

# Multigrid Regularized Image Reconstruction for Limited-Data Tomography

Rafał Zdunek\*

*Institute of Telecommunications, Teleinformatics and Acoustics  
Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland  
e-mail: Rafał.Zdunek@pwr.wroc.pl*

(Rec. January 11, 2007)

**Abstract:** Limited-data tomography, to which electromagnetic geotomography belongs, is analyzed in this paper. In this technique, a discrete forward projection model may be expressed by a rank-deficient system of linear equations whose the nullspace is non-trivial. This means that some image components may fall into the nullspace, and hence the minimal-norm least-square solution, to which many image reconstructions methods converge, may be different from the true one. The Algebraic Reconstruction Technique (ART), Simultaneous Iterative Reconstruction Technique (SIRT), or Conjugate Gradients Least Squares (CGLS) are examples of such methods. In this paper, we deal with the question of how to partially recover the missing image components. First, we analyze the advantages of using the iterative Tikhonov regularization and the Maximum A Posteriori (MAP) algorithm with Gibbs prior. Then, we conclude that the missing (nullspace) image components can be partially recovered if the MAP algorithm is implemented through a multigrid technique. The results, which are presented for synthetic noise-free and noisy data, demonstrate the validity of our assumption. The problem of estimating the regularization and scaling parameters in the MAP algorithm is also addressed.

**Key words:** limited-data tomography, multigrid image reconstruction, hyperparameter estimation, electromagnetic geotomography

## I. INTRODUCTION

In limited-data tomography, the angular range of rays along which an object is probed is limited. This is the case, for example, in electromagnetic geotomography [1-6] that belongs to a class of borehole imaging techniques. Synthetic data from this type of tomography are used in our experiments.

Assuming some field simplifications, a forward projection model, which is a limited data Radon transform, is linear. The limitation of the angular range of rays makes the model underdetermined, and consequently an ill-posed problem has to be solved. In a discrete form of the model the equivalent system of linear equations is rank-deficient and ill-conditioned. This implies a nullspace of the system is non-trivial, and the minimal-norm least-square solution ( $x_{LS}$ ) may not be the true solution. In practice, this means that many iterative image reconstruction methods such as: Algebraic Reconstruction Technique (ART), Simultaneous Iterative Reconstruction Technique (SIRT), Conjugate Gradients Least Squares (CGLS), or LSQR, [1, 7-10] which

are convergent to  $x_{LS}$ , reconstruct only a more or less degraded version of the true image. Also some direct methods, such as the Truncated Singular Value Decomposition (TSVD) or Tikhonov regularization (a standard form when a discrete regularization operator is expressed by an identity matrix) [7, 9] give approximative solutions to  $x_{LS}$ . Thus, some image components, which fall into a nullspace, cannot be recovered from data without using additional information on the true solution.

There are many possibilities of tackling this problem. One of them is to add some desired image components from a nullspace to  $x_{LS}$  [11]. A numerical study of a nullspace in inverse problems of borehole tomography was thoroughly investigated by e.g. Koltracht et al. [12]. Nevertheless, there are no clear rules how to properly combine the vectors spanning a nullspace to form a missing part of the solution.

Another approach for making the reconstruction more efficient is to make the dimension of a nullspace smaller by changing a layout of measurement points. A nullspace

\*Temporary address: Laboratory for Advanced Brain Signal Processing, Brain Science Institute, RIKEN, 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan, e-mail: zdunek@brain.riken.jp

substantially contracts if an examined object is accessible for taking measurements from three edges at least [12]. This was confirmed in practice by Sanny and Sassa [13].

In a statistical approach, the missing components of the image can be inferred from, e.g. a distribution of pixel values in the images reconstructed in preceding iterations. This is the One-Step-Late (OSL) technique which is commonly applied in the Maximum A Posteriori (MAP) algorithms (see, e.g. [14-17]), whereby regularization for very noisy data is efficient. The inference may be performed on the basis of interactions between pixels in a given neighborhood. The information about roughness of the image is then incorporated into a probabilistic inverse model under the form of prior  $p(x)$  (e.g. the Gibbs prior or an entropy function). In our considerations, we use the MAP algorithm with the Gibbs prior, which was proposed, among others, by Geman and McClure [14], Green [15], Hebert and Leahy [16].

The aim of this paper is to propose a technique for better recovering the missing (nullspace) image components in limited-data tomography. Such components cannot be recovered with typical (well-known) image reconstruction methods. In the paper, we propose the technique that combines the benefits from the above-mentioned last two approaches to pursue the goal. Our novel approach uses a multigrid implementation of the MAP algorithm with the Gibbs prior. The multigrid technique has been widely discussed in the context of its application to tomographic image reconstruction, but to the best knowledge of the author, never as the technique for pursuing such a goal. This is the main novelty of this paper.

A multigrid implementation of the Expectation-Maximization (EM) algorithm in Position Emission Tomography (PET) was proposed by Ranganath et al. [18]. In this technique, the whole iterative process is divided into sub-processes (grids) in which the image resolution increases successively. The first sub-process starts from any initial guess. In the next sub-process, the iterations start from the final solution of the preceding sub-process, extrapolated to a higher resolution. This technique substantially decreases the overall computational cost because each successive sub-process has a faster convergence rate.

Pan and Yagle [19] presented a numerical study of some iterative image reconstruction algorithms in multigrid implementations applied to PET and SPECT (Single Photon Emission Computed Tomography). Their analysis is limited to two sub-processes: a coarse grid (low resolution) and a fine grid (high resolution). They concluded that multigrid reconstruction is efficient for images with local smoothness. Otherwise, it may slow down a convergence rate.

The results obtained by Pan and Yagle [19] have motivated our present study. However, we used a multigrid

technique not only for improving a convergence rate of reconstruction but also for recovering the missing components from a nullspace. Preliminary experiments showed that the rank of the discrete linear forward model increases when the coarse grid has a denser ray-coverage [20]. Thus changing the measurement setup, the coarse grid's high-frequency components can be recovered better, and then, exported to a fine grid. However, to make the reconstruction from noisy data efficient, the standard regularization (basic Tikhonov regularization with an identity matrix as a regularization operator), which filters out some high-frequency components from the solution, cannot be applied here. Thus we suggest using the MAP algorithm with Gibbs prior for the reconstruction.

The remainder of this paper is organized as follows. In the next section, we present the mathematical background to our problem. In the third section, the Iterative Tikhonov Regularization (ITR) and the MAP algorithm with Gibbs prior are characterized. Some simplified tools for estimation of the parameters in this algorithm are proposed in Section IV. In the simulations, synthetic noise-free and noisy data generated from an original image were used to test our propositions. The results are discussed in Section V. Finally, the short summary is presented.

