Digital Breast Tomosynthesis Reconstruction with Detector Blur and Correlated Noise

Jiabei Zheng, Jeff Fessler, Heang-Ping Chan

University of Michigan

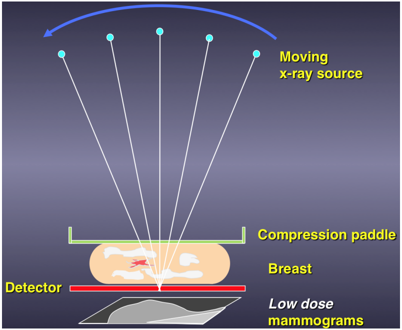
CT Meeting

July 10, 2016

Outline

1. Background
2. Reconstruction method
   - Formulating the reconstruction problem
   - Solving the reconstruction problem
3. Phantom and patient results
   - DBT system
   - Experimental phantom
   - Patient case
4. Conclusion and future work

Introduction to digital breast tomosynthesis (DBT)



- DBT has been developed to deal with overlapping tissue in mammogram.
- Goal: improve DBT reconstruction by modeling detector blur and correlated noise.
- A first step towards systematic MBIR for DBT.
- Hope to improve image quality for both subtle microcalcifications and mass spiculations.

Outline

Combining Lambert-Beer Law and Detector Blur

Measurement model combines detector blur and Lambert-Beer law:

$$\bar{\boldsymbol{Y}}_i = I_0 \boldsymbol{B}_i \, \mathrm{e}^{-\boldsymbol{A}_i \boldsymbol{f}}$$

- $\bar{\boldsymbol{Y}}_i$: expected projection view for the ith view angle.
- $\boldsymbol{f}$: unknown 3D attenuation image.
- $\boldsymbol{A}_i$: forward projector for ith view angle.
- $\boldsymbol{B}_i$: the blurring operation:
  - Allowed to be projection view angle dependent.
  - Assumed linear shift-invariant within each projection.
  - Determined from published system MTF [1]
- I_0: expected projection value in absence of imaged object
  (Can be a constant or a diagonal matrix for nonuniform flux.)
- monoenergetic approximation
- $\mathrm{e}^{-\boldsymbol{x}}$ for vector $\boldsymbol{x}$ denotes element-wise exponentiation

Measurement model approximation for DBT

- We prefer the reconstruction problem to have a quadratic data-fit term.
- The non-diagonal blur matrix $\boldsymbol{B}_i$ before the exponential is complicating.
- In DBT, we assume that the image $\boldsymbol{f}$ is composed of two parts:

$$\boldsymbol{f} = \boldsymbol{f}_{\text{background}} + \boldsymbol{f}_{\text{signal}}$$

  - low-frequency background $\boldsymbol{f}_{\text{background}}$ is approximately uniform within the support of the blurring kernel: $\boldsymbol{B}_i \boldsymbol{A}_i \boldsymbol{f}_{\text{background}} \approx \boldsymbol{A}_i \boldsymbol{f}_{\text{background}}$
  - small structures $\boldsymbol{f}_{\text{signal}}$ (such as MC) contribute very little to the projection value: $\boldsymbol{A}_i \boldsymbol{f}_{\text{signal}} \ll \mathbf{1}$
- Combining yields the simpler approximation (for DBT, not CT):

$$\bar{\boldsymbol{Y}}_i = I_0 \boldsymbol{B}_i \, \mathrm{e}^{-\boldsymbol{A}_i \boldsymbol{f}} \approx I_0 \, \mathrm{e}^{-\boldsymbol{B}_i \boldsymbol{A} \boldsymbol{f}} \, .$$

  (*cf.* exponential edge-gradient effect [2])
- Thus the expected log-transformed projection is approximately linear:

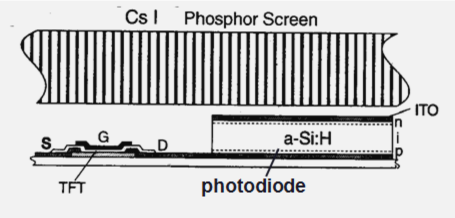
$$\boldsymbol{y}_i \triangleq \log(I_0 / \boldsymbol{Y}_i) \Longrightarrow \bar{\boldsymbol{y}}_i \approx \boldsymbol{B}_i \boldsymbol{A}_i \boldsymbol{f}.$$

Correlated noise: The Covariance Matrix

Cost function needs the measurement covariance matrix.

