A signal detection model for quantifying over-regularization in non-linear image reconstruction

Emil Y. Sidky\textsuperscript{1}, John Paul Phillips\textsuperscript{1}, Weimin Zhou\textsuperscript{2}, Greg Ongie\textsuperscript{3}, Juan P. Cruz-Bastida\textsuperscript{1}, Ingrid S. Reiser\textsuperscript{1}, Mark A. Anastasio\textsuperscript{2}, and Xiaochuan Pan\textsuperscript{1}

\textsuperscript{1}Department of Radiology, The University of Chicago, 5841 S. Maryland Ave., Chicago, IL 60637, USA
\textsuperscript{2}Department of Bioengineering, University of Illinois at Urbana-Champaign, 1406 W. Green St., Urbana, IL 61801, USA
\textsuperscript{3}Department of Mathematical and Statistical Sciences, Marquette University, 1313 W. Wisconsin Ave., Milwaukee, WI 53233, USA

Version typeset December 10, 2020

Author to whom correspondence should be addressed. email: sidky@uchicago.edu

Abstract

Purpose: Many useful image quality metrics for evaluating linear image reconstruction techniques do not apply to or are difficult to interpret for non-linear image reconstruction. The vast majority of metrics employed for evaluating non-linear image reconstruction are based on some form of global image fidelity, such as image root mean square error (RMSE). Use of such metrics can lead to over-regularization in the sense that they can favor removal of subtle details in the image. To address this shortcoming, we develop an image quality metric based on signal detection that serves as a surrogate to the qualitative loss of fine image details.

Methods: The metric is demonstrated in the context of a breast CT simulation, where different equal-dose configurations are considered. The configurations differ in the number of projections acquired. Image reconstruction is performed with a non-linear algorithm based on total variation constrained least-squares (TV-LSQ). The resulting images are studied as a function of three parameters: number of views acquired, total variation constraint value, and number of iterations. The images are evaluated visually, with image RMSE, and with the proposed signal-detection based metric. Loss of signal detectability through the image reconstruction process is taken as a quantitative measure of loss of fine details in the image.

Results: Loss of signal detectability is seen to correlate well with the blocky or patchy appearance due to over-regularization with TV-LSQ, and this trend runs counter to the image RMSE metric, which tends to favor the over-regularized images.

Conclusions: The proposed signal detection based metric provides an image quality assessment that is complimentary to that of image RMSE. Using the two metrics in concert may yield a useful prescription for determining CT algorithm and configuration parameters when non-linear image reconstruction is used.
# Contents

I. Introduction 1

II. Methods 3
   II.A. Breast CT simulation .......................... 3
   II.B. TV-LSQ image reconstruction ..................... 4
   II.C. SKE/BKE signal detection model ..................... 5
      II.C.1. Hybrid-CHO ................................ 9
   II.D. Test phantom for visual correspondence .............. 12

III. Results 13
   III.A. Signal detectability as a function of iteration number 14
   III.B. Signal detectability as a function of $N_{\text{views}}$ and $\gamma$ 16
   III.C. Estimation of subject TV and its impact on IQ metric trends 18

IV. Discussion 20

V. Conclusion 21

A Appendix: The CPPD algorithm for TV-LSQ 22
1. **Introduction**

The effort in developing non-linear image reconstruction algorithms for X-ray computed tomography (CT) has been steadily increasing over the past couple of decades. The non-linearity arises from incorporation of some forms of prior information in the reconstruction process or some forms of physics modeling. For example, edge-preserving regularization and spectral response modeling both yield an image reconstruction algorithm that yields images that depend non-linearly on the CT data\(^1,2\). Exploitation of sparsity or transform sparsity also involves non-linear image reconstruction\(^3,4,5,6\). Most recently, deep-learning based data processing is being investigated for generating tomographic images directly from CT projection data using convolutional neural networks (CNNs)\(^7,8\). Such CNNs also process the tomographic data in a non-linear fashion.

