Extended Krylov Subspace Methods for Transient Wavefield Problems

R. F. Remis
Laboratory of Electromagnetic Research
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands

Abstract — We present an extended Krylov subspace method to solve multiscale transient electromagnetic wavefield problems. A basis of an extended Krylov subspace is generated by iterating with the system matrix and its inverse. We show that such a basis can be computed very efficiently via three-term recurrence and CG-type updating formulas by exploiting specific symmetry properties of the system matrix, which are related to energy conservation and reciprocity. Multiscale transmission line and full electromagnetic wavefield problems are considered, and numerical experiments illustrate the performance of the method.

1. INTRODUCTION

In this paper we present an Extended Krylov Subspace method (EKS method, see [3]) to efficiently simulate electromagnetic wavefield propagation. In particular, we apply the EKS method to time-domain transmission line problems and transient electromagnetic wavefield problems are considered as well.

Starting point for both types of problems is a first-order finite-difference state-space representation obtained after discretizing the transmission line equations or Maxwell’s equations in space using standard two-point finite-differences on a nonuniform staggered grid. The finite-difference representation is written in terms of a so-called system matrix and the field solution for transmission lines or Maxwell’s equations is essentially given by a temporal convolution of the source vector and the system matrix evolution operator (matrix exponent).

For problems encountered in practice, the order of the system matrix can be very large and direct evaluation of the evolution operator is simply not feasible. Fortunately, we do not need the evolution operator on its own; only its action on the source vector is required. We therefore approximate the wavefield quantities by elements taken from an extended Krylov subspace. Such a space is generated by the source vector, the system matrix, and an inverse of the system matrix. Loosely speaking, the idea is that by appending a standard Krylov basis with vectors consisting of powers of an inverse of the system matrix, we approximate early- and late-time responses simultaneously.

To generate a basis of the extended Krylov space, we obviously need an inverse of the system matrix. For transmission line problems we can show that this inverse exists if (part of) the transmission line contains some loss mechanism. Moreover, we have found an explicit expression for this inverse and there is no need to numerically solve a system of equations at every step. For electromagnetic wavefield problems the inverse of the system matrix does not exist in general. Here, however, we can work with its so-called group inverse. We have found an explicit expression for this inverse as well.

Having the necessary inverses available, we still need to generate a basis of the extended Krylov space. We show that this can be done very efficiently by exploiting the symmetry properties of the system matrix and its (group) inverse. In particular, a (J-)orthogonal basis of the extended Krylov subspace can be generated via short Lanczos/CG-type recurrence relations. The expansion coefficients follow from a Galerkin procedure and having found these coefficients we can construct the EKS field approximations. Numerical experiments illustrate the performance of the method.

2. BASIC EQUATIONS

In this section, we briefly review the semidiscrete finite-difference systems for (multiconductor) transmission lines and \(E\)-polarized electromagnetic wavefields in two dimensions. In both cases, we arrive at these systems by discretizing the spatial coordinate(s) on a possibly nonuniform Yee-grid and using two-point finite-difference formulas for the spatial derivatives.
2.1. The Semidiscrete Transmission Line System

To fix ideas, we consider a transmission line of length $L$ with a position dependent per-unit-length (p.u.l.) capacitance $c$, inductance $\ell$, and resistance $r$. The $z$-direction is taken as the direction of propagation and the line is terminated at $z=L$ by a load resistance $R_{ld}$ in series with a load inductance $L_{ld}$. A voltage generator is present at the near end of the line ($z=0$) and it consists of a voltage source $v_s(t)$ in series with a source resistance $R_s$.

For lines of the above type, the semidiscrete transmission line system can be written in the form

$$(\mathbf{A} + \mathbf{I}\partial_t)v = v_s(t)s,$$  \hspace{1cm} (1)

where $\mathbf{f} = [v^T, i^T]^T$ is the field vector and $\mathbf{v}$ and $\mathbf{i}$ are column vectors containing the finite-difference approximations of the voltage and the current, respectively. The total number of unknown time-dependent finite-difference approximations is denoted by $n$. Furthermore, $\mathbf{s}$ is the source vector and $\mathbf{A}$ is the so-called system matrix. It can be shown that if $R_s > 0$, or $R_{ld} > 0$, or the p.u.l. resistance $r > 0$ at least on some part of the transmission line, then the system matrix is nonsingular and an explicit expression for its inverse can be given as well (see [6]). There is no need to numerically solve a system of equations to compute the matrix-vector product $\mathbf{A}^{-1}\mathbf{x}$ for a given vector $\mathbf{x}$. Finally, the system matrix is $J$-symmetric because of reciprocity. More precisely, matrix $\mathbf{A}$ satisfies the symmetry relation

$$\mathbf{A}^T\mathbf{J} = \mathbf{JA} \quad \text{with} \quad \mathbf{J}^T = \mathbf{J}.$$  \hspace{1cm} (2)

For the transmission line problem considered here, matrix $\mathbf{J}$ is diagonal and not positive definite. An explicit expression for matrix $\mathbf{J}$ can be found in [6]. The inverse of the system matrix is also $\mathbf{J}$-symmetric, since $J$-symmetry is closed under inversion. Finally, we mention that semiconductor transmission lines can be handled essentially by tensoring the equations for a single transmission line.

