldsymbol{x}/|\boldsymbol{x}|$, j_ℓ is the ℓth spherical Bessel function and Y_ℓ^m is the spherical harmonic of index ℓ and momentum m. The same holds for each of the directions $\boldsymbol{d}_j$ and thus we can write

$$\exp(\mathrm{i}k\boldsymbol{x}\cdot\boldsymbol{d}) - \sum_{n=1}^{p_K} c_n \exp(\mathrm{i}k\boldsymbol{x}\cdot\boldsymbol{d}_n) \approx 4\pi\sum_{\ell=0}^{L_K}\sum_{m=-\ell}^{\ell}\mathrm{i}^\ell j_\ell(k|\boldsymbol{x}|)Y_\ell^m(\hat{\boldsymbol{x}})\overline{\left(Y_\ell^m(\boldsymbol{d}) - \sum_{n=1}^{p_K} c_n Y_\ell^m(\boldsymbol{d}_n)\right)}$$

to accuracy $(1+\sum_{n=1}^{p_K}|c_n|)\epsilon$ when $k|\boldsymbol{x}|$ is large and where the c_n, $n = 1,\ldots,p_K$ are suitable expansion coefficients. The right-hand side will vanish if we can choose the coefficients c_n such that

$$Y_\ell^m(\boldsymbol{d}) = \sum_{n=1}^{p_K} c_n Y_\ell^m(\boldsymbol{d}_n)$$

for $0 \leqslant \ell \leqslant L_K$ and $-\ell \leqslant m \leqslant \ell$. For a given choice of $\boldsymbol{d}$ and approximating directions this is a system of $(L_K+1)^2$ equations in p_K unknowns (the c_n's). We thus can choose $p_K = (L_K+1)^2$ and then choose the directions $\{\boldsymbol{d}_n\}_{n=1}^{p_K}$ so that they form a fundamental system for the given set of spherical harmonics (this is always possible since the spherical harmonics are linearly independent) [26]. This guarantees the invertibility of the matrix that would arise if we wished to compute the coefficients c_n. Hence the above equation is solvable. Various tables of fundamental sets of directions for $L_K = 1,\ldots,29$ are given in [21] (optimal with respect to different criteria). Note that the need for a fundamental set of directions is also suggested by the statement of the error estimate in Theorem 3.1 where only certain choices of directions give a good error estimate.

There remains the possibility that $(1+\sum_{n=1}^{p_K}|c_n|)$ might grow very rapidly with p_K and hence make the above estimates meaningless. We use the directions $\{\boldsymbol{d}_n\}_{n=1}^{p_K}$ (from $p_K = 4$ to $p_K = 121$ or $L_K = 1,\ldots,10$) for the optimal spherical codes from the website [22]. For these directions we have computed the expansion coefficients c_n, $n = 1,\ldots,p_K$ for $p_K = (L_K+1)^2$ and $L_K = 1,\ldots,10$ tested at 10,000 randomly chosen points on the unit sphere. The results are shown in Table 1 and show that, apart from $L_K = 4$, the factor generally does not grow rapidly for the values of L_K used here. We have no explanation for the large value for $L_K = 4$.

In conclusion, if we compute L_K via (14) and set $p_K = (L_K+1)^2$ we can approximate the trace of the plane wave on the boundary of the element K to accuracy approximately ϵ provided the element is non-degenerate and $k|\boldsymbol{x}|$ is large on the faces of the element (i.e. for example if the element contains a sphere that many wavelengths across) and the directions are chosen to be a fundamental set. In practice we use estimate (14) to motivate a heuristic for calculating L_K even for elements that are small compared to the wavelength of the

Table 1
The maximum of the factor $(1+\sum_{n=1}^{p_K}|c_n|)$ appearing in the error analysis as a function of L_K computed from 10,000 randomly chosen direction vectors $\boldsymbol{d}$ on the unit sphere

Order L_K	1	2	3	4	5	6	7	8	9	10
Dimension p_K	4	9	16	25	36	49	64	81	100	121
Factor $(1+\sum_{n=1}^{p_K}\lvert c_n\rvert)$	3	4.2	4.6	2×10^{10}	20	11	16	25	58	39

radiation, but this needs to be improved. Note that, even for large elements, this procedure only concerns the local approximation of plane waves and does not guarantee good dispersion error or good approximation error for more general fields.

If kh_K is large, due to the slower than logarithmic growth of W, we have that

$$L_K \leqslant Ckh_K + 1/2$$

for any fixed $C > 1$ and for kh_K large enough. This avoids the task of computing W. Motivated by this, we allow the user to choose coefficients A_j, $j = 0, 1, 2$, then compute

$$p_K = \text{round}(A_2(kh_K)^2 + A_1 kh_K + A_0), \tag{15}$$

where "round" corresponds to the nearest integer value of the given function. This gives the number of directions on K and the actual directions are drawn from the table of spherical codes as mentioned above. No restriction is placed on the choice of p_K other than the practical restriction that $3 \leqslant p_K \leqslant 130$ so adjacent elements may have widely different number of basis elements. We shall present more details of this approach and some numerical tests in Section 4.

3.2. Adding a PML layer

The first order absorbing boundary condition obtained by setting $Q = 0$ and $\sigma = \sqrt{\mu_0}/\sqrt{\epsilon_0}$ in (3) requires the absorbing boundary to be far from the scatterer to obtain reasonable accuracy. It is thus desirable to be able to use more efficient mesh termination methods. In [15] we showed how to implement the perfectly matched layer (PML) in the UWVF for acoustics. A similar approach can be taken for the Maxwell system and we now outline that approach here. The PML is applied to the Maxwell system in free space where $\epsilon = \epsilon_0$ and $\mu = \mu_0$. For simplicity we shall assume that the PML is applied when $|x_i| = x_{0,i} > 0$ so the standard Maxwell system governs the field in the box $|x_i| \leqslant x_{0,i}$, $i = 1, 2, 3$ which contains the scatterer. The PML will occupy the region outside the box and within the box $-x_{0,i} - L_i < x_i < x_{0,i} + L_i$, $i = 1, 2, 3$. Thus the PML has thickness L_i in the direction x_i along the ith coordinate axis. Our experience with the acoustic UWVF is that, unlike standard finite element methods, the UWVF works well with a constant absorption in the layer. As we shall see this allows the analytic calculation of certain integrals in the theory, and the constant PML does not cause unacceptable reflections at the PML boundary. A key assumption is that the planes $x_i = \pm x_{0,i}$, $i = 1, 2, 3$ at the interface between the PML and the ordinary Maxwell region of the computational domain are the union of faces of elements (i.e. the planes coincide with boundaries between elements – any element is either entirely in the PML or entirely in the Maxwell part of the computational domain).

We noted in the introduction that in general the UWVF cannot handle graded materials (i.e. having parameters that vary continuously in space). However, because the PML is based on a change of variables, we shall see that spatially varying PML parameters can be used in the UWVF.

In order to define the PML we use a complex stretching of the spatial variables [6,19] so that we define

$$\tilde{x}_i = \begin{cases} x + \frac{\mathrm{i}\sigma_0}{k} \frac{|x_i - x_{0,i}|}{L_i} & \text{for } |x_i| > x_{0,i}, \\ x_i & \text{for } |x_i| < x_{0,i} \end{cases} \tag{16}$$

for $i = 1, 2, 3$ where $k = \omega\sqrt{\epsilon_0\mu_0}$, L_i is the previously defined thickness of the absorbing layer in the ith direction and $\sigma_{i,0} > 0$ is the constant PML absorption parameter in the ith direction. Replacing x_i, formally, by $\tilde{x}_i$ in (1), (2) defines the non-physical electromagnetic field denoted $\tilde{\boldsymbol{E}}$ and $\tilde{\boldsymbol{H}}$ which satisfy the Maxwell system with respect to the "tilde" variables:

$$-\mathrm{i}\omega\epsilon_0\tilde{\boldsymbol{E}} - \tilde{\nabla} \times \tilde{\boldsymbol{H}} = 0,$$
$$-\mathrm{i}\omega\mu_0\tilde{\boldsymbol{H}} + \tilde{\nabla} \times \tilde{\boldsymbol{E}} = 0,$$

in the PML where $\tilde{\nabla}\times$ denotes the curl in "tilde" variables. We now use the definition of the "tilde" variables to change variables back to real coordinates x_i, $i = 1, 2, 3$. Define $d_i = 1 + \mathrm{i}\sigma_{i,0}/(kL_i)$ and let the matrices A and B be given by

$$A = \begin{pmatrix} 1/(d_2 d_3) & 0 & 0 \\ 0 & 1/(d_1 d_3) & 0 \\ 0 & 0 & 1/(d_1 d_2) \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix}$$

then after the change of variables the above equations become

$$-\mathrm{i}\omega\epsilon_0 \tilde{\boldsymbol{E}} - A\nabla \times B\tilde{\boldsymbol{H}} = 0,$$
$$-\mathrm{i}\omega\mu_0 \tilde{\boldsymbol{H}} + A\nabla \times B\tilde{\boldsymbol{E}} = 0.$$

Defining the computed fields in the PML denoted, in an abuse of notation, by $\boldsymbol{E} = B\tilde{\boldsymbol{E}}$ and $\boldsymbol{H} = B\tilde{\boldsymbol{H}}$ we obtain the following system for the non-physical electromagnetic field in the PML:

$$-\mathrm{i}\omega\epsilon_0 A^{-1}B^{-1}\boldsymbol{E} - \nabla \times \boldsymbol{H} = 0, \tag{17}$$
$$-\mathrm{i}\omega\mu_0 A^{-1}B^{-1}\boldsymbol{H} + \nabla \times \boldsymbol{E} = 0. \tag{18}$$

Thus defining the non-physical anisotropic electromagnetic parameters in the PML by $\epsilon_{\mathrm{B}} = \epsilon_0 A^{-1}B^{-1}$ and $\mu_{\mathrm{B}} = \mu_0 A^{-1}B^{-1}$ we see that in the PML the fields satisfy the Maxwell system (1), (2) with ϵ and μ replaced by ϵ_{B} and μ_{B}, respectively (the subscript B refers to J.P. Bérenger who first proposed the PML in 1996 [1]). Note that now μ_{B} is complex valued, and ϵ_{B} and μ_{B} are symmetric but not Hermitian matrices.

Within the PML region, the derivation of the UWVF in Section 2 still applies. The only change is to allow μ to be complex so that the adjoint problem becomes

$$-\mathrm{i}\omega\overline{\epsilon_{\mathrm{B}}}\boldsymbol{\xi}^K - \nabla \times \boldsymbol{\psi}^K = 0, \tag{19}$$
$$-\mathrm{i}\omega\overline{\mu_{\mathrm{B}}}\boldsymbol{\psi}^K + \nabla \times \boldsymbol{\xi}^K = 0, \tag{20}$$

in each element K in the PML. The same boundary condition can be used on the outer surface of the PML, and since the matrix B is continuous within the PML, the same matching condition and flux computation can be used across inter-element boundaries.