## II. BACKGROUND

In electromagnetic geotomography, the discrete linear forward projection model is described by the following system of linear equations:

$$Ax = b; \quad (1)$$

where  $A = [a_{ij}] \in \mathbb{R}^{M \times N}$  is a coefficient matrix,  $x = [x_1, \dots, x_N]^T \in \mathbb{R}^N$  is an image vector,  $b = [b_1, \dots, b_M]^T \in \mathbb{R}^M$  is a data vector,  $M$  is a number of rays,  $N$  is a number of pixels, and  $M \geq N$ . Element  $x_j$  is the attenuation coefficient of the  $j$ -th pixel,  $a_{ij}$  is the length of the  $i$ -th ray in the  $j$ -th pixel, and  $b_i$  is total attenuation measured along the  $i$ -th ray.

For a numerical study of the problem, the spectral properties of matrix  $A$  are essential. So let  $A$  have the following Singular Value Decomposition (SVD) [7]:

$$A = U \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^T = \sum_{i=1}^N u_i \sigma_i v_i^T, \quad (2)$$

where

$$U = [u_1, \dots, u_{r(A)}, u_{r(A)+1}, \dots, u_M] \in \mathbb{R}^{M \times M}, \quad (3)$$

$$V = [v_1, \dots, v_{r(A)}, v_{r(A)+1}, \dots, v_M] \in \mathbb{R}^{N \times N}, \quad (4)$$

are matrices with orthonormal columns, where  $r(A)$  stands for the rank of  $A$ . Diagonal matrix  $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_{r(A)}, 0, \dots, 0) \in \mathbb{R}^{N \times N}$  has  $r(A)$  singular values which are nonnegative and appear in a nonincreasing order. The reciprocals of singular values constitute a spectrum of  $A$  with high-frequency components corresponding to low singular values and low-frequency components corresponding to high singular values. Vectors  $\{u_1, \dots, u_{r(A)}\}$  and  $\{v_{r(A)+1}, \dots, v_N\}$  span the orthogonal basis of range  $R(A)$  and nullspace  $N(A)$ , respectively; whereas  $\{v_1, \dots, v_{r(A)}\}$  and  $\{u_{r(A)+1}, \dots, u_M\}$  constitute range  $R(A^T)$  and nullspace  $N(A^T)$  of  $A^T$ . This holds for  $\mathbb{R}^M = R(A) + N(A^T)$  and  $\mathbb{R}^N = R(A^T) + N(A)$ .

Let

$$\text{LSS}(A; b) = \left\{ x \in \mathbb{R}^N : \min_x \|Ax - b\|_2 \right\} \quad (5)$$

be the set of all the least-squares solutions to (1) and let  $x_{LS}$  be the solution of minimal-norm. Thus:

$$\text{LSS}(A; b) = x_{LS} + N(A). \quad (6)$$

If  $A$  is a full rank matrix, then  $N(A)$  has only the zero-value vector, and if system (1) is consistent ( $b \in R(A)$ ),  $x_{LS}$  is its true solution. In this case, many methods, such as ART (based on Kaczmarz method), SIRT (based on Landweber iterations), CGLS, LSQR, Tikhonov regularization (basic form), and TSVD [7, 9, 10] converge to  $x_{LS}$ , provided that the initial guess ( $x_0$ ) has no components in  $N(A)$ . This explains the fact why the methods work so well in some tomographic applications [19, 18] with consistent systems of the full rank. In real applications, data are corrupted with noise and the systems of linear equations are inconsistent ( $b \notin R(A)$ ). In this case, the methods give approximations which are in a certain distance from  $x_{LS}$ , and the distance is more or less ‘‘controllable’’, usually by regularization. A numerical analysis of the distance for ART-like algorithms is presented in [21].

In limited-data tomography,  $A$  is usually a rank-deficient matrix, that is,  $r(A) < \min\{M, N\}$  and so  $N(A)$  is non-trivial. For any  $x$  we have  $x = x_r + x_n$ , where  $x_r \in R(A^T)$  and  $x_n \in N(A)$ . The forward projection of  $x$  onto  $\mathbb{R}^M$  is as follows:  $Ax = A(x_r + x_n) = b_r + 0 = b_r$ , where  $b_r \in R(A)$ .

This proves that we cannot recover all the image components from data  $b$  without using additional information on the true solution.

### III. IMAGE RECONSTRUCTION METHODS

There are many methods for solving systems of linear equations, which are commonly applied to image reconstruction in tomography. However, we focus only on two

groups of methods: the ones which converge to  $x_{LS}$  and the others that give a different solution e.g. some statistical methods. Among the methods of the first group, we decided to consider only the Iterative Tikhonov Regularization (ITR) which is a particular case of the generalized Landweber iterations [7, 9, 22]. Its benefits are a high convergence rate, a moderate computational cost, and a good control over noisy perturbation – efficient regularization. Nevertheless, the results presented in [23] show that this method loses its robustness for noisy data in electromagnetic geotomography. Thus we present here the results obtained by the ITR only for a comparison. In the present study, the stress is put on one of the statistical methods that works very well in tomographic applications [24, 25], particularly in SPECT and PET, where data are very noisy. This is the MAP method with the Gibbs prior.

#### III.1. Iterative Tikhonov Regularization

The iterative formula of the ITR method is as follows:

$$x_{k+1} = x_k + \left( A^T A + \alpha^2 I_N \right)^{-1} A^T (b - Ax_k), \quad (7)$$

where  $x_k$  is the solution after the  $k$ -th iteration and  $I_N \in \mathbb{R}^{N \times N}$  is an identity matrix. The spectral filtration in (7) is additionally controlled by tuning regularization parameter  $\alpha^2$ .

For noisy data we have:  $b = b_r + p$ , where  $b_r \in R(A)$  and perturbations  $p = p_r + p_n$ , where  $p_r \in R(A)$  and  $p_n \in N(A^T)$ . Perturbation  $p_n$  has no effect on the image. For the initial guess the following decomposition is true:  $x_0 = x_0^{(r)} + x_0^{(n)}$ , where  $x_0^{(r)} \in R(A^T)$  and  $x_0^{(n)} \in N(A)$ . Considering (2), iterations (7) can be expressed in the following form:

$$\begin{aligned} x_{k+1} = & x_0^{(n)} + \sum_{i=1}^{r(A)} f_i \frac{u_i^T b_r}{\sigma_i} v_i + \sum_{i=1}^{r(A)} \frac{f_i}{\sigma_i} (u_i^T p_r) v_i + \\ & + \sum_{i=1}^N (1 - f_i) (v_i^T x_0^{(r)}) v_i, \end{aligned} \quad (8)$$

where  $f_i$  are filter factors defined as:

$$f_i = 1 - \left( 1 - \frac{\sigma_i^2}{\sigma_i^2 + \alpha^2} \right)^k. \quad (9)$$

Factors  $f_i$ ,  $f_i / \sigma_i$  and  $(1 - f_i)$  denote data, noise and initial guess gain, respectively.

Note that the filter factors in (9) depend not only on the regularization parameter  $\alpha$  and the singular values  $\sigma_i$  but also on the number of iterations  $k$ . This implies that the regularization can be controlled by both parameters ( $\alpha$  and

$k$ ), however, the control over the number of iterations, so-called the truncation of iterations, is very difficult to perform, so we restrict only to the regularization with the parameter  $\alpha$ .

