Multigrid Reconstruction with Block-Iterative Updates for Breast Tomosynthesis

Koen Michielsen and Johan Nuyts*
Department of Imaging and Pathology, division of Nuclear Medicine & Molecular Imaging, KU Leuven, Leuven 3000, Belgium and Medical Imaging Research Center, KU Leuven, Leuven 3000, Belgium
(Dated: October 5, 2015)

Purpose: We wish to evaluate the possible advantages of using a multigrid approach to maximum-a-posteriori reconstruction in digital breast tomosynthesis together with block-iterative updates in the form of either plane-by-plane updates or ordered subsets.

Methods: We previously developed a penalized maximum likelihood reconstruction algorithm with resolution model dedicated to breast tomosynthesis[1]. This algorithm was extended with ordered subsets and multigrid updates, and the effects on the convergence and on limited angle artifact appearance were evaluated on a mathematical phantom and patient data. To ensure a fair comparison, the analysis was performed at the same computational cost for all methods. To assess convergence and artifact creation in the phantom reconstructions, we looked at posterior likelihood, sum of squared residuals, contrast of identical calcifications at different positions, and the standard deviation between the contrasts of these calcifications. For the patient cases we calculated posterior likelihood, measured the signal difference to noise ratio of subtle microcalcifications, and visually evaluated the reconstructions.

Results: We selected multigrid sequences scoring in the best 10% of the four evaluated parameters, except for the reconstructions with subsets where a low standard deviation of the contrast was incompatible with the three other parameters. In further evaluation of phantom reconstructions from noisy data and patient data, we found improved convergence and a reduction in artifacts for our chosen multigrid reconstructions compared to the single grid reconstructions with equivalent computational cost, although there was a diminishing return for an increasing number of subsets.

Conclusions: Multigrid reconstruction improves upon reconstruction with a fixed grid when evaluated at a fixed computational cost. For multigrid reconstruction, using plane-by-plane updates or applying ordered subsets resulted in similar performance.

I. INTRODUCTION

Digital breast tomosynthesis (DBT) is a recent pseudo three dimensional imaging technique that reconstructs the breast volume as a series of planes (which we define as parallel to the detector, unless specified otherwise), with small in-plane pixel spacing (typically between 50 and 150 μm) and a much larger inter-plane spacing (typically 1 mm). This reduces the effect of overlapping normal tissue structures on the visualization of low contrast lesions, such as masses in dense breasts, which are difficult to detect in projection images[2]. A recent overview of clinical studies examining the effect of DBT on cancer detection rates by Houssami and Skaane[3] concludes that while addition of 2-view DBT to standard mammography improves cancer detection rates, there is still insufficient evidence to justify a change from digital mammography to DBT. It should be noted that all systems included in the referenced studies used filtered backprojection (FBP) for reconstruction, and at this point there have been several studies showing the potential of iterative algebraic and maximum likelihood (ML) methods to improve upon this method[4–6].

Reconstruction of DBT projections is a challenging problem since the acquisition process is far away from ideal conditions required for exact inversion. The most important deviation is the limited angle setup, with systems listed by Sechopoulos[7] as currently in clinical use or under development having an angular sampling range between 11° and 50°. This means that in theory an infinite number of tissue distributions could explain the acquired measurement and thus could be considered valid reconstructions. Most of these systems also acquire projection images with a continuously moving x-ray source instead of a step-and-shoot approach, which results in additional blurring of the data in the tube travel direction[8]. These and any other system specific considerations need to be taken into account when adapting a reconstruction algorithm to these systems.

Filtered backprojection is the most common reconstruction method and was adapted for use in DBT by Mertelmeier et al.[9], with further refinement of the filter by Orman et al.[10]. Work by Ludwig et al.[11] and Erhard et al.[12] showed that it is possible to design reconstruction filters to include desirable features from iterative methods and more recent developments by Abdurahman et al.[13] add statistical artifact reduction and super-sampling to further increase image quality. In contrast to the direct inversion in FBP, iterative methods optimize a cost function, typically either to find the least squares solution as in SART[14] or POCS[15], or to find the maximum likelihood solution as for the ML convex algorithm from Lange and Fessler[16] first introduced in DBT by Wu et al.[17].

The main assumptions for iterative methods are that the optimal solution to a cost function with a more accurate forward model and more complete prior knowledge will result in a better reconstruction than the solution to

a less accurate model. In addition, when noise suppression is incorporated into the cost function (as in penalized likelihood or maximum-a-posteriori reconstruction), it is assumed that iterating till convergence is beneficial. A distinct disadvantage of the iterative methods is their high computational cost, even if the reconstruction is stopped before convergence. Wu[17] for example finds that reconstruction with nonregularized ML resulted in patient images with sufficient feature contrast and detail after eight iterations which is roughly equivalent to 24 times the computational cost of FBP, considering that each ML iteration contains one forward projection and two backprojection operations, and FBP only includes a single backprojection.

This high computational cost can be offset by performing the same calculations more efficiently, for example by parallel execution on a dedicated graphical processing unit (GPU), or alternatively by devising more efficient algorithms that need fewer calculations to perform the same reconstruction. In most instances these two methods will be combined.

Iterative image reconstruction is a problem well suited to parallelization since all operations can be performed independently on all image voxels[18] and practical applications have demonstrated significant acceleration[19–21]. The main limitation of GPUs is the small amount of available memory which might limit their use for large datasets. This is one of the reasons why some implementations will only offload the most computationally expensive parts of the algorithm (such as the forward and backprojections) to the GPU, in order to get a most of the acceleration while avoiding the memory limitations.

Although memory size and computational power keep growing year after year, the relatively small GPU memory will likely remain the main bottleneck as imaging modalities evolve towards ever larger data sizes at the same time[22]. Therefore methods that perform better for the same computational cost will always be useful. Block-iterative methods (discussed in section II B) are frequently used for this purpose, whereas multigrid methods (section II D) are less popular, but we will show their usefulness for DBT reconstruction in this paper. Other options, which we did not explore in this work, can accelerate convergence by using more efficient update steps to optimize the cost function[23–26]. New update strategies for DBT were proposed by Sidky et al.[27] and Park et al.[28], and for computed tomography by Ramani and Fessler[29] and Kim et al.[30].

In this work we will examine whether a multigrid approach can accelerate the convergence of our block-iterative maximum likelihood reconstruction which includes an acquisition dependent resolution model and a combined quadratic and total variation regularization. The existing and new variants that we use are first presented separately and then combined with a multigrid algorithm that can use specific acquisition models and update strategies for each grid size. We will find the multigrid sequence with the fasted convergence for each block-iterative method based on reconstructions from simulated projection data, and then further evaluate this selection on simulated data with added noise and a small set of patient data.

II. MATERIALS AND METHODS

A. Maximum Likelihood Reconstruction

Maximum likelihood reconstruction is formulated as an optimization problem that includes the forward acquisition model. With this approach, the analytic inversion is avoided and replaced by an iterative numerical inversion. The accuracy and complexity of the forward model will determine the properties of the algorithm. A simple form of the estimated transmission scan $\hat{y}_i$ can be written as

$$\hat{y}_i = b_i e^{-\sum_j l_{ij}\mu_j}, \tag{1}$$

with μ_j the linear attenuation in voxel j of the reconstruction volume, l_{ij} the intersection length between projection line i and voxel j, and b_i the unattenuated value for projection line i. Using this model and assuming Poisson noise in the measured data, it is possible to calculate the log-likelihood L for the measured projection data y_i:

$$L = \sum_i y_i \ln \hat{y}_i - \hat{y}_i - \ln(y_i!). \tag{2}$$

Using equations (1) and (2), we determine the update step $\Delta\vec{\mu}$ needed to iteratively maximize log-likelihood L starting from image $\vec{\mu}^0$. To this end we approximate L with its second order Taylor expansion at the current reconstruction $\vec{\mu}^n$ and then apply optimization transfer to a separable surrogate function for each voxel j[31–34]. This results in the following update equation for reconstruction volume $\vec{\mu}$:

$$\mu_j^{n+1} = \mu_j^n + \frac{-\alpha_j \frac{\partial L}{\partial \mu_j}|_{\vec{\mu}=\vec{\mu}^n}}{\sum_k \alpha_k \frac{\partial^2 L}{\partial \mu_j \partial \mu_k}|_{\vec{\mu}=\vec{\mu}^n}}, \tag{3}$$

with k a second index over all image voxels, and $\alpha_j \geq 0$ a design parameter introduced by Fessler et al.[31] that can be chosen freely.