2.2. The Semidiscrete Maxwell System

We consider $E$-polarized fields in a two-dimensional configuration that is invariant in the $z$-direction. The media in this configuration are characterized by the position dependent permittivity $\varepsilon$ and permeability $\mu$ and an external electric-current source $J^\text{ext} = w(t)J^\text{sp}(x)$ generates the electromagnetic field. The scalar function $w$ is called the source signature and it vanishes for $t<0$. The electric and magnetic field strength vanish for $t<0$ as well.

The semidiscrete Maxwell system can also be written in a form as given by Eq. (1). For electromagnetic wavefields, the field vector is given by $\mathbf{f} = [\mathbf{h}_x^T, \mathbf{e}_z^T, \mathbf{h}_y^T]^T$, where $\mathbf{h}_x$ contains all the time-dependent finite-difference approximations of the $x$-component of the magnetic field strength, and the finite-difference approximations of the $z$-component of the electric field strength and the $y$-component of the magnetic field strength are stored in the vectors $\mathbf{e}_z$ and $\mathbf{h}_y$, respectively. The total number of unknowns is again denoted by $n$.

The source vector $\mathbf{s}$ is given by $\mathbf{s} = [0^T, (j^{\text{sp}})^T, 0]^T$, where $(j^{\text{sp}})$ is a finite-difference approximation of $J^\text{sp}$. Finally, for Maxwell’s equations we have $v_x(t) = -w(t)$ and $\mathbf{A}$ is the Maxwell system matrix. This matrix satisfies the symmetry relations

$$\mathbf{A}^T\mathbf{W}^\text{en} = -\mathbf{AW}^\text{en} \quad \text{(energy conservation)} \quad \text{and} \quad \mathbf{A}^T\mathbf{J} = \mathbf{JA} \quad \text{(reciprocity)}.$$  \hspace{1cm} (3)

The first relation states that the system matrix is skew-symmetric with respect to the diagonal and positive definite energy matrix $\mathbf{W}^\text{en}$. We refer to $\mathbf{W}^\text{en}$ as an energy matrix, since $\frac{1}{2}\mathbf{f}^T\mathbf{W}^\text{en}\mathbf{f}$ is a finite-difference approximation of the stored electromagnetic energy in the configuration. Using the skew-symmetry of matrix $\mathbf{A}$, we can show that energy is conserved.

The second relation in Eq. (3) states that the Maxwell system matrix is $J$-symmetric. Again, this symmetry relation holds because of reciprocity and here too matrix $\mathbf{J}$ is diagonal and not positive definite. Furthermore, $\frac{1}{2}\mathbf{f}^T\mathbf{J}\mathbf{f}$ is a finite-difference approximation of the field Lagrangian (see [2]). Explicit expressions for matrix $\mathbf{A}$, $\mathbf{W}^\text{en}$, and $\mathbf{J}$ can be found in [7].

To construct a basis of a standard extended Krylov subspace, the inverse of the system matrix is required. The problem with the system matrix for Maxwell’s equations is that its inverse does not exist. The system matrix is singular and we therefore resort to the Drazin (group) inverse of the system matrix. This inverse, denoted by $\mathbf{A}^D$, is uniquely defined by the following conditions (see [1]):

$$\mathbf{A}^D\mathbf{A} = \mathbf{A}^D, \quad \mathbf{A}\mathbf{A}^D = \mathbf{A}^D\mathbf{A}, \quad \text{and} \quad \mathbf{A}^{p+1}\mathbf{A}^D = \mathbf{A}^p,$$  \hspace{1cm} (4)

where $p$ is the index of matrix $\mathbf{A}$. Recall that the index of matrix $\mathbf{A}$ is defined as the smallest nonnegative integer $p$ such that $\text{rank}(\mathbf{A}^p) = \text{rank}(\mathbf{A}^{p+1})$. If the index of a matrix is equal to one,
then its Drazin inverse is called the group inverse and it is denoted by $A^\#$. The Maxwell system matrix has an index $p = 1$ and therefore it has a group inverse $A^\#$. Since the source vector is in the range of the system matrix, we also have $AA^\#s = A^\# As = s$.