Physics of CsI phosphor / a:Si Active Matrix Flat Panel Detector:

X-ray photons $\Rightarrow$ Visible light photons $\Rightarrow$ Electronic signal (measured)



Quantum noise depends on detector blur but electronic noise does not.

Covariance matrix for the ith projection view is *non-diagonal* due to blur:

$$\boldsymbol{K}_i = \boldsymbol{B}_i \boldsymbol{K}_i^{\mathrm{q}} \boldsymbol{B}_i' + \boldsymbol{K}_i^{\mathrm{r}}$$

- $\boldsymbol{K}_i^{\mathrm{q}}$: diagonal covariance of quantum noise
- $\boldsymbol{K}_i^{\mathrm{r}}$: diagonal covariance of readout noise

(See related CT work of Tilley, Siewerdsen, Stayman [3], [4], [5], [6], [7].)

DBT Reconstruction Problem

Assuming $\boldsymbol{y}_i$ has approximately a Gaussian distribution: $\boldsymbol{y}_i \sim \mathcal{N}(\bar{\boldsymbol{y}}_i, \boldsymbol{K}_i)$ leads to a regularized reconstruction problem with non-diagonal weighting:

$$\begin{aligned}
\hat{\boldsymbol{f}} &= \arg\min_{\boldsymbol{f}} \frac{1}{2} \sum_{i=1}^{m} \|\boldsymbol{y}_i - \boldsymbol{B}_i \boldsymbol{A}_i \boldsymbol{f}\|^2_{\boldsymbol{K}_i^{-1}} + R(\boldsymbol{f}) \\
&= \arg\min_{\boldsymbol{f}} \frac{1}{2} \sum_{i=1}^{m} \|\boldsymbol{S}_i \left(\boldsymbol{y}_i - \boldsymbol{B}_i \boldsymbol{A}_i \boldsymbol{f}\right)\|^2_2 + R(\boldsymbol{f})
\end{aligned}$$

- Regularizer: $R(\boldsymbol{f}) = \beta \sum_k \psi([\boldsymbol{C}\boldsymbol{f}]_k)$
  - $\boldsymbol{C}\boldsymbol{f}$ computes 2D (in-plane) finite differences
  - edge-preserving hyperbola potential: $\psi(z) = \delta^2(\sqrt{1 + (z/\delta)^2} - 1)$
- Inverse matrix square root of noise covariance:

$$\boldsymbol{S}_i \triangleq \boldsymbol{K}_i^{-1/2} = (\boldsymbol{B}_i \boldsymbol{K}_i^{\mathrm{q}} \boldsymbol{B}_i' + \boldsymbol{K}_i^{\mathrm{r}})^{-1/2}.$$

This non-diagonal term is the computational challenge.

Outline

Implementing S_i efficiently in DBT

- Noise covariance $\boldsymbol{K}_i = \boldsymbol{B}_i \boldsymbol{K}_i^{\mathrm{q}} \boldsymbol{B}_i' + \boldsymbol{K}_i^{\mathrm{r}}$ is non-diagonal.
- Implementing $\boldsymbol{S}_i = \boldsymbol{K}_i^{-1/2} \, \boldsymbol{S}_i$ is challenging in general, particularly in body CT where bones etc. cause very nonuniform noise.
- In DBT, compressed breasts have fairly uniform thickness, mainly composed of soft tissue.
- Key idea: we approximate quantum noise as a constant for all detector elements for each projection view:

$$\boldsymbol{K}_i^{\mathrm{q}} = \sigma_{i,\mathrm{q}}^2 \boldsymbol{I}.$$

- We also assume all detector elements have similar readout noise variance for each projection view:

$$\boldsymbol{K}_i^{\mathrm{r}} = \sigma_{i,\mathrm{r}}^2 \boldsymbol{I}.$$

- Thus the non-diagonal noise covariance matrix simplifies:

$$\boldsymbol{K}_i \approx \sigma_{i,\mathrm{q}}^2 \boldsymbol{B}_i \boldsymbol{B}_i' + \sigma_{i,\mathrm{r}}^2 \boldsymbol{I}.$$

Implementing S_i efficiently for DBT (cont.)