At the boundary between the PML and Maxwell (vacuum) regions the same matching of fields between adjacent elements can also be performed. This is because the boundaries between the PML and Maxwell regions are coordinate planes that are the union of faces in the mesh. For example suppose we have one tetrahedron K in the vacuum region, and another K' in the PML meeting at a common face on the surface $x_1 = x_{1,0}$. Across the interface (i.e. on f) the change of variables approach implies that $\boldsymbol{E}|_K \times \boldsymbol{n}^K = \tilde{\boldsymbol{E}}_{K'} \times \boldsymbol{n}^K$. But $d_2 = d_3 = 1$ in K' and only $d_1 \neq 1$. Since only tangential components of the field are continuous across f we have also $\boldsymbol{E}|_K \times \boldsymbol{n}^K = (B\tilde{\boldsymbol{E}}_{K'}) \times \boldsymbol{n}^K = \boldsymbol{E}_{K'} \times \boldsymbol{n}^K$. Similarly $\boldsymbol{H} \times \boldsymbol{n}^K$ is also continuous across f. Thus the flux matching Eq. (9) holds between K and K' since both sides only involve tangential components of the relevant fields. We can conclude that the UWVF equation (11) holds throughout the Maxwell region and the PML regions of the computational domain, provided the modified adjoint equations (19) and (20) are used in calculating F_K for elements in the PML.

The PML can be discretized, because, by construction, plane wave solutions of the adjoint system (19) and (20) can be derived from standard plane wave solutions in the "tilde" coordinates via the change of variables (16).

3.3. Point sources

In Section 4 we shall investigate fields originating from an electric dipole source at the point $\boldsymbol{x}_0 \in \Omega$. The dipole point source can be defined as the solution of the Maxwell system

$$-\mathrm{i}\omega\epsilon\boldsymbol{E} - \nabla \times \boldsymbol{H} = \boldsymbol{j} \quad \text{in } \Omega, \tag{21}$$
$$-\mathrm{i}\omega\mu\boldsymbol{H} + \nabla \times \boldsymbol{E} = 0 \quad \text{in } \Omega, \tag{22}$$

where $\boldsymbol{j} = I\boldsymbol{a}\delta_{x_0}$ and where I and $\boldsymbol{a}$, $|\boldsymbol{a}| = 1$ denote the amplitude and polarization of the dipole. In addition, δ_{x_0} denotes the Dirac delta function. Following the procedure of Section 2, it is easy to show that the right-hand

side of Eq. (21) provides a term to the right-hand side of the UWVF equation (11). For the point source in the element $\boldsymbol{x}_0 \in K$, the additional term is simply as

$$-2\int_K \boldsymbol{j}\cdot\overline{\boldsymbol{\xi}^K} = -2I\boldsymbol{a}\cdot\overline{\boldsymbol{\xi}^K(\boldsymbol{x}_0)}. \tag{23}$$

An alternative approach to modeling point sources is to use the explicitly known fundamental solution for Maxwell's equations as part of the basis and so take into account the singularity exactly (assuming the singularity is away from the boundary of the element containing it). This method might remove the need to refine the mesh near the singularity. We have not investigated this alternative.

3.4. Reconstruction within elements

Since the UWVF method provides an approximation for the function $\mathscr{X}^K$ (which is a function of $\boldsymbol{E}$ and $\boldsymbol{H}$) on each element face ∂K, but not a direct solution for the electric field $\boldsymbol{E}$ and the magnetic field $\boldsymbol{H}$, a post-processing step is needed to resolve $\boldsymbol{E}$ and $\boldsymbol{H}$ within elements. The approximation $\mathscr{X}_h^K$ for the function $\mathscr{X}^K$ is constructed by using plane waves $\boldsymbol{\phi}_\ell^{F,K}$ and $\boldsymbol{\phi}_\ell^{G,K}$ which in turn are solutions of the adjoint Maxwell system (4) and (5). It is clear that in the absence of absorption, the permittivity ϵ and the wave number k are real valued and the adjoint system (4) and (5) is the same as the physical Maxwell equations (1) and (2). Consequently, for real valued ϵ, the plane wave basis functions $\boldsymbol{\phi}_\ell^{F,K}$ and $\boldsymbol{\phi}_\ell^{G,K}$ are solutions of the local Maxwell equations in the corresponding element K. Therefore, it is easy to observe that for any element K in a non-absorbing medium, the approximation $\boldsymbol{u}_h^K$ for $\boldsymbol{u}^K$ is

$$\boldsymbol{u}_h^K = \sum_{\ell=1}^{p_K} x_\ell^K \boldsymbol{\phi}_\ell^{F,K} + x_{M_k+\ell}^K \boldsymbol{\phi}_\ell^{G,K}. \tag{24}$$

Resolving the field $\boldsymbol{u}^K$ for elements in an absorbing medium (i.e. ϵ is complex valued) or within the PML requires a different approach. The method used here is analogous with the UWVF post-processing technique introduced for the Helmholtz problem in [15]. Namely, we want to approximate the fields $\boldsymbol{E}$ and $\boldsymbol{H}$ in a plane wave basis which is a solution of the actual Maxwell system (1) and (2) (or (17) and (18) in the PML), rather than using the adjoint plane wave basis of the discrete UWVF. Therefore, we define a new set of plane waves basis functions as

$$\hat{\boldsymbol{\phi}}_\ell^{F,K} = \begin{pmatrix} \hat{\boldsymbol{\xi}}_\ell^{F,K} \\ \hat{\boldsymbol{\psi}}_\ell^{F,K} \end{pmatrix} \quad \text{and} \quad \hat{\boldsymbol{\phi}}_\ell^{G,K} = \begin{pmatrix} \hat{\boldsymbol{\xi}}_\ell^{G,K} \\ \hat{\boldsymbol{\psi}}_\ell^{G,K} \end{pmatrix},$$

where the pairs $(\hat{\boldsymbol{\xi}}_\ell^{F,K}, \hat{\boldsymbol{\psi}}_\ell^{F,K})$ and $(\hat{\boldsymbol{\xi}}_\ell^{G,K}, \hat{\boldsymbol{\psi}}_\ell^{G,K})$ are solutions of the physical Helmholtz system (1) and (2) ((17) and (18) in the PML) so that

$$\begin{aligned} \hat{\boldsymbol{\xi}}_\ell^{F,K} &= \sqrt{\mu^K}\boldsymbol{F}_\ell^K \exp(\mathrm{i}\hat{k}\boldsymbol{d}_\ell^K\cdot\boldsymbol{x}), \quad \hat{\boldsymbol{\psi}}_\ell^{F,K} = \mathrm{i}\sqrt{\epsilon^K}\boldsymbol{F}_\ell^K \exp(\mathrm{i}\hat{k}\boldsymbol{d}_\ell^K\cdot\boldsymbol{x}), \\ \hat{\boldsymbol{\xi}}_\ell^{G,K} &= \sqrt{\mu^K}\boldsymbol{G}_\ell^K \exp(\mathrm{i}\hat{k}\boldsymbol{d}_\ell^K\cdot\boldsymbol{x}), \quad \hat{\boldsymbol{\psi}}_\ell^{G,K} = -\mathrm{i}\sqrt{\epsilon^K}\boldsymbol{G}_\ell^K \exp(\mathrm{i}\hat{k}\boldsymbol{d}_\ell^K\cdot\boldsymbol{x}), \end{aligned}$$

where $\hat{k} = \omega\sqrt{\mu^K\epsilon^K}$. For the elements in the PML, the permittivity ϵ^K and permeability μ^K are taken as the modified parameters $\epsilon^K = \epsilon_{\mathrm{B}}$ and $\mu^K = \mu_{\mathrm{B}}$ which also leads to a matrix valued wave number $\hat{k}$.

The next step is to compute new coefficients y_ℓ^K, $1 \leqslant \ell \leqslant 2p_K$ for this new basis corresponding to the coefficients x_ℓ^K of the discrete UWVF problem. In particular, in each element in an absorbing medium (a physical interpretation of the PML is an anisotropic absorbing medium) we want to approximate the solution $\boldsymbol{u}^K$ as

$$\boldsymbol{u}_h^K = \sum_{\ell=1}^{p_K} y_\ell^K \hat{\boldsymbol{\phi}}_\ell^{F,K} + y_{M_k+\ell}^K \hat{\boldsymbol{\phi}}_\ell^{G,K}. \tag{25}$$

In the context of the UWVF, the coefficients y_ℓ^K for the element K are naturally obtained as a solution of the equation

$$\vec{y}_K = \hat{D}_K^{-1} D_K \vec{x}_K, \tag{26}$$

where the vectors $\vec{y}_K$ and $\vec{x}_K$ contain the coefficients y_ℓ^K and x_ℓ^K, $1 \leqslant \ell \leqslant 2p_K$ for the element K. The $2p_K \times 2p_K$ matrices $\hat{D}_K$ and D_K are assembled as the diagonal blocks of D in the discrete UWVF equation (13). However, due to the two types of plane wave bases involved in the post-processing, the matrix $\hat{D}_K$ is computed using integrals of the form

$$\int_{\partial K} (L^{K,+}\hat{\boldsymbol{\phi}}_\ell^{F,K}) \cdot \overline{(L^{K,+}\hat{\boldsymbol{\phi}}_{\ell'}^{F,K})}\, \mathrm{d}A,$$

i.e. using the physical (non-adjoint) plane waves only. Similar integrals are needed for the basis functions involving "G" basis functions as well as both "F" and "G" functions. The integrals for the matrix D_K include both adjoint and non-adjoint basis functions being of the form

$$\int_{\partial K} (L^{K,+}\boldsymbol{\phi}_\ell^{F,K}) \cdot \overline{(L^{K,+}\hat{\boldsymbol{\phi}}_{\ell'}^{F,K})}\, \mathrm{d}A.$$

In essence, the post-processing step for a non-absorbing medium is trivial since the solution for $\boldsymbol{E}$ and $\boldsymbol{H}$ can be extended directly within elements using the same plane basis functions and coefficients needed to approximate the UWVF function $\mathscr{X}$, see Eq. (24). For elements in an absorbing medium or in the PML, the extension of the solution in the elements can be computed similarly (see Eq. (25)) but by defining a new non-adjoint plane wave basis and by resolving coefficients for the new basis using a relatively simple UWVF-type fitting (26).