While non-linear image reconstruction may allow for accurate image reconstruction in CT systems involving low-dose illumination or sparse sampling, the resulting image characteristics can depend strongly on the scanned object. This object dependence presents a difficult challenge for developing meaningful image quality metrics needed to guide algorithm parameter selection in a non-subjective fashion. As a result, much work on non-linear image reconstruction techniques present images resulting from algorithms where the parameters are tuned by eye. Such an approach may be fine for an initial introduction of a new image reconstruction algorithm or if the CT system/reconstruction parameter space is limited enough where it is feasible to tune by eye. The tune by eye method, however, blunts the impact of advanced image reconstruction because such image reconstruction techniques themselves involve numerous parameters and they aim to broaden the scope of possible CT system configurations – enlarging the parameter space of CT hardware. Attempting to perform comparisons between different non-linear image reconstruction algorithms only complicates the matter further. With a large parameter space, the tune-by-eye method becomes impractical.

Avoiding the subjective tune-by-eye method, many researchers in advanced CT image reconstruction turn to one of three image fidelity metrics in their simulations: root-mean-square-error (RMSE), peak signal-to-noise ratio (PSNR), or structural similarity index (SSIM). These metrics are useful, in a simulation setting, because they present a measure of how close a reconstructed image is to a ground truth image. This information in turn
is useful for investigating the underlying inverse problem. When considering clinical imaging tasks that rely on viewing subtle image features, optimizing system and reconstruction parameters on these global image fidelity metrics can lead to significantly over-regularized images.

One problem is that these image fidelity metrics do not provide a sense of image resolution, noise level, or noise quality. PSNR, from its name, would seem to provide information on the image noise level, but what is called “noise” in PSNR is actually the difference between the reconstructed and truth images, and this difference includes both image noise and deterministic artifacts from either unmodeled non-stochastic physics or insufficient sampling. For non-linear image reconstruction algorithms, concepts such as the point-spread function and the noise power spectrum do not have a simple and direct interpretation as they do for linear systems theory. For example, in non-linear image reconstruction, the resulting image cannot be interpreted as a convolution of a reconstructed point-like object and the underlying true object function. As a result, they are used rarely in the evaluation of non-linear image reconstruction.

In order to prevent over-smoothing by optimizing non-linear image reconstruction solely on image RMSE, an image quality metric is needed that is sensitive to subtle features in the image and that is easy to interpret. To develop such a metric, we turn to signal detection theory, and investigate the use of the ideal observer for a simple signal-known-exactly/background-known-exactly detection task. Signal detection theory has been investigated in the context of evaluation of image reconstruction algorithms. For the present work, the signal is chosen to be a point-like object and its amplitude is set so that it is at the limit of detectability in the CT data space. It is known, that image reconstruction or any other image processing operations cannot increase signal detectability (see pages 829-30 in Barrett and Myers), but it is possible that image reconstruction can reduce signal detectability. Quantifying this loss of detectability is precisely what we would like to use as a measure of over-regularization. Having such a measure would allow optimization of image RMSE under the condition that signal detection is constrained to be at or above a desired set level and thus prevent over-regularization.

The setting for developing this metric, here, is a dedicated breast CT simulation where image reconstruction is performed by total-variation (TV), least-squares optimization (TV-
II. METHODS

The TV-LSQ algorithm is non-linear and it allows for accurate image reconstruction from sparse-view CT data under ideal noiseless conditions. When TV-LSQ is employed for noisy, realistic data it is often reported that the images are patchy or blocky, and one solution to avoid this subject quality is to generalize the TV-norm\textsuperscript{19,20}. For the present work, however, we argue that the patchiness resulting from use of TV regularization can also be a side effect of over-regularization due to parameter optimization using image RMSE. Using the proposed signal detectability metric can help to disallow parameter settings that cause over-regularization and, specifically, the patchy appearance from over-regularization with the TV-norm.

