Proceedings of ALGORITMY 2016
pp. 333–342

NOTES ON PERFORMANCE OF BIDIAGONALIZATION-BASED NOISE LEVEL ESTIMATOR IN IMAGE DEBLURRING *

IVETA HNĚTYNKOVÁ†, MARIE KUBÍNOVÁ‡, AND MARTIN PLEŠINGER§

Abstract. Image deblurring represents one of important areas of image processing. When information about the amount of noise in the given blurred image is available, it can significantly improve the performance of image deblurring algorithms. The paper [11] introduced an iterative method for estimating the noise level in linear algebraic ill-posed problems contaminated by white noise. Here we study applicability of this approach to image deblurring problems with various types of blurring operators. White as well as data-correlated noise of various sizes is considered.

Key words. image deblurring, linear ill-posed problem, noise, noise level estimate, Golub–Kahan iterative bidiagonalization

AMS subject classifications. 15A29, 65F10, 65F22

1. Introduction. When recording digital images, some form of blurring often occurs, e.g., when camera lens is out of focus, light conditions are not perfect, the object is moving etc. In such a case the information from a particular image pixel is spread to surrounding pixels resulting in a lower-quality image. As an additional problem, the recorded image can contain unknown errors in form of variations of pixel density usually referred to as noise with different properties based on its origin. Image deblurring methods aim to reconstruct the true sharp image while suppressing the influence of noise, by using a mathematical model of the blurring process; see, e.g., [8, Chapters 1 and 3].

Let B, $X \in \mathbb{R}^{l\times m}$ represent the blurred noisy image and its unknown sharp counterpart, respectively. In many applications, the blurring process is linear or can be well approximated by a linear model, which is an assumption we will follow. In that case we can model the blurring process as

$$Ax \approx b\,, \qquad A \in \mathbb{R}^{n\times n}\,, \qquad b \in \mathbb{R}^{n}\,, \tag{1.1}$$

where A is a linearized (e.g., discretized) blurring operator, b and x are vectorized forms of B and X (obtained by stacking the columns of the matrix into a single vector), respectively, and $n = lm$. The right-hand side of the linear algebraic problem above can be formally written as

$$b = b^{\text{exact}} + b^{\text{noise}},$$

where b^{exact} is the unknown smooth noise-free right-hand side and b^{noise} represents unknown noise. We refer to the quantity

$$\delta_{\text{noise}} \equiv \frac{\|b^{\text{noise}}\|}{\|b^{\text{exact}}\|} \tag{1.2}$$

*This work has been supported by the GAČR grant No. P201/13–06684S.

†Charles University in Prague, Faculty of Mathematics and Physics, Prague, and Institute of Computer Science, AS CR, Prague, Czech Republic (hnetynkova@cs.cas.cz).

‡Charles University in Prague, Faculty of Mathematics and Physics, Prague, and Institute of Computer Science, AS CR, Prague, Czech Republic (kubinova@cs.cas.cz).

§Department of Mathematics, Technical University of Liberec, Liberec, and Institute of Computer Science, AS CR, Prague, Czech Republic (martin.plesinger@tul.cz).

as the *noise level*. Since noise is supposed to be small compared to noise-free data, $\delta_{\text{noise}} \ll 1$ is a realistic assumption.

Properties of this model have been widely analyzed; see [6] for a summary, [14] and [8] for applications in image processing, and also [11], [3], [10] for the behavior in the context of Krylov subspace methods. In particular, it is known that the singular values of A usually decay gradually to zero without a noticeable gap and the singular vectors of A represent increasing frequencies. The model satisfies the discrete version of the so-called *Picard condition* meaning that on average the sizes of projections of b^{exact} onto the left singular subspaces of A decrease faster than the singular values. On the other hand, b^{noise} typically does not satisfy such a condition. Consequently, linear image deblurring models (1.1) represent a typical example of an *ill-posed problem*; see [8, Chapter 5].

It is well known that information about the amount of noise can significantly improve performance of image deblurring methods; see for example [13], [2], [1], [7], and also [8, Chapter 6]. Such information is however rarely available. The paper [11] introduced an inexpensive method for estimating the unknown white noise level in linear ill-posed algebraic problems with a smoothing operator A. The estimate is obtained by the Golub–Kahan iterative bidiagonalization [4], a short recurrence based Krylov subspace method. It relies on the assumptions that the model (1.1) satisfies the discrete Picard condition, A has the smoothing property, the left singular vectors of A represent increasing frequencies, and b^{exact} is smooth. Because the method needs only evaluation of matrix-vector products, it can take advantage of a specific structure of A often present in image deblurring problems; see [8, Chapter 4].