3.5. Iterative solution and parallelization

In his thesis Cessenat [4] suggests to solve the UWVF equation (13) by writing it as

$$(I - D^{-1}C)\vec{x} = D^{-1}\vec{f} \tag{27}$$

and applying a damped Richardson scheme. Note that D^{-1} is easy to compute since D is block diagonal, so the action of $D^{-1}C$ on any vector can be computed at the expense, essentially, of multiplying by C. We have found that the stabilized bi-conjugate gradient scheme (BiCGStab) is faster, and we use that method in all the examples presented here. An alternative, and perhaps attractive, approach is to solve $(D - C)\vec{x} = \vec{f}$ using BiCGStab with an approximation of D^{-1}, as a left preconditioner. This might help avoid the adverse effects of the high condition number of D. An interesting question is how to obtain a better left preconditioner – but we do not consider that problem here.

The UWVF has been parallelized using the same technique used to parallelize the acoustic UWVF in [14]. A domain decomposition strategy is used. The mesh is decomposed into collections of elements using METIS (a graph partitioning algorithm, see [17]). Because the number of basis functions per element differs widely, the predicted number of basis functions per element is used to weight the METIS graph nodes to improve load balancing. Note that elements are only connected through faces which simplifies the connectivity graph and decreases the number of elements in one sub-domain that are connected to another compared to a FEM solution.

Once the mesh is partitioned, each partition is sent to a processor (using MPI) and the processor performs the conditioning check and adjusts the number of unknowns per element (as described in Section 3.1). The local matrix D is also computed. This requires no communication. Then the matrix C is computed requiring communication to determine the directions on elements neighboring each partition (through faces). Finally, $\vec{f}$ is computed locally. Then the bi-conjugate gradient scheme is parallelized in the usual way using parallel matrix multiply.

4. Numerical results

To investigate the UWVF method for simulating problems involving the time-harmonic Maxwell problems, we study the method for several different model cases for which the exact solution is known. In the first case we approximate the field emitted by an electric dipole in free-space. Since the domain of the problem is physically unbounded, an absorbing boundary condition (ABC) is need on the exterior boundary of the computational

domain. We compare two ABCs of which the first is obtained by choosing $Q = 0$ in (3). This condition is referred to as ABC in the following sections. The second method to truncate the domain is the perfectly matched layer (PML) outlined in Section 3.2.

Second, we approximate the field emitted by the dipole in an inhomogeneous medium. In particular, when the dipole is located over a layered material for which an exact solution is available via the Sommerfeld integral [19]. The third model problem is the scattering of a plane wave from a perfectly conducting sphere. In this case, the principal interest is in the computation of the far-field pattern and in the efficiency of the parallelized UWVF code. However, prior to proceeding to specific model cases we shall outline the method for selecting a stable basis for the discrete UWVF. Finally, we provide some solutions using the NASA Almond and compare the results to results in the literature.

In all simulations, the accuracy is reported using a discrete L^2 norm so that the relative error for the electric field $\boldsymbol{E}$ is computed as

$$\text{Error} = \sqrt{\frac{\sum_{j=1}^{N} |\boldsymbol{E}_j - \boldsymbol{E}_j^{\text{ex}}|^2}{\sum_{j=1}^{N} |\boldsymbol{E}_j^{\text{ex}}|^2}},$$

where $\boldsymbol{E}_j$ and $\boldsymbol{E}_j^{\text{ex}}$ are the UWVF approximation and an exact solution of the problem at vertices referred by the subscript j.

4.1. The choice of basis

Since it is known from the previous UWVF studies that the method can suffer from instability if the plane wave basis is not carefully chosen [16], we begin this study by examining a method for selecting a basis on each element (i.e. a possibly different number of directions on each element) which leads to stable solution of the UWVF problem. The stability of the problem is in particular importance since we use the Bi-CGstab iteration for solving the UWVF matrix system.

As was noted in 2D UWVF simulations of the Helmholtz equation in [16], if the number of basis directions is too large, then the matrix $(I - D^{-1}C)$ may become ill-conditioned. It was also observed that by controlling the condition number of matrix blocks D_K element by element it is possible to have control over the conditioning of the overall UWVF matrix system [16]. More precisely, a large user specified tolerance is set and the computation of the matrix blocks D_K is begun element by element by using a relatively small number of basis functions per element. After the assembly of D_K, its condition number is computed (note that this involves only the small local matrix D_K). If the condition number is below the predetermined tolerance, then the local basis dimension (number of directions on K) is increased and the matrix D_K is recomputed. This procedure is repeated until the largest number of basis functions giving a condition number below the tolerance found. Our experience is that this approach ensures that the iterative scheme for solving (27) converges provided the tolerance is not chosen too large. Accuracy can be improved by choosing the tolerance larger within the overall constraint of requiring the iterative method converge.

While in 2D it is possible to make a relatively poor initial guess for the basis dimension, due to the wider range of possible basis dimension in 3D, a better approach for 3D problems is needed. The analysis of Section 3.1 showed that when kh_K is large, the approximation error for p_K plane waves is a relatively simple function of $k_K h_K$. On the other hand, numerical experiments for the 3D Helmholtz problem in [14] show that by constraining the condition number of D_K leads to almost an linear relation relationship between the scaled wave number $k_K h_K$ and the basis dimension p_K. Despite the fact that the relationship between the conditioning and the error is not yet properly understood, we focus on the controlling the conditioning since ill-conditioning leads to divergent Bi-CG iterations.

Let k_{re} denote the real part of the wave number k. In Fig. 1, the basis dimension p_K is plotted as a function of $k_{\text{re}} h_K^{\text{av}}$ when the maximum condition number of the matrix blocks D_K is limited by the tolerances 10^5, 10^7 and 10^9. The element size parameter h_K^{av} is defined as a mean distance of the element vertices from its centroid by

$$h_K^{\mathrm{av}} = \frac{1}{4}\sum_{j=1}^{4} |\boldsymbol{x}_{\mathrm{CM}}^K - \boldsymbol{x}_j^K|,$$

where $\boldsymbol{x}_{\mathrm{CM}}^K$ is the position of the centroid of the tetrahedron K and $\boldsymbol{x}_j^K$, $j = 1,\ldots,4$ are the coordinates of the vertices.

As in the Helmholtz case [14], we see an almost linear relationship between the basis dimension and the element size scaled wave number $k_{\mathrm{re}} h_K^{\mathrm{av}}$. Motivated by this observation and our accuracy considerations in Section 3.1, we suppose that the basis dimension can be approximated by using the quadratic polynomial (15) where the coefficients A_2, A_1 and A_0 are computed using a least-squares fit to data of Fig. 1 and are listed in Table 2.

We note that since the estimate (15) uses the real part of the wave number only, the absorption is not taken into account. However, numerical simulations show that the absorption has a strong effect on the condition number of D_K. More precisely, the condition number increases with the increasing absorption coefficient. Therefore, for a given condition number limit less basis functions are allowed in elements occupied by an absorbing medium. Since we want to investigate the performance of the PML (which generates an absorption), the estimate (15) for the elements in the PML is misleading. Consequently, the basis estimate (15) is used mainly for choosing the initial basis for the conditioning based selection. In particular, to give a fair comparison between the PML and ABC, the same strategy for choosing the basis must be used. We shall show, however, in Section 4.6 that in the absence of absorption, the estimates of (15) are directly applicable.

4.2. Electric dipole in free-space

The first model problem we investigate is to compute the field due to an electric dipole obtained by using Eq. (23) as the right-hand side of the UWVF. In this section we shall assume that $\mu = 1$ so that, in the free-space, the exact solution of the problems is

$$\boldsymbol{E}_{\mathrm{ex}} = \mathrm{i}\omega I \phi(\boldsymbol{x}, \boldsymbol{x}_0)\boldsymbol{a} - \frac{I}{\mathrm{i}\omega\epsilon}\nabla_x(\nabla_x\phi \cdot \boldsymbol{a}), \tag{28}$$

where

$$\phi(\boldsymbol{x}, \boldsymbol{x}_0) = \frac{\exp(\mathrm{i}\omega\sqrt{\epsilon}|\boldsymbol{x} - \boldsymbol{x}_0|)}{4\pi|\boldsymbol{x} - \boldsymbol{x}_0|}.$$

The geometry and the mesh used in the free-space dipole simulations are shown in Fig. 2. To avoid possible spurious accuracy due to symmetry, the point source is located at the point $(0.2, 0.2, 0.2)$ of the $1.0 \times 1.0 \times 1.0$ cube centered at the origin. The cube is surrounded by 0.1 thick PML. Due to the presence of the singularity at the location of the point source, the mesh is refined near the point $(0.2, 0.2, 0.2)$.

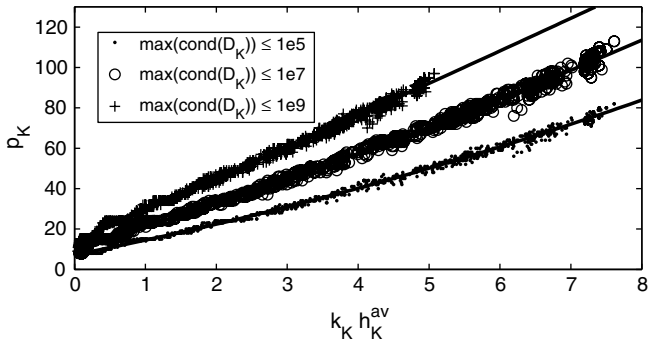

Fig. 1. The number of basis functions p_K as a function of $k_K h_K^{\mathrm{av}}$ when the basis dimension is chosen by constraining the maximum condition number of D_K.

Table 2
Parameters for the basis polynomials

Max(Cond(D_K))	A_2	A_1	A_0
10^5	0.3175	6.9819	7.1573
10^7	0.2578	10.7795	10.0676
10^9	0.0053	15.5926	13.5771

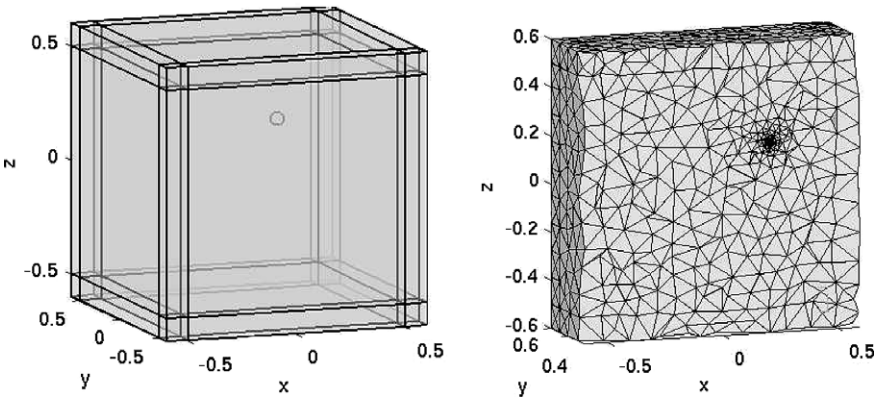

Fig. 2. *Left*: The domain enclosing the point source at $(0.2, 0.2, 0.2)$ (mark by a small sphere). The actual region of interest is the $1.0 \times 1.0 \times 1.0$ cube surrounded by a 0.1 thick PML. *Right*: The mesh for the problem consisting of 22,620 elements and 4430 vertices. The mesh is refined near the location of the point source so that elements size increases with distance from the source r as $5r^2$.