Considering the behavior of singular values of  $A$ , factors  $\{f_i\}$  form the magnitude characteristics of a low-pass filter while factors  $\{(1-f_i)\}$  behave inversely [19]. We can then conclude that using a multigrid implementation of the ITR method, only high-frequency components of a coarse grid can be imported to a fine grid. This implies that a fine grid's image can be enhanced by some missing components from  $N(A)$ , provided that they can be recovered in reconstruction of a coarse grid's image. Unfortunately, for noisy data, a higher value of  $\alpha$  must be set to minimize perturbation errors (e.g. [9]) and the spectrum from which the image is recovered is strongly limited. In such a case, high frequency components in a coarse grid cannot be recovered with the ITR method. This explains the fact why we obtained such poor results in [23] for noisy data. Nevertheless, the idea presented there has become the motivation for the research in the present study.

### III.2. MAP Algorithm with Gibbs Prior

The MAP algorithm, which is discussed in the present paper, uses the Expectation-Maximization (EM) procedure, first proposed by Dempster et al. [26], for maximization of a likelihood function. Shepp and Vardi [24] were leaders in using the EM approach to image reconstruction in emission tomography. They maximized the likelihood function expressed in terms of a joint probability of independent Poisson distributions:

$$p(b|x) = \prod_{i=1}^M \frac{\left( \sum_{j=1}^N a_{ij} x_j \right)^{b_i}}{b_i!} \exp \left\{ - \sum_{j=1}^N a_{ij} x_j \right\}. \quad (10)$$

As the result the following EM algorithm was derived:

$$x_j^{(k+1)} = \frac{x_j^{(k)}}{\sum_{i=1}^M a_{ij}} \sum_{i=1}^M \frac{a_{ij} b_i}{\sum_{j=1}^N a_{ij} x_j^{(k)}}. \quad (11)$$

In transmission tomography, algorithm (11) was first applied by Lange and Carson [25]. Nevertheless, the EM algorithm does not recover image components from  $N(A)$ , and moreover, it can be easily shown that it converges to  $x_{LS}$  for a Gaussian likelihood function.

Additional information in a form of some prior  $p(x)$  can be easily added to likelihood function (10) employing the

Bayes' theorem. Thus using the EM procedure, posterior function  $p(x|b)$  is maximized. Assuming  $p(x)$  is defined by the Gibbs statistics

$$p(x) = \frac{\exp\{-\beta U(x)\}}{\int_x \exp\{-\beta U(x)\} dx}, \quad (12)$$

where  $U(x)$  is a total energy function, the final form of the MAP algorithm with Gibbs prior [14, 16, 15] is as follows:

$$x_j^{(k+1)} = \frac{x_j^{(k)} c_j^{(k)}}{\sum_{i=1}^M a_{ij} + \beta \frac{\partial}{\partial x_j} U(x) \Big|_{x_j=x_j^{(k)}}}, \quad (13)$$

where

$$c_j^{(k)} = \sum_{i=1}^M \frac{a_{ij} b_i}{\sum_{j=1}^N a_{ij} x_j^{(k)}},$$

and  $\beta$  is a regularization parameter. The total energy function usually has the form:

$$U(x) = \sum_j \sum_{n \in N_j} w_{j,n} V(x_j, x_n, \delta) \quad (14)$$

where  $N_j$  is the nearest neighborhood of pixel  $j$ ,  $w_{j,n}$  is a weighting factor between pixels  $j$  and  $n$ , and  $V(x_j, x_n, \delta)$  is a clique energy dependent on scaling factor  $\delta$ .

Weighting factors  $\{w_{j,n}\}$  are usually defined according to the Markov Random Field (MRF), however, taking into account a strong limitation of an angular range of rays in borehole tomography, we took the same values of  $\{w_{j,n}\}$  as the ones experimentally determined in [20]. Thus, we set

$$w_{j,n} = \left\{ 3; 1/3; 1/\sqrt{3^2 + (1/3)^2} \right\}$$

for pixels adjacent along a horizontal, vertical and diagonal line, respectively.

There are various suggestions for defining function  $V(x_j, x_n, \delta)$ . Some of them are discussed by Lalush and Tsui [27] in the context of application to SPECT. An analytical study of the convergence properties of the MAP algorithm with respect to general properties of clique energies was presented by Lange [17]. Considering the results presented in [20], we decided to select only the following Green's proposal [15] as the most effective in our application:

$$V(x_j, x_n, \delta) = \delta \ln \left[ \cosh \left( \frac{x_j - x_n}{\delta} \right) \right]. \quad (15)$$

Differentiating (15) with respect to  $x_j$ , we obtained the simplified function

$$\frac{\partial}{\partial x_j} U(x) = \sum_{n \in N_j} w_{j,n} \tanh\left(\frac{x_j - x_n}{\delta}\right)$$

which was then directly applied to (13).

#### IV. ESTIMATION OF PARAMETERS

To make reconstruction efficient, regularization parameters must be estimated prior to reconstruction. For regularization methods convergent to  $x_{LS}$  estimation of the parameter is not so difficult because it leads to finding an optimum point in which regularization errors and perturbations errors are well balanced. This point corresponds to the corner on the L-curve or it is the minimum of the Generalized Cross-Validation (GCV) criterion [7, 9]. Parameter  $\alpha$  in (7) can be estimated only for the first iteration, thus we can directly employ  $V(\mu)$  [7, p212] as for the Tikhonov regularization.

For the MAP algorithm with Gibbs prior, the problem is much more complicated since the two parameters must be estimated and the information about the norm of the solution is not so useful. This is because the algorithm converges to the solution different than  $x_{LS}$ . For rank-deficient systems, parameter-choice criteria based only on measurements of the 2-norm of a residual vector are not sufficient, and hence, the Ordinary Cross-Validation (OCV) technique is useless in our application. Thus only statistical tools, which maximize a marginal likelihood, seem to be efficient here. There are many statistical methods for estimating parameters  $\beta$  and  $\delta$ , but we mention only two, that seem to be generalized versions.

Higdon et al. [28] presented the method for fully Bayesian estimation of Gibbs hyperparameters (parameter  $\delta$  in (13)). In this method, the relative values of Gibbs partition functions (the denominator in (12)) are estimated through Markov Chain Monte Carlo (MCMC) sampling, and then, these values are used to sample from joint posterior distributions on the image. Their results show that the method works very well in emission tomography, however, the computational cost is rather high.

Zhou et al. [29] analyzed another approach. They pointed out that the optimum value of  $\beta$  balances the expectations with respect to the prior and posterior densities, that is, it holds  $\mathcal{E}[U(x)|y, \beta] = \mathcal{E}[U(x)|\beta]$ . However, the equation could not be solved so easily and such numerical tools like an adaptive quadrature method for numerical computation of integrals and the Newton-Raphson algorithm were used. This approach also requires a high computational cost.