The paper is organized as follows. Section 2 summarizes the main ideas of the noise level estimation presented in [11]. Section 3 studies its applicability to image deblurring problems with various types and amount of noise. Spatially invariant as well as spatially variant blur is considered. Section 4 concludes the paper.

2. Iterative noise level estimate. The Golub–Kahan iterative bidiagonalization starting with the vectors $w_0 \equiv 0$, $s_1 \equiv b/\beta_1$, where $\beta_1 \equiv \|b\| \neq 0$, computes for $j = 1, 2, \ldots$

$$\begin{aligned}
&\widetilde{w}_j = A^T s_j - \beta_j w_{j-1} && \text{(orthogonalization step)}\\
&\alpha_j = \|\widetilde{w}_j\|\,, \qquad w_j = \frac{1}{\alpha_j}\,\widetilde{w}_j && \text{(normalization step)}\\
&\widetilde{s}_{j+1} = A w_j - \alpha_j s_j && \text{(orthogonalization step)}\\
&\beta_{j+1} = \|\widetilde{s}_{j+1}\|\,, \quad s_{j+1} = \frac{1}{\beta_{j+1}}\,\widetilde{s}_{j+1} && \text{(normalization step)}
\end{aligned}$$

until $\alpha_j = 0$ or $\beta_{j+1} = 0$, or the dimension n of the problem is reached. Assume that the process does not terminate before the step k. Then the left bidiagonalization vectors $s_1, \ldots, s_k$ represent an orthonormal basis of the Krylov subspace

$$\mathcal{K}_k(AA^T, b) \equiv \operatorname{span}\{b, AA^T b, \ldots, (AA^T)^{k-1} b\},$$

and the right bidiagonalization vectors $w_1, \ldots, w_k$ represent an orthonormal basis of the Krylov subspace

$$\mathcal{K}_k(A^T A, A^T b) \equiv \operatorname{span}\{A^T b, A^T AA^T b, \ldots, (A^T A)^{k-1} A^T b\}.$$

Denote

$$L_k \equiv \begin{bmatrix} \alpha_1 & & & \\ \beta_2 & \alpha_2 & & \\ & \ddots & \ddots & \\ & & \beta_k & \alpha_k \end{bmatrix} \in \mathbb{R}^{k\times k} \tag{2.1}$$

the bidiagonal matrix of the normalization coefficients, representing a projection (restriction) of the operator A onto the above defined k-dimensional Krylov subspaces,

$$L_k = [s_1, \ldots, s_k]^T A\, [w_1, \ldots, w_k];$$

see [4], [15]. Let $p_1^{(k)}$ be the left singular vector corresponding to the smallest singular value of L_k.[1]

In [11] it was described how *white noise* from the right-hand side b propagates in the Golub–Kahan iterative bidiagonalization, particularly in the left bidiagonalization vectors; see also [10], [12]. While the starting vector s_1 is smooth (since it is dominated by the scaled b^{exact}), during the bidiagonalization process, as k increases, the left bidiagonalization vectors s_k become more and more dominated by the high-frequency part of propagated noise b^{noise}. This is caused by projecting out the low-frequency components (arising mostly from b^{exact} and partly also from the low-frequency part of b^{noise}) in order to achieve orthogonality among the bidiagonalization vectors. The iteration where the most high-frequency dominated vector is obtained is called the *noise revealing* iteration and is denoted by k_{noise}. After this iteration, a part of noise is projected out resulting in a smoother left bidiagonalization vector. Analysis of this phenomenon in [11] allowed to derive two quantities estimating the noise level: The *cumulative (amplification) ratio*

$$\varphi_k \equiv \prod_{j=1}^{k} \frac{\alpha_j}{\beta_{j+1}} \tag{2.2}$$

and *the size of the first entry* of $p_1^{(k)}$, i.e.,

$$|(p_1^{(k)}, e_1)|, \tag{2.3}$$

where $(\,\cdot\,,\,\cdot\,)$ denotes the standard inner product and $e_1 = [1, 0, \ldots, 0]^T \in \mathbb{R}^k$. It was proved that (2.2) and (2.3) both (on average) decrease until k_{noise}. After this iteration, the cumulative ratio increases while the size of the first entry of $p_1^{(k)}$ begins to almost stagnate. This allows to detect the iteration k_{noise} in which the best noise level approximation is obtained; see [11] and also [17]. Note that both estimators are relatively cheap to compute. Since noise usually propagates rapidly, k_{noise} is very small in comparison to n. The bidiagonalization coefficients α_j, β_{j+1} are directly available, computation of the singular vector $p_1^{(k)}$ for a small bidiagonal matrix L_k can be performed efficiently.