In all simulations for this problem, the basis is chosen by limiting the condition number of the matrix blocks D_K. The initial guess for the basis is made using the polynomials of Fig. 1. During the assembly of matrix D, the largest number of plane waves which give the condition number of D_K below the predetermined limit is chosen. The Bi-CGstab is terminated when the relative residual is below 10^{-5}.

Fig. 3 shows that all three condition number limits used in Fig. 1 lead to convergent Bi-CGstab iteration when the angular frequency is $\omega = 10\pi$ and the ABC is used. Subsequent simulations show that the condition numbers below 10^5 give a sufficiently large basis dimension for accuracy in the frequency range used in this study (to be quantified shortly). Therefore, it is used in the rest of the simulations. Hence, the number of basis functions in each element is approximately the same as shown in the lowest graph of Fig. 1, depending on the local wave number and size of the element. We note, however, that the actual number of basis functions can have small variation between elements, despite the same local wave number k_K and element size parameter h_K^{av}, since the conditioning is also affected by the shape of the elements. And as noted earlier, the absorption in the PML elements reduces the number of basis functions as compared to the estimates of Fig. 1 which are computed for a non-absorbing medium.

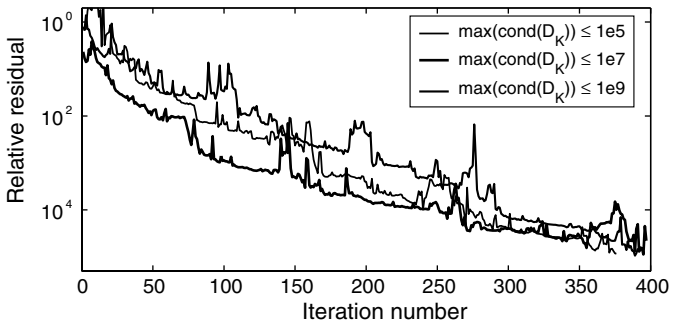

Fig. 3. The convergence of the Bi-CGstab for different condition numbers at $\omega = 10\pi$.

4.2.1. The effect of the PML

One objective of this study is to investigate the performance of the PML in the UWVF as a method for eliminating spurious reflections arising from the truncation of wave problems on unbounded domains. Fig. 4 shows the solution for the free-space dipole at the angular frequency $\omega = 30\pi$.

The effect of the PML decay parameter σ_0 on the accuracy of the UWVF approximation is investigated in Fig. 5. Since the boundary condition on the exterior boundary of the PML is Eq. (3) with $Q = 0$ and $\boldsymbol{g} = 0$, the case $\sigma_0 = 0$ corresponds to the low order absorbing boundary condition referred to as ABC in this study. Simulation are computed for three different angular frequencies $\omega = 10\pi, 20\pi$ and 30π. The corresponding wavelengths are $\lambda = 0.2, 0.1$ and 0.0667, so the thickness of the PML in terms of wavelengths is $0.5\lambda, \lambda$ and $3/2\lambda$. The same figure also shows the number of Bi-CGstab iterations needed to reach the relative residual below 10^{-5}.

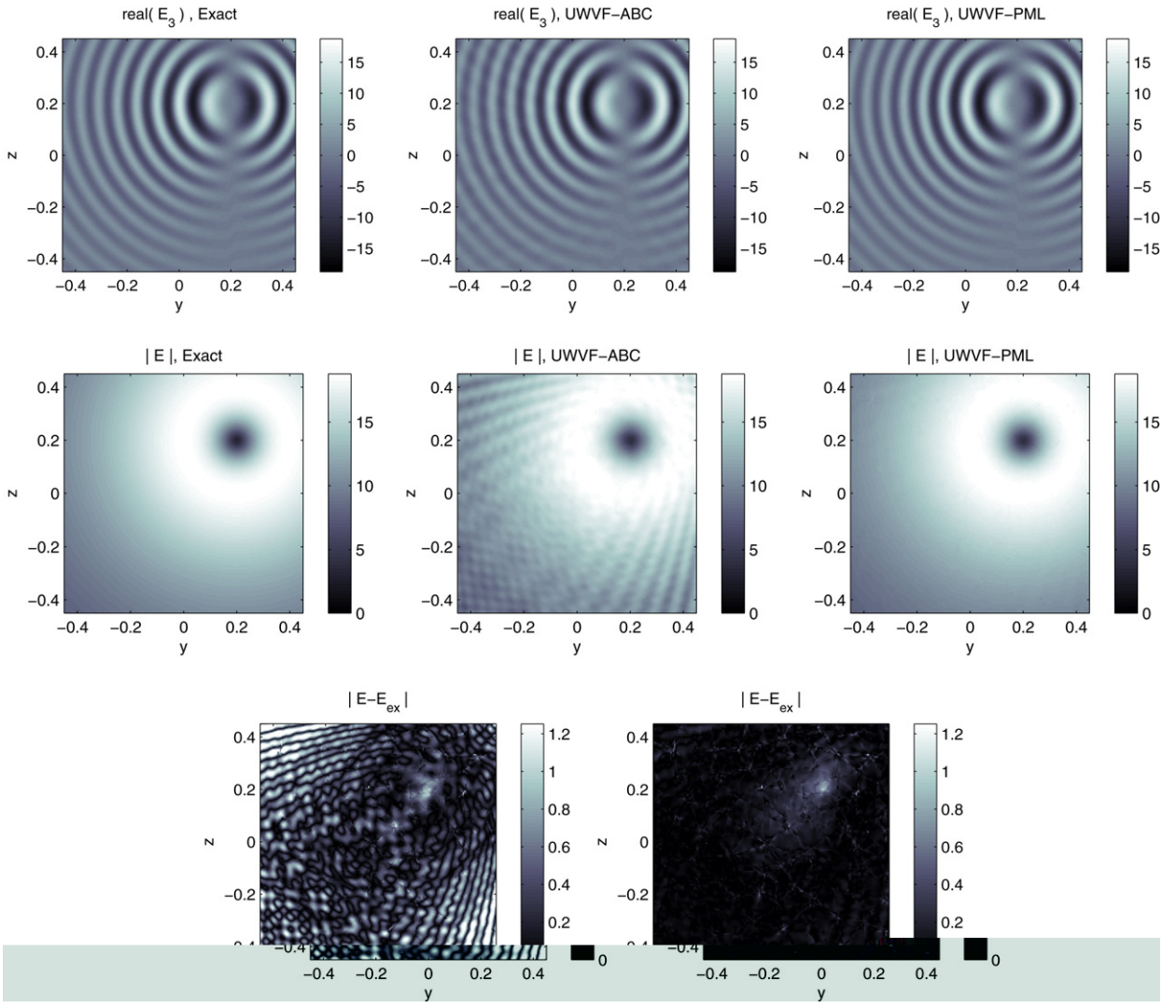

Fig. 4. UWVF approximation for the free-space dipole at $\omega = 30\pi$. The exact solution is computed using Eq. (28). The UWVF-ABC solution corresponds to the absorbing boundary condition (3) with $Q = 0$ and $\boldsymbol{g} = 0$. The approximation with the PML is for the decay parameter $\sigma_0 = 2$. The top row shows the real parts of the z-component of the electric field. The middle row shows the full amplitude of the electric filed $\boldsymbol{E}$. The bottom row presents the distribution of the error in the UWVF-ABC and UWVF-PML solutions. The errors for the ABC and PML approximations are 6.69% and 2.63%, respectively.

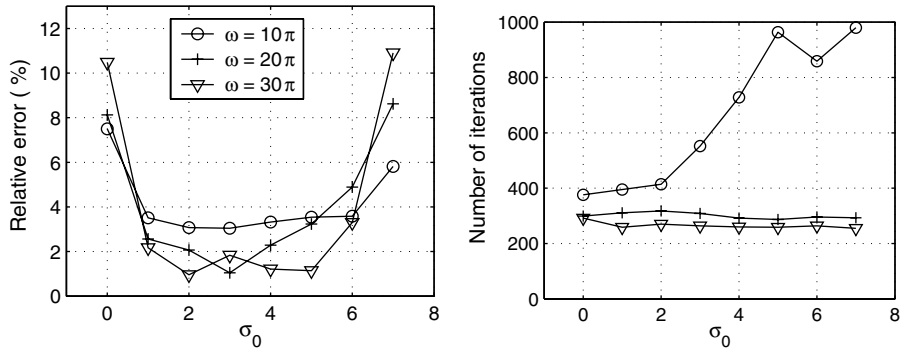

Fig. 5. *Left*: Error as a function of the PML decay parameter σ_0 at the point $(-0.2,-0.2,-0.2)$. *Right*: The number of BiCGstab iterations as a function of σ_0 in the same problem.

These results suggest that the UWVF approximation using the PML becomes unstable at low frequency causing an increased number of iterations. This could, perhaps, be helped by taking a thicker layer. On the other hand, there is a window of values for σ_0 which improve the accuracy of the UWVF-PML in comparison to the UWVF-ABC. When the decay parameter is too large, the error increases. This is caused by the reduced number of basis functions in the PML elements resulting from the condition number based criterion for choosing the basis. More precisely, we limit the maximum condition number of blocks D_K below 10^5. At large σ_0 this criterion is met or even exceeded when the number of plane waves in the PML elements is only three.

The instability of the UWVF-PML approximation at low frequencies is seen more clearly in Fig. 6 in which the error and the number of degrees of freedom (DOF) are plotted as a function of the frequency. The axis on the top of the error plot shows the ratio of the maximum element size $h_{\max}$ and the wavelength λ. The PML leads to poorer accuracy than the ABC at the lowest frequency $\omega = 5\pi$ after which the performance of the PML improves. Results suggest that the use of the PML at higher frequencies has two advantages, First, it reduces the error. Second, due to the smaller number of basis functions needed for the elements in the PML, it also reduces the size of the problem (of course, the PML would also allow us to reduce the size of the computational domain which would further reduce the size of the problem).

