

Reconstruction of 3-D Microwave Images based on a Block-BiCGStab Algorithm

Corentin Friedrich^{1,2}, Sébastien Bourguignon¹, Jérôme Idier¹ and Yves Goussard²

¹ LUNAM Université, École Centrale de Nantes / IRCCyN UMR CNRS 6597, 1 rue de la Noë, 44321 Nantes, France

² École Polytechnique de Montréal, Dept. of Electrical Engineering, C.P. 6079, H3C 3A7, Montréal QC, Canada

E-mail: ¹surname.name@irccyn.ec-nantes.fr, ²yves.goussard@polymtl.ca

Abstract. In 3-D microwave imaging, gradient-based optimization algorithms usually make use of the so-called stabilized version of the biconjugate gradient iterative method (BiCGStab) in order to solve multiple linear systems. We propose to use a block version of BiCGStab to jointly solve the multiple right-hand side linear systems. Illuminations are partitioned in subgroups, which makes the method more efficient. The reconstruction process is studied on realistic simulated data and illustrates the efficiency of the method compared to BiCGStab.

1. Introduction

The development of microwave imaging algorithms has received a lot of interest in the last few years with applications such as biomedical imaging and geoscience [1, 2]. Solving the inverse scattering problem under realistic conditions presents several difficulties, a critical one being the reduction of the computational cost associated with three-dimensional problems. Nonlinear inversion methods usually rely on iterative algorithms to reconstruct the dielectric properties of the unknown object. A regularized data misfit cost function is minimized using a gradient-based optimization procedure, such as the Distorted Born iterative method (DBIM) [3] or the Gauss-Newton method [2, 4]. At each iteration, the computation of a descent direction and of a step size requires the evaluation of the objective function and of its gradient. Such computations rely on the resolution of linear systems equal to the number of illuminations. In practice, exact inversion of such systems becomes computationally prohibitive for large 3-D problems and approximate solutions are obtained with iterative methods, such as the conjugate gradient method [5], the quasi minimal residual (QMR) method [6] or the stabilized version of the biconjugate gradient stabilized (BiCGStab) method [1, 7].

Different methods have been proposed to solve efficiently such multiple systems, such as the marching-on-anything technique [4] or a massive parallelization scheme [2]. Since the linear systems share the same system matrix, block versions of iterative algorithms are appropriate, which simultaneously solve linear systems with multiple right-hand sides. In [8], a block-QMR algorithm was proposed for microwave imaging. In [9], a block version of BiCGStab was developed for random system inversion, which was shown to outperform the Block-QMR algorithm. In [10], we proposed a Partial-Block BiCGStab procedure to solve multiple forward

problems in 3-D microwave imaging, which partitions the sources in several subgroups and uses Block-BiCGStab on each group. In this paper, the Partial-Block BiCGStab algorithm is applied for the computation of the objective function and its gradient at each iteration of the inversion process.

In Section 2, the 3-D formulation is presented. The inverse problem is introduced in Section 3 where the Block-BiCGStab algorithm and the Partial-Block version are detailed. Finally, the computational costs of the different methods are compared on simulated data in Section 4.

2. Formulation of the 3-D Problem

An unknown inhomogeneous object of complex permittivity $\epsilon(\mathbf{r})$ is included in a volume V , where \mathbf{r} denotes the spatial position in V . We assume a homogeneous background medium with complex permittivity ϵ_b and magnetic permeability μ_0 , that is the vacuum permeability. We suppose that all objects and media are nonmagnetic. Successive known harmonic incident electric fields $\mathbf{E}_i^{\text{inc}}(\mathbf{r})$, $i = 1, \dots, N_S$ illuminate the object with angular frequency ω . The scattered fields $\mathbf{E}_i^{\text{scat}}$ are measured at locations \mathbf{r}_j^R , $j = 1, \dots, N_R$. In the following, the time dependence term $e^{-j\omega t}$ is omitted. The observation equation links the data with the total electric field \mathbf{E}_i in the domain V . It reads [11]:

$$\mathbf{E}_i^{\text{scat}}(\mathbf{r}_j^R) = (k_b^2 + \nabla\nabla\cdot) \int_V g(\mathbf{r}_j^R, \mathbf{r}') \chi(\mathbf{r}') \mathbf{E}_i(\mathbf{r}') d\mathbf{r}', \quad \forall j = 1, \dots, N_R \quad (1)$$

where k_b and χ respectively denote the wavenumber $k_b = \omega^2 \mu_0 \epsilon_b$ and the contrast function $\chi(\mathbf{r}) = (\epsilon(\mathbf{r}) - \epsilon_b) / \epsilon_b$. The $\nabla\nabla\cdot$ term represents the gradient of the divergence. The homogeneous Green's function $g(\mathbf{r}, \mathbf{r}')$ is defined as $g(\mathbf{r}, \mathbf{r}') = \exp(jk_b \|\mathbf{r} - \mathbf{r}'\|) / 4\pi \|\mathbf{r} - \mathbf{r}'\|$. Note that the integral in (1) is a convolution with the kernel $g(\mathbf{r}, \mathbf{r}')$. The total field $\mathbf{E}_i(\mathbf{r})$ in the domain V is ruled by the electric field integral equation (EFIE) or domain equation:

$$\mathbf{E}_i(\mathbf{r}) = \mathbf{E}_i^{\text{inc}}(\mathbf{r}) + (k_b^2 + \nabla\nabla\cdot) \int_V g(\mathbf{r}, \mathbf{r}') \chi(\mathbf{r}') \mathbf{E}_i(\mathbf{r}') d\mathbf{r}', \quad \forall \mathbf{r} \in V. \quad (2)$$

In order to obtain a discrete model, we assume that the volume V is a cuboid divided into N cubic voxels of dimensions N_x , N_y and N_z . Let $\mathbf{e}_i^{\text{inc}}$ denote the vector of length $3N$ containing the components of the incident field discretized on each voxel n with center \mathbf{r}_n , $n = 1, \dots, N$. Similarly, let \mathbf{e}_i denote the discretized total electric field. The contrast function $\chi(\mathbf{r})$ is assumed to be constant on each voxel. The vector \mathbf{x} of length N contains the unknown contrast on all voxels. The domain equation (2) is discretized using the method of moments [12], which yields:

$$\mathbf{e}_i = \mathbf{e}_i^{\text{inc}} + \mathbf{G}\mathbf{X}\mathbf{e}_i, \quad (3)$$

where \mathbf{X} is a $3N \times 3N$ diagonal matrix whose diagonal contains three replicas of the vector \mathbf{x} and \mathbf{G} is a $3N \times 3N$ convolution matrix corresponding to the discretization of $(k_b^2 + \nabla\nabla\cdot)g(\mathbf{r}, \mathbf{r}')$. The discretization procedure is detailed in [13]. Eq. (3) can be written as a linear system:

$$\mathbf{e}_i^{\text{inc}} = \mathbf{L}_x \mathbf{e}_i, \quad \text{with } \mathbf{L}_x = \mathbf{I}_{3N} - \mathbf{G}\mathbf{X} \quad (4)$$

where \mathbf{I}_{3N} is the identity matrix of size $3N$. Determining the total field \mathbf{e}_i in (4) corresponds to solving the forward problem.

The observation equation (1) is discretized similarly. The three components of the scattered fields $\mathbf{E}_i^{\text{scat}}$ measured at the N_R receivers are stored in the vector \mathbf{y}_i of size $3N_R$. The discretized observation equation reads

$$\mathbf{y}_i = \mathbf{G}_o \mathbf{X} \mathbf{e}_i + \mathbf{b}. \quad (5)$$

The $N_R \times N$ matrix \mathbf{G}_o is constructed similarly to \mathbf{G} . The vector \mathbf{b} is a perturbation term accounting for noise and model errors. Combining the domain equation (4) and the observation equation (5), the measurements \mathbf{y}_i can be expressed as functions of the contrast \mathbf{x} according to:

$$\mathbf{y}_i = \mathbf{G}_o \mathbf{X} \mathbf{L}_x^{-1} \mathbf{e}_i^{\text{inc}} + \mathbf{b}. \quad (6)$$