Since both above mentioned methods are very computationally expensive, we used in the simulations the parameters which were roughly approximated by very simplified computations. Instead of maximizing marginal likelihood function  $p(y|\delta)$ , we compute the joint pseudo-prior density

$$p(x|\delta) = \prod_{j=1}^N p(x_j|x_n, n \in N_j) \quad (16)$$

$$\propto \exp\{-\beta U(x)\},$$

where the partition function (the denominator in (12)) is canceled. The motivation for such a choice was the Maximum Pseudo-Likelihood (MPL) function presented by Zhou et al. [29] and used for computation of  $\beta$ . In our experiments, we choose such a value of  $\delta$  which has the highest impact on the total energy function  $U(x)$  in (14) for  $\beta=1$ . Then, the influence of  $U(x)$  on the image is reduced by tuning the regularization parameter  $\beta$ . Thus the optimum value of  $\delta$  is determined by the minimum of the joint pseudo-prior density (16) for  $\beta=1$ .

Parameter  $\beta$  was estimated with respect to the curve of function  $U(x)$  in (14). Our experiments show that its minimum corresponds to the optimum value of  $\beta$ . This happens because  $U(x)$  measures roughness in the image. For low values of  $\beta$ , the solution is dominated by likelihood function (10) and this implies that the roughness in the solution tends to rise in case of noisy data. For high value of  $\beta$ , the prior dominates in the solution and the parasite effects of oversmoothing may take place. Although  $U(x)$  may be small in the preceding iteration, a high value of  $\beta$  excessively emphasizes the roughness, and in a current iteration the image may be less smooth.

#### V. RESULTS

The results were assessed by means of quality and quantity criteria. The former concerns displaying the solution on a 2D patch plot. The later measures the following relative error:  $\|x - x_{\text{exact}}\|_2 / \|x_{\text{exact}}\|_2$ , where  $x$  is the reconstructed solution and  $x_{\text{exact}}$  is referred to as an original image.

Synthetic data were generated for the original image shown in Fig. 1 (left). Its resolution is  $32 \times 32$  pixels and hence a total number of pixels is 1024. The smooth-edge objects in the image imitate possibly real inhomogeneities in an examined area, e.g. voids, cracks, etc. in rocks. The smoothness is due to the assumption that the density to be reconstructed (e.g. attenuation coefficient, phase shift) does not change abruptly. The sharp-edge objects are useful only in simulations to show interactions between adjacent pix-

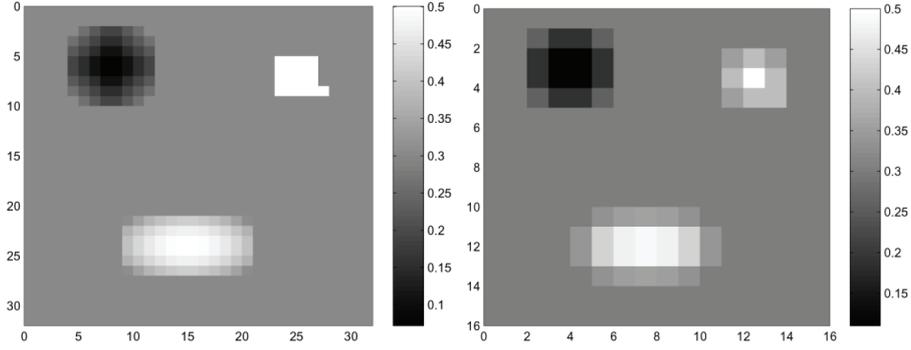


Fig. 1. Original images with resolution of: a fine grid (left), a coarse grid (right)

els. The image was considered as a fine grid. To monitor a relative error of a coarse grid, the fine grid was transformed into an image of  $16 \times 16$  pixels, which is illustrated in Fig. 1 (right). A total number of pixels in a coarse grid is 256. The transformation was performed by averaging a square of four pixels.

Let us assume that the original image is associated with a typical electromagnetic geotomographic area which is a 32 m square, thus a pixel is  $1 \times 1$  m or  $2 \times 2$  m for a fine grid or a coarse grid, respectively. The area is probed along multiple straight-line rays that are traced between transmitter points regularly spaced along one borehole and receiver points regularly spaced along the other borehole. Let us also consider three measurement setups which differ in the density of ray-coverage. There are 16, 32 and 64 transmitter and receiver points for the first, second and third measurement setup, respectively. Thus the corresponding total numbers of rays are 256, 1024 and 4096. The coefficient matrices associated with the coarse grid are denoted as follows:  $A_{\text{normal}}^{(1)} \in \mathbb{R}^{256 \times 256}$ ,  $A^{(1)} \in \mathbb{R}^{1024 \times 256}$  and  $A_{\text{dense}}^{(1)} \in \mathbb{R}^{4096 \times 256}$ . For the fine grid only the second measurement setup was adopted and  $A^{(2)} \in \mathbb{R}^{1024 \times 1024}$  is a coefficient matrix associated with it.

The singular values of the above coefficient matrices associated with the coarse grid are plotted in Fig. 2. Singular values of rank-deficient problems [9] can be separated into two groups: large singular values and very small ones. Increasing the density of ray-coverage, we observe that the former cluster is more and more numerous. This implies that the numerical rank of the system (1) increases and its nullspace contracts. Thus the second approach for the treatment of a nullspace discussed in the Introduction is met. We can also conclude that for practical reasons 1024 rays are sufficient for a considerable improvement in the reconstruction of the missing components, especially as the difference in the rank of matrices  $A^{(1)}$  and  $A_{\text{dense}}^{(1)}$  is slight.

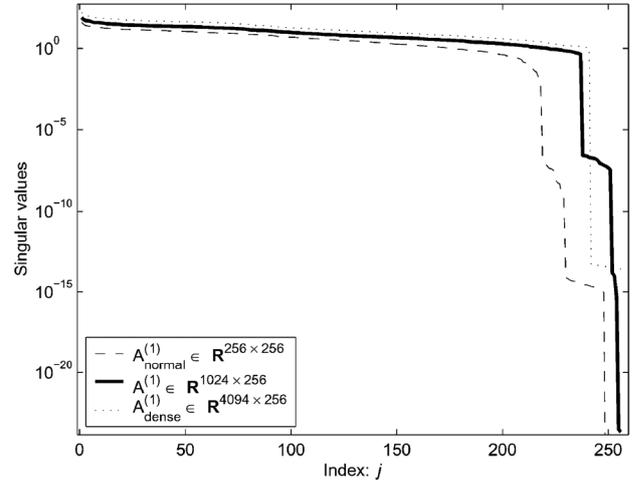


Fig. 2. Singular values of the coefficient matrices with various ray-coverages

Figure 3 presents solutions  $x_{LS}$  which were computed as follows:

$$x_{LS} = V_{\text{range}} V_{\text{range}}^T x_{\text{exact}}, \quad (17)$$

where  $v_{\text{range}} = [v_1, \dots, v_{r(A)}] \in \mathbb{R}^{N \times r(A)}$  for corresponding matrix  $A$ . Figures 3 (left), (middle) and (right) are associated with matrices  $A_{\text{normal}}^{(1)}$ ,  $A^{(1)}$  and  $A^{(2)}$ , respectively. There is a big difference between the figures referred to as (left) and (middle). Figure 3 (middle) is nearly the same as the original one (Fig. 1 (right)). This proves the efficiency of using the second measurement setup.