4.2.2. Field near the singularity

Since the field of the dipole has a singularity at its origin, it is important to investigate the error of the UWVF approximations as a function of the distance from the singularity. Fig. 7 presents the error for approx-

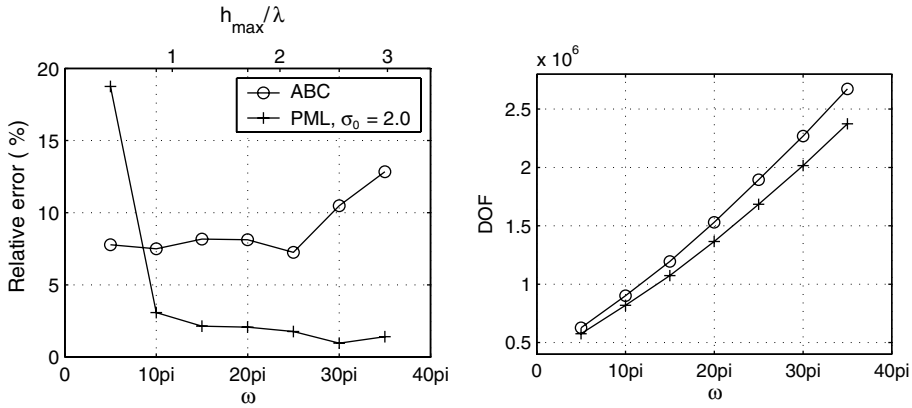

Fig. 6. *Left*: Error as a function of the angular frequency ω at the point $(-0.2,-0.2,-0.2)$. The top axis of the figure shows the number of wavelengths per largest element edge length in the non-PML region $h = 0.17$. *Right*: The number of degrees of freedom (DOF) as the function of ω.

imating the solution using the UWVF-ABC and the UWVF-PML along the diagonal of the cubic computational domain. The error is plotted on the line from the point $(-0.5, -0.5, -0.5)$ to $(0.5, 0.5, 0.5)$. As is to be expected, the error peaks strongly at location of the dipole where the true solution is unbounded. While the solution with the ABC has wavy spurious reflections, the error for the PML is smoother within the computational domain. However, the PML error also increases rapidly near the corner of the actual computational domain. This suggests that the PML still induces weak reflections back to the computational domain.

4.3. Layered media

The UWVF method can easily be used for problems in an inhomogeneous medium. Fig. 8 shows the UWVF approximations for a dipole source above a layered medium. The upper domain $z > 0$ has $\epsilon = 1$ and in the region $z < 0$ $\epsilon = 2$. Results are shown for $\omega = 30\pi$. As in the case of the homogeneous medium, the use of the PML reduces spurious reflections from the exterior boundary. This can be shown by comparing the results for the ABC and the PML with a analytical solution of the problem which is outlined in [19]. In Fig. 9, we plot the field $|\boldsymbol{E}|$ along the line $z = -0.2$ in the $x = 0$ plane. The wavy spurious reflection of the ABC are almost extinguished when the PML is used. Consequently, the error is reduced from 3.4% to 0.8%.

4.4. Scattering from a sphere

The third model problem we study is the scattering of a plane wave from a perfectly conducting sphere with radius $R = 0.5$. The actual region of interest is a cube with side length 0.55. This domain is surrounded by a 0.2 units thick PML. The mesh used in all simulations of this section is shown in Fig. 10. It has maximum length

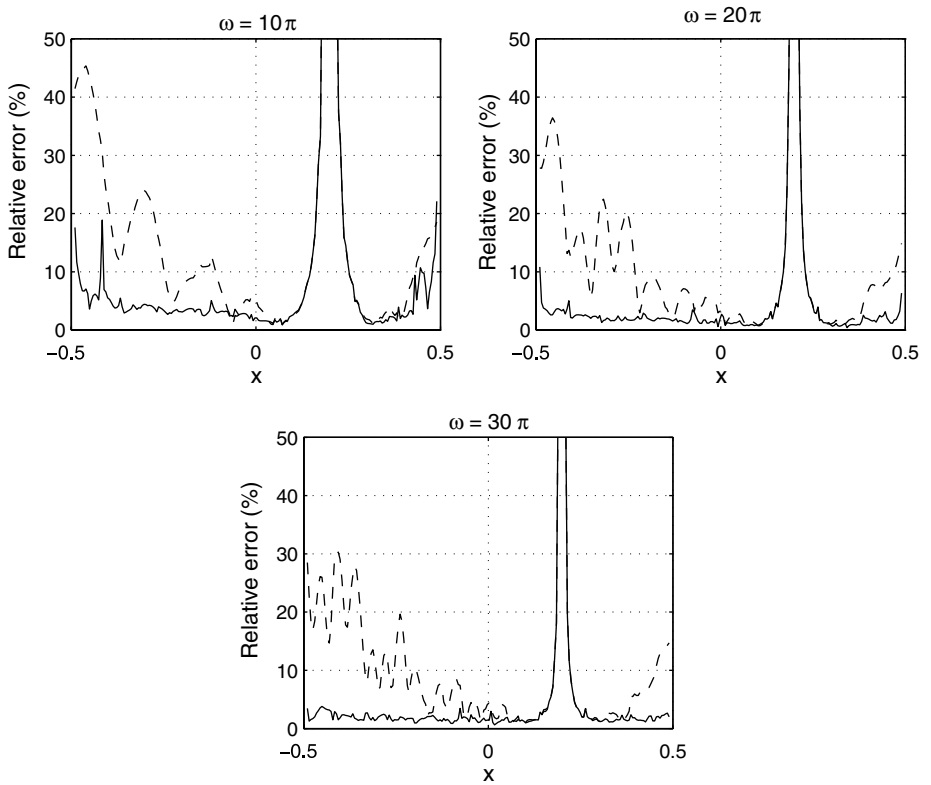

Fig. 7. The error on the diagonal of the computational domain (along the line from $(-0.5, -0.5, -0.5)$ to $(0.5, 0.5, 0.5)$). The dashed line is the error for the ABC and solid line shows the error when the PML is used. The PML simulations are computed using the decay parameter $\sigma_0 = 2.0$. The peak in the error is at the location of the dipole at $(0.2, 0.2, 0.2)$.

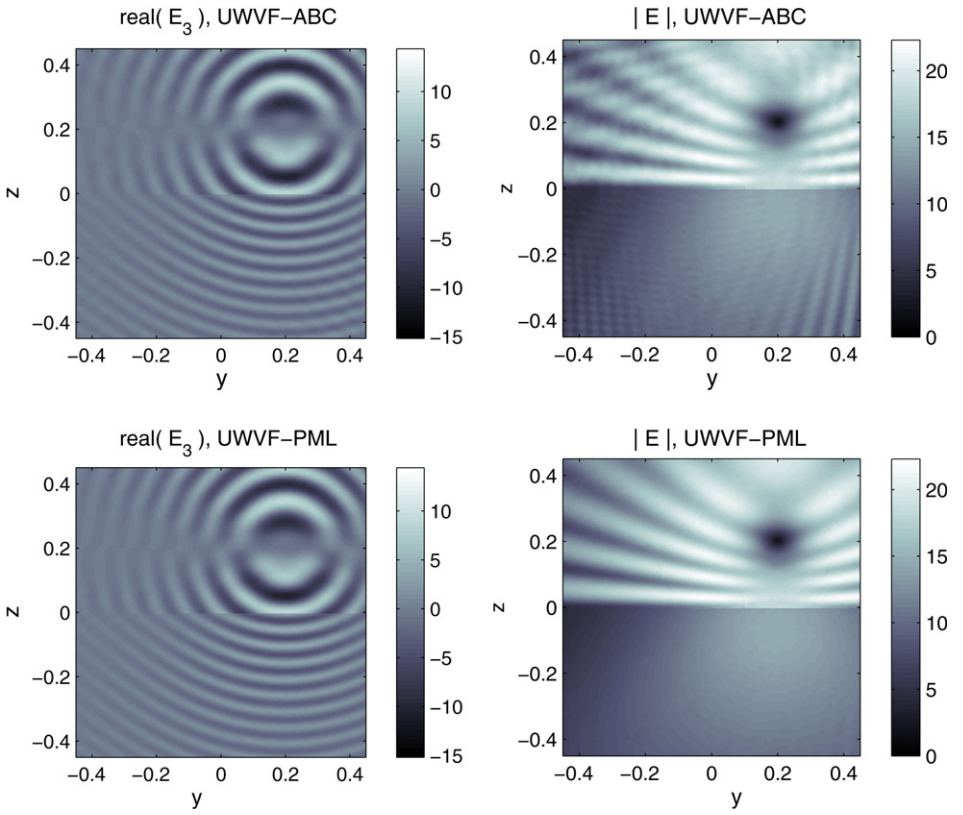

Fig. 8. The UWVF-approximations for a dipole in a layered medium when $\omega = 30\pi$. The PML improves the solution.

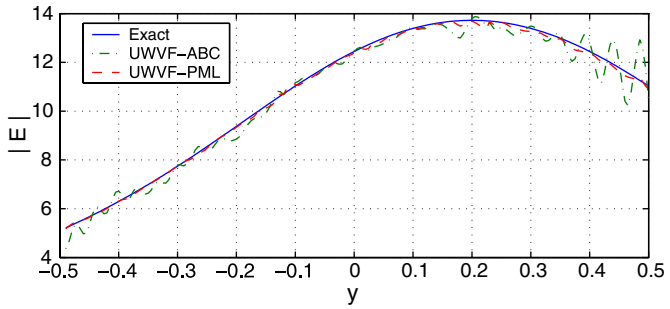

Fig. 9. The field $|\boldsymbol{E}|$ along the line $z = -0.2$ in the $x = 0$ plane. The PML reduces the error from 3.4% to 0.8%. The exact solution for the problem is formulated in [19].

of an element edge $h_{\max} = 0.196$. To ensure an accurate geometric representation of the surface of the sphere, the mesh is refined near the scatterer. We compute solution at $\omega = 30\pi$ which gives $\lambda/h_{\max} = 0.340$, i.e. approximately three wavelengths per element.

The problem can be decomposed into the scattered part $\boldsymbol{E}_{\text{sc}}$ and the incident part $\boldsymbol{E}_{\text{in}}$. The incident field is a y-direction polarized plane wave propagating in the direction of positive x-axis. The problem is formulated for the scattered field only. In the UWVF, we set $Q = 1$ and $\boldsymbol{g} = (1/2)\boldsymbol{n} \times \boldsymbol{E}_{\text{in}}$ (see Eq. (3)) on the surface of the sphere. On the exterior boundary we have $Q = 0$, $\sigma = \sqrt{\mu_0}/\sqrt{\epsilon_0}$ and $\boldsymbol{g} = 0$ which corresponds to the low-order ABC. We also compute the electric far-field pattern $\boldsymbol{E}_{\infty}$ defined in spherical coordinates (r, θ, ϕ) as

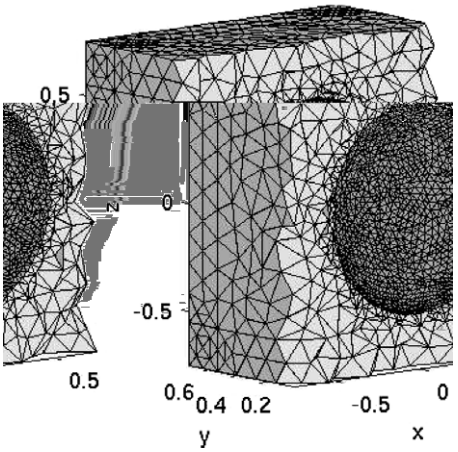

Fig. 10. The mesh used in the scattering from the sphere simulations. The mesh consists of 12,176 vertices and 45,461 tetrahedra. To better approximate the sphere, the mesh is relatively fine on near the surface of the scatterer.

$$\boldsymbol{E}(r,\theta,\phi) \approx \boldsymbol{E}_\infty(\theta,\phi)\frac{e^{i\omega\sqrt{\mu_0\epsilon_0}r}}{r},$$

when $r \to \infty$. The computation of the far-field from the UWVF approximation is presented in [4]. In all far-field simulations of this study, the integrals for the resolving the far-field are computed over the exterior boundary of the computational domain.