We point out that the patchy appearance for over-regularization with the TV-norm is a somewhat subjective assessment, and therefore the claim that the proposed metric characterizes patchiness quantitatively is also subjective and cannot be proven mathematically. We do attempt to design the simulation so that the subjectivity is limited as much as possible, but in the end the utility of the proposed metric can only be demonstrated by showing metric correspondences with images and it is left to the observer to decide whether this correspondence is useful or not.

In Sec. II, we present the parameters of the breast CT simulation, the details of the TV-LSQ algorithm, and the channelized Hotelling observer (CHO) for the signal-known-exactly/background-known-exactly (SKE/BKE) detection task. The results, presented in Sec. III, demonstrate the correspondences between the proposed signal detection metric and reconstructed images for select parameter settings of the breast CT simulation and TV-LSQ algorithm. The results are discussed in Sec. IV, and finally, we conclude the paper in Sec. V.

II. Methods

II.A. Breast CT simulation

For the studies presented here, we consider a fixed dose simulation, where the number of projections is varied while keeping the total patient exposure constant. The configuration is 2D circular, fan-beam scanning and is representative of the mid-plane slice of a 3D circular cone-beam scan. The mean continuous data function, $g$, is modeled as the X-ray transform
II. METHODS

II.B. TV-LSQ image reconstruction

of the object function

\[ g(\theta, \xi) = Pf(\theta, \xi) = \int_0^\infty f(r_0(\theta) + t \hat{\phi}(\theta, \xi)) dt, \tag{1} \]

where \( f \) represents the continuous object function; \( Pf \) is the continuous X-ray transform of \( f \); \( \theta \) indicate the view angle of the X-ray source; \( \xi \) is the detector bin location on a linear detector; \( r_0(\theta) \) indicates the X-ray source position; and the unit vector \( \hat{\phi} \) points from the X-ray source to the detector bin indicated by \( \xi \), accordingly \( \hat{\phi} \) is a function of \( \theta \) and \( \xi \). The data function is sampled at a variable number of views \( N_{\text{views}} \) and 512 detector bins. The noise level in the measured transmission data is specified by fixing the total number of incident photons to

\[ N_{\text{photons}} = 10^{10}. \]

In the simulations we consider varying \( N_{\text{views}} \) between 128 and 512, and for the maximum end of this range the number of incident photons along each measured ray is approximately 38,000 photons, which is on the low end of actual breast CT systems\textsuperscript{16,21}. To model noise due to the detection of finite numbers of quanta, a Poisson distribution is assumed in the X-ray transmission measurements. Accounting for the logarithm processing needed to arrive at the line-integration model, Eq. (1), we model the noisy discrete data with a Gaussian distribution with mean

\[ \bar{g}_\ell = g(\theta_\ell, \xi_\ell), \tag{2} \]

and variance

\[ \text{Var}(g_\ell) = \left( \frac{N_{\text{photons}}}{N_{\text{views}}} \exp(-g(\theta_\ell, \xi_\ell)) \right)^{-1}, \tag{3} \]

where \( \ell \) is an index for each of the transmission rays in the projection data. It is clear from Eq. (3) that the noise variance decreases with decreasing numbers of views, and there is a tradeoff between \( N_{\text{views}} \) and signal-to-noise ratio in each projection.

II.B. TV-LSQ image reconstruction

In order to formulate the TV-LSQ optimization the continuous data model in Eq. (1) is discretized, taking the form of a large linear system

\[ g = Xf, \]
where the pixelized 512×512 image is represented by \( f \); X-ray projection becomes the matrix \( X \); and the \( N_{\text{views}} \times 512 \) data is denoted by \( g \). Because we consider \( N_{\text{views}} \leq 512 \) this linear system can be under-determined. The TV-LSQ optimization problem is formulated as

\[
\hat{f} = \arg \min_f \frac{1}{2} \| g - X f \|_2^2 \text{ such that } \| (|Df|_{\text{mag}}) \|_1 \leq \gamma,
\]