We now use the example from [11] to illustrate the behavior of both estimators on the problem `shaw` from the Regularization Toolbox [5] in MATLAB. This problem represents a one-dimensional image restoration model obtained as a quadrature

[1] Note that we use the notation introduced in [11].

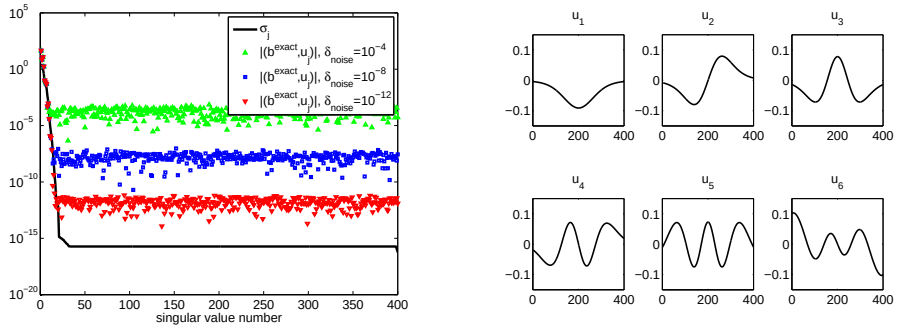

FIG. 2.1. *Illustration of the violation of the discrete Picard condition for the problem* `shaw(400)` *with white noise and the noise levels* $\delta_{\text{noise}} = 10^{-4}$, 10^{-8}, *and* 10^{-12} *(left). Increasing frequencies in the first six left singular vectors of* A *for the problem* `shaw(400)` *(right).*

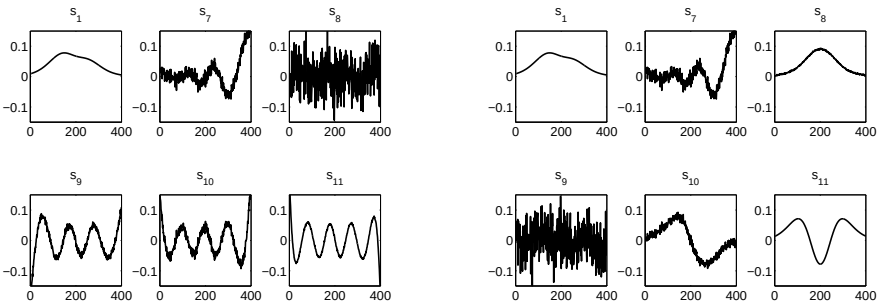

FIG. 2.2. *Comparison of several left bidiagonalization vectors* s_k *computed by the Golub–Kahan iterative bidiagonalization implemented with (left) and without (right) reorthogonalization, for the problem* `shaw(400)` *with white noise and the noise level* $\delta_{\text{noise}} = 10^{-4}$.

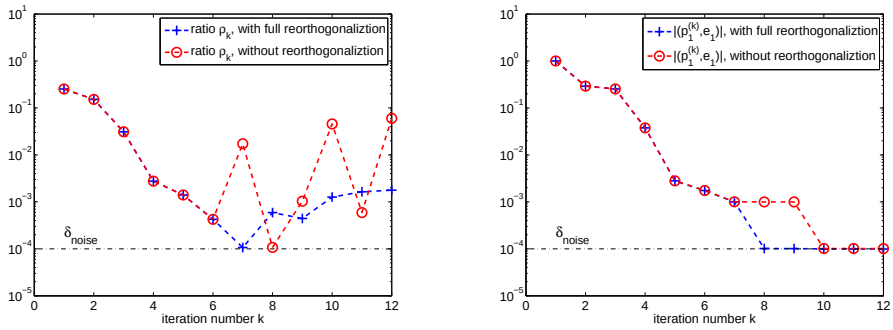

FIG. 2.3. *Estimates obtained by cumulative amplification ratios (left), and sizes of the first entry of* $p_1^{(k)}$ *(right), for the problem* `shaw(400)` *with white noise and the noise level* $\delta_{\text{noise}} = 10^{-4}$. *Computations were performed with and without reorthogonalization.*