Fig. 11 shows the near-field UWVF-PML approximations with $\sigma_0 = 2.0$ for the scattered field $\boldsymbol{E}_{\text{sc}}$. The total field $\boldsymbol{E}_{\text{sc}} + \boldsymbol{E}_{\text{in}}$ is shown on the right. The electric far-field pattern for the same problem is plotted in Fig. 12. The figure shows UWVF approximations using the ABC and the PML which both compare well with the analytical Mie series solution (the PML solution is almost indistinguishable from the exact solution).

The error in the far-field and the number of degrees of freedom (DOF) as a function of the angular frequency ω is shown in Fig. 13. The axis on the top of the error plot shows the ratio of the maximum element size $h_{\max}$ and the wavelength λ.

4.5. NASA almond

The last model problem is the scattering of a plane wave from the NASA almond for which experimental data is published in [27]. For this example we report dimensional lengths chosen to correspond to the

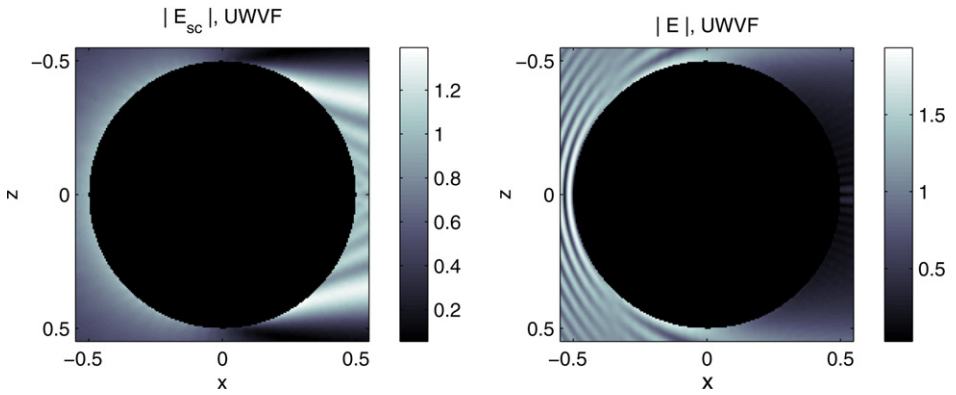

Fig. 11. The UWVF-PML approximation for the scattering problem at $\omega = 30\pi$. The amplitude of the scattered field is shown on the left. On the right is the amplitude of the total field (giving an idea of the radius of the sphere in wavelengths).

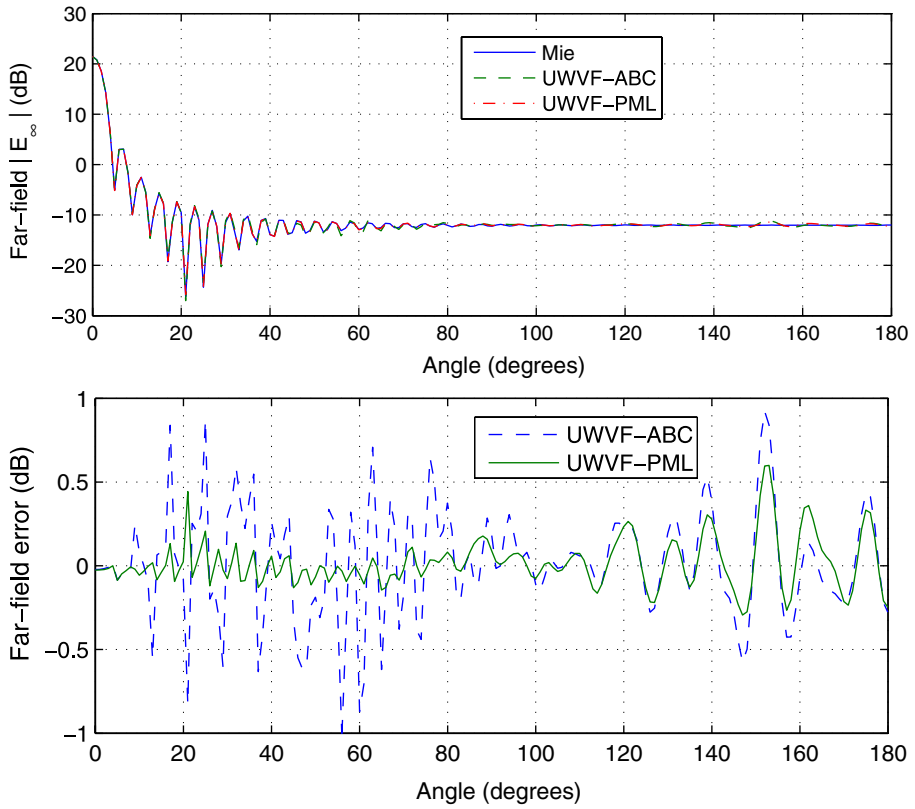

Fig. 12. *Top*: The far-field pattern $|\boldsymbol{E}_\infty|$ for the scattered field from the sphere at $\omega = 30\pi$. *Bottom*: The corresponding difference between the UWVF approximations and Mie series solution in decibels (i.e. $20(\ln|\boldsymbol{E}_\infty| - \ln|\boldsymbol{E}_\infty^{\mathrm{Mie}}|)$) shows that the UWVF approximations are in good agreement with the Mie series solution.

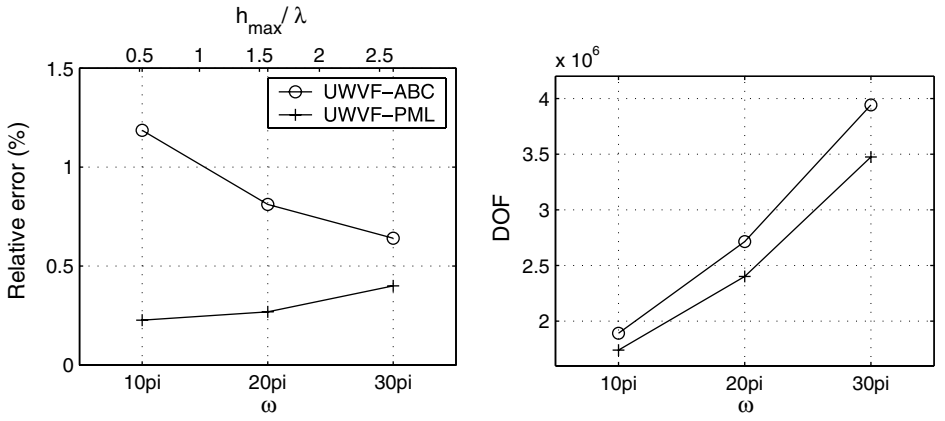

Fig. 13. *Left*: Error in the far-field pattern as the function of the angular frequency ω. *Right*: The number of degrees of freedom (DOF) as a function of ω. Note that the same mesh is used. Thus as ω increases the geometric error in representing the sphere will dominate and we expect the PML and ABC solutions to have roughly the same error at high frequency.

experimental measurements. The perfectly conducting metallic almond-shaped scatterer is 25.2 cm long. The coordinate system is chosen so that the longest dimension of the almond is in the x-direction and the smallest in the z-direction. The computational domain is $40 \times 40 \times 40\ \mathrm{cm}^3$ cube. The object is illuminated by plane

waves which propagate in (x, y)-plane and are vertically polarized (i.e. the VV-polarization in [27]). Hence, the formulation of the problem for the scattered field $\boldsymbol{E}_{\mathrm{sc}}$ is identical with that of the sphere in the previous section.

For each plane wave, the back-scattered radar cross section (RCS) is computed. We simulate two of the measured frequencies. At 1.19 GHz, the wavelength is $\lambda = 25.2$ cm and 9.92 GHz corresponds to $\lambda = 3.0$ cm.

In Fig. 15 we compare UWVF approximation for the RCS with the simulation computed using the electromagnetics module of a commercial finite element solver called COMSOL Multiphysics 3.2 (Comsol AB, Stockholm, Sweden) [10]. The UWVF approximations are computed on the meshes of Fig. 14. We use the same fine UWVF mesh (with $h_{\max} = 0.072$) also for the finite element method (FEM) simulations using COMSOL Multiphysics 3.2. In that case $\lambda/h_{\max} = 3.5$ which can be expected to be near the minimum that can be handled by COMSOL Multiphysics due to the limitations of the underlying first order edge (vector) edge element method. An additional FEM simulation is computed using a mesh with $h_{\max} = 4.0$ cm which gives $\lambda/h_{\max} = 6.3$ (the mesh consists of 72,559 tetrahedra and 14,306 vertices). All UWVF and FEM results are computed using the same low-order ABC on the truncated exterior boundary (i.e. $Q = 0$, $\sigma = \sqrt{\mu_0}/\sqrt{\epsilon_0}$ and $\boldsymbol{g} = 0$ in Eq. (3)).

A comparison of the results with Woo's experimental data (measured data is extracted every 10 degrees from the enlarged Fig. 3 of [27]) shows that all simulations are in good agreement with the measurements. However, despite the equal accuracy in the surface triangulation and the same statement of the problem, the COMSOL and UWVF solutions differ slightly near the minima at 45° and 135° angles. A detailed inspection with Fig. 3 of Ref. [27] suggests that the location of the minima and the value of the RCS at the minima are captured more accurately by the UWVF. The refinement of the mesh improved the accuracy of the FEM approximation but further refinement was impossible due to the limited memory in our workstation (COMSOL Multiphysics was run on a 3.0 GHz Pentium 4 processor having 3.0 GB RAM).

It is evident that the UWVF is not very efficient for solving problems at low frequencies (such as the almond problem at 1.19 GHz) when the detailed representation of the geometry requires the use of a mesh containing many small elements. Despite the dense mesh near the scatterer, the UWVF basis should consists of at least three plane waves per element which inevitably increases the overall number of degrees of freedom. In the coarse mesh, the number of plane waves in the UWVF basis varies from 3 to 22, the corresponding DOF being 222,236. The solution required 476 MB memory.