In the experiments, noisy data were generated as follows:  $b_{\text{noisy}} = b_{\text{exact}} + N(\mu_0, \sigma^2)$ , where  $b_{\text{exact}} = Ax_{\text{exact}} \in R(A)$ ,  $N(\mu_0, \sigma^2)$  is a Gaussian noise with mean  $\mu_0 = 0$  and variance  $\sigma^2$  chosen in this way to obtain noisy data with a desired value of the Signal-to-Noise Ratio (SNR), where  $\text{SNR} = 20 \log(\|b_{\text{exact}}\|_2 / \|b_{\text{noisy}} - b_{\text{exact}}\|_2)$ . We used the noisy data for which  $\text{SNR} = 25$  dB and  $\text{SNR} = 30$  dB.

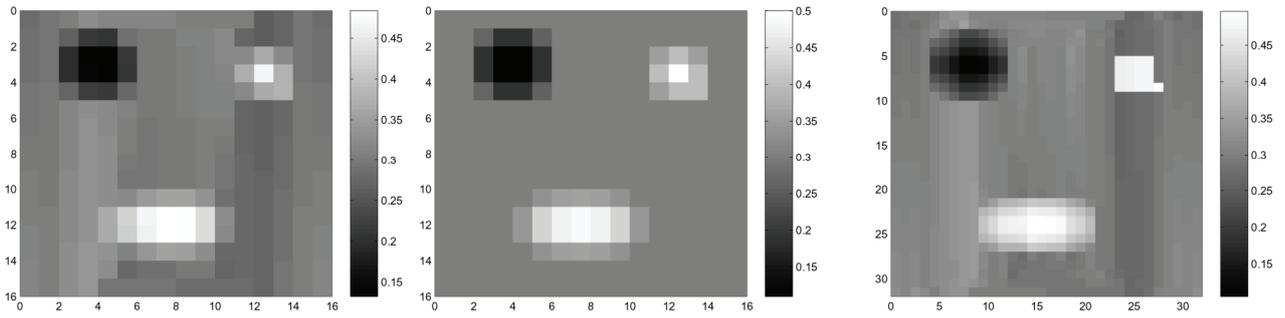


Fig. 3. Projections of the original images onto  $R(A^T)$  for:  $A_{\text{normal}}^{(1)} \in \mathbb{R}^{256 \times 256}$  (left),  $A^{(1)} \in \mathbb{R}^{1024 \times 256}$  (middle);  $A^{(2)} \in \mathbb{R}^{1024 \times 1024}$  (right)

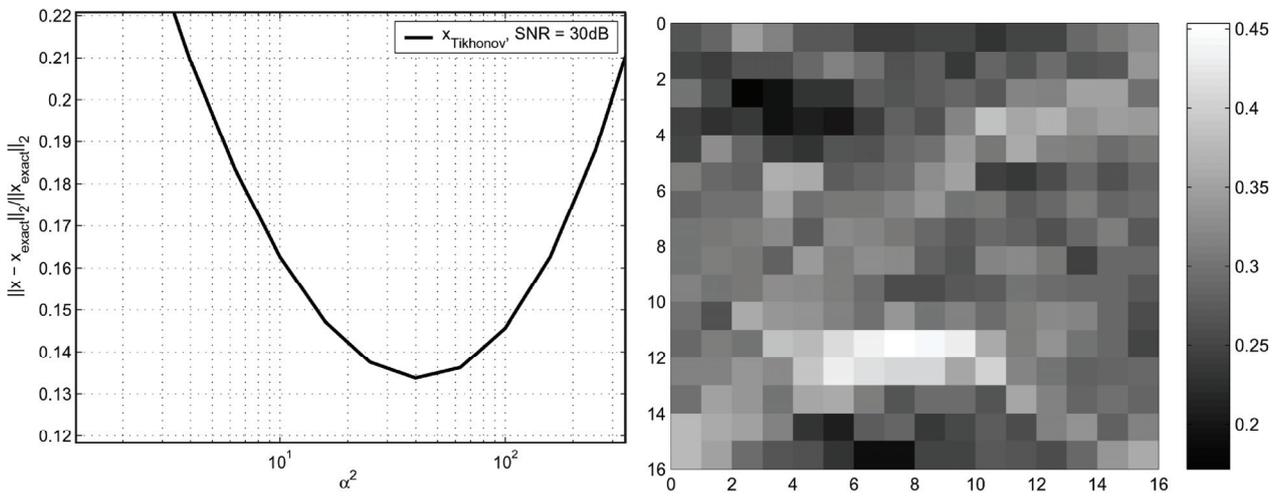


Fig. 4. Tikhonov regularization: relative error versus  $\alpha^2$  for noisy data with SNR = 30 dB (left); the image reconstructed for  $\alpha^2 = 40$  and SNR = 30 dB (right)

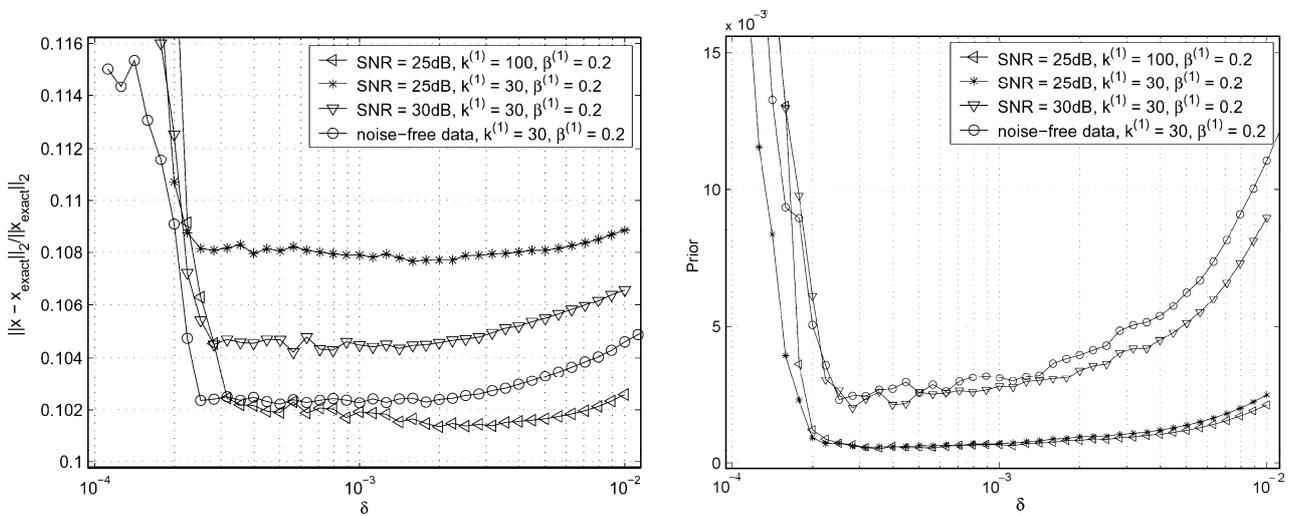


Fig. 5. Estimation of  $\delta$  with: relative error (left); prior probability (right)