Calculating the derivatives of L for model $\hat{y}_i$ in equation (1) results in the following expression for the update in every iteration:

$$\Delta\mu_j = \frac{\alpha_j \sum_i l_{ij}(\hat{y}_i - y_i)}{\sum_i l_{ij}\hat{y}_i \sum_k l_{ik}\alpha_k}. \tag{4}$$

In this equation we can see that the specific choice of α_j influences the convergence speed of individual voxels j. Choosing a high value for a particular voxel j will increase the convergence speed in that voxel at the expense of other voxels that are influenced by the increase of α_j in the denominator. For uniform convergence we

can choose $\alpha_j = 1$ for all j which produces the maximum likelihood for transmission (MLTR) update from our previous work[32]. Another frequent choice is setting $\alpha_j = \mu_j^n$, creating an update step proportional to the current attenuation estimate and thus increasing convergence speed for high attenuation regions and slowing it down elsewhere. This is equivalent to the ML convex algorithm of Lange and Fessler[16].

Like all Bayesian methods, the cost function in equation (2) can be extended with prior information[35], allowing for additional constraints, which is a definite advantage in DBT where reconstruction is an ill posed problem due to the limited angular range of the projections. In this work we use both the quadratic[36] and total variation TV_{l_1} [37,38] smoothing priors.

The quadratic prior can be added as a penalty to the log-likelihood in equation (2):

$$L = \sum_i y_i \ln \hat{y}_i - \hat{y}_i - \ln(y_i!) - \frac{\beta_Q}{4} \sum_{j,k} w_{jk}(\mu_j - \mu_k)^2, \quad (5)$$

with quadratic penalty strength β_Q, and symmetric neighbor weights $w_{jk} = w_{kj}$. This modification is compatible with the optimization transfer to surrogate functions, and results in the following update:

$$\Delta\mu_j = \frac{\alpha_j \sum_i l_{ij}(\hat{y}_i - y_i) + \alpha_j \beta_Q P'(\vec{\mu})}{\sum_i l_{ij}\hat{y}_i \sum_k l_{ik}\alpha_k + \alpha_j \beta_Q P''}, \quad (6)$$

with

$$P'(\vec{\mu}) = \sum_k w_{jk}(\mu_j - \mu_k), \quad (7)$$

$$P'' = -2 \sum_k w_{jk}. \quad (8)$$

The total variation prior is implemented as an additional smoothing step after each update of the reconstruction volume. At that point we apply the GP algorithm derived by Beck and Teboulle[38], with the regularization parameter λ chosen as the desired strength β_{TV} of the prior divided by the denominator of the update step that was just applied:

$$\lambda = \frac{\beta_{TV}}{\sum_i l_{ij}\hat{y}_i \sum_k l_{ik}\alpha_k + \alpha_j \beta_Q P''}. \quad (9)$$

Which gives a combined cost function of:

$$L = \sum_i y_i \ln \hat{y}_i - \hat{y}_i - \ln(y_i!) - \frac{\beta_Q}{4} \sum_{j,k} w_{jk}(\mu_j - \mu_k)^2 - 4\beta_{TV} \sum_{j,k} w_{jk} |\mu_j - \mu_k| \quad (10)$$

We set the prior strengths $\beta_Q = 10^4$ and $\beta_{TV} = 2$, and choose a $3 \times 3 \times 3$ element neighborhood w where only the four direct neighbors in the plane parallel to the detector are set to 0.25 and all other elements are set to 0. This combination of priors yielded a good compromise between resolution and noise on realistic data and was chosen in consultation with an experienced radiologist after evaluating a range of settings on a small set of patient cases. The accelerated reconstruction reported on in this paper facilitates further task based optimization of the prior parameters, which is ongoing work but outside the scope of this paper.

B. Block-Iterative Methods

One of the main disadvantages of ML iterative reconstruction methods is the high number of iterations and corresponding computation time that is needed to produce high quality images. For this reason algorithms are typically modified to include accelerated update strategies such as ordered subsets[39–42], or image-block updates[31,34,43–45].

The MLTR update with subsets (OSTR) can be implemented by limiting the sum over projection rays i in equation (4) to only include rays from the selected projection angles in subset S. When including a smoothing prior, the backprojections need to be reweighed with the proportion of the total number of projection angles N_A and the number of projection angles in the current subset N_S to preserve the relative strength between the data and prior:

$$\Delta\mu_j = \frac{\alpha_j \frac{N_A}{N_S} \sum_i l_{ij}(\hat{y}_i - y_i) + \alpha_j \beta_Q P'(\vec{\mu})}{\frac{N_A}{N_S} \sum_i l_{ij}\hat{y}_i \sum_k l_{ik}\alpha_k + \alpha_j \beta_Q P''}. \quad (11)$$

No attempt was made to optimize the order of the subsets, which were ordered such that each subset had maximum angular difference with its predecessor, and if possible also with all previous subsets, following the advice of Hudson and Larkin[39]. The chosen angles for each subset are listed in table I.

	Subsets S
2	(1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25); (2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24)
5	(1, 6, 11, 16, 21); (5, 10, 15, 20, 25); (3, 8, 13, 18, 23); (2, 7, 12, 17, 22); (4, 9, 14, 19, 24)
12	(1, 13, 25); (12, 24); (6, 18); (9, 21); (3, 15); (8, 20); (2, 14); (7, 19); (11, 23); (5, 17); (10, 22); (4, 16)
25	1; 25; 13; 7; 19; 4; 16; 10; 22; 9; 21; 8; 20; 6; 18; 5; 17; 3; 15; 2; 14; 24; 11; 23; 12

TABLE I. Subset update ordering.

The image-block version of the MLTR algorithm, where the blocks are chosen as planes parallel to the detector ($\mathrm{MLTR_p}$) can be implemented by limiting the voxels μ_j we will update in equation (3) to those within a single plane, i.e. choosing $\alpha_j = 1$ for the voxels inside that plane and $\alpha_j = 0$ for the voxels outside that plane, and repeating this for all planes in the reconstruction

volume. The update equation for plane P then becomes

$$\Delta\mu_j = \frac{\sum_i l_{ij}\left(\hat{y}_i - y_i\right) + \beta_Q P'(\vec{\mu})}{\sum_i l_{ij}\hat{y}_i \sum_{k\in P} l_{ik} + \beta_Q P''} : j \in P. \tag{12}$$

In this equation, the accelerated convergence can be understood by analyzing the denominator. The sum over all voxels k, $\sum_k l_{ik}$ in equations (4) and (6), is reduced to a sum over the voxels in a single plane P, $\sum_{k\in P} l_{ik}$. This means the update step will increase by a factor roughly equal to the number of planes in the reconstruction volume, which is typically between 20 and 80 for breast tomosynthesis, depending on the patient. Splitting the updates in this way does not actually change the computational cost of one update of the complete volume, since for each full iteration the same number of intersection lengths l_{ij} need to be calculated. Because the voxels that are updated simultaneously are only weakly coupled, convergence should not be adversely affected by this simultaneous update[45].

C. Reconstruction with Resolution Recovery

It is possible to use different reconstruction models for each of the image blocks[46], which we did in previous work[1] by including a resolution model with plane-dependent point spread function, to compensate for blur caused by the motion of the x-ray tube during the image acquisition[8]. This required replacing the model in equation (1) by the model in equation (13), with p the plane index in the reconstruction volume, and A^p the smoothing kernel that models the resolution in plane p.

$$\hat{y}_i = b_i \prod_p \sum_n A^p_{in} e^{-\sum_{j\in p} l_{nj}\mu_j} \tag{13}$$

We can then calculate the update step for this model from equation (3), again with $\alpha_j = 1$ for the voxels inside the updated plane, and $\alpha_j = 0$ for the voxels outside that plane. Relying on the assumption that the intersection lengths vary little over the range of smoothing kernel A^P_{in} results in the MLTR_{pr} update step[1]:

$$\Delta\mu_j = \frac{\sum_i l_{ij}\psi^P_i \sum_n A^P_{in}\frac{\hat{y}_n - y_n}{\bar{\psi}^P_n} + \beta_Q P'(\vec{\mu})}{\sum_i l_{ij}\psi^P_i \sum_{k\in P} l_{ik} \sum_n A^P_{in}\frac{\hat{y}_n}{\bar{\psi}^P_n} + \beta_Q P'(\vec{\mu})} \tag{14}$$

for $j \in P$, and with:

$$\psi^p_i = e^{-\sum_{j\in p} l_{ij}\mu_j}, \tag{15}$$

$$\bar{\psi}^p_i = \sum_n A^p_{in}\psi^p_n. \tag{16}$$

Using the resolution model in equation (13) instead of the model in equation (1) improved the detectability of simulated microcalcifications[1], but required a good initialization to counteract limited angle artifacts.

In a preliminary study for this work we examined the effect of the update order when starting from a good initialization, and found that it is more efficient to apply the image updates consistently from bottom to top rather than alternating the order. Since this good initialization is one of the motivations to use the multigrid method, we used the bottom to top updates for all MLTR_{r} and MLTR_{pr} methods in this work.