However, a great advantage of the UWVF is that a single mesh can be used for simulations over a wide range of frequencies by increasing the number of basis functions along with the frequency (as shown in Fig. 1). To demonstrate this feature of the method, we solve the RCS for the almond at 9.92 GHz using the coarser mesh of Fig. 14. The results are shown in Fig. 16. We note that even though the solution for the coarse mesh has several wavelengths per element size ($\lambda/h_{\max} = 0.21$), the method can accurately resolve

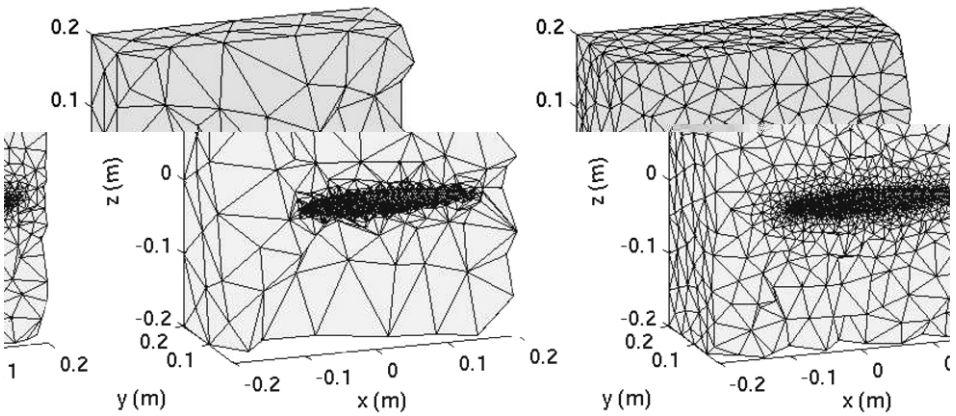

Fig. 14. Two meshed used in the NASA almond simulations. The coarse mesh consists of 12,715 tetrahedra and 3658 vertices. The fine mesh is constructed from 44,903 tetrahedra and 9130 vertices. Despite the large difference in the number of elements, the surface discretization is almost equal for both meshes. The surface of the almond is represented using 5486 triangles in the dense mesh and 5278 triangles in the coarse mesh.

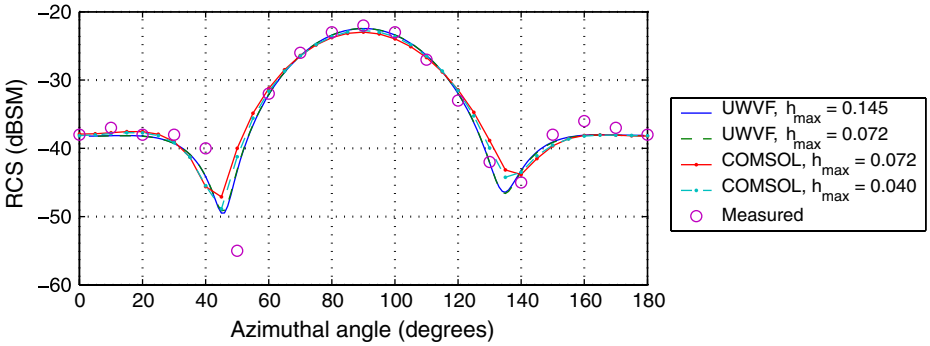

Fig. 15. A comparison of the UWVF and COMSOL Multiphysics 3.2 simulations for the NASA almond at 1.19 GHz (VV-polarization). Experimental data in every 10 degrees is extracted from Fig. 3 of Ref. [27].

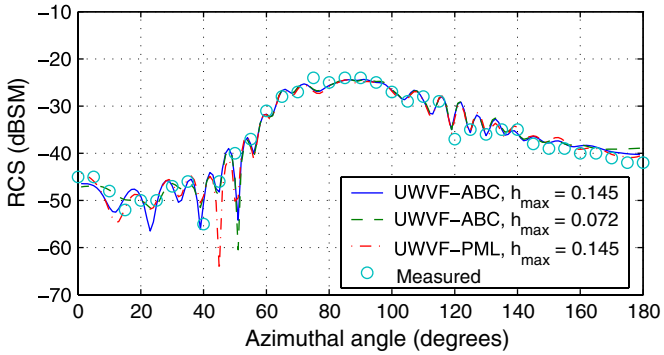

Fig. 16. UWVF approximations for the RCS from the NASA almond at 9.92 GHz (VV-polarization). Experimental data for the problem can be found from Fig. 6 of [27].

the RCS. In this case, the number of basis functions varies from 4 to 130 and the number of DOF is 657,538. The solution needed 7.73 GB memory. To verify the convergence of the solution, we also plot the same data that is computed on the fine mesh of Fig. 14. The two solutions are almost indistinguishable between angles 55° and 180° but differ in angles below 55° when the back scattered signal is extremely weak.

Due to poor PML solutions at low frequencies, the PML could not be used for the UWVF simulations at 1.19 GHz. However, the PML is fully functional at higher frequencies and therefore we also plot the RCS computed using PML in Fig. 16. The mesh used with the PML is relatively coarse ($h_{\max} = 14.5$ cm) and it is constructed by surrounding the $40 \times 40 \times 40$ cm^3 computational domain by a 5 cm thick PML. The mesh consisted of 14,211 tetrahedra and 3943 vertices. The decay parameter is set to $\sigma_0 = 2.0$. The approximation with the PML is again in good agreement with the two UWVF-ABC solutions but differs most in the angles below 55°.

A comparison of UWVF approximations with the experimental data (measured data is extracted in five degrees spacing from the enlarged Fig. 6 of Ref. [27]) shows that all simulations compare well with measurements in angles larger than 55°. At smaller angles the back scattered wave weakens and the differences in the UWVF approximations increases. It is difficult to judge which one the three simulated RCS best corresponds to the measurements. However, the UWVF-PML simulation gives the best approximation for the RCS when the target is illuminated from the angle zero. In addition, the location of the minima in the RCS are somewhat correctly captured using the PML. However, the deepness of the minima differs from the measurements. The differences in two UWVF-ABC solutions can arise from the approximation error or in small differences in the surface representation.

4.6. Parallel efficiency and the CPU-time

In the final part of this study, we investigate the parallel efficiency of the UWVF code. We also examine the distribution of the CPU-time between different sub-procedures of the UWVF method and compare the overall CPU-time with that of a commercial finite element solver.

The problem used here is scattering from the conducting sphere at $\omega = 10\pi$. The computation of the far-field pattern is not included into the computation time since it is done as a post-processing step on a single processor. In Fig. 17, we plot the CPU-time as a function of processor number when the solution is computed using the ABC and PML. In both cases, the scalability compares well with the ideal speed-up $1/np$ where np is the number of processors.

For Fig. 17, the UWVF-ABC solution is computed using the number basis functions which is obtained directly from the estimate of Fig. 1 for the condition number 10^5 (i.e. the number of basis functions is not changed during the assembly of matrix D). This approach is feasible when there is no strongly absorbing materials or the PML in the computational domain. The accuracy of the solution is almost equal when the basis is chosen using the estimate of Fig. 1 or when the basis is chosen by constraining the condition number during the assembly of matrix D. The errors for these two approaches are 1.33% and 1.19%, respectively. The corresponding DOFs are 1,658,908 and 1,892,382. When the PML is used, the error is reduced to 0.23% but then the basis must be selected during assembly which leads to 1,740,922 DOFs.

Fig. 18 shows how the CPU-time is distributed between different sub-procedures during the computation. The major difference between the ABC and PML simulations is in the assembly of the matrix D. This is because the fastest way of assembling the matrices is to use the pre-selected number of basis functions (the figure in the middle). In the two other graphs, the number of basis functions for each element is chosen during the assembly of D by constraining the condition number of the matrix blocks D_K (see Section 4.1). The initial number of basis functions is chosen using the estimates of Fig. 1. Since the estimate does not take into account the effect of absorption in the PML on the number of basis functions, the initial basis must be corrected during the assembly more for the PML than in the case of the ABC, which is seen in the longer CPU-time.

Finally, we compare the CPU-time needed by the UWVF method and a commercial finite element solver (COMSOL Multiphysics 3.2). To be able to compute the simulation in a single processor, the frequency of the scattering problem is lowered to $\omega = 2\pi$, otherwise the problem is the same as in the previous studies of this section. Since only the ABC is readily available in COMSOL, it is also used in the UWVF simulations. In Fig. 19 we show the UWVF and FEM approximations for the far-field pattern. The same plot also includes the exact Mie series solution of the problem. Corresponding numbers of degrees of freedom, memory requirement, CPU-time and error are reported in Table 3.

To use all advanced features of COMSOL Multiphysics, the FEM solution is computed using multigrid accelerated GMRES (generalized minimal residual) iteration which is the fastest and most memory efficient solver available in COMSOL. The DOF shown in Table 3 is associated with the fine mesh of the multigrid

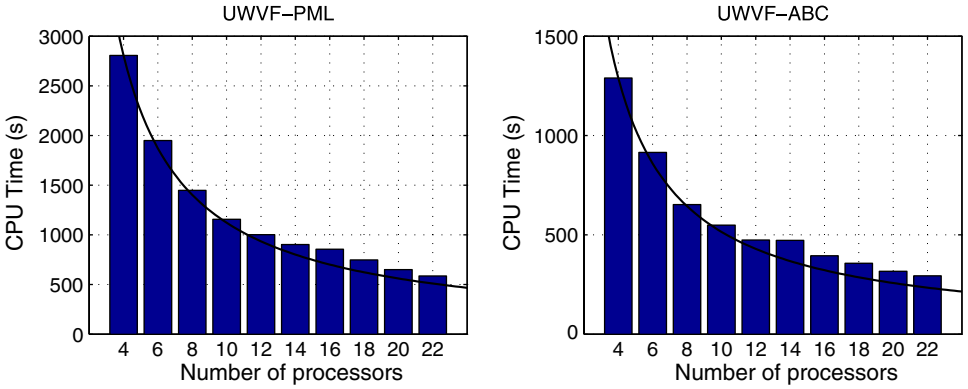

Fig. 17. The CPU-time as a function of the number of processors. *Left*: The computation times when the condition number basis selection and the PML is used at $\omega = 10\pi$. *Right*: Efficiency for the direct use of the estimate (15) and the ABC. Solid lines show the ideal speed-up $1/np$ where np is the number of processors.