- Now we can simplify inverting the noise covariance matrix:

$$\boldsymbol{K}_i \approx \sigma_{i,\mathrm{q}}^2 \boldsymbol{B}_i \boldsymbol{B}_i' + \sigma_{i,\mathrm{r}}^2 \boldsymbol{I}.$$

- For periodic boundary conditions, the blur matrix is circulant and diagonalizable by a DFT:

$$\boldsymbol{B}_i = \boldsymbol{Q}^{-1} \boldsymbol{H}_i \boldsymbol{Q}.$$

  - $\boldsymbol{Q}$ is the 2D discrete Fourier Transform (DFT) matrix.
  - $\boldsymbol{H}_i$ is the blur frequency response for the ith view.
- The square-root inverse of the noise covariance simplifies:

$$\boldsymbol{S}_i = \boldsymbol{K}_i^{-1/2} = \boldsymbol{Q}^{-1} (\sigma_{i,\mathrm{q}}^2 \boldsymbol{H}_i \boldsymbol{H}_i' + \sigma_{i,\mathrm{r}}^2 \boldsymbol{I})^{-1/2} \boldsymbol{Q}.$$

- Multiplying $\boldsymbol{S}_i$ by a vector is a simple (high-pass) filter using FFTs.
- No iterative method for matrix inversion required for DBT!
  - $\sigma_{i,\mathrm{q}}^2$ estimated using a Lucite slab of appropriate thickness.
  - $\sigma_{i,\mathrm{r}}^2$ estimated from the dark current image of the detector.

Overall cost function

- Overall cost function for DBT image reconstruction:

$$\hat{\boldsymbol{f}} = \arg\min_{\boldsymbol{f}} \Psi(\boldsymbol{f}), \quad \Psi(\boldsymbol{f}) \triangleq \frac{1}{2}\sum_{i=1}^{m} \|\boldsymbol{S}_i\left(\boldsymbol{y}_i - \boldsymbol{B}_i\boldsymbol{A}_i\boldsymbol{f}\right)\|_2^2 + R(\boldsymbol{f})$$
$$= \frac{1}{2}\|\tilde{\boldsymbol{y}} - \tilde{\boldsymbol{A}}\boldsymbol{f}\|_2^2 + R(\boldsymbol{f})$$

- Prewhitened projection data via FFT-based filtering: $\tilde{\boldsymbol{y}} \triangleq \begin{bmatrix} \boldsymbol{S}_1\boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{S}_m\boldsymbol{y}_m \end{bmatrix}$.
- Prewhitened system matrix (for analysis): $\tilde{\boldsymbol{A}} \triangleq \begin{bmatrix} \boldsymbol{S}_1\boldsymbol{B}_1\boldsymbol{A}_1 \\ \vdots \\ \boldsymbol{S}_m\boldsymbol{B}_m\boldsymbol{A}_m \end{bmatrix}$.
- Hessian matrix for data-fit term: $\tilde{\boldsymbol{A}}'\tilde{\boldsymbol{A}} = \sum_{i=1}^{m} \boldsymbol{A}_i'\boldsymbol{B}_i'\boldsymbol{S}_i'\boldsymbol{S}_i\boldsymbol{B}_i\boldsymbol{A}_i$.