Instead of the plane-by-plane updates, it is equally possible to use subsets by sequentially calculating the updates for each plane without updating transmission scan $\hat{y}$, using $\sum_k l_{ik}$ instead of $\sum_{k\in P} l_{ik}$ in the denominator, and reweighing the backprojection by ${}^{N_A}/_{N_S}$ as for the OSTR update in equation (11). This gives us the OSTR_{r} update:

$$\Delta\mu_j = \frac{\frac{N_A}{N_S}\sum_{i\in S} l_{ij}\psi^{p(j)}_i \sum_n A^{p(j)}_{in}\frac{\hat{y}_n - y_n}{\bar{\psi}^{p(j)}_n} + \beta_Q P'(\vec{\mu})}{\frac{N_A}{N_S}\sum_{i\in S} l_{ij}\psi^{p(j)}_i \sum_k l_{ik} \sum_n A^{p(j)}_{in}\frac{\hat{y}_n}{\bar{\psi}^{p(j)}_n} + \beta_Q P''}, \tag{17}$$

with S the currently selected subset and $p(j)$ the plane containing voxel j.

D. Multigrid Reconstruction

Multigrid reconstruction, which was first introduced by Ranganath et al.[47] in PET imaging, can accelerate convergence by starting with a coarse grid in the reconstruction volume and then decreasing the voxel spacing in multiple steps during the iterations, initializing the volume with the interpolated values from the previous grid. Further experiments by Pan and Yagle[48] found that this method only accelerates convergence of high frequency components compared to a fixed grid when the volume is locally smooth. Changing the pixel spacing in the projection data together with the grid changes in the image domain, further reduces the computational cost of each iteration[49], and was successfully applied in tomosynthesis reconstruction by Chen and Barner[50].

These methods use the same acquisition model and optimization strategy for all grid sizes, which is sufficient for simple models, but not optimal for more complex models such as equation (13), where the smoothing kernel will be negligible at coarse grid sizes. Next to this, the typical reconstruction grid in DBT is anisotropic, with in-plane voxel spacing identical to the pixels spacing of the detector, and a patient dependent number of 1 mm planes, so that resizing the grid equally in all directions is not practical.

There is a lot of freedom in a multigrid algorithm with regard to when to switch to a finer grid or another update step. Due to the different cost functions in each step, and the non-linearity of iterative methods it is hard to predict which sequence will result in the best reconstruction. Therefore we examine a selection of the parameter space and evaluate the reconstructions to find the schemes that result in the best convergence, measured by: a) the likelihood calculated from equations (10) and (13), b) good and artifact-free convergence in image domain, measured

by the sum of squared residuals (SSR) between the reconstructed image and the original phantom image, and c) good and uniform convergence resulting in position independent visualization of the microcalcifications, which is measured by the contrast and the standard deviation of the contrast measured in different locations.

We set the following constraints when searching the parameter space:

a) A sequence will either use plane-by-plane MLTR updates or ordered subsets, but not both since our simulation experiments indicated that combining them adversely affects convergence.

b) We want to fix the total computational cost of the reconstruction to the equivalent of four full resolution MLTR_{pr} updates and the reconstruction should finish with at least two full resolution MLTR_{pr} or OSTR_{r} updates.

c) Reconstructions start at 8×8 rebinning with at least four MLTR iterations before switching to the MLTR_{p} and MLTR_{pr} updates, or can start with any number of MLTR or OSTR updates before switching to OSTR_{r} updates.

d) The resolution model should not be used in grids rebinned coarser than 2×2 since the width of the convolution mask in the resolution model is typically less than two pixels at 4×4 rebinning, and once the switch is made to use the resolution model, it should be used for all following updates.

We decided to determine the theoretical computational cost of the algorithms at different grid sizes instead of measuring real calculation times because the implementations of the different update steps are not optimized to the same degree. The relative computational cost is proportional to the total number of projections and backprojections, and the number of projection lines in each projection or backprojection. For each of the algorithms, switching to a coarser grid, with 2×2 rebinning in both projection data and reconstruction volume, reduces the number of projection lines, and thus the computational cost, by a factor of four.

The number of projections and backprojections for each of the algorithms is as follows: the MLTR update in equation (4) contains two backprojections ($\sum_i l_{ij}$), and two projections ($\sum_j l_{ij}$). The first projection is used to calculate forward model $\hat{y}$ and the second one to calculate the projection weights $\sum_j l_{ij}\alpha_j$, which are the same in every iteration and can thus be precalculated. This results in a total of three forward or backprojection operations for each iteration.

The MLTR_{p} update in equation (12) has the same number of forward and backprojections as the MLTR update, although they are split between the different planes. Here it is not feasible to precalculate the weighs $\sum_{j \in p} l_{ij}\alpha_j$ since this needs to be done separately for all planes p and requires keeping too much data in computer memory. Thus each iteration has four forward or backprojection operations. Switching to the MLTR_{pr} updates in equation (14) adds another projection to calculate each ψ^p and $\bar{\psi}^p$, resulting in a total of five forward or backprojection operations.

The OSTR update in equation (11) has the same number of forward and backprojections as MLTR, and can also precalculate projection weights $\sum_j l_{ij}\alpha_j$, resulting in a total cost of three forward or backprojections for a full iteration over all subsets. The cost of the OSTR_{r} updates in equation (17) is the same as the MLTR_{pr} updates, except for the fact that the projection weights can be precalculated, resulting in a total cost of four instead of five forward or backprojections.

We should note that for the subset updates the smoothing prior is now applied to the entire volume after each subset. This means that this cost can become significant when using many subsets. Therefore we add an extra cost of $^1/_{25^{\text{th}}}$ of a projection (roughly corresponding to one operation on the reconstruction volume) for each subset when using 5, 12 or 25 subsets.

All relative computational costs for the five updates and four grid sizes that will be used are shown in table II.

rebinning factor	iteration cost		
	MLTR OSTR	MLTR_{p} OSTR_{r}	MLTR_{pr}
1×1	192	256	320
2×2	48	64	80
4×4	12	16	20
8×8	3	4	5

TABLE II. Relative computational cost of one iteration of the included algorithms (without smoothing prior).

E. Phantom & Simulation

Since evaluating all possible multigrid iterations schemes on full sized data would not be feasible, we perform the optimization and evaluation on a simulated acquisition of a three dimensional phantom. This simulation was performed using the geometry in figure 1. The X-axis is placed at chest side of the detector, going left to right, the Y-axis goes front-to-back along the detector, and the Z-axis goes from the detector to the 0° position of the x-ray source. The center of rotation was set 608.5 mm below the x-ray source (r_{s}) and 47 mm above the detector (r_{d}). The phantom was placed 17 mm above the detector (d_{pd}).

The phantom was created by taking a region of $2048 \times 48 \times 600$ cubic voxels (85 μm) out of an ellipsoid with axes of 1024, 1024, and 640 voxels, and filling this area with white noise filtered by a power law filter[51,52] $f(\nu) = \kappa/\nu^{\beta}$, with frequency ν, $\beta = 3$ and $\kappa = 10^{-5}$ mm^{-1}. The values inside the phantom were then rescaled between the attenuation of fat and glandular breast tissue. Three sets of 15 high contrast calcifications modeled as spheres with

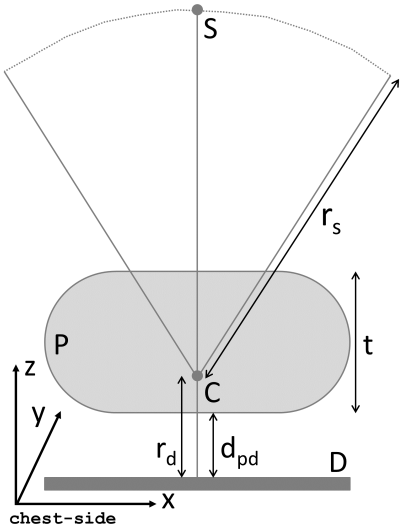

FIG. 1. The simulation geometry (not to scale), with x-ray source S, detector D, phantom P, and center of rotation C.

150 μm diameter were created in a $5 \times 5 \times 5$ μm^3 grid and inserted at 10 mm, 25 mm, and 40 mm above the bottom of the phantom. One slice in the X-Z plane of the phantom is shown in figure 2.

FIG. 2. Slice of the optimization phantom with the locations of the inserted calcifications, shown with exaggerated contrast.

Scatterfree projection data were generated for 25 angles distributed equally between -25° and 25° from the vertical axis, with a photon energy of 20 keV. For each source position α, nine subsources were simulated between $\alpha - 0.115°$ and $\alpha + 0.115°$ to represent the pulsed exposures from a continuously moving x-ray source. The detector consisted of 3584×64 pixels of 85×85 μm^2, which were supersampled by a factor of five during the simulation, and then rebinned to the original size. Blank scan value b_i was set to 2000 for all projection lines i. No noise was added to the projections unless specified otherwise. In those cases Poisson noise is generated in each detector pixel using the pixel value in the noiseless projection data as mean for the Poisson distribution. This noise level is roughly 20% higher than in the patient cases we used in the evaluation. With these projection data, the phantom was reconstructed to a final volume of $2048 \times 48 \times 51$ voxels for all methods, with an in-plane pixel size of 85×85 μm^2 and 1 mm plane separation.