discretization of a first kind Fredholm integral equation on the integration intervals $[-\frac{\pi}{2}, \frac{\pi}{2}]$. Here, the smoothing kernel is given by

$$K(s,t) = \Big(\cos(s) + \cos(t)\Big)^2 \left(\frac{\sin(u)}{u}\right)^2, \quad \text{where} \quad u = \pi\Big(\sin(s) + \sin(t)\Big);$$

see [16] for the description of the model. The linear problem with the size $n = 400$ is

contaminated by white noise generated by `randn(400,1)` rescaled to obtain a particular noise level δ_{noise}. Figure 2.1 (left) shows the sizes of the projections (b, u_j), where u_j are the left singular vectors of the matrix A corresponding to the singular values ordered in the nonincreasing order. We see how the presence of noise of various noise level results in the violation of the discrete Picard condition for the subspaces corresponding to smaller singular values. Figure 2.1 (right) presents increasing frequencies in the left singular vectors of A.

Figure 2.2 illustrates how white noise reveals in the left bidiagonalization vectors for $\delta_{\text{noise}} = 10^{-4}$. The vectors in the left part are computed with full double reorthogonalization in the Golub–Kahan iterative bidiagonalization in order to simulate exact arithmetic. Clearly, the frequencies increase before they become maximal in s_8, computed in the iteration $k_{\text{noise}} = 7$; see the algorithm above. The right part shows the delay in noise revealing caused by reappearance of a smooth vector, as a result of the loss of orthogonality in the bidiagonalization implemented without reorthogonalization. However, the effect is still present and $k_{\text{noise}} = 8$. Figure 2.3 compares estimates obtained by cumulative amplification ratios (left) and by sizes of the first entry of $p_1^{(k)}$ (right). We see that both estimators give very accurate and comparable results for computation with as well as without reorthogonalization. It is worth noting that oscillations in the cumulative ratio computed without reorthogonalization can cause difficulties in automatic detection of k_{noise}. Thus, in the following we restrict ourselves only to the estimate (2.3). Analysis of the methods detecting the point of stagnation is out of the scope of this paper; see [17] for some ideas.

3. Performance for 2D image deblurring problems. Robustness of the estimator (2.3) is studied on a sharp testing picture X of size 167×250 pixels, i.e., $n = 41\,750$; see Figure 3.1. The experiments are performed in MATLAB, with the use of functions from the Image Processing Toolbox.

3.1. Spatially invariant blur. First we consider a standard spatially invariant blurring model, where blurring of each individual pixel in X is characterized by a given point-spread-function (PSF); see [8, Chapter 3]. Presented results include models for a Gaussian blur, motion blur, and disc blur. Using the function `fspecial`, we construct two PSFs with smaller and larger support for each type of blur, giving in total six testing PSFs; see Figure 3.2. Note that since we only need to perform matrix-vector multiplications, we do not form the corresponding blurring matrix A explicitly. The blurred images B are computed by the 2D convolution using the function `conv2` with the parameter `valid`, i.e., only the part computed without the zero-padded edges is returned. Multiplication by the matrix A in the bidiagonalization is performed by the function `conv2` with the parameter `same` representing zero boundary conditions. For testing purposes, the image B is contaminated by noise using the function `imnoise` with four different parameter settings. We consider two types of noise: white noise with Gaussian distribution (parameter `gaussian`), and uniformly distributed data-correlated noise (parameter `speckle`). Variances $\sigma^2 = 10^{-2}$ and $\sigma^2 = 5 \cdot 10^{-6}$ give two different noise levels δ_{noise} for each type of noise.

Figure 3.3 provides similar information as Figure 2.1, here for the matrix A corresponding to the larger Gaussian PSF and white noise. Again we see the violation of the discrete Picard condition. The so-called left singular images (reshaped left singular vectors) of the blurring matrix A tend to be dominated by higher frequencies, i.e., more oscillations appear in both vertical and horizontal directions, as k increases; see [8] for details.