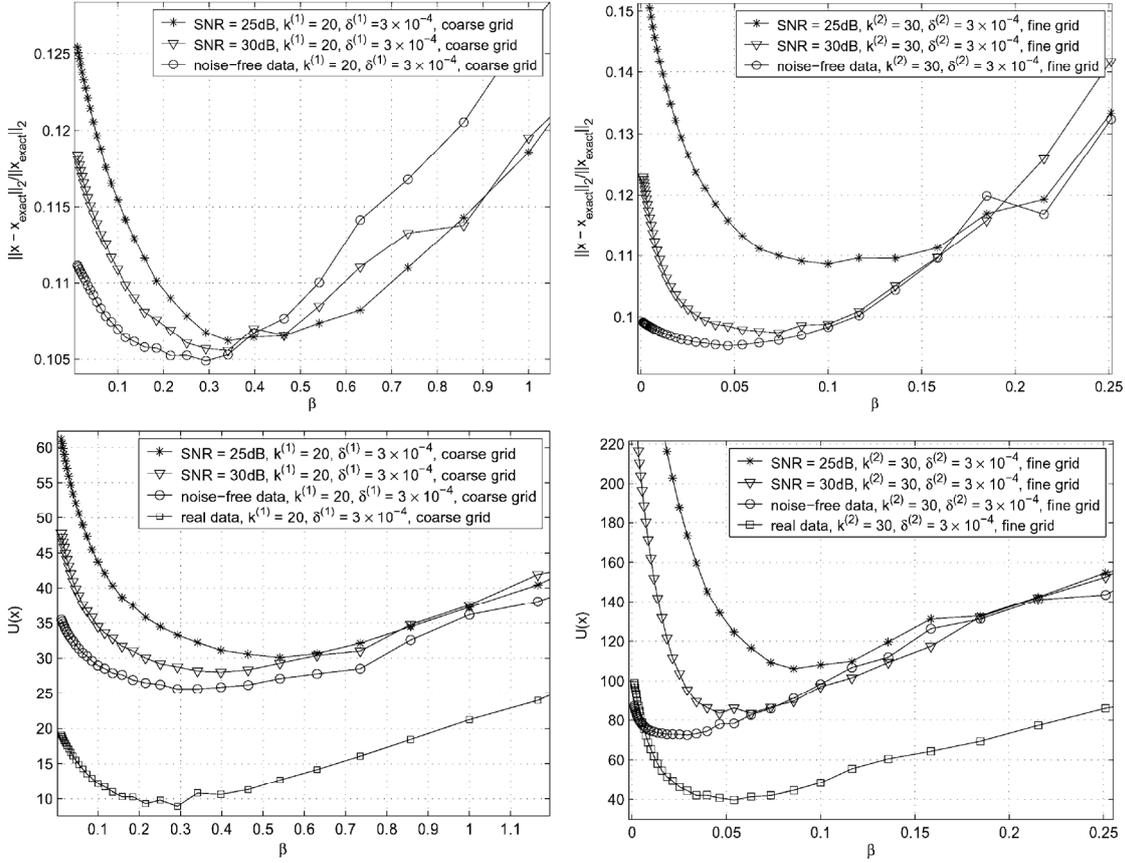


Fig. 6. Estimations of  $\beta$ : relative errors (top row);  $U(x)$  (bottom row); for a coarse grid (left); for a fine grid (right)

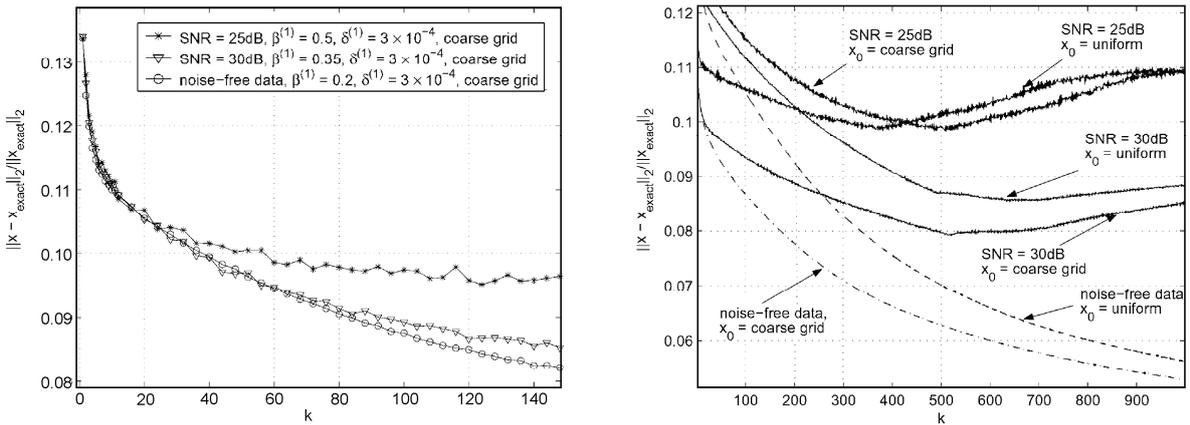


Fig. 7. Relative errors for: a coarse grid (left); a fine grid (right)

Figure 4 illustrates the results obtained with the Tikhonov regularization ( $x_0 = 0$  and  $k = 1$  in (7)). Parameter  $\alpha^2 = 40$ , which minimizes the relative error (Fig. 4 (left))

for the noisy data with SNR = 30 dB, was used for the reconstruction of the image shown in Fig. 4 (right). We checked that for more iterations, a higher value of  $\alpha^2$