F. Phantom & Patient Evaluation

After the parameter search, we compare the best MLTR$_{\mathrm{pr}}$ and OSTR$_{\mathrm{r}}$ multigrid sequences to 4 iterations of the MLTR$_{\mathrm{pr}}$ algorithm with under-relaxation in the first two iterations[1] and 4 full iterations of OSTR$_{\mathrm{r}}$(25) (OSTR$_{\mathrm{r}}$ with 25 subsets) or 5 full iterations of OSTR$_{\mathrm{r}}$(5) or OSTR$_{\mathrm{r}}$(12).

For the phantom data, we use the same criteria for the parameter search (section II D) for comparison, with the addition of the signal difference to noise ratio (SDNR) of the calcifications in reconstructions from noisy data.

After the phantom evaluation, we reconstruct data from five patient cases using the same methods. Here we calculate the likelihood, measure SDNR for a selection of microcalcifications, and perform a visual evaluation. Because the reconstruction model in equation (13) assumes a mono-energetic acquisition without scatter, all patient data were corrected for scatter[53] and beam hardening[54] before reconstruction.

III. RESULTS

A. Convergence of Ordered Subsets in DBT

We evaluated the four choices of subsets listed in table I together with MLTR and MLTR$_{\mathrm{p}}$ by reconstructing the phantom described in section II E with 1000 iterations, after initialization by 10 MLTR iterations. The difference between the upper bound of the likelihood, defined as

$$L_{max} = \sum_i y_i \ln y_i - y_i - \ln(y_i!) \tag{18}$$

and the posterior likelihood L from equation (10), is plotted in figure 3 to compare the convergence of the six methods. For less than 100 iterations, using more subsets increases the speed of convergence, and MLTR$_{\mathrm{p}}$ performs similarly to OSTR(5) (OSTR with 5 subsets). The value of $L_{max} - L$ after 1000 iterations is listed in table III, where MLTR$_{\mathrm{p}}$ shows the best convergence, followed by OSTR(5), OSTR(2), MLTR, OSTR(12), and OSTR(25).

Algorithm	$L_{max} - L$
MLTR	$1.261{\cdot}10^5$
MLTR$_{\mathrm{p}}$	$1.255{\cdot}10^5$
OSTR(2)	$1.256{\cdot}10^5$
OSTR(5)	$1.255{\cdot}10^5$
OSTR(12)	$1.264{\cdot}10^5$
OSTR(25)	$1.295{\cdot}10^5$

TABLE III. $L_{max} - L$ for MLTR, MLTR$_{\mathrm{p}}$, and OSTR after 1000 iterations.

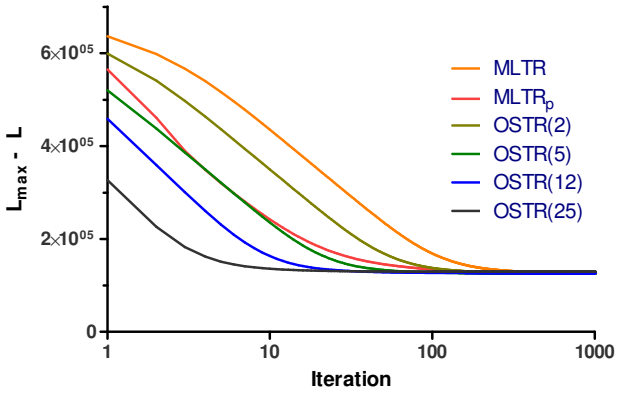

FIG. 3. Difference between the upper bound of the log-likelihood L_{max} and posterior log-likelihood L as a function of iterations for the MLTR, MLTR_p, and OSTR algorithms.

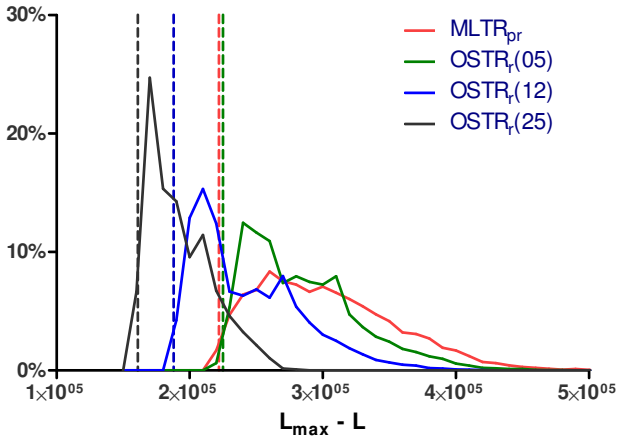

FIG. 4. Histogram of $L_{max} - L$ from phantom reconstructions using all multigrid sequences. The vertical lines indicate the selected sequences from table IV.

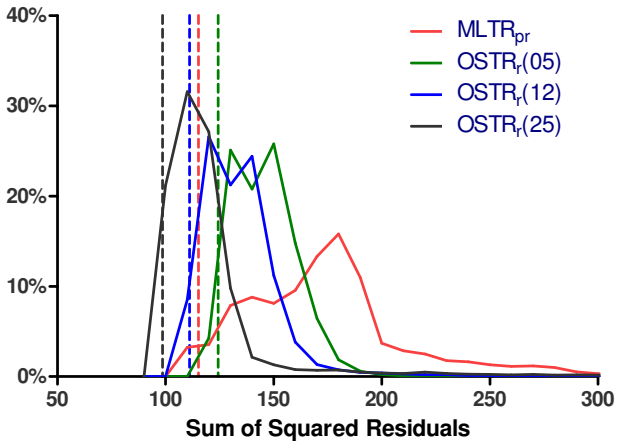

FIG. 5. Histogram of the sum of squared residuals from phantom reconstructions using all multigrid sequences. The vertical lines indicate the selected sequences from table IV.

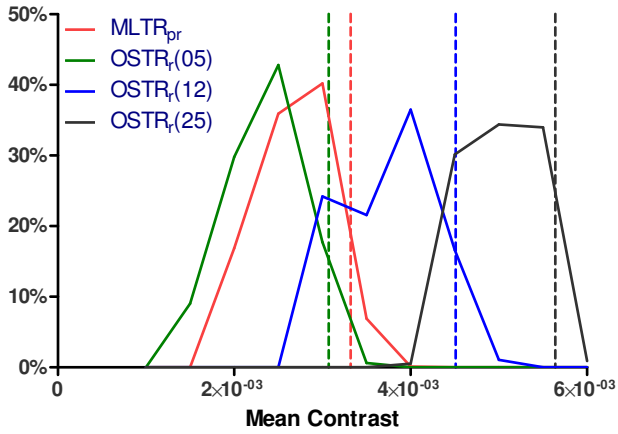

FIG. 6. Histogram of the mean contrast of the calcifications in the phantom from reconstructions using all multigrid sequences. The vertical lines indicate the selected sequences from table IV.

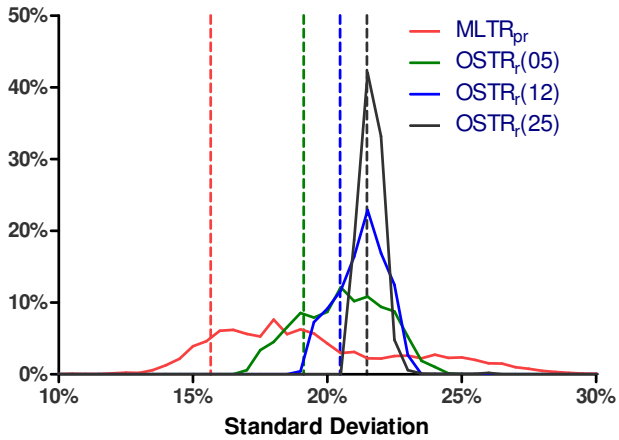

FIG. 7. Histogram of the standard deviation of the mean contrast of calcifications in three locations in the phantom from reconstructions using all multigrid sequences. The vertical lines indicate the selected sequences from table IV.

B. Choice of the Multigrid Sequence

Following the restrictions listed in section II D, we evaluated 8041 multigrid sequences for MLTR_pr, 6339 for $\text{OSTR}_\text{r}(5)$, 4383 for $\text{OSTR}_\text{r}(12)$, and 2613 for $\text{OSTR}_\text{r}(25)$. Figures 4, 5, 6, and 7 show the histograms of: the difference between the upper bound of the log-likelihood L_{max} and posterior log-likelihood L; the sum of squared residuals between the reconstruction and the original phantom; the mean contrast of the calcifications in the phantom; and the standard deviation of the mean contrast of calcifications in three locations in the phantom respectively.