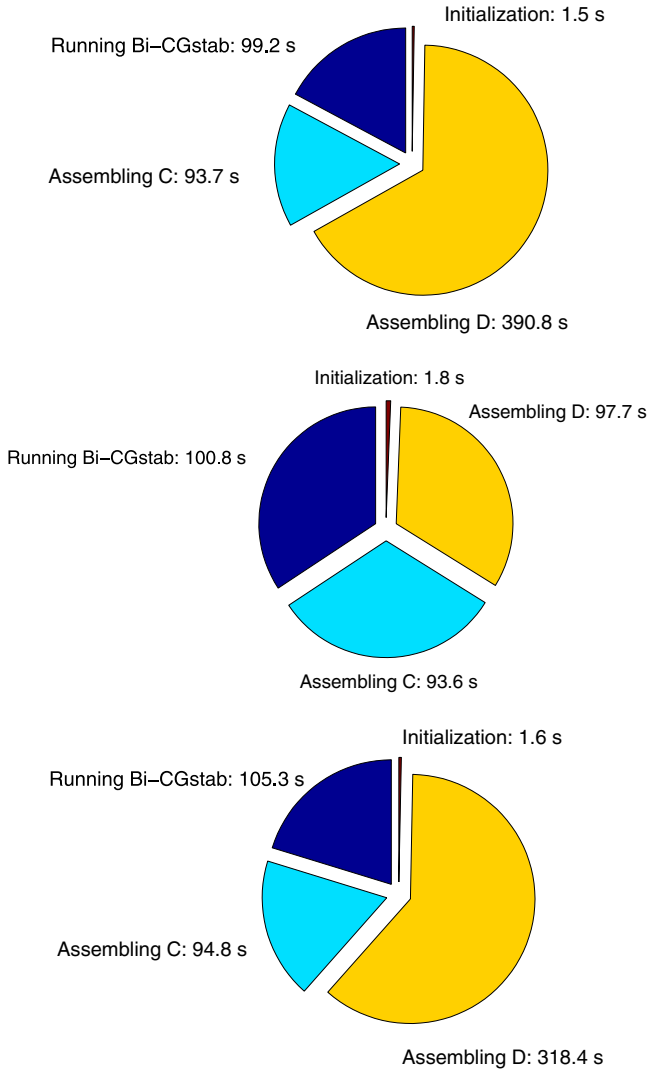

Fig. 18. The distribution of the CPU time for two strategies for choosing the basis. *Top*: The computation time when the condition number basis selection and the PML is used at $\omega = 10\pi$ with 22 processors. *Mid*: The same plot for the direct use of the estimate (15) and the ABC. *Bottom*: The distribution of the CPU-time when the condition number basis selection and the ABC is used.

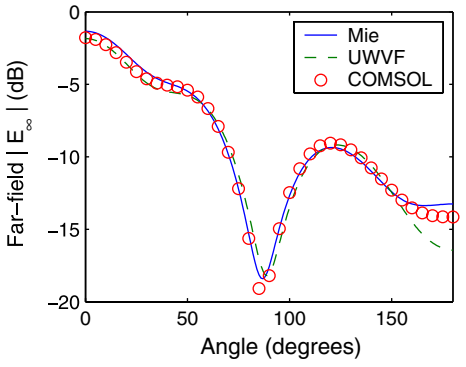

Fig. 19. Scattered far-field from the sphere at $\omega = 2\pi$.

Table 3
Comparison of the UWVF and FEM

Method	DOF	MEM (Gb)	CPU-time (s)	Error (%)
FEM	247,863	1.11	371	4.58
UWVF	35,690	0.11	212	7.54

iteration. The coarse mesh consisted of 25,290 elements. Since the ABC is used, the pre-selected number of basis functions for the UWVF is used (corresponding the condition number 10^5 in Fig. 1).

The UWVF leads to a reduced need for memory and CPU-time but, with the given meshes, results in a slightly larger error than the FEM. The reason for the larger error is an inaccurate geometric representation of the surface of the sphere. Since the UWVF allows the use of very coarse meshes, only 1122 elements is used for this problem. Hence, the coarse triangulation of the surface of the sphere provides an additional approximation error. The origin of this error can be verified by increasing the number of basis function in the UWVF mesh which does not reduce the error of the approximation. The error could be decreased, for example, by using curved elements near the surface as was done in 2D simulations in Ref. [16]. This possibility, however, is not investigated in this study.

5. Conclusion

We have investigated the feasibility of using the ultra weak variational formulation (UWVF) to solve time-harmonic Maxwell problems. In the first part of the paper, the UWVF was shown to be a upwind discontinuous Galerkin (DG) method with a special choice of basis functions. Namely, when the basis is a solution of the local adjoint Maxwell equation.

The main goal of this study was to present a practical implementation of the UWVF method. A parallel procedure for solving the UWVF approximations was outlined. In addition, we introduced a method for choosing the number plane wave basis functions for the UWVF, so that the resulting matrix equation is solvable using standard iterative methods (we used stabilized bi-conjugate gradient iteration). The idea is to constrain the condition number of local (element-wise) matrix system which provides a control over the conditioning of the whole UWVF matrix equation. We showed that in the absence of absorption, a stable number of basis functions can be estimated for each element based on the local wave number and the element size. Numerical examples in 3D showed the feasibility of the proposed computational method.

Potential future developments of the UWVF include:

(1) Theoretical estimates for the error and condition number for the UWVF (our method was based on the use of experimental simulations for providing the information on the condition number and accuracy).
(2) Reduction in the number of directions in the plane wave basis by altering also the directions of the plane waves element by element (only the number of equidistributed directions as was varied in this study). This will require the use of information of the dominant directions of the solution.
(3) A better parametrization of surfaces or interfaces in the computational mesh to avoid the need for using dense meshes near boundaries. This would also reduce the size of the problem by reducing the number of elements in the mesh. For example curved elements have been used on a circular scatterer in 2D UWVF simulation [16].
(4) To improve the accuracy of the UWVF near singularities. Despite the refinement of the mesh near singularities, the UWVF approximations failed near the point source. It may be possible to improve the accuracy there by using singular basis functions (such as the Bessel functions) or by coupling the UWVF with a standard polynomial based discontinuous Galerkin method which can be used near singularities.

Acknowledgments

This work was supported in part by the US AFOSR under Grant No. FA9550-05-1-0127. This material is based upon work supported by the National Science Foundation under Grant No. DMS-0322583. The

authors thank Houssem Haddar who supplied the layered medium code for the exact solution in that case. Tomi Huttunen thanks the Measurement and Sensor Laboratory (University of Oulu) in Kajaani for the collaboration in the field of electromagnetism and for the financial support from their Measurepolis-program.

References

[1] J. Bérenger, Perfectly matched layer for the FDTD solution of wave–structure interaction problems, IEEE Trans. Antennas Propag. 44 (1996) 110–117.
[3] Q. Carayol, F. Collino, Error estimates in the fast multipole method for scattering problems part 1: truncation of the Jacobi–Anger series, ESIAM: Mathematical Modeling and Numerical Analysis 38 (2004) 371–394.
[4] O. Cessenat, Application d'une nouvelle formulation variationnelle aux équations d'ondes harmoniques, Problèmes de Helmholtz 2D et de Maxwell 3D, Ph.D. Thesis, Université Paris IX Dauphine, 1996.
[5] O. Cessenat, B. Després, Using plane waves as base functions for solving time harmonic equations with the Ultra Weak Variational Formulation, J. Comput. Acoust. 11 (2003) 227–238.
[6] W.C. Chew, W.H. Weedon, A 3D perfectly matched medium from modified Maxwell's equations with stretched coordinates, Microwave Opt. Technol. Lett. 7 (1994) 599–604.
[8] D. Colton, R. Kress, Integral Equation Methods in Scattering Theory, Wiley, New York, 1983.
[9] D. Colton, P. Monk, Herglotz wave functions in inverse electromagnetic scattering theory, in: M. Ainsworth, P. Davies, D. Duncan, P. Martin, B. Rynne (Eds.), Topics in Computational Wave Propagation: Direct and Inverse Problems, Lecture Notes in Computational Science and Engineering, vol. 31, Springer, Berlin, 2003, pp. 267–294.
[10] Homepage of COMSOL Multiphysics software, 2006. Available from: <http://www.comsol.com>.
[11] E. Darrigrand, P. Monk, Coupling of the ultra-weak variational formulation and an integral representation using a fast multipole method in electromagnetism, J. Comput. Appl. Math., in press.
[12] E. Darve, The fast multipole method: numerical implementation, J. Comput. Phys. 160 (2000) 195–240.
[13] R. Hiptmair, Coupling of finite elements and boundary elements in electromagnetic scattering, SIAM J. Numer. Anal. 41 (2003) 919–944.
[14] T. Huttunen, J. Kaipio, P. Monk, Parallelized UWVF method for 3D Helmholtz problems, in: 4th ECCOMAS, Jyvskyl, Finland, July 2004.
[15] T. Huttunen, J. Kaipio, P. Monk, The perfectly matched layer for the ultra weak variational formulation of the 3D Helmholtz equation, Int. J. Numer. Meth. Eng. 61 (2004) 1072–1092.
[16] T. Huttunen, P. Monk, J. Kaipio, Computational aspects of the ultra weak variational formulation, J. Comput. Phys. 182 (2002) 27–46.
[17] G. Karypis, V. Kumar, A fast and high quality multilevel scheme for partitioning irregular graphs, SIAM J. Sci. Comput. 20 (1999) 359–392.
[18] P. Lesaint, P. Raviart, On a finite element method for solving the neutron transport equation, in: C. deBoor (Ed.), Mathematical Aspects of Finite Element Methods in Partial Differential Equations, Academic Press, New York, 1974, pp. 89–123.
[19] P. Monk, Finite Element Methods for Maxwell's Equations, Oxford University Press, Oxford, 2003.
[20] V.V. Nikol'skiy, T.I. Lavrova, The method of minimum autonomous blocks and its application to wavequide diffraction problems, Radio Eng. Electron. Phys. 23 (1978) 1–10.
[21] I. Sloan, R. Womersley, Extremal systems of points and numerical integration on the unit sphere, Adv. Comput. Math. 21 (2004) 107–125.
[22] N. Sloane, Tables of spherical codes (with collaboration of R.H. Hardin, W.D. Smith and others) published electronically at http://www.research.att.com/njas/packings, 2000.
[23] Y.O. Shlepnev, Trefftz finite elements for electromagnetics, IEEE Trans. Microwave Theory 50 (2002) 1328–1339.
[24] T. Weiland, A discretization method for the solution of Maxwell's equations for six-component fields, Electron. Commun. AEUE 31 (1977) 116–120.
[25] T. Weiland, Numerical solution of Maxwell's equation for static, resonant and transient problems, in: T. Berceli (Ed.), Studies in Electrical and Electronic Engineering 28B, URSI International Symposium on Electromagnetic Theory Part B, Elsevier, New York, 1986, pp. 537–542.
[26] R. Womersley, I. Sloan, How good can polynomial interpolation on the sphere be? Adv. Comput. Math. 14 (2000) 195–226.
[27] A.C. Woo, H.T.G. Wang, M.J. Schuh, M.L. Sanders, Benchmark radar targets for the validation of computational electromagnetics programs, IEEE Antennas Propag. 35 (1993) 84–89.