The group inverse of the system matrix is skew-symmetric with respect to $W^{en}$ and it is $J$-symmetric as well. An explicit expression for the group inverse can be given and it shows that the solution of Poisson’s equation is required to evaluate its action on a given vector. Specifically, evaluating the action of the group inverse on the source vector requires the solution of a single Poisson equation and $A^\#s$ is a finite-difference approximation of the Biot-Savart law [5].

3. EXTENDED KRYLOV SUBSPACE APPROXIMATIONS

In the extended Krylov method, the field vector $f$ is approximated by elements drawn from the Krylov subspace

$$
K^{k,m} = \text{span}\{A^{-k+1}s, A^{-k+2}s, \ldots, A^{-1}s, s, As, \ldots, A^{m-1}s\},
$$

where $A^{-1}$ has to be replaced by $A^\#$ if electromagnetic wavefield problems are considered. Notice that the maximal dimension of $K^{k,m}$ is $d = k + m - 1$ and we are interested in extended Krylov approximations for which $d \ll n$.

Constructing the approximations is a two-step procedure. We first generate a basis for $K^{k,m}$ and expand the Krylov approximations in terms of the basis vectors. The second step consists of determining the time-dependent expansion coefficients.

By exploiting the symmetry properties of the system matrix (matrix $A$ is $J$-symmetric or skew-symmetric or both), we can efficiently construct the basis vectors of $K^{k,m}$. The skew-symmetry of the system matrix can be used to construct an orthogonal basis for $K^{k,m}$ via three-term recurrence and CG-type update equations, while if we exploit the $J$-symmetry of the system matrix then a $J$-orthogonal basis for $K^{k,m}$ can be constructed in much the same way as is done for the skew-symmetric case. We note that the algorithm based on $J$-symmetry may suffer from breakdown, since matrix $J$ is not positive definite. In our numerical work we have never detected a breakdown of the algorithm, but we do not have a proof that no breakdown can occur either (in exact or finite-precision arithmetic).

Let us now give a brief description of how we generate a ($J$-)orthogonal basis for $K^{k,m}$. We follow the approach outlined in [3]. A different (perhaps more flexible) approach is presented in [4].

Let the columns $w_i$ of the $n$-by-$d$ matrix $W_d$ form a ($J$-)orthogonal basis of $K^{k,m}$. We expand the Krylov approximations in terms of these basis vectors. Specifically, the extended Krylov approximation of order $d$ is written as

$$
f_d(t) = w_1c_1(t) + w_2c_2(t) + \ldots + w_dc_d(t) = W_d c(t),
$$

where $c = [c_1(t), c_2(t), \ldots, c_d(t)]^T$ is the vector of expansion coefficients. We now partition matrix $W_d$ as

$$
W_d = (Q_k, V_{m-1}) \quad \text{where} \quad Q_k = (q_1, q_2, \ldots, q_k) \quad \text{and} \quad V_{m-1} = (v_1, v_2, \ldots, v_{m-1}).
$$

The vectors $q_i$, $i = 1, 2, \ldots, k$, are computed with $A^{-1}$ (or $A^\#$) as iteration matrix, while the vectors $v_i$, $i = 1, 2, \ldots, m - 1$, are generated with the system matrix $A$. Specifically, the basis vectors are constructed as follows:

1. Run $k$ steps of the Lanczos algorithm with matrix $A^{-1}$ (or $A^\#$) and take the source vector $s$ as a starting vector to obtain $q_1, q_2, \ldots, q_k$;
2. Determine $v_1$ and $v_2$ by orthogonalizing $As$ against $Q_k$ and $Av_1$ against $v_1$ and $Q_k$;
3. Given the vectors $v_1$ and $v_2$, the remaining vectors $v_3, v_4, \ldots, v_{m-1}$ can be computed via a three-term recurrence relation by exploiting the $J$-symmetry or skew-symmetry of the system matrix.

Notice that in step 2 all basis vectors $q_i$ are required. Storage of $Q_k$ can be avoided, however, by rewriting $v_1$ and $v_2$ in terms of an auxiliary vector (cf. [3]). This auxiliary vector can be computed recursively by including a CG-type update formula in step 1. Storage of $Q_k$ is avoided in this way.

After completion of the above algorithm, we have the summarizing equation

$$
AW_d = W_d R_d + t_{m,m-1} v_m e_d^T,
$$
where \( t_{m,m-1} \) is a recurrence coefficient from step 3 and \( e_d \) is the \( d \)th column of the \( d \)-by-\( d \) identity matrix. Furthermore, matrix \( R_d \) is a matrix of order \( d \) with a block structure

\[
R_d = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix},
\]

where \( R_{11} \) is a dense square matrix of order \( k \), \( R_{22} \) is a square and tridiagonal matrix of order \( m-1 \), and all elements of \( R_{12} \) (\( R_{21} \)) are zero, except for the elements in its first column (row).