Algorithm: Separable Quadratic Surrogates (SQS)

- Cost function: $\Psi(\boldsymbol{f}) = \frac{1}{2}\|\tilde{\boldsymbol{y}} - \tilde{\boldsymbol{A}}\boldsymbol{f}\|_2^2 + R(\boldsymbol{f})$
- Both the quadratic data-fit term and the regularizer are convex.
- To apply SQS we need upper bounds on their Hessians [8].
- As usual for SQS: $\nabla^2 R(\boldsymbol{f}) \preceq \beta \boldsymbol{C}'\boldsymbol{C} \preceq \beta\,\mathrm{diag}\left\{|\boldsymbol{C}|'\,|\boldsymbol{C}|\,\boldsymbol{1}\right\} = 8\beta\boldsymbol{I}$.
- Need to find diagonal majorizing matrix $\boldsymbol{D}$ such that $\tilde{\boldsymbol{A}}'\tilde{\boldsymbol{A}} \preceq \boldsymbol{D}$.
- Then modified SQS algorithm for minimizing DBT cost function is:

$$\boldsymbol{f}^{(n+1)} = \boldsymbol{f}^{(n)} - [\boldsymbol{D} + 8\beta\boldsymbol{I}]^{-1}\nabla\Psi(\boldsymbol{f}^{(n)}).$$

- We use ordered subsets (OS), with one view at a time (ala SART), to accelerate early convergence.
- We call this the SQS-DBCN method,
  where DBCN stands for Detector Blur and Correlated Noise.

Diagonal majorizer for Hessian $\tilde{A}'\tilde{A}$

- The usual choice would be $D = \text{diag}\left\{|\tilde{A}|'|\tilde{A}|\mathbf{1}\right\}$.
- Implementing this would be difficult due to negative values in S_i.
- Instead, note that because $H_i H_i' \preceq I$:

$$\begin{aligned} B_i' S_i' S_i B_i &= Q^{-1} H_i' (\sigma_{i,\mathrm{q}}^2 H_i H_i' + \sigma_{i,\mathrm{r}}^2 I)^{-1} H_i Q \\ &\preceq Q^{-1} ((\sigma_{i,\mathrm{q}}^2 + \sigma_{i,\mathrm{r}}^2)^{-1} I) Q = (\sigma_{i,\mathrm{q}}^2 + \sigma_{i,\mathrm{r}}^2)^{-1} I. \end{aligned}$$

That inequality leads to the following diagonal majorizer:

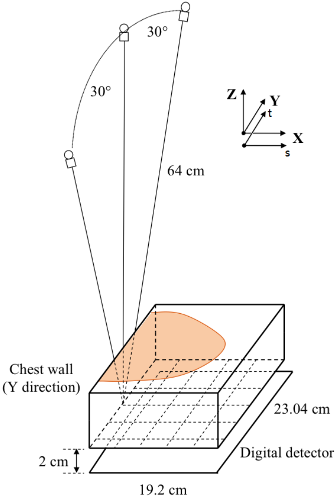
$$\tilde{A}'\tilde{A} = \sum_{i=1}^{m} A_i' B_i' S_i' S_i B_i A_i \preceq \sum_{i=1}^{m} (\sigma_{i,\mathrm{q}}^2 + \sigma_{i,\mathrm{r}}^2)^{-1} A_i' A_i \preceq D$$

$$D \triangleq \sum_{i=1}^{m} (\sigma_{i,\mathrm{q}}^2 + \sigma_{i,\mathrm{r}}^2)^{-1} \,\text{diag}\{A_i' A_i \mathbf{1}\}.$$

- This diagonal majorizer is as easy to implement as usual SQS case.

Outline

Geometry of the DBT System

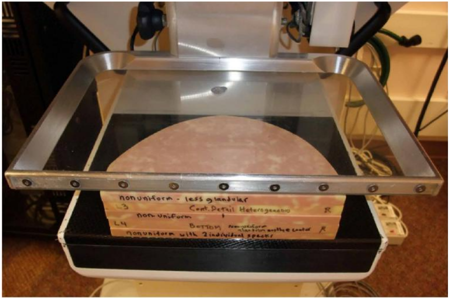


- GE GEN2 prototype DBT system
- 21 projections within $\pm 30°$ sequentially with $3°$ increment.
- Used central 9 views over $\pm 12°$
- Detector resolution = 1920×2304
- Detector pixel size = 0.1mm
- voxels: dx = dy = 0.1mm, dz = 1mm.