FIG. 3.1. *Sharp testing image X of size 167×250 pixels.*

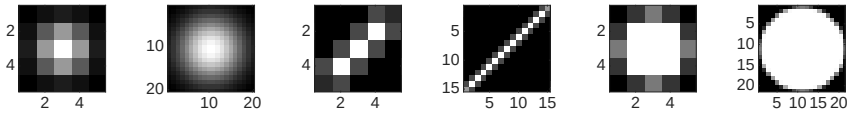

FIG. 3.2. *Considered PSFs defining the blurring matrix A. From the left: two Gaussian, two motion and two disc PSFs.*

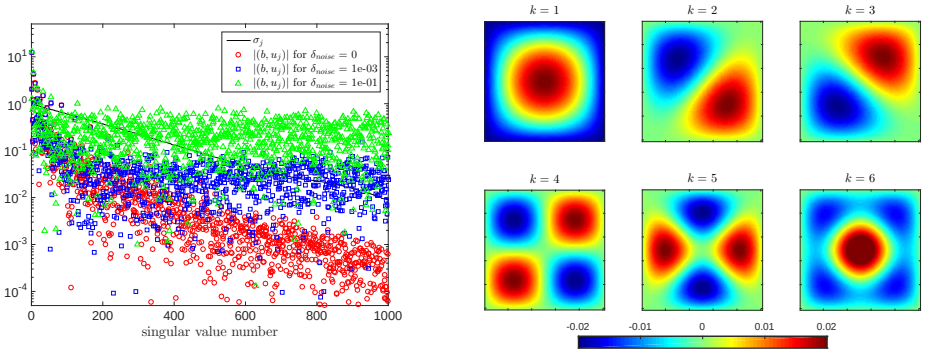

FIG. 3.3. *Illustration of violation of the discrete Picard condition for the 2D image deblurring problem with larger Gaussian blur and white noise of various noise levels (left). Increasing frequencies in the first six left singular vectors of the corresponding blurring matrix A printed as 2D images (right).*

Figure 3.4 shows blurred noisy images together with the corresponding noise level estimates obtained for models with the smaller Gaussian blur, for the four above described noise settings. Bidiagonalization with and without reorthogonalization is used. First, we observe that the overall behavior of the estimator does not significantly depend on the reorthogonalization, except for the fact that for lower noise levels the delay in noise revealing results in the increase of the computational cost, since more iteration steps are required to obtain a reasonable estimate (note the different scaling of the x axis in the second and fourth row). This problem is not present in experiments with more realistic higher noise level (the first and third row), where noise propagates quickly. This is a positive message since for larger images, reorthogonalization cannot be performed because of its enormous computational cost and memory requirements. Furthermore, we see that the expected stagnation in the estimator allowing to detect the noise revealing iteration is more significant for higher noise levels. Figure 3.5 is the counterpart of Figure 3.4 for larger Gaussian blur. We observe clear stagnation

in all curves. The behavior of the estimator for higher noise levels is generally very similar to Figure 3.4, however we see that in Figure 3.5 the lower noise levels are overestimated.

Summarizing, the best results are obtained for large blur and large noise levels, shown in the first and third row of Figure 3.5 making the estimator more successful on complicated problems. However, for noise below a certain level problems appear. For smaller amount of blur the noise revealing iteration can not be perfectly detected, while for larger amount of blur the noise level is overestimated. This is more significant in experiments with data-correlated noise, which is not surprising as the estimator is based on revealing of the *high-frequency part of noise* in the vectors s_k, thus it does not take into account the smooth part of noise. For white noise, this does not represent a complication. However, even though data-correlated noise is white-noise like, its behavior partly resembles behavior of the smooth noise-free data in b.

Figure 3.6 gives noise level estimates for two variants of the motion and disc blur (specified on the top) with four different noise settings (specified on the left). All results were obtained by the Golub–Kahan iterative bidiagonalization implemented without reorthogonalization. We see similar results as in experiments with the Gaussian blur. Consequently, the blur type has generally minor impact on the performance of the estimator.