where \( D \) is the finite differencing approximation to the image gradient; \( |\cdot|_{\text{mag}} \) is the pixelwise magnitude of the spatial gradient vector \( Df \); \( \| (|Df|_{\text{mag}}) \|_1 \) is the image total variation (TV); and \( \gamma \) is the TV constraint value. When the data \( g \) are generated from a test image \( f_{\text{test}} \) with no noise added, the test image can be recovered with TV-LSQ choosing \( \gamma = \| Df_{\text{true}} \|_1 \) for sparse-view sampling with \( N_{\text{views}} < 512 \). The degree of under-sampling permitted depends on the sparsity in the gradient magnitude image (GMI) \( |Df_{\text{test}}|_{\text{mag}} \). This possibility of accurate image reconstruction for sparse-view CT enables the consideration of the CT configurations described in the breast CT simulation.

The TV-LSQ optimization problem can be efficiently solved by the Chambolle-Pock primal-dual (CPPD) algorithm\(^{23,24,25}\). For completeness we provide the pseudocode for this algorithm in Appendix A. We do consider early stopping of the CPPD algorithm and allow the total number of iterations, \( N_{\text{iter}} \), to vary from 10 to 500. At \( N_{\text{iter}} = 500 \) the TV-LSQ problem is solved to a high degree of numerical accuracy for all scan configurations considered in this work.

In total, three parameters are varied in the breast CT simulation: \( N_{\text{views}} \), \( N_{\text{iter}} \), and the TV constraint \( \gamma \). Even under this restricted simulation with three parameters specifying the image, it is difficult to tune-by-eye; not only is the parameter space too large but the image qualities are difficult to compare. As will be seen, quantitative image fidelity metrics such as image RMSE, alone, may not provide a reasonable objective means of image comparison and optimization, particularly when small subtle signals are the features of interest.

## II.C. SKE/BKE signal detection model

To provide an objective metric that characterizes the preservation of subtle details in the TV-LSQ reconstructed images, signal detection theory is employed to measure the loss of signal detectability for an ideal observer model. The design of the detectability metric involves the following steps: select the signal properties such that it is on the border of
II.C. SKE/BKE signal detection model

METHODS

Figure 1: (Left) Background image used for the signal-known-exactly/background-known-exactly detection task. The gray scale window is $[0.174, 0.253]$ cm$^{-1}$. (Right) Central 128×128 ROI of the mean difference of 200 filtered back-projection (FBP) reconstructed noise realizations from the signal-present and signal-absent sinograms. The size of this central ROI is indicated with the yellow box on the background image. The gray scale window is $[-0.0075, 0.02]$ cm$^{-1}$.

Detectability in the sinogram data domain; generate multiple realizations of signal-present and signal-absent sinograms; perform TV-LSQ reconstruction of all data realizations; divide the resulting image set into training and testing data; train the signal-present/signal-absent classifier; and finally, measure the image domain detectability with the testing images. The data model and data signal detection task is set up so that the ideal observer performance can be analytically computed. In this way, the data domain detectability serves as a precisely known upper bound to the image domain detectability. The loss in detectability, passing through image reconstruction, provides a quantitative measure that is an indication of loss of fine details in the image and may reflect the subjective property of image patchiness.

The images in Fig. 1 illustrate the detection task employed in this work. The background disk attenuation is representative of fat tissue and is set to 0.194 cm$^{-1}$. The ring at the edge represents the skin-line with attenuation 0.233 cm$^{-1}$. The phantom is defined on a 2048×2048 grid and is 18×18 cm$^2$ in physical dimensions. The pixel size is chosen much smaller than the detector resolution so that the phantom can be regarded as quasi-continuous. Projection of this background image yields the mean background sinogram. The signal is defined as a Gaussian function with full-width-half-maximum of 100 microns (the reconstructed image grid uses a pixel size of 350 microns) and amplitude of 0.04 cm$^{-1}$. Projection of the background plus signal yields the mean signal-present sinogram. To appreciate the difficulty of the detection task, we also show in Fig. 1 the mean difference image of both hypotheses over 200 noise realizations, reconstructed by FBP for $N_{\text{views}} = 512$. The reconstruction grid is a 512×512 pixel array. It is apparent that the speckle noise is still
visible even after averaging over 200 realizations; the signal would not be visible in the reconstructed image from a single noise realization.