- Initialized with uniform image: $\boldsymbol{f}^{(0)} = 0.05/$mm

Outline

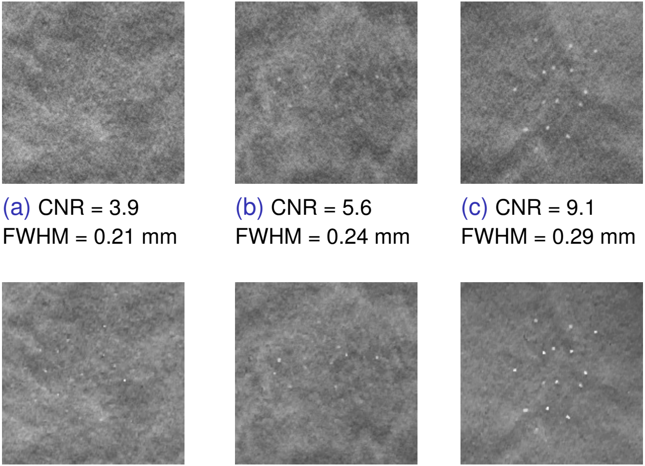
Experimental phantom



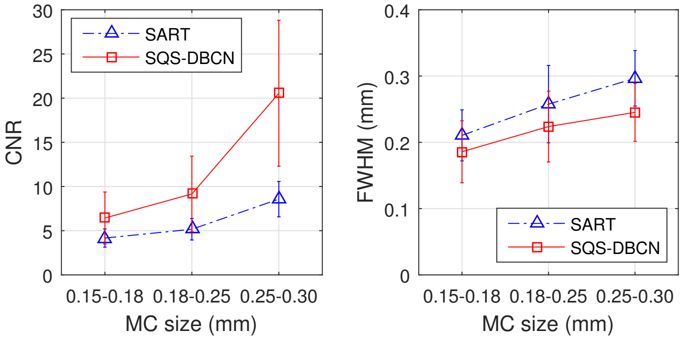
- Stack of five 1-cm-thick 50% adipose/50% glandular heterogeneous slabs that mimic the composition and parenchymal pattern of the breast.
- Clusters of calcium carbonate specks of three nominal size ranges (0.25-0.30mm, 0.18-0.25mm, and 0.15-0.18mm), sandwiched at random locations between the slabs to simulate MCs of different conspicuities.
- SQS-DBCN reconstructions compared with SART (3 iterations) [9], [10].

Reconstructed microcalcification (MC) comparison

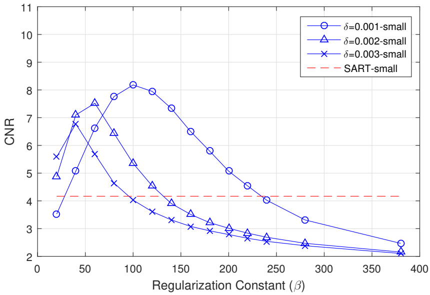
	0.15-0.18mm MC	0.18-0.25mm MC	0.25-0.30mm MC
SART (3 iterations)	(a) CNR = 3.9 FWHM = 0.21 mm	(b) CNR = 5.6 FWHM = 0.24 mm	(c) CNR = 9.1 FWHM = 0.29 mm
SQS-DBCN $\beta = 80$ $\delta = 0.002/\mathrm{mm}$	(d) CNR = 6.9 FWHM = 0.17 mm	(e) CNR = 8.4 FWHM = 0.22 mm	(f) CNR = 20.6 FWHM = 0.25 mm

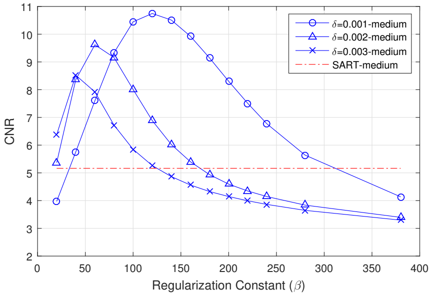


Comparison of CNR and FWHM



- Results averaged over 5 clusters of each size in the phantom:
  - 49 of 0.15-0.18mm MCs
  - 66 of 0.18-0.25mm MCs
  - 64 of 0.25-0.30mm MCs
- SQS-DBCN generally enhanced CNR and decreased FWHM, *i.e.*, the MCs appear sharper.