3. Three-dimensional reconstruction

3.1. Optimization of a regularized objective function

Solving the inverse problem consists in determining the contrast function of an unknown object embedded in the volume V from measurements of the scattered field for N_S incident fields. Using the discretized model (6), the problem amounts to estimating the discretized contrast \mathbf{x} from data $\mathbf{y}_1, \dots, \mathbf{y}_{N_S}$. This problem is ill-posed, therefore we define the solution as the minimizer of a regularized cost function:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \mathcal{F}(\mathbf{x}), \text{ with } \mathcal{F}(\mathbf{x}) = \mathcal{F}^D(\mathbf{x}) + \lambda \mathcal{R}(\mathbf{x}) \quad (7)$$

where $\mathcal{F}^D(\mathbf{x}) = \sum_{i=1}^{N_S} \|\mathbf{y}_i - \mathbf{G}_o \mathbf{X} \mathbf{L}_x^{-1} \mathbf{e}_i^{\text{inc}}\|^2$ is the quadratic data misfit and $\mathcal{R}(\mathbf{x})$ is a regularization term operating on the differences between adjacent voxels:

$$\mathcal{R}(\mathbf{x}) = \sum_{k,\ell,m} \varphi(\mathbb{X}_{k,\ell,m} - \mathbb{X}_{k-1,\ell,m}) + \varphi(\mathbb{X}_{k,\ell,m} - \mathbb{X}_{k,\ell-1,m}) + \varphi(\mathbb{X}_{k,\ell,m} - \mathbb{X}_{k,\ell,m-1}) \quad (8)$$

where $\{\mathbb{X}_{k,\ell,m}\}_{k,\ell,m}$ represents the 3-D arrangement of \mathbf{x} . We use the “ $\ell_2\ell_1$ ” function $\varphi(u) = \sqrt{\delta^2 + |u|^2}$, which tends to preserve edges in the reconstructed 3D-image. Moreover, for any $\delta \neq 0$, the cost function is continuously differentiable on \mathbb{R}^N and gradient-based optimization methods can be used to solve (7). The parameter λ controls the trade-off between the data misfit and the regularization function. In this paper, parameters λ and δ are set empirically. We consider the local minimization of criterion (7) with the L-BFGS algorithm [14]. Each iteration of such algorithm involves one computation of the gradient and possibly several computations of the objective function for the line search step. The gradient reads:

$$\nabla_{\mathbf{x}} \mathcal{F} = - \sum_{i=1}^{N_S} \text{diag} \{ \mathbf{L}_x^{-1} \mathbf{e}_i^{\text{inc}} \}^\dagger (\overline{\mathbf{L}_x})^{-1} \left[\mathbf{G}_o^\dagger (\mathbf{y}_i - \mathbf{G}_o \mathbf{X} \mathbf{L}_x^{-1} \mathbf{e}_i^{\text{inc}}) \right] + \nabla_{\mathbf{x}} \mathcal{R}, \quad (9)$$

where the symbol \dagger denotes the Hermitian adjoint, the overline denotes the complex conjugation and $\text{diag}\{\mathbf{u}\}$ stands for the diagonal matrix with diagonal \mathbf{u} .

3.2. Block algorithms for linear system resolutions

The evaluation of the cost function \mathcal{F} requires the computation of N_S forward problems $\mathbf{L}_x^{-1} \mathbf{e}_i^{\text{inc}}$, $i = 1, \dots, N_S$. The computation of the gradient also requires the resolution of N_S systems with matrix $\overline{\mathbf{L}_x}$. For realistic 3-D problems, exact inversion of such linear systems is computationally prohibitive: \mathbf{L}_x is a full matrix with $3N \times 3N$ elements. Consequently, the systems are generally solved by iterative methods, such as the BiCGStab algorithm [1, 2], which must be called N_S times for the computation of \mathcal{F} and N_S times for the resolution of systems with normal matrix $\overline{\mathbf{L}_x}$ in (9). The most computationally demanding steps in BiCGStab are the matrix-vector multiplications with the matrix \mathbf{L}_x , which essentially correspond to convolution operations, that can be efficiently implemented by Fast Fourier Transform algorithms.