The selected multigrid sequences for MLTR_pr and for OSTR_r with 5, 12, and 25 subsets are shown as vertical lines in figures 4, 5, 6, and 7, and are listed in table IV. The chosen MLTR_pr scores in the top 10% for the four examined parameters, but for OSTR_r there were no se-

Binning	Multigrid Sequences			
8×8	5×MLTR	19×OSTR(5)	9×OSTR(12)	4×OSTR(25)
4×4	11×MLTR$_p$	9×OSTR(5)	4×OSTR(12)	3×OSTR(25)
2×2	7×MLTR$_p$	7×OSTR(5)	7×OSTR(12)	5×OSTR(25)
1×1		1×OSTR(5)	1×OSTR(12)	1×OSTR(25)
1×1	2×MLTR$_{pr}$	2×OSTR$_r$(5)	2×OSTR$_r$(12)	2×OSTR$_r$(25)

TABLE IV. Overview of the selected multigrid sequences, shown as vertical lines in figures 4, 5, 6, and 7.

quences where all four examined parameters were in the top 10%, so the standard deviation of the mean contrast was not taken into account when selecting the sequence.

To make sure these choices remained valid for noisy data, we repeated the analysis for a selection of sequences using 10 noise realizations of the projection data, and found that the average likelihood, sum of squared residuals, and mean contrast from the noise realizations correlated well (R> 0.95) with the results from the noiseless reconstructions. This was not the case when looking at the standard deviation between the means of the three groups of calcifications. Here we found correlations of 0.94 for MLTR$_{pr}$, and 0.73, 0.48, and -0.06 for OSTR$_r$ with 5, 12, and 25 subsets respectively.

C. Phantom Evaluation

Now we compare the selected multigrid (MG) reconstructions to the original single grid (SG) versions. Table V lists the parameters used to choose the multigrid sequence and the signal difference to noise ratio (SDNR) for reconstructions with noise. In these reconstructions we see that the noise level increases from 3.1% for MG MLTR$_{pr}$ to 3.3%, 4.6%, and 5.4% for MG OSTR$_r$ with 5, 12, and 25 subsets respectively. Therefore we add an extra set of OSTR reconstructions with adjusted β_Q and β_{TV} to create reconstructions with matching noise levels.

In table V we see that $L_{max} - L$ and the sum of squared residuals consistently score better for the multigrid version of the algorithms, with the largest gains for the MLTR$_{pr}$ method, and reduced benefits for an increasing number of subsets in the MLTR$_r$ methods. The mean contrast of the included calcifications improves in the noiseless multigrid reconstructions, except for 25 subsets, where the contrast remains unchanged. For reconstructions with noise, contrast improves when using 5 subsets, and changes of 2% or less for the other methods. The standard deviation between the mean contrast of the calcifications in different locations in the phantom are all very similar, with only MG MLTR$_{pr}$ scoring much lower with 15.7%. Average signal difference to noise ratio (SDNR) was calculated for the calcifications in the reconstructions with noise. Only MLTR$_{pr}$ shows a noticeable change with an increase of 15% when switching to multigrid. For the OSTR$_r$ methods SDNR is lowered with 3% or less when switching to multigrid reconstruction.

The tabulated residuals are indicative of the strength of reconstruction artifacts, but not their distribution. Figure 8 shows part of a coronal slice of a reconstruction from a sinogram with noise for the single and multigrid versions of the MLTR$_{pr}$ and OSTR$_r$(5) methods. Here we see that SG MLTR$_{pr}$ overestimates the attenuation close to the detector, which is much reduced in the multigrid variant and not present in the OSTR$_r$(5) images. The main difference between MLTR$_{pr}$ and OSTR$_r$(5) can be seen in the bottom left of the phantom where OSTR$_r$(5) underestimates the attenuation close to the phantom edge to a greater degree than MLTR$_{pr}$. There were almost no visual difference between the artifacts shown for MG OSTR$_r$(5) and those in MG OSTR$_r$(12) and MG OSTR$_r$(25) which were not shown. The structure of the artifacts for SG OSTR$_r$(12) and SG OSTR$_r$(25) is similar to that in SG OSTR$_r$(5) but with decreasing intensity.

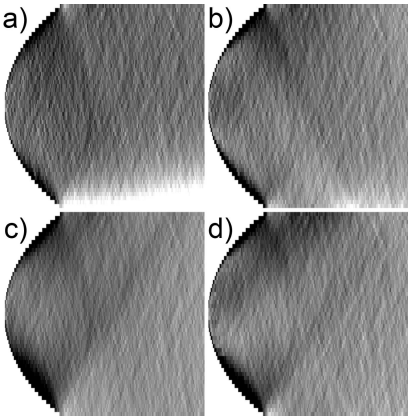

FIG. 8. Artifacts in the coronal plane (the XZ plane in figure 1) for one noise realization of the phantom reconstructions. a) SG MLTR$_{pr}$, b) MG MLTR$_{pr}$, c) SG OSTR$_r$(5), d) MG OSTR$_r$(5).

D. Patient Evaluation

We continue the evaluation of the multigrid methods by reconstructing five patient cases. The average of the difference between the upper bound of the likelihood and the posterior likelihood is listed in table VI and shows that switching from single to multigrid improves convergence for these patient cases. The average SDNR from 14 subtle microcalcifications (between 2 and 5 in each patient case) is listed in table VII. MLTR$_{pr}$ shows the largest gain, with a decreasing advantage for multigrid reconstruction with an increasing number of subsets.

Figures 9 and 10 show an irregular opacity and a cluster of microcalcifications respectively. Here we see switching to multigrid improves contrast most for MLTR$_{pr}$ and OSTR$_r$(5), and that there is little visual difference between single and multigrid versions of OSTR$_r$(12) and OSTR$_r$(25). Figure 11 illustrates the reconstruction artifacts for the different methods. The

Algorithm	$L_{max}-L$ SG	$L_{max}-L$ MG	Residuals SG	Residuals MG	Mean Contrast SG	Mean Contrast MG	Standard Deviation SG	Standard Deviation MG	Mean SDNR SG	Mean SDNR MG
Results from reconstructions without noise:										
$MLTR_{pr}$	$8.66{\cdot}10^5$	$2.22{\cdot}10^5$	339	117	$2.10{\cdot}10^{-3}$	$3.32{\cdot}10^{-3}$	21.4%	15.7%		
$OSTR_r(5)$	$5.99{\cdot}10^5$	$2.25{\cdot}10^5$	181	124	$2.16{\cdot}10^{-3}$	$3.07{\cdot}10^{-3}$	21.8%	19.1%		
$OSTR_r(12)$	$4.61{\cdot}10^5$	$1.88{\cdot}10^5$	144	111	$4.07{\cdot}10^{-3}$	$4.51{\cdot}10^{-3}$	21.7%	20.5%		
$OSTR_r(25)$	$3.46{\cdot}10^5$	$1.61{\cdot}10^5$	113	98.6	$5.65{\cdot}10^{-3}$	$5.64{\cdot}10^{-3}$	21.1%	21.5%		
Results from reconstructions with noise:										
$MLTR_{pr}$	$3.33{\cdot}10^6$	$2.98{\cdot}10^6$	412	134	$2.74{\cdot}10^{-3}$	$2.69{\cdot}10^{-3}$	22.0%	21.8%	1.31	1.50
$OSTR_r(5)$	$3.17{\cdot}10^6$	$2.99{\cdot}10^6$	193	140	$2.49{\cdot}10^{-3}$	$2.68{\cdot}10^{-3}$	23.0%	21.9%	1.43	1.43
$OSTR_r(12)$	$3.04{\cdot}10^6$	$2.95{\cdot}10^6$	174	142	$3.72{\cdot}10^{-3}$	$3.65{\cdot}10^{-3}$	23.9%	23.7%	1.41	1.39
$OSTR_r(25)$	$2.96{\cdot}10^6$	$2.94{\cdot}10^6$	155	144	$4.30{\cdot}10^{-3}$	$4.31{\cdot}10^{-3}$	25.1%	25.2%	1.39	1.38
Results from reconstructions with noise and modified β_Q and β_{TV}:										
$OSTR_r(5)$	$3.20{\cdot}10^6$	$3.01{\cdot}10^6$	194	140	$2.31{\cdot}10^{-3}$	$2.60{\cdot}10^{-3}$	22.6%	21.5%	1.45	1.44
$OSTR_r(12)$	$3.21{\cdot}10^6$	$3.13{\cdot}10^6$	164	139	$2.80{\cdot}10^{-3}$	$2.73{\cdot}10^{-3}$	20.5%	20.7%	1.55	1.51
$OSTR_r(25)$	$3.28{\cdot}10^6$	$3.25{\cdot}10^6$	138	134	$2.66{\cdot}10^{-3}$	$2.67{\cdot}10^{-3}$	19.7%	21.0%	1.50	1.48

TABLE V. Comparison between the single grid (SG) and the selected multigrid (MG) versions of the reconstruction algorithms.