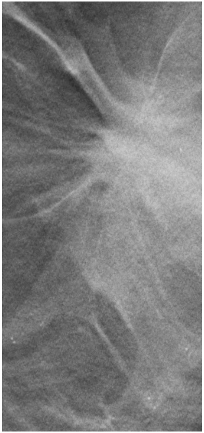
- Average CNR vs. β, δ
- Small and medium MC sizes
- Red lines indicate SART

- For different MC sizes, CNR-optimal β, δ varies
- Reducing δ improves max CNR
- Proposed SQS-DBCN outperforms SART over a large range of parameters.

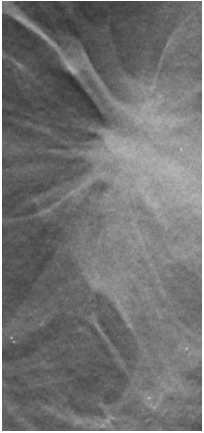
Outline

1. Background
2. Reconstruction method
   - Formulating the reconstruction problem
   - Solving the reconstruction problem
3. Phantom and patient results
   - DBT system
   - Experimental phantom
   - Patient case
4. Conclusion and future work

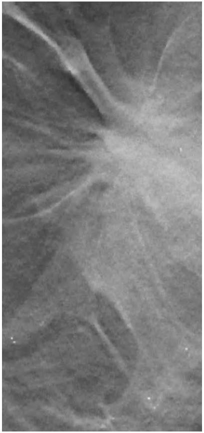
Comparison of reconstructed images for human subject DBT



(a) SART,
3 iterations



(b) SQS-DBCN
$\beta = 80,$
$\delta = 0.002/\mathrm{mm}$



(c) SQS-DBCN
$\beta = 120,$
$\delta = 0.001/\mathrm{mm}$

MC CNR increases from (a) to (c). However, spiculations and tissue textures become more patchy and artificial in (c). Need better figures of merit.

Conclusion

- Proposed DBT reconstruction method incorporates detector blur and a correlated noise model.
  A step towards developing model-based iterative reconstruction for DBT.
- Computationally efficient algorithm adds just one 2D FFT pair per view per iteration.
  - Update per view $\approx$ 2.7 seconds with 8 threads and modified SF [11]
  - One 2D FFT pair $\approx$ 0.03 seconds $\Longrightarrow$ 1% overhead
  - SQS-DBCN for 9 views and 10 iterations $\approx$ 5 min
- Both quantitatively and visually the new SQS-DBCN method can better enhance MCs compared with the SART while preserving the image quality of spiculations and tissue texture, if parameters are chosen well.
- SQS-DBCN method relies on good parameter selection and accurate estimation of noise variance.
- Future work: develop an adaptive parameter selection method, improve estimation of noise variances, generalize model to relax the assumptions.

Bibliography

[1] R. García-Mollá, R. Linares, and R. Ayala, "A study of DQE dependence with beam quality, GE Senographe Essential detector for mammography," in *World congr. on med. phys. and biomed. engin.*, 2009, 886–9.

[2] P. M. Joseph and R. D. Spital, "The exponential edge-gradient effect in x-ray computed tomography," *Phys. med. biol.*, vol. 26, no. 3, 473–87, May 1981.

[3] J. W. Stayman, W. Zbijewski, S. Tilley, and J. H. Siewerdsen, "Generalized least-squares CT reconstruction with detector blur and correlated noise models," in *Proc. spie 9033 medical imaging 2014: Phys. med. im.*, 2014, p. 909 335.

[4] S. Tilley, J. H. Siewerdsen, and J. W. Stayman, "Iterative CT reconstruction using models of source and detector blur and correlated noise," in *Proc. 3rd intl. mtg. on image formation in x-ray ct*, 2014, 363–7.

[5] S. Tilley, J. Siewerdsen, and J. W. Stayman, "Generalized penalized weighted least-squares reconstruction for deblurred flat panel CBCT," in *Proc. intl. mtg. on fully 3d image recon. in rad. and nuc. med*, 2015, 236–9.