In this paper, we propose a method to solve all the linear systems jointly. Indeed, the operator \mathbf{L}_x does not depend on the illuminations and the linear systems can be aggregated. The forward problems involved in the computation of both the cost function and its gradient read:

$$\mathbf{E} = \mathbf{L}_x^{-1} \mathbf{E}^{\text{inc}} \quad (10)$$

with $\mathbf{E}^{\text{inc}} = [\mathbf{e}_1^{\text{inc}}, \dots, \mathbf{e}_{N_S}^{\text{inc}}]$ and $\mathbf{E} = [\mathbf{e}_1, \dots, \mathbf{e}_{N_S}]$. The computation of the gradient also requires a second set of system resolutions with matrix $\overline{\mathbf{L}}_{\mathbf{x}}$, that can also be aggregated as:

$$\overline{\mathbf{L}}_{\mathbf{x}}^{-1} \left[\mathbf{G}_o^{(3)\dagger} \left(\mathbf{Y} - \mathbf{G}_o^{(3)} \mathbf{X}^{(3)} \mathbf{L}_{\mathbf{x}}^{-1} \mathbf{E}^{\text{inc}} \right) \right] \quad (11)$$

with $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_{N_R}]$. We propose to use the Block-BiCGStab algorithm, which is a generalization of the BiCGStab algorithm to multiple right-hand side systems [9]. It is described in Algorithm 1 for the joint resolution of the forward problems in (10). The notation $\langle \mathbf{T}, \mathbf{S} \rangle_F$ denotes the Frobenius product between matrices \mathbf{T} and \mathbf{S} , that is, the trace of matrix $\mathbf{T}^\dagger \mathbf{S}$. If $N_S = 1$, Block-BiCGStab reduces to BiCGStab. In the following, the implementation of the computation of the total fields in (10) is detailed. The resolution of (11) is performed similarly.

Algorithm 1 Block-BiCGStab algorithm

- 1: For an initial guess $\mathbf{E}^{(0)}$, $\mathbf{R}^{(0)} = \mathbf{E}^{\text{inc}} - \mathbf{L}_{\mathbf{x}} \mathbf{E}^{(0)}$,
 $\mathbf{P}^{(0)} = \mathbf{R}^{(0)}$
 - 2: Choose an arbitrary $3N \times N_S$ matrix $\tilde{\mathbf{R}}_0$
 - 3: $k = 0$
 - 4: **repeat**
 - 5: $\mathbf{V} = \mathbf{L}_{\mathbf{x}} \mathbf{P}^{(k)}$
 - 6: solve $(\tilde{\mathbf{R}}_0^\dagger \mathbf{V}) \mathbf{A} = \tilde{\mathbf{R}}_0^\dagger \mathbf{R}^{(k)}$
 - 7: $\mathbf{S} = \mathbf{R}^{(k)} - \mathbf{V} \mathbf{A}$
 - 8: $\mathbf{T} = \mathbf{L}_{\mathbf{x}} \mathbf{S}$
 - 9: $\omega = \langle \mathbf{T}, \mathbf{S} \rangle_F / \langle \mathbf{T}, \mathbf{T} \rangle_F$
 - 10: $\mathbf{E}^{(k+1)} = \mathbf{E}^{(k)} + \mathbf{P}^{(k)} \mathbf{A} + \omega \mathbf{S}$
 - 11: $\mathbf{R}^{(k+1)} = \mathbf{S} - \omega \mathbf{T}$
 - 12: solve $(\tilde{\mathbf{R}}_0^\dagger \mathbf{V}) \mathbf{B} = -\tilde{\mathbf{R}}_0^\dagger \mathbf{T}$
 - 13: $\mathbf{P}^{(k+1)} = \mathbf{R}^{(k+1)} + (\mathbf{P}^{(k)} - \omega \mathbf{V}) \mathbf{B}$
 - 14: $k \leftarrow k + 1$
 - 15: **until** $\forall i, \|\mathbf{r}_i^{(k)}\| / \|\mathbf{e}_i^{\text{inc}}\| < \text{tolerance}$, where $\mathbf{r}_i^{(k)}$ is
the i -th column of $\mathbf{R}^{(k)}$
-