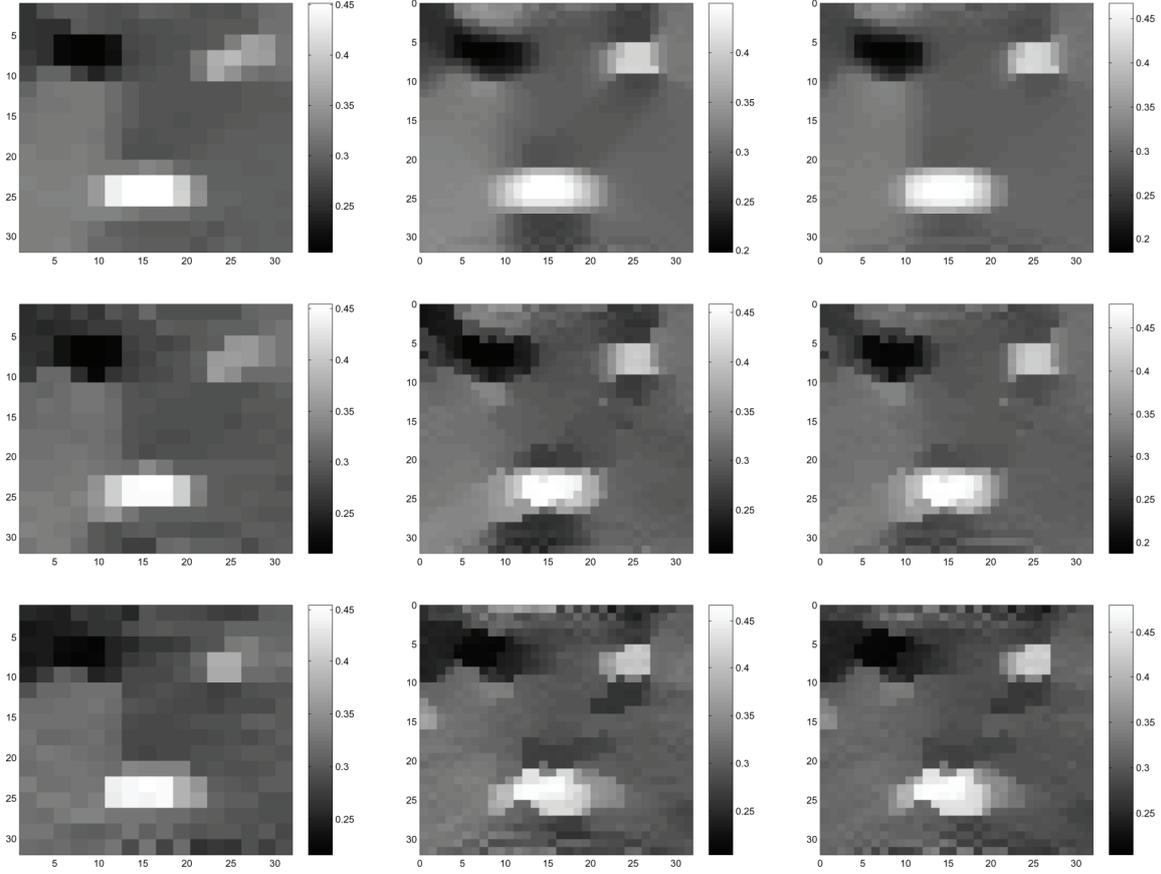


Fig. 8. Images reconstructed within 100 iterations from: noise-free data (top row); noisy data with SNR = 30 dB (middle row); noisy data with SNR = 25 dB (bottom row); a coarse grid (left column); a single grid with resolution of a fine grid (middle column); a fine grid (right column)

should be taken, but the final image remained almost unchanged. Thus this is the best result we could obtain with the standard Tikhonov regularization. However, note that we use only a basic form of the Tikhonov regularization, i.e. where the regularization operator is defined by an identity matrix. In general, the regularization term can be expressed, e.g. by a Total Variation (TV) functional which may include the information on local smoothness. In this case, the generalized Tikhonov regularization may give a competitive solution to the MAP algorithm with Gibbs prior.

The results reconstructed with the MAP algorithm with Gibbs prior are presented in Figs. 5-8. The superscript (1) or (2) refers to the coarse or fine grid, respectively. The initial guess for the coarse grid was defined as  $x_0 = \zeta e_N$ , where  $\zeta = (\sum_i b_i) / (\sum_i \sum_j a_{i,j})$  and  $e_N = [1, \dots, 1]^T \in \mathbb{R}^N$ . We checked that  $x_0 \in R(A)$ . Let us label such an initial guess as “uniform”. The reconstructions of the fine grid started from either the extrapolated final solution of a coarse grid or the uniform initial guess (defined as above). In case of the later, a single grid only was taken into consideration.

Figure 5 refers to the estimation of  $\delta$  for the coarse grid. The optimal parameters determined by means of the relative error criterion (Fig. 5 (left)) correspond to the parameters inferred from the plots of prior probability (16) illustrated in Fig. 5 (right). The plots show that the optimum does not depend considerably on the SNR and a number of iterations, therefore, the same value of  $\delta$  was used for the fine grid.

The curves of the relative error versus  $\beta$  are delineated in Fig. 6 (top) for both grids. The corresponding curves of  $U(x)$  are plotted in Fig. 6 (bottom). In this case, we also obtained a very good agreement in matching the respective minima of the curves for all kinds of data.

For parameters  $\delta$  and  $\beta$  determined in this way, the relative errors versus iterations are presented in Fig. 7 for the coarse grid (left) and the fine grid (right). The images reconstructed within 100 iterations are shown in Fig. 8 in the following layout: the coarse grid's images are in the left column, the single grid's images with the resolution of the fine grid (for “ $x_0 = \text{uniform}$ ”) are in the middle column and

the fine grid's images are in the right column; the images reconstructed from noise-free, less noisy, and more noisy data are presented in the first, second, and third row.

All the algorithms were implemented in MATLAB with using vectorization as much as possible. We assess the computational cost taking into consideration elementary flops. The computational cost of a single iterative step in the EM algorithm (for  $A \in \mathbb{R}^{M \times N}$ ) can be expressed as follows:  $(2MN + N)_M + (M + N)_D + (2MN)_A$ , where subscripts  $M$ ,  $D$  and  $A$  stand for elementary multiplications, divisions and additions. Hence we obtain  $4MN + 2N + M$  elementary flops. For the MAP algorithm with Gibbs prior, the computational cost can be roughly determined as  $(2MN + 10N)_M + (M + 9N)_D + (2MN + 9N)_A + (8N)_S + (8N)_E$ , where subscripts  $S$  and  $E$  denote subtractions and evaluations of build in function  $\tanh(\cdot)$ . Unfortunately, this algorithm cannot be fully vectorized and some loops must be used. This prolongs the computational time of a single iteration, and moreover, the evaluation of function lasts much longer than the computation of the other flop. Thus only the measurements of elapsed times may be more credible, however, they are strongly hardware-dependant. The elapsed times measured in milliseconds per one iteration on the computer with a dual core 3 GHz CPU are as follows: 4.3 and 21.2 for the EM algorithm with respective  $A^{(1)}$  and  $A^{(2)}$ , and 23.2 and 92 for the MAP-Gibbs algorithm with  $A^{(1)}$  and  $A^{(2)}$ , respectively.

## VI. CONCLUSIONS

The paper demonstrates an additional benefit of the multigrid technique for limited-data tomography besides its a high convergence rate, i.e. a possibility of exporting some image components from a coarse grid to a fine grid, whereby a high resolution image can be enhanced by some missing components from a nullspace, provided that they are recovered by a coarse grid. To meet this condition, a coarse grid must be reconstructed using the same set of data as for a fine grid (dense ray-coverage) and the MAP algorithm with Gibbs prior should be used to increase resistance of reconstruction to noisy perturbations.

The benefits from covering the examined object with rays more densely is demonstrated in Figs. 2 and 3 (compare the left and middle images in Fig. 3).

The superiority of the MAP algorithm with Gibbs prior over, e.g. the Tikhonov regularization is also easily noticeable (compare the image in Fig. 4 with the images in Fig. 8 (second row)).