[6] S. Tilley, J. H. Siewerdsen, W. Zbijewski, and J. W. Stayman, "Nonlinear statistical reconstruction for flat-panel cone-beam CT with blur and correlated noise models," in *Proc. spie 9783 medical imaging 2016: Phys. med. im.*, 2016, 97830R.

[7] S. Tilley, J. H. Siewerdsen, and J. W. Stayman, "Model-based iterative reconstruction for flat-panel cone-beam CT with focal spot blur, detector blur, and correlated noise," *Phys. med. biol.*, vol. 61, no. 1, 296–319, Jan. 2016.

[8] H. Erdoğan and J. A. Fessler, "Ordered subsets algorithms for transmission tomography," *Phys. med. biol.*, vol. 44, no. 11, 2835–51, Nov. 1999.

[9] A. H. Andersen and A. C. Kak, "Simultaneous algebraic reconstruction technique (SART): A superior implementation of the ART algorithm," *Ultrasonic imaging*, vol. 6, no. 1, 81–94, Jan. 1984.

[10] 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.*, vol. 33, no. 10, 3781–95, Oct. 2006.

[11] Y. Long, J. A. Fessler, and J. M. Balter, "3D forward and back-projection for X-ray CT using separable footprints," *Ieee trans. med. imag.*, vol. 29, no. 11, 1839–50, Nov. 2010.

Supported in part by National Institutes of Health grant R01 CA151443.

Backup: derivation

Assume:

1. $\boldsymbol{f} = \boldsymbol{f}_{\text{background}} + \boldsymbol{f}_{\text{signal}}$
2. smooth back-ground: $\boldsymbol{B}\boldsymbol{A}\boldsymbol{f}_{\text{background}} \approx \boldsymbol{A}\boldsymbol{f}_{\text{background}}$
3. low-contrast details: $\boldsymbol{A}\boldsymbol{f}_{\text{signal}} \ll \mathbf{1}$
4. unity DC response of blur: $\boldsymbol{B}\mathbf{1} = \mathbf{1}$

$$
\begin{aligned}
\boldsymbol{B}\,\mathrm{e}^{-\boldsymbol{A}\boldsymbol{f}} &= \boldsymbol{B}\,\underbrace{\mathrm{e}^{-\boldsymbol{A}(\boldsymbol{f}_{\text{signal}}+\boldsymbol{f}_{\text{background}})}}_{[1]} \\
&= \boldsymbol{B}\,\mathrm{e}^{-\boldsymbol{A}\boldsymbol{f}_{\text{signal}}}\,\mathrm{e}^{-\boldsymbol{A}\boldsymbol{f}_{\text{background}}} \\
&\approx \boldsymbol{B}\,\underbrace{(\mathbf{1} - \boldsymbol{A}\boldsymbol{f}_{\text{signal}})}_{[3]}\,\underbrace{\mathrm{e}^{-\boldsymbol{B}\boldsymbol{A}\boldsymbol{f}_{\text{background}}}}_{[2]} \\
&= \underbrace{(\mathbf{1} - \boldsymbol{B}\boldsymbol{A}\boldsymbol{f}_{\text{signal}})}_{[4]}\,\mathrm{e}^{-\boldsymbol{B}\boldsymbol{A}\boldsymbol{f}_{\text{background}}} \\
&\approx \underbrace{\mathrm{e}^{-\boldsymbol{B}\boldsymbol{A}\boldsymbol{f}_{\text{signal}}}}_{[3]}\,\mathrm{e}^{-\boldsymbol{B}\boldsymbol{A}\boldsymbol{f}_{\text{background}}} = \mathrm{e}^{-\boldsymbol{B}\boldsymbol{A}(\boldsymbol{f}_{\text{signal}}+\boldsymbol{f}_{\text{background}})} = \mathrm{e}^{-\boldsymbol{B}\boldsymbol{A}\boldsymbol{f}}
\end{aligned}
$$

For a 0.2 mm MC with $\mu = 1.5/\mathrm{mm}$, $\boldsymbol{A}\boldsymbol{f}_{\text{signal}} = 0.3 \ll 1$

Noise standard deviation from lucite slab