The data domain ideal observer detectability is computed as a signal-to-noise ratio (SNR) for detection, see Sec. 13.2.8 in Barrett and Myers\(^9\), which is straight-forward for the data model specified in Eqs. (2) and (3). For additive Gaussian noise, using the small signal approximation, the ideal observer and ideal linear observer are equivalent. The ideal linear observer performance is computed by first solving for the Hotelling template

\[ w_{\text{data}} = \frac{\bar{g}_{\text{sig}+} - \bar{g}_{\text{sig}-}}{\text{Var}(g_{\text{sig}-})}, \]

where

\[ \bar{g}_{\text{sig}+} = Pf_{\text{sig}+}, \]
\[ \bar{g}_{\text{sig}-} = Pf_{\text{sig}-}, \]

and the small signal approximation is assuming

\[ \text{Var}(g_{\text{sig}+}) \approx \text{Var}(g_{\text{sig}-}). \]

The SNR for detection in the data domain is computed from the dot product of the Hotelling template and the signal projection data

\[ \text{SNR}_{\text{data}}^2 = w_{\text{data}}^T (\bar{g}_{\text{sig}+} - \bar{g}_{\text{sig}-}). \]

The SNR metric can be converted to a receiver operating characteristic (ROC) area-under-the-curve (AUC), or equivalently a percent-correct (PC) on a two-alternative-forced-choice (2-AFC) observer experiment (page 823 in Barrett and Myers\(^9\))

\[ \text{PC}_{\text{data}} = \text{AUC}_{\text{data}} = \frac{1}{2} + \frac{1}{2} \text{erf} \left( \frac{\text{SNR}_{\text{data}}^2}{2} \right). \]

For the equal-dose breast CT simulation at the specified noise level and the given signal properties, the signal detectability in the data domain corresponds to

\[ \text{PC}_{\text{data}} = 86.57\%, \]

where the range of possible performance values are 50%, corresponding to guessing on the 2-AFC experiment, to 100%, a 2-AFC perfect score. That the ideal observer performance is
significantly less than 100% in the data domain is intended by design. This design requirement is why it is necessary to use the subtle signal shown in Fig. 1.

As pointed out in Sec. 13.2.6 of Barrett and Myers\textsuperscript{9}, image reconstruction can only maintain or lose signal detectability with the ideal observer, and as a result the ideal observer is not commonly used for assessing tomographic images after reconstruction. Essentially, from the ideal observer perspective, image reconstruction should not be performed at all. Constrained by the fact that human observers can interpret reconstructed images much more easily than sinograms, there is still potentially useful knowledge to be gained from the ideal observer in assessing the efficiency of the image reconstruction algorithm; namely, it can address the question of how well the separability between signal-present and signal-absent hypotheses is preserved in passing through image reconstruction. In other words, does the image reconstruction algorithm wipe out the signal in the detection task? This is a particularly relevant question for recent efforts in non-linear image reconstruction where strong assumptions are being exploited to obtain tomographic images for sparse sampling conditions or low-dose scanning. The image-domain ideal observer performance is also useful in that it provides a theoretical upper bound on human observer performance, and no amount of post-processing will allow this bound to be exceeded.