Algorithm	$L_{max}-L$ ($\times 10^8$) SG	SG$_\beta$	MG	MG$_\beta$
$MLTR_{pr}$	1.904		1.755	
$OSTR_r(5)$	1.804	1.809	1.758	1.759
$OSTR_r(12)$	1.797	1.823	1.783	1.805
$OSTR_r(25)$	1.851	1.883	1.843	1.874

TABLE VI. Average $L_{max}-L$ from 5 patients for single and multigrid reconstructions with fixed β (SG and MG) and with adjusted β (SG$_\beta$ and MG$_\beta$).

Algorithm	Mean SDNR SG	SG$_\beta$	MG	MG$_\beta$
$MLTR_{pr}$	3.27		3.84	
$OSTR_r(5)$	3.22	3.37	3.62	3.66
$OSTR_r(12)$	3.35	3.65	3.50	3.83
$OSTR_r(25)$	3.28	3.69	3.28	3.74

TABLE VII. Average SDNR of 14 subtle calcifications from 5 patients for single and multigrid reconstructions with fixed β (SG and MG) and with adjusted β (SG$_\beta$ and MG$_\beta$).

largest difference can be seen between the single and multigrid versions of $MLTR_{pr}$ where the attenuation in the single grid reconstruction is underestimated in most of the volume due to the overestimation close to the detector. The differences between single and multigrid are much smaller for the $OSTR_r$ reconstructions, but the improvement for multigrid is clear from visual evaluation. While the artifact in the MG $OSTR_r$ images are visually

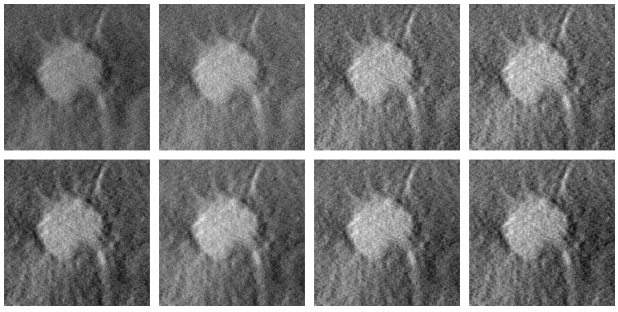

FIG. 9. Irregular opacity for single (top row) and multigrid (bottom row) variants of (from left to right) $MLTR_{pr}$, $OSTR_r(5)$, $OSTR_r(12)$, and $OSTR_r(25)$.

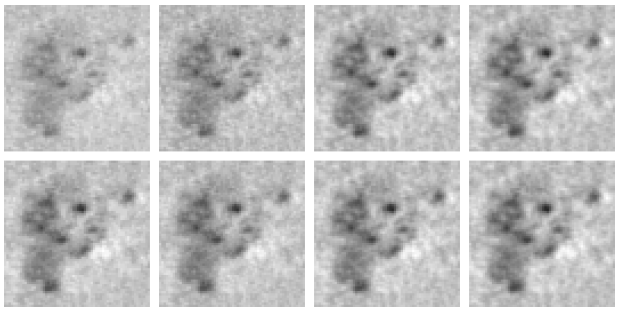

FIG. 10. Microcalcification cluster for single (top row) and multigrid (bottom row) variants of (from left to right) $MLTR_{pr}$, $OSTR_r(5)$, $OSTR_r(12)$, and $OSTR_r(25)$, with inverted grayscale.

similar, but with varying intensity, there are clear differences with the MG $MLTR_{pr}$ reconstruction, shown by the arrows in figure 11. Closer to the detector side (top row), the breast edge is better represented for MG $MLTR_{pr}$, while the reverse is true closer to the top of the reconstruction volume (bottom row). In the center of the volume (middle row), the differences are relatively small, with one artifact appearing for increasing number of subsets near the top of the image.

Next to the changes in contrast and artifacts, we also see a difference in noise levels between the reconstructions in figures 9, 10, and 11. The average noise levels in the five patient cases are 4.1%, 3.9%, 4.3%, and 4.4% for $MLTR_{pr}$, $OSTR_r(5)$, $OSTR_r(12)$, and $OSTR_r(25)$ respectively when using the modified prior strength that resulted in equal noise levels in the phantom reconstructions.

The patient reconstructions with adjusted β were also evaluated by an experienced radiologist specialized in breast imaging. She found the SG $OSTR_r(12)$, SG $OSTR_r(25)$, MG $OSTR_r(12)$ and MG $OSTR_r(25)$ reconstructions to be identical from a diagnostic point of view, but all containing too much noise, making it harder to spot microcalcifications. The multigrid versions of $MLTR_{pr}$ and $OSTR_r(5)$ did improve the contrast of soft tissue structures and microcalcifications while maintain-

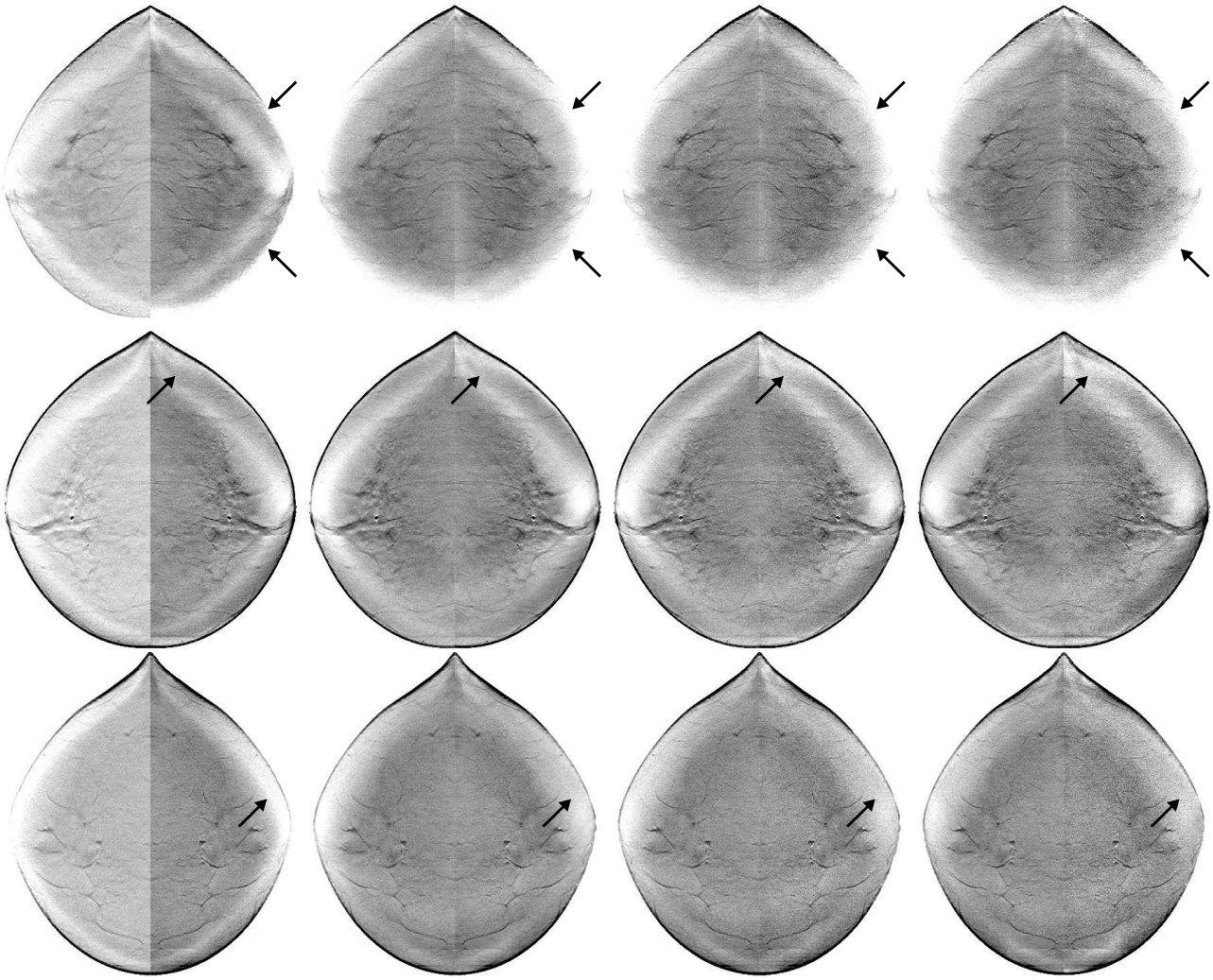

FIG. 11. Reconstruction planes at 15 (top row), 30 (middle row), and 45 mm (bottom row) from the detector cover for single (left facing) and multigrid (right facing) variants of (from left to right) $\mathrm{MLTR_{pr}}$, $\mathrm{OSTR_r}(5)$, $\mathrm{OSTR_r}(12)$, and $\mathrm{OSTR_r}(25)$. The window is set from 0.03 to 0.06 $\mathrm{mm^{-1}}$ with inverted grayscale.

ing a manageable amount of noise. These two multigrid reconstructions were also considered diagnostically equal and were preferred over the other methods. The image artifacts were not considered a fundamental problem since the information in those areas can still be examined by changing the window settings, but there is a strong preference not to have to do this because it significantly increases the time needed to read each reconstruction volume.