The second step consists of determining the expansion coefficients. To this end, we substitute the Krylov approximation of Eq. (6) in the semidiscrete system of Eq. (1), use Eq. (8), and apply a (pseudo-)Galerkin procedure. The expansion coefficients are found as

\[
c(t) = \|s\| \int_{\tau=0}^{t} v_s(\tau) \exp[-R_d(t-\tau)]e_1 \, d\tau \quad \text{for} \quad t > 0,
\]

where \( \|s\| \) is the 2-norm of the source vector (we normalize each basis vector such that it has unit 2-norm). With the expansion coefficients at our disposal, the extended Krylov approximation becomes

\[
f_d(t) = \|s\| W_d \int_{\tau=0}^{t} v_s(\tau) \exp[-R_d(t-\tau)]e_1 \, d\tau \quad \text{for} \quad t > 0.
\]

This expression is evaluated at specified receiver positions and therefore we only need to store rows of matrix \( W_d \) that correspond to these receiver locations.

4. NUMERICAL EXPERIMENTS

To illustrate the performance of the extended Krylov method, we solve a multiscale transmission line problem for which an exact solution is known. The transmission line has a length \( L = 0.1 \) m and consists of a single active conductor only. The p.u.l. capacitance and inductance are given by \( c = 0.2 \cdot 10^{-7} \) F/m and \( \ell = 0.05 \cdot 10^{-7} \) H/m, respectively, and the source signature is given by the derivative of a shifted Gaussian with a center frequency of 1.5 GHz (a 1.5 GHz monocycle). This pulse has a pulse width of about one nanosecond. Since we are interested in the late-time behavior of the Krylov approximations, we set the p.u.l. resistance to zero (\( r = 0 \)) and we short-circuit the line at its far end (\( R_{ld} = 0 \) and \( L_{ld} = 0 \)). Finally, we take \( R_s = 10 \) m\( \Omega \). The Gaussian pulse will now bounce back and forth and decays very slowly. The source is switched on at \( t = 0 \) and the length of the time interval of observation is one microsecond. The order of the system matrix is \( n = 20000 \).

Since the observation interval is 1000 times the pulse width of the source, it is difficult to illustrate how the Krylov method approximates the exact response on the complete time interval of observation. We therefore show the computed and exact response on a subinterval only running from \( t = 245 \) ns to \( t = 265 \) ns. In Figure 1, the exact response (solid line) and the Krylov approximation of order \( d = 80 \) with \( k = 40 \) and \( m = 41 \) are shown. We observe that the arrival times are already captured, but the amplitudes are not correct. This is observed over the complete interval of observation. Increasing the order of the approximation to \( d = 120 \) using \( k = 60 \) and \( m = 61 \), we obtain the Krylov approximation as shown in Figure 2 (dashed line). The solid line signifies the

```
Figure 1: The voltage \( V \) at \( z = L/2 \) from \( t = 245 \) ns to \( t = 265 \) ns. The solid line signifies the exact result, the dashed line is the Krylov approximation of order \( d = 80 \) with \( k = 40 \) and \( m = 41 \).
```
Figure 2: The voltage \( V \) at \( z = L/2 \) from \( t = 245 \) ns to \( t = 265 \) ns. The solid line signifies the exact result, the dashed line is the Krylov approximation of order \( d = 120 \) with \( k = 60 \) and \( m = 61 \).

exact response and we observe that now both the arrival times and the amplitudes are modeled correctly. This is also observed on the remaining part of the observation interval. This convergence behavior turns out to be typical. All numerical experiments that we have carried out so far indicate that arrival times are approximated very well on a large time interval of observation already after a few iterations of the extended Krylov method (using \( m = k + 1 \)). Subsequent iterations correct for the amplitudes and the resulting Krylov approximation is accurate on large time intervals of interest after only a small number of iterations.

5. CONCLUSIONS

In this paper we have presented an extended Krylov subspace method for multiscale transient electromagnetic wavefield problems. A basis of the Krylov space \( K^{k,m} \) is generated by iterating with the system matrix and its (group) inverse. More precisely, first \( k \) iterations are carried out with the inverse of the system matrix to obtain the first \( k \) basis vectors, and subsequently the basis vectors related to powers of the system matrix are computed. Having this basis available, we obtain a basis for \( K^{k,m+1} \) by simply carrying out an additional iteration with matrix \( A \). However, it seems that it is not possible to efficiently construct a basis for \( K^{k+1,m} \) starting from an already computed basis of \( K^{k,m} \). A different approach in which this problem is avoided is presented in [4]. Furthermore, in all our numerical work we take \( m = k + 1 \), but this is just a choice. Other relations between \( k \) and \( m \) may lead to faster convergence and in future work we plan to investigate this issue further.

ACKNOWLEDGMENT

The research reported in this paper is supported by the Dutch Technology Foundation (STW). This support is gratefully acknowledged.

REFERENCES