3.2. Spatially variant blur. In addition to spatially invariant blur, we investigate the behavior of the estimator (2.3) for a special type of spatially variant blur: a rotational blur recently studied in the context of image deblurring in [9]. Consider a sharp image represented by the central part of size 167×167 of the image from Figure 3.1 in order to avoid the large black areas appearing at the edges when rotating the whole rectangular image. The code to construct the rotational blurring operator has been provided by Per Christian Hansen, and it is identical to the code used in [9]. We consider three different blurs: rotation by 10°, rotation by 20°, and tilt by 10°. All the resulting blurred images are corrupted by additive white noise with two different variances $\sigma^2 = 10^{-3}$ and $\sigma^2 = 5 \cdot 10^{-6}$. The noisy images together with noise level estimates computed by the Golub–Kahan iterative bidiagonalization without reorthogonalization for all settings are shown in Figure 3.7.

The results are very similar to results for the spatially invariant blurring. The estimates for large noise level are accurate (left). Especially in case of strong blurring (rotation by 20°), the curve stagnates very close to the actual noise level. For the smaller noise level, the stagnation is not so significant.

4. Conclusions. Presented paper has studied performance of the noise level estimator proposed in [11], which is based on the iterative Golub–Kahan bidiagonalization, on image deblurring problems. Implementations with and without reorthogonalization have been compared. We have demonstrated that reorthogonalization does not improve the quality of the estimate, although for small noise levels we would need more iterations to obtain estimate of the same accuracy as by the algorithm with reorthogonalization. We have shown that the performance of the estimator does not significantly depend on the particular type of blur but it is generally more successful on problems with higher noise levels. For smaller noise levels, the expected stagnation of the estimator has been rather slow, making the detection of the noise revealing iteration (where the best estimate should be obtained) complicated. For data-correlated noise of lower noise level, the estimator has not been reliable, as it underestimated some noise levels while it overestimated the others. Further analysis of the observed

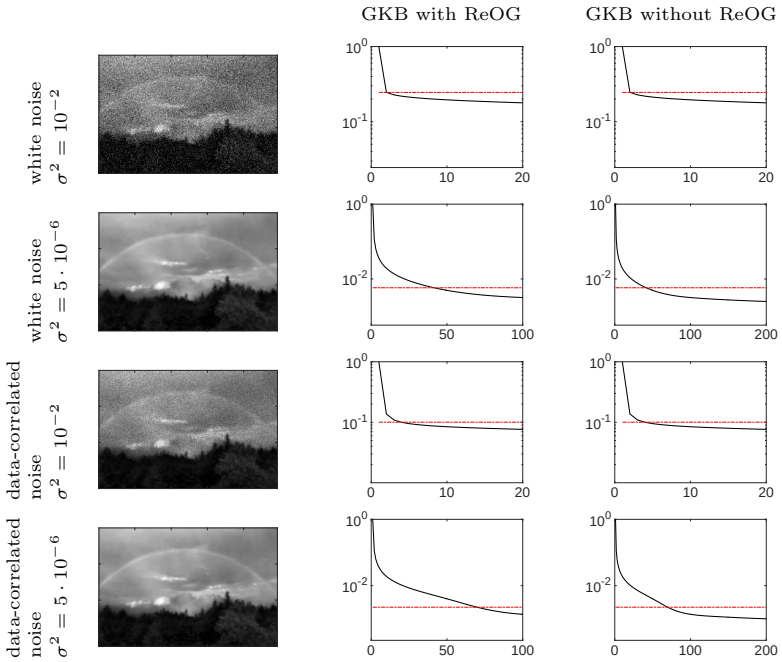

FIG. 3.4. *Blurred noisy images and the corresponding noise level estimates for smaller Gaussian blur and four considered noise settings (specified on the left). Computed by the Golub–Kahan bidiagonalization (GKB) with and without reortogonalization (ReOG) (specified on the top). Red line represents the exact noise level.*