IV. DISCUSSION

We started our evaluation of multigrid reconstruction for DBT by comparing the relative convergence speed of $\mathrm{MLTR_p}$ and OSTR (figure 3). Although only $\mathrm{MLTR_p}$ was guaranteed to converge, differences after 1000 iterations were very small thanks to the addition of the smoothing priors that strongly reduce the limit cycle solution for the OSTR reconstructions. Even though these methods converge to slightly different solutions, we do not expect this to be noticeable when stopping after a few iterations.

The selected multigrid sequences in table IV seem like logical choices, with several iterations at each grid level before moving to the next size. When looking at the full data, we see that suboptimal sequences concentrate too many iterations at a single grid size rather than using a more equal distribution. This type of sequences would not be considered when setting a performance measure to decide when to switch grids as in the original work of Ranganath[47]. By fixing the computational cost instead, we could compare all methods on an equal footing but at the cost of including many suboptimal possibilities in the first step of our analysis.

This computational cost was chosen to allow an $\mathrm{MLTR_{pr}}$ reconstruction on current hardware (e.g. nVidia Tesla C2075) within 5 minutes, which is reasonable for

the current clinical use where images are evaluated before the patient leaves. This cost will of course need to be re-evaluated when the tomosynthesis system changes, the reconstruction workstation is upgraded with a new CPU or GPU, or for other clinical settings, such as for example a screening environment where images are read in batches several days after the examination.

While we used reconstructions from noiseless data to choose the multigrid sequences, we found good correlations between these parameters and the average of the same parameter calculated from multiple reconstructions from noisy data, with the exception of the standard deviation of the mean contrast of the calcifications in three different locations in the phantom. Here we found good correlation for MG $\mathrm{MLTR_{pr}}$, which produces a large range of values, both in reconstructions from noiseless data (figure 7) and in reconstructions from noisy data (histogram not shown). The MG $\mathrm{OSTR_r}$ reconstructions also show the same behavior in the noiseless and noisy reconstructions, with results getting more clumped together for increasing subsets. This makes the lack of correlation less of a problem, since the exact choice of multigrid sequence will only have a small effect on the resulting value of the standard deviation. We found that there were no examined multigrid sequences for $\mathrm{OSTR_r}$ that scored in the lowest 10% of values for this standard deviation while at the same time scoring in the best 10% for the other parameters, and we decided to ignore this parameter when choosing the multigrid sequence. In the end this doesn't seem like a bad choice when we compare the scores of the standard deviation for multigrid $\mathrm{MLTR_{pr}}$ and $\mathrm{OSTR_r}$ with modified beta in table V, where the values lie in a small range from 20.7% to 21.8%. From these results it appears that the choices we examined in this paper had little influence on this standard deviation, but despite this we think it is a valid performance measure to expect identical calcifications in different parts of the phantom to be reconstructed with the same contrast.

If we compare the histograms in figures 4, 5, and 6 it seems that the chosen multigrid sequence with 25 subsets will easily outperform the $\mathrm{MLTR_{pr}}$ sequence and those with 5 and 12 subsets. After introducing noise in the projection data and adjusting the prior weights to produce similar noise levels in the reconstructions, only small differences remain. With different values of β_Q and β_{TV}, the algorithms are optimizing different cost functions and it is no longer possible to directly compare the posterior likelihood. The results indicate that even with the multigrid acceleration, the convergence is still incomplete, and more so if fewer subsets are used. Since high frequencies are slower to converge, this early stopping of the iterations has a smoothing effect. Since MG $\mathrm{OSTR_r}(25)$ has the strongest convergence, it needs more smoothing by the prior to obtain a matched noise level.

In general we find the same conclusions when looking at the patient data, with only relatively small differences between the four multigrid sequences with adjusted prior strength. On closer examination there are some noticeable differences such as the ranking of the average posterior likelihood for the multigrid sequences with the original fixed prior strength (third column in table VI) which shows that $\mathrm{MLTR_{pr}}$ has reached better convergence than the three $\mathrm{OSTR_r}$ reconstructions, and that convergence decreases with an increasing number of subsets, which is opposite to the results for the phantom. We also find that the modified prior strength is not entirely effective for the patient cases. The range of the average noise levels is reduced from 4.1%–6.9% to 3.9%–4.4%, but this differences is clearly visible in the patient images, where the $\mathrm{OSTR_r}(12)$ and $\mathrm{OSTR_r}(25)$ reconstructions are considered too noisy for easy diagnosis.

We suspect that the reason for this discrepancy between phantom and patient reconstructions is due to the differences in noise composition and data consistency. The simulated projection data contain pure Poisson noise and are fully consistent, while the patient projection data contain other noise contributions from the detector and possible data inconsistencies from imperfect scatter and beam hardening corrections and possible patient motion.

V. CONCLUSIONS

In this work we examined the combination of a multigrid reconstruction strategy with block-iterative updates both in the form of plane-by-plane updates and subsets while optimizing a likelihood cost function with a position dependent resolution model. We found improved convergence and a reduction in artifacts for our chosen multigrid reconstructions compared to the single grid reconstructions with equivalent computational cost, although there was a diminishing return for an increasing number of subsets. Ordered subsets provided a stronger acceleration, but when comparing the multigrid reconstructions at matched noise levels, plane-by-plane updating and using ordered subsets produced very similar performance.

ACKNOWLEDGMENTS

We would like to thank Prof. Dr. C. Van Ongeval for her help in evaluating the patient reconstructions, the reviewers for their helpful comments, and Siemens Healthcare for their financial support.

* johan.nuyts@uzleuven.be

[1] K. Michielsen, K. Van Slambrouck, A. Jerebko, and J. Nuyts, "Patchwork reconstruction with resolution modeling for digital breast tomosynthesis," Med. Phys. **40**, 031105 (2013).

[2] R.E. Bird, T.W. Wallace, and B.C. Yankaskas, "Analysis of cancers missed at screening mammography," Radiology **184**, 613–617 (1992).

[3] N. Houssami and P. Skaane, "Overview of the evidence on digital breast tomosynthesis in breast cancer detection," Breast **22**, 101–108 (2013).

[4] T. Wu, R.H. Moore, E.A. Rafferty, and D.B. Kopans, "A comparison of reconstruction algorithms for breast tomosynthesis," Med. Phys. **31**, 2636–2647 (2004).

[5] Y. Zhang, H.P. Chan, B. Sahiner, J. Wei, M.M. Goodsitt, L.M. Hadjiiski, J. Ge, and C. Zhou, "A comparative study of limited-angle cone-beam reconstruction methods for breast tomosynthesis," Med. Phys. **33**, 3781–3795 (2006).

[6] D. Van de Sompel, M. Brady, and J. Boone, "Task-based performance analysis of FBP, SART and ML for digital breast tomosynthesis using signal CNR and channelised hotelling observers," Med. Image Anal. **15**, 53–70 (2010).

[7] I. Sechopoulos, "A review of breast tomosynthesis. part i. the image acquisition process," Med. Phys. **40**, 014301 (2013).

[8] N.W. Marshall and H. Bosmans, "Measurements of system sharpness for two digital breast tomosynthesis systems," Phys. Med. Biol. **57**, 7629–7650 (2012).

[9] T. Mertelmeier, J. Orman, W. Haerer, M.K. Dudam, M.J. Flynn, and J. Hsieh, "Optimizing filtered backprojection reconstruction for a breast tomosynthesis prototype device," Proc. SPIE **6142**, 61420F (2006).

[10] J. Orman, T. Mertelmeier, and W. Haerer, "Adaptation of image quality using various filter setups in the filtered backprojection approach for digital breast tomosynthesis," in *Digital Mammography*, edited by S.M. Astley, M. Brady, C. Rose, and R. Zwiggelaar (Springer, Berlin, 2006), pp. 175–182.

[11] J. Ludwig, T. Mertelmeier, H. Kunze, and Wo. Härer, "A novel approach for filtered backprojection in tomosynthesis based on filter kernels determined by iterative reconstruction techniques," in *Digital Mammography*, edited by E.A. Krupinski (Springer, Berlin, 2008), pp. 612–620.

[12] K. Erhard, M. Grass, S. Hitziger, A. Iske, and T. Nielsen, "Generalized filtered back-projection for digital breast tomosynthesis reconstruction," Proc. SPIE **8313**, 831320 (2012).

[13] S. Abdurahman, F. Dennerlein, A. Jerebko, A. Fieselmann, and T. Mertelmeier, "Optimizing high resolution reconstruction in digital breast tomosynthesis using filtered back projection," in *Breast Imaging*, edited by H. Fujita, T. Hara, and C. Muramatsu (Springer, Berlin, 2014), pp. 520–527.