Algorithm 2 Partial-Block BiCGStab algorithm

- 1: Build subsets $\mathbf{E}^{\text{inc}(p)}$ for $p = 1 \dots P$ by partitioning
the N_S sources.
 - 2: **for** $p = 1 \dots P$ **do**
 - 3: Run the Block-BiCGStab algorithm on $\mathbf{E}^{\text{inc}(p)}$
and $\mathbf{E}^{(p)}$ instead of \mathbf{E}^{inc} and \mathbf{E} .
 - 4: **end for**
-

The Block-BiCGStab method is an iterative algorithm, where each iteration requires the resolution of two linear systems of size N_S with matrix $\tilde{\mathbf{R}}_0^\dagger \mathbf{V}$ at steps 6 and 12 of Algorithm 1. These systems are relatively small, therefore the corresponding matrices can be computed and system resolutions can be performed exactly. However, during the iterative process, the matrix $\tilde{\mathbf{R}}_0^\dagger \mathbf{V}$ may become ill-conditioned, which leads to instabilities. This may happen in particular when two sources, say with indices i and j , are very close. In that case, the incident fields $\mathbf{e}_i^{\text{inc}}$ and $\mathbf{e}_j^{\text{inc}}$ are highly correlated and this property is propagated into the quantities $\mathbf{R}^{(k)}$ and \mathbf{V} . To overcome the instabilities, we propose a procedure called Partial-Block BiCGStab. To avoid ill-conditioned matrices, the N_S sources are divided into P subgroups such that all sources in one subgroup p are as far as possible from each other. Then, matrix \mathbf{E}^{inc} is split into P submatrices $\mathbf{E}^{\text{inc}(p)}$ in which the columns are weakly correlated. Similarly, matrix \mathbf{E} is split into submatrices $\mathbf{E}^{(p)}$. Then, the Block-BiCGStab is applied independently on the P subproblems. This procedure is summarized in Algorithm 2. For $P = 1$, Partial-Block BiCGStab reduces to Block-BiCGStab. For $P = N_S$, it reduces to the BiCGStab algorithm. **The choice of P is important and some information about its influence are given in [10].** The resolution of the multiple forward problems is faster for Partial-Block BiCGStab than for BiCGStab and is even better for high contrasted objects [10]. Here, the Partial-Block BiCGStab procedure is used to compute the cost function and the gradient as well in order to speed up the reconstruction algorithm using L-BFGS.

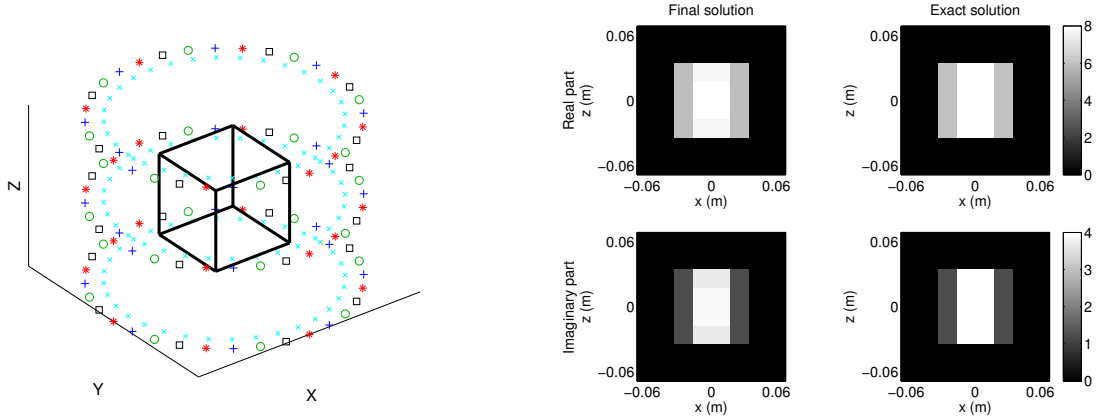


Figure 1. Left: configuration setup with 96 sources and 96 receivers (cyan x symbols) distributed on three coaxial circles. The black cube represents the volume of interest V . The four different symbols ($+$, $*$, \circ , \square) represent the four subgroups of sources. Right: cross-section of the object at $y = 0$. First column: real and imaginary parts of the obtained solution. Second column: true object.