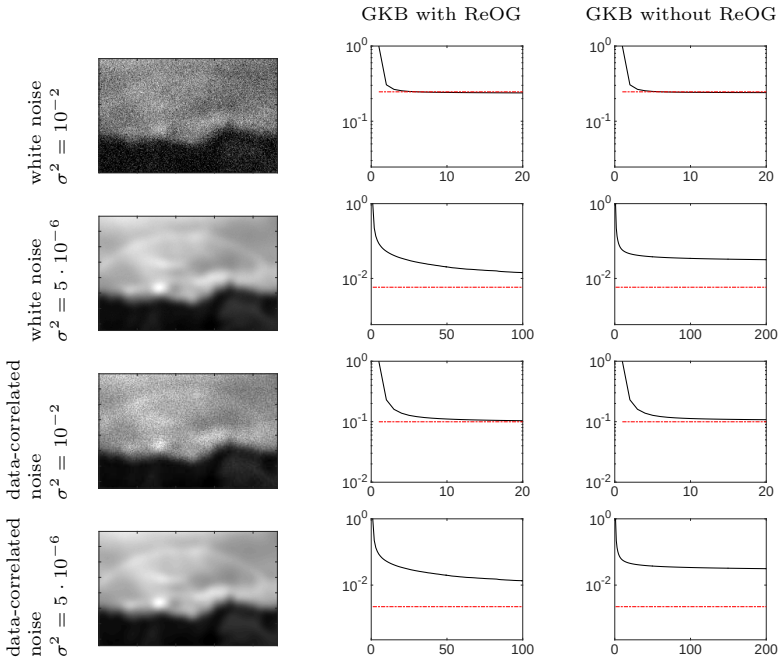

FIG. 3.5. *Results similar as in Figure 3.4 computed for the model with the larger Gaussian blur.*

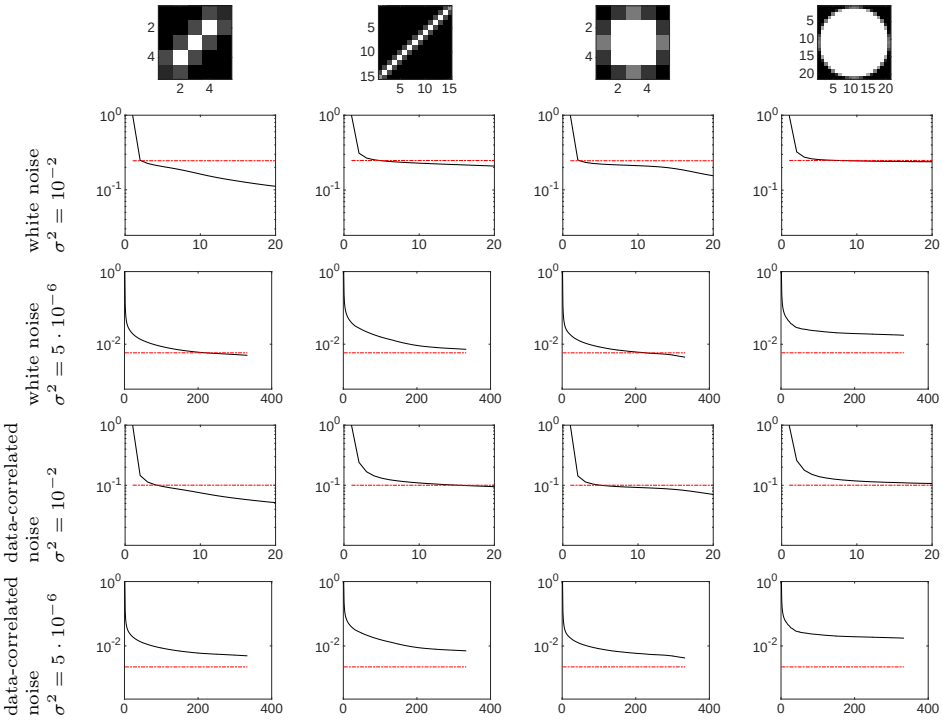

FIG. 3.6. *Noise level estimates for models with the motion and disc blur (specified on the top) and four considered noise settings (specified on the left), computed without reortogonalization.*