[14] Y. Zhang, H.P. Chan, B. Sahiner, J. Wei, M.M. Goodsitt, L.M. Hadjiiski, J. Ge, C. Zhou, M.J. Flynn, and J. Hsieh, "Tomosynthesis reconstruction using the simultaneous algebraic reconstruction technique (SART) on breast phantom data," Proc. SPIE **6142**, 614249 (2006).

[15] E.Y. Sidky, X. Pan, I.S. Reiser, R.M. Nishikawa, R.H. Moore, and D.B. Kopans, "Enhanced imaging of microcalcifications in digital breast tomosynthesis through improved image-reconstruction algorithms," Med. Phys. **36**, 4920–4932 (2009).

[16] K. Lange and J.A. Fessler, "Globally convergent algorithms for maximum a posteriori transmission tomography," IEEE Trans. Image Process. **4**, 1430–1438 (1995).

[17] T. Wu, A. Stewart, M. Stanton, T. McCauley, W. Phillips, D.B. Kopans, R.H. Moore, J.W. Eberhard, B.H. Opsahl-Ong, L.T. Niklason, and M.B. Williams, "Tomographic mammography using a limited number of low-dose cone-beam projection images," Med. Phys. **30**, 365–380 (2003).

[18] G. Pratx and L. Xing, "GPU computing in medical physics: A review," Med. Phys. **38**, 2685–2697 (2011).

[19] F. Xu and K. Mueller, "Accelerating popular tomographic reconstruction algorithms on commodity PC graphics hardware," IEEE Trans Nucl. Sci. **52**, 654–663 (2005).

[20] I. Goddard, T. Wu, S. Thieret, A. Berman, H. Bartsch, M.J. Flynn, and J. Hsieh, "Implementing an iterative reconstruction algorithm for digital breast tomosynthesis on graphics processing hardware," Proc. SPIE **6142**, 61424V (2006).

[21] D. Matenine, Y. Goussard, and P. Després, "GPU-accelerated regularized iterative reconstruction for few-view cone beam CT," Med. Phys. **42**, 1505–1517 (2015).

[22] A. Eklund, P. Dufort, D. Forsberg, and S.M. LaConte, "Medical image processing on the GPU – Past, present and future," Med. Image Anal. **17**, 1073–1094 (2013).

[23] J. Barzilai and J.M. Borwein, "Two-Point step size gradient methods," IMA J. Numer. Anal. **8**, 141–148 (1988).

[24] Y. Nesterov, "Smooth minimization of non-smooth functions," Math. Program. **103**, 127–152 (2005).

[25] T. Goldstein and S. Osher, "The split bregman method for L1-Regularized problems," SIAM J. Imag. Sci. **2**, 323–343 (2009).

[26] A. Chambolle and T. Pock, "A First-Order Primal-Dual algorithm for convex problems with applications to imaging," J. Math. Imaging Vis. **40**, 120–145 (2010).

[27] E.Y. Sidky, J.H. Jorgensen, and X. Pan, "Convergence of iterative image reconstruction algorithms for digital breast tomosynthesis," in *2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record*, edited by B. Yo (IEEE, Piscataway NJ, 2012), pp. 3394–3396.

[28] Y. Park, H. Cho, U. Je, D. Hong, M. Lee, C. Park, H. Cho, S. Choi, and Y. Koo, "Compressed-sensing (CS)-based digital breast tomosynthesis (DBT) reconstruction for low-dose, accurate 3D breast x-ray imaging," J. Korean Phys. Soc. **65**, 565–571 (2014).

[29] S. Ramani and J.A. Fessler, "A splitting-based iterative algorithm for accelerated statistical x-ray CT reconstruction," IEEE Trans. Med. Imaging **31**, 677–688 (2012).

[30] D. Kim, S. Ramani, and J.A. Fessler, "Combining ordered subsets and momentum for accelerated x-ray CT image reconstruction," IEEE Trans. Med. Imaging **34**, 167–178 (2015).

[31] J.A. Fessler, E.P. Ficaro, N.H. Clinthorne, and K. Lange, "Grouped-coordinate ascent algorithms for penalized-likelihood transmission image reconstruction," IEEE Trans. Med. Imaging **16**, 166–175 (1997).

[32] J. Nuyts, B. De Man, P. Dupont, M. Defrise, P. Suetens, and L. Mortelmans, "Iterative reconstruction for helical CT: a simulation study," Phys. Med. Biol. **43**, 729–737

(1998).
[33] K. Lange, D.R. Hunter, and I. Yang, "Optimization transfer using surrogate objective functions," J. Comput. Graph. Stat. **9**, 1–20 (2000).
[34] K. Van Slambrouck and J. Nuyts, "Reconstruction scheme for accelerated maximum likelihood reconstruction: The patchwork structure," IEEE Trans. Nucl. Sci. **61**, 173–181 (2014).
[35] T. Hebert and R.M. Leahy, "A generalized EM algorithm for 3-D bayesian reconstruction from poisson data using gibbs priors," IEEE Trans. Med. Imaging **8**, 194–202 (1989).
[36] E. Mumcuoğlu, R.M. Leahy, and S.R. Cherry, "Bayesian reconstruction of PET images: methodology and performance analysis," Phys. Med. Biol. **14**, 1777–1807 (1996).
[37] A. Sawatzky, C. Brune, F. Wubbeling, T. Kosters, K. Schafers, and M. Burger, "Accurate EM-TV algorithm in PET with low SNR," in *IEEE Nuclear Science Symposium Conference Record* edited by P. Selin (IEEE, Piscataway NJ, 2008), pp. 5133–5137.
[38] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," IEEE Trans. Image Process. **18**, 2419–2434 (2009).
[39] H.M. Hudson and R.S. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," IEEE Trans. Med. Imaging **13**, 601–609 (1994).
[40] H. Erdogan and J.A. Fessler, "Ordered subsets algorithms for transmission tomography," Phys. Med. Biol. **44**, 2835–2851 (1999).
[41] C. Kamphuis and F.J. Beekman, "Accelerated iterative transmission CT reconstruction using an ordered subsets convex algorithm," IEEE Trans. Med. Imaging **17**, 1101–1105 (1998).
[42] G. Wu, J.G. Mainprize, and M.J. Yaffe, "Characterization of a constrained paired-view technique in iterative reconstruction for breast tomosynthesis," Med. Phys. **40**, 101901 (2013).
[43] K. Sauer and C. Bouman, "A local update strategy for iterative reconstruction from projections," IEEE Trans. Signal Process. **41**, 534–548 (1993).
[44] J.A. Fessler and A.O. Hero, "Space-alternating generalized expectation-maximization algorithm," IEEE Trans. Signal Process. **42**, 2664–2677 (1994).
[45] J.A. Fessler and D. Kim, "Axial block coordinate descent (ABCD) algorithm for x-ray CT image reconstruction," in *Proceedings of the 11th international Meeting on Fully Three-dimensional Image Reconstruction in Radiology and Nuclear Medicine* edited by M. Kachelriess and M. Rafecas (www.fully3d.org, 2011), pages 262–265.
[46] K. Van Slambrouck and J. Nuyts, "Metal artifact reduction in computed tomography using local models in an image block-iterative scheme, " Med. Phys. **39**, 7080–7093 (2012).
[47] M.V. Ranganath, A.P. Dhawan, and N. Mullani, "A multigrid expectation maximization reconstruction algorithm for positron emission tomography," IEEE Trans. Med. Imaging **7**, 273–278 (1988).
[48] T.S. Pan and A.E. Yagle, "Numerical study of multigrid implementations of some iterative image reconstruction algorithms," IEEE Trans. Med. Imaging **10**, 572–588 (1991).
[49] A. Raheja, T.F. Doniere, and A.P. Dhawan, "Multiresolution expectation maximization reconstruction algorithm for positron emission tomography using wavelet processing," IEEE Trans. Nucl. Sci. **46** 594–602 (1999).
[50] P. Chen and K.E. Barner, "Three-dimensional multiresolution statistical reconstruction for tomosynthesis," in *IEEE International Symposium on Biomedical Imaging: Macro to Nano (Conference Record)* edited by R.M. Leahy and C. Roux (IEEE, Piscataway NJ, 2004), pp. 559–562.
[51] K.G. Metheany, C.K. Abbey, N. Packard, and J.M. Boone, "Characterizing anatomical variability in breast CT images," Med. Phys. **35**, 4685–4694 (2008).
[52] E. Engstrom, I. Reiser, and R. Nishikawa, "Comparison of power spectra for tomosynthesis projections and reconstructed images," Med. Phys. **36**, 1753–1758 (2009).
[53] B. Liu, T. Wu, R.H. Moore, and D.B. Kopans, "Monte carlo simulation of x-ray scatter based on patient model from digital breast tomosynthesis," Proc. SPIE **6142**, 61421N (2006).
[54] G.T. Herman, "Correction for beam hardening in computed tomography," Phys. Med. Biol. **24**, 81–106 (1979).