4. Simulation Results

In this section, we evaluate the computation time and the convergence of our method compared to BiCGStab and Block-BiCGStab. First, the resolution of the multiple forward problems are studied. Then, results on the reconstruction process are presented. We consider a simulated experiment presented in Fig. 1 (left) that can be found in the literature, with $N_S = 96$ sources at frequency 2.45 GHz embedded in the air of permittivity $\epsilon_b = 1$. The sources are placed on three coaxial circles at heights $z = -\lambda_b$, 0 and λ_b , and radius $1.5\lambda_b$, where λ_b is the wavelength of the incident fields. On each circle, $N_S/3 = 32$ sources are equally distributed. A lossy cube of size $0.5\lambda_b$ with contrast $\chi_1 = 6 + 1.2j$ is embedded in a cubic volume V of size λ_b . Inside the lossy cube, a cuboid is inserted with square cross-section of $0.25\lambda_b$, height $0.5\lambda_b$ and contrast $\chi_2 = 8 + 4j$. The volume V is discretized into $8 \times 8 \times 8$ voxels. The 96 receivers are placed on three coaxial circles at the same heights as the sources and with radius $1.3\lambda_b$.

Synthetic data are simulated on a grid of $16 \times 16 \times 16$ voxels and zero-mean white Gaussian noise is added with signal-to-noise ratio equal to 20 dB. The reconstruction procedure is implemented with Matlab on a personal computer with 8 Go RAM and eight cores clocked at 3.40 GHz (four hyperthreaded cores). The reconstruction process is parallelized by simply executing loops with Matlab `parfor` instruction. For BiCGStab, the $2N_S$ system resolutions are sent in parallel to the different processors. The Block-BiCGStab algorithm is not easily parallelizable and therefore is not parallelized. For the Partial-Block BiCGStab algorithm, the resolutions of the different subproblems are also sent in parallel to the different processors. For Partial-Block BiCGStab, we divide the set of sources into four subgroups (see Fig. 1 left). **For Block-BiCGStab and Partial-Block BiCGStab, $\tilde{\mathbf{R}}_0$ is set to $\mathbf{L}_x \mathbf{R}^{(0)}$.**

Fig. 2 plots the evolution of the relative residual errors for the resolution of the 96 forward problems using the three algorithms: BiCGStab, Block-BiCGStab and Partial-Block BiCGStab. The tolerance is set to 10^{-6} , which is a strict value. Two initializations are tested: $\mathbf{E}^{(0)} = \mathbf{E}^{\text{inc}}$, and $\mathbf{E}^{(0)}$ close to the exact solution of (10) (which can be computed directly for such reasonable size problem). We note that regardless of the initialization, BiCGStab and Partial-Block BiCGStab converge in nearly the same number of iterations, but Partial-Block BiCGStab is the most efficient. On the contrary, the Block-BiCGStab algorithm converges much faster if it is initialized close to the solution. The computation times for the three resolutions with