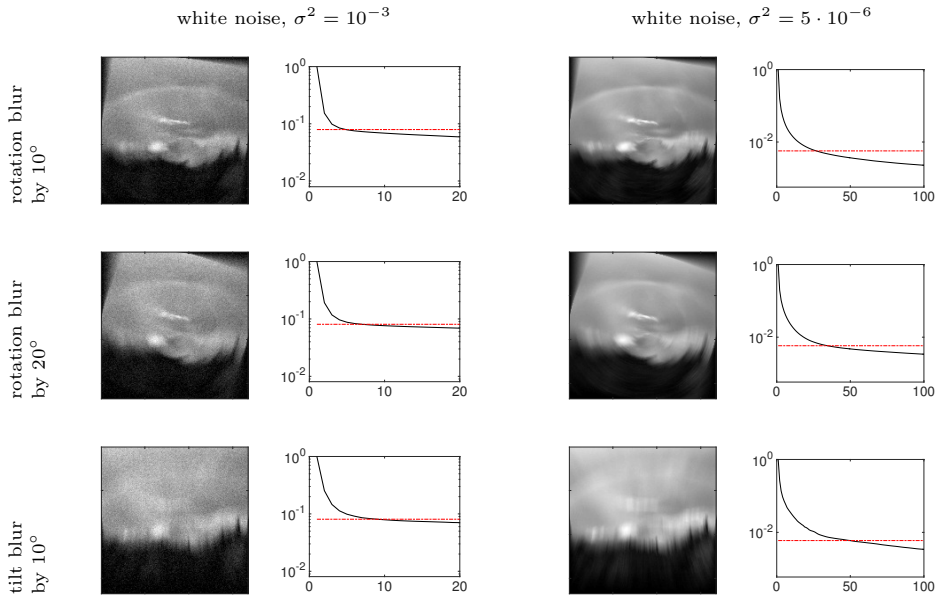

FIG. 3.7. *Noise level estimates for models with the rotational blur and white noise with two different variances (specified on the top), computed without reortogonalization.*

behavior and related issues is out of the scope of this paper and will be presented elsewhere.

REFERENCES

[1] P. Blomgren and T. F. Chan, *Modular solvers for constrained image restoration problems using the discrepancy principle*, Numer. Linear Algebra Appl., 9 (2002), pp. 348–358.

[2] L. Desbat and D. Girard, *The "minimum reconstruction error" choice of regularization parameters: Some more efficient methods and their application to deconvolution problems*, SIAM J. Sci. Comput., 16 (1995), pp. 1387–1403.

[3] S. Gazzola, P. Novati, and M. R. Russo, *On Krylov projection methods and Tikhonov regularization*, ETNA, 44 (2015), pp. 83–123.

[4] G. H. Golub and W. Kahan, *Calculating the singular values and pseudo-inverse of a matrix*, SIAM J. Numer. Anal., Ser. B 2 (1965), pp. 205–224.

[5] P. C. Hansen, *Regularization Tools Version 4.1 (for MATLAB Version 7.3). A MATLAB package for analysis and solution of discrete ill-posed problems* (available at `http://www.imm.dtu.dk/~pcha/Regutools`).

[6] ———, *Rank-Deficient and Discrete Ill-Posed Problems, Numerical Aspects of Linear Inversion*, SIAM Publications, Philadelphia, PA, 1998.

[7] P. C. Hansen and T. K. Jensen, *Noise propagation in regularizing iterations for image deblurring*, ETNA, 31 (2008), pp. 204–220.

[8] P. C. Hansen, J. G. Nagy, and D. P. O'Leary, *Deblurring Images: Matrices, Spectra, and Filtering*, SIAM Publications, Philadelphia, PA, 2006.

[9] P. C. Hansen, J. G. Nagy, and K. Tigkos, *Rotational image deblurring with sparse matrices*. BIT Numerical Mathematics, 54 (2013), pp. 649–671.

[10] I. Hnětynková, M. Kubínová, and M. Plešinger, *On noise propagation in residuals of Krylov subspace iterative regularization methods*, submitted.

[11] I. Hnětynková, M. Plešinger, and Z. Strakoš, *The regularizing effect of the Golub–Kahan iterative bidiagonalization and revealing the noise level in the data*, BIT Numerical Mathematics, 49 (2009), pp. 669-696.

[12] M. Michenková, *Regularization techniques based on the least squares method*, Diploma thesis, Charles University in Prague, 2013.

[13] V. A. Morozov, *On the solution of functional equations by the method of regularization*, Soviet Math. Dokl., 7 (1966), pp. 414–417.

[14] F. Natterer, *The Mathematics of Computerized Tomography*, John Wiley & Sons and Teubner, Stuttgart, 1986.

[15] C. C. Paige, *Bidiagonalization of matrices and solution of linear equations*, SIAM J. Numer. Anal., 11 (1974), pp. 197–209.

[16] C. B. Shaw, Jr., *Improvements of the resolution of an instrument by numerical solution of an integral equation*, J. Math. Anal. Appl., 37 (1972), pp. 83–112.

[17] K. Vasilík, *Linear algebraic modeling of problems with noisy data*, Diploma thesis, Charles University in Prague, 2011.