The phenomena of passing on some missing image components from the coarse grid to the fine grid is well illustrated in Fig. 7 (right). The curves referred to as “ $x_0 =$

coarse grid” reach lower values of the relative error than the curves plotted for “ $x_0 = \text{uniform}$ ”, except for the very noisy data with  $\text{SNR} = 25$  dB. Thus we can conclude that for the very noisy data, the multigrid implementation of this algorithm does not recover the missing image components. Nevertheless, the general results obtained for the very noisy data (Fig. 8 (third row)) are satisfactory and are much better than the ones obtained with the standard Tikhonov regularization. Figure 8 also suggests that in the fine grid's images (right column) the parasite smearing effects from inhomogeneities are less visible than in the images reconstructed for “ $x_0 = \text{uniform grid}$ ” (middle column). These effects, which appear also in Fig. 3 (left and right), result from a lack of high-frequency (nullspace) image components.

Figures 5 and 6 show that the proposed simplified tools for estimation of the hyperparameters in the MAP algorithm with Gibbs prior work well for all the kinds of data.

Summing up, the results presented in this paper have confirmed our theoretical considerations and have demonstrated that the multigrid implementation of the MAP algorithm with Gibbs prior is very efficient in limited-data tomography. This can be considered as a new and important argument for using a multigrid implementation of the MAP algorithms (beside of its fast convergence properties) for image reconstruction in limited-data tomography.

## References

- [1] K. A. Dines and R. J. Lytle, *Computerized geophysical tomography*, Proc. IEEE **67**, 1065-1073 (1979).
- [2] D. L. Lager and R. J. Lytle, *Determining a subsurface electromagnetic profile from high frequency measurements by applying reconstruction technique algorithms*, Radio Science **12**, 249-260 (1977).
- [3] R. J. Lytle, E. F. Laine, D. L. Lager and D. J. Davis, *Using cross borehole electromagnetic probing to locate high contrast anomalies*, Geophysics **44**, 1667-1676 (1979).
- [4] N. Pendock, *Radio-wave tomography for geological mapping: Underground and obscured object imaging and detection*, SPIE **1942**, 96-104 (1993).
- [5] S. F. Somerstein et al., *Radio-frequency geotomography for remotely probing the interiors of operating mini- and commercial-sized oil-shale retorts*, Geophysics **49**, 1288-1300 (1984).
- [6] A. Pralat and R. Zdunek, *Electromagnetic geotomography – selection of measuring frequency*, IEEE Sensors Journal **5**, 242-250 (2005).
- [7] Å Björck, *Numerical Methods for Least-Squares Problems*, SIAM, Philadelphia (1996).
- [8] Y. Censor, *Finite series-expansion reconstruction methods*, Proc. IEEE **71**, 409-419 (1983).
- [9] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems*, SIAM, Philadelphia, 1998.
- [10] C. Paige C and M. Saunders, *LSQR: an algorithm for sparse linear equations and sparse least squares*, ACM Trans. Math. Software **8**, 43-71 (1982).

- [11] P. S. Rowbotham and R. G. Pratt, *Improved inversion through use of the null space*, *Geophysics* **62**, 869-883, (1997).
- [12] I. Koltracht, P. Lancaster and D. Smith, *The structure of some matrices arising in tomography*, *Linear Algebra Appl.* **130**, 193-218 (1990).
- [13] T. A. Sanny and K. Sassa, *Detection of fault structure under a near-surface low velocity layer by seismic tomography: synthetic studies*, *Journal of Applied Geophysics* **35**, 117-131 (1996).
- [14] S. Geman and D. McClure, *Bayesian image analysis: an application to single photon emission tomography*, *Proc. Amer. Statist. Assoc. Stat. Comp. Sect.* 12-18 (1985).
- [15] P. J. Green, *Bayesian reconstructions from emission tomography data using a modified EM algorithm*, *IEEE Trans. Med. Imaging* **9**(1), 84-93 (1990).
- [16] T. Hebert and R. Leahy, *A generalized EM algorithm for 3D Bayesian reconstruction from Poisson data using Gibbs priors*, *IEEE Trans. Med. Imag.* **8**, 194-202 (1989).
- [17] K. Lange, *Convergence of EM image reconstruction algorithms with Gibbs smoothing*, *IEEE Trans. Med. Imag.* **9**(4), 439-446 (1990).
- [18] M. V. Ranganath, A. P. Dhawan and N. Mullani, *A multi-grid expectation maximization reconstruction algorithm for positron emission tomography*, *IEEE Trans. Med. Imag.* **7**, 273-278 (1988).
- [19] T. S. Pan and A. E. Yagle, *Numerical study of multigrid implementations of some iterative image reconstruction algorithms*, *IEEE Trans. Med. Imag.* **10**, 572-588 (1991).
- [20] R. Zdunek, *Optimization of methods of reconstructing images of electromagnetic wave attenuation coefficient distribution in earth*, PhD thesis, Wroclaw University of Technology, Wroclaw, Poland, February 2002. in Polish.
- [21] C. Popa and R. Zdunek, *Kaczmarz extended algorithm for tomographic image reconstruction from limited-data*, *Mathematics and Computers in Simulation* **65**(6), 579-598 (2004).
- [22] A. Neumaier, *Solving ill-conditioned and singular linear systems: A tutorial on regularization*, *SIAM Rev.* **40**, 636-666 (1998).
- [23] R. Zdunek and A. Pralat, *Multigrid implementation of some regularized image reconstruction methods for limited-data tomography*, *Industrial Process Tomography 553-558*, Banff, Canada, September 2003.
- [24] L. A. Shepp and Y. Vardi, *Maximum likelihood reconstruction for emission tomography*, *IEEE Trans. Med. Imag.* **M11**, 113-122 (1982).
- [25] K. Lange and R. Carson, *EM reconstruction algorithms for emission and transmission tomography*, *J. Comp. Assisted Tomo.* **8**(2), 306-316 (1984).
- [26] A. P. Dempster, N. M. Laird, and D. B. Rubin, *Maximum likelihood from incomplete data via the EM algorithm*, *Journal of the Royal Statistical Society* **39**(1), 1-38 (1977).
- [27] D. S. Lalush and B. M. W. Tsui, *Simulation evaluation of Gibbs prior distributions for use in maximum a posteriori SPECT reconstructions*, *IEEE Trans. Med. Imaging* **11**(2), 267-275 (1992).
- [28] D. M. Higdon et al., *Fully Bayesian estimation of Gibbs hyperparameters for emission computed tomography data*, *IEEE Trans. Med. Imag.* **16**(5), 516-525 (1997).
- [29] Z. Zhou, R. M. Leahy and J. Qi, *Approximate maximum likelihood hyperparameter estimation for Gibbs priors*, *IEEE Trans. Med. Imag.* **6**, 844-861 (1997).



**RAFAL ZDUNEK** received the MSc and PhD degrees in telecommunications from Wroclaw University of Technology, Poland, in 1997 and 2002, respectively. Since 2002, he has been a Lecturer in Institute of Telecommunications, Teleinformatics and Acoustics, Wroclaw University of Technology, Poland. In 2004, he was a Visiting Associate Professor in the Institute of Statistical Mathematics, Tokyo, Japan. Since 2005, he has worked as Research Scientist in Brain Science Institute, RIKEN, Saitama, Japan. His area of interests includes numerical methods and inverse problems in application to nonnegative matrix factorization, blind source separation, tomographic image reconstruction, and WCDMA network optimization. He has published over 40 journal and conference papers.