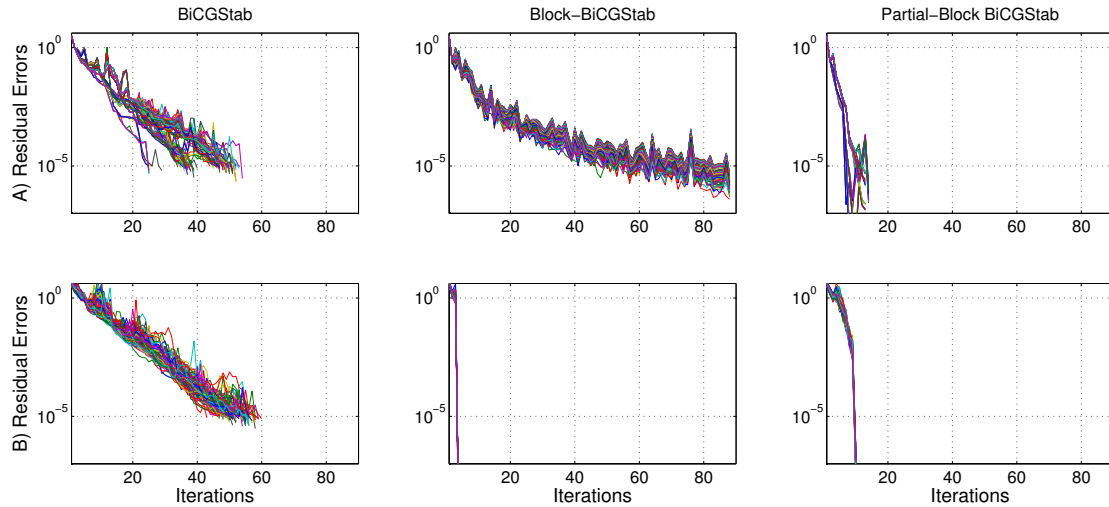


Figure 2. Residual errors for the 96 forward problems as a function of iterations for BiCGStab, Block-BiCGStab and Partial-Block BiCGStab. First row: algorithms are initialized with the incident fields $\mathbf{E}^{(0)} = \mathbf{E}^{\text{inc}}$. Second row: algorithms are initialized close to the solution.

$\mathbf{E}^{(0)} = \mathbf{E}^{\text{inc}}$ are respectively 17 s, 38 s and 6.5 s. With parallelization, the cost of BiCGStab decreases to 4.3 s and that of Partial-Block BiCGStab to 2.2 s.

In the reconstruction process, the three algorithms were implemented to solve (10) and (11) involved in the cost function and its gradient. The hyperparameter λ was set to 10^{-5} and δ to 10^{-2} . At iteration k of the L-BFGS algorithm, the resolutions of (10) and (11) are initialized with the solutions obtained at iteration $k - 1$. Implementations based on the BiCGStab, Block-BiCGStab and Partial-Block BiCGStab algorithms converge respectively in 855 s, 1004 s and 573 s. The tolerance set to 10^{-6} on the residual errors for the three algorithms leads to good accuracy in the resolutions of the linear systems.

A perspective is to solve inverse problems of larger size. We will also study the influence of the tolerance: a weak tolerance generates strong approximations in the system resolutions and could introduce instabilities in the convergence of the reconstruction algorithm.

References

- [1] Zhang Z Q, Liu Q H, Xiao C, Ward E, Ybarra G and Joines W T 2003 *IEEE Trans. Biomed. Eng.* **50** 1180–1189
- [2] Abubakar A, Habashy T M, Pan G and Li M K 2012 *IEEE Trans. Ant. Propag.* **60** 2431–2441 ISSN 0018-926X
- [3] Winters D W, Van Veen B D and Hagness S C 2010 *IEEE Trans. Ant. Propag.* **58** 145–154 ISSN 0018-926X
- [4] De Zaeytijd J, Franchois A, Eyraud C and Geffrin J M 2007 *IEEE Trans. Ant. Propag.* **55** 3279–3292
- [5] Zwamborn P and van den Berg P M 1992 *IEEE Trans. Microwave Theory Tech.* **40** 1757–1766
- [6] Wang C F and Jin J M 1998 *IEEE Trans. Microwave Theory Tech.* **46** 553–558 ISSN 0018-9480
- [7] Van der Vorst H A 1992 *SIAM J. Stat. Sci. Comp.* **13** 631–644
- [8] Fang Q, Meaney P M, Geimer S D, Streltsov A V and Paulsen K D 2004 *IEEE Trans. Medical Imaging* **23** 475–484
- [9] El Guennouni A, Jbilou K and Sadok H 2003 *Electron. Trans. Numer. Anal.* **16** 2
- [10] Friedrich C, Idier J, Bourguignon S and Goussard Y 2015 *Accep. for pub. Proc. 9th European Conf. Antennas Propag.*
- [11] Pastorino M 2010 *Microwave Imaging* (John Wiley & Sons)
- [12] Harrington R F and Harrington J L 1996 *Field Computation by Moment Methods* (Oxford University Press)
- [13] Abubakar A and van den Berg P M 2004 *J.Comput. Phys.* **195** 236–262 ISSN 0021-9991
- [14] Bertsekas D P 1999 *Nonlinear Programming* (Belmont, MA: Athena Scientific)