For computing the image-domain detectability, we employ the 2-AFC PC figure-of-merit for the ideal observer in the image domain because it is easy to interpret; the 2-AFC test intuitively connects image ensemble properties with single image noise realizations; and we have a hard theoretical upper bound that it cannot exceed $\text{PC}_{\text{data}} = 86.57\%$. This last property that, $\text{PC}_{\text{image}} \leq \text{PC}_{\text{data}}$,

also naturally provides a measure for the loss of signal-detectability passing through image reconstruction. To provide an accurate and precise estimate of $\text{PC}_{\text{image}}$, 4000 noisy data realizations of both signal-present and signal-absent hypotheses are generated. All of the data realizations are reconstructed with the TV-LSQ algorithm. Half of the resulting images under each hypothesis are used to train an ideal-observer classifier, and the remaining half of the images is used to generate the $\text{PC}_{\text{image}}$ metric and its error bars. (Because $\text{PC}_{\text{image}}$ is computed from noise realizations, it is necessary to work with a small signal due to its inherent uncertainty. If the data domain PC is close to 100%, the resulting drop in going
to the image domain PC may be too small to be significant.) The large number of image realizations leads to a high precision, and the accuracy results from surveying a number of classifiers including both ideal linear observer and ideal observer estimators. For the ideal linear observer, we have investigated the channelized Hotelling observer\textsuperscript{26} with different channel formulations and a single-layer neural network (SLNN)\textsuperscript{27}. For the ideal observer, several implementations of a convolutional neural network (CNN)\textsuperscript{27} have been explored. We have found that a hybrid-CHO yields $\text{PC}_\text{image}$ equal to the results, within error bars, from the NN classifiers over the range of simulation parameters investigated. We present the hybrid-CHO because of its relative simplicity, but the equivalence of the hybrid-CHO with the SLNN and CNN is significant because the hybrid-CHO exploits approximate rotational symmetry in the detection task while the SLNN and CNN do not. This approximate symmetry allows for a reduction in the number of channels needed for the hybrid-CHO, and the equivalence with the NN-based observers indicates that the reduced set of channels in the hybrid-CHO is not compromising performance of the hybrid-CHO as an observer model.

\textbf{II.C.1. Hybrid-CHO}

The theory for estimation of the CHO $\text{PC}_\text{image}$ and its variance is covered in Gallas and Barrett\textsuperscript{26} and Chen \textit{et al.}\textsuperscript{28}. The hybrid-CHO developed here exploits approximate rotational symmetry that results from use of a small rotationally-symmetric signal and uniform angular sampling in the sinogram. Because the detection task design is approximately rotationally symmetric, it lends itself well to the use of standard Laguerre-Gauss channels\textsuperscript{26}, which are circularly symmetric. The Laguerre-Gauss channels on their own, however, do not provide an optimal basis because of the small size of the signal in combination with the fact that the image is discretized on a Cartesian grid. To account for both of these aspects of the CT imaging set-up, we propose a hybrid channel set composed of Laguerre-Gauss channels combined with single-pixel channels at the location of the signal. The observer model is referred to as a hybrid-CHO, reflecting this hybrid channel set.

The data for computing the hybrid-CHO performance consist of the central 128x128 region of pixels from each of the 512x512 image realizations; thus there are a total of 4,000 signal-present and signal-absent 128x128 ROIs for training and testing the hybrid-CHO. The
II.C. SKE/BKE signal detection model

The continuous definition of the Laguerre-Gauss channels is

\[ L_n = \sum_{k=0}^{n} \binom{n}{k} (-1)^k x^k, \]

\[ u_n(r|a) = \sqrt{\frac{2}{a}} \exp \left( -\frac{\pi r^2}{a^2} \right) L_n \left( \frac{2\pi r^2}{a^2} \right), \]

where the radius \( r \) is defined \( r^2 = x^2 + y^2 \); \( x, y \) indicate location on the 128x128 ROI; and the units of \( x \) and \( y \) are scaled so that \( (x, y) = (0, 0) \) is the center of the ROI and \( (x, y) = (1, 1) \) is the upper right corner of the ROI. The parameters of the Laguerre-Gauss channels are the order \( n \) and Gaussian radial decay parameter \( a \), which is specified in the same scaled units as \( r \). The discrete representation of the Laguerre-Gauss channels is obtained by evaluating \( u_n(r|a) \) at the center of each of the pixels in the ROI. The single-pixel channels are defined as

\[ u(s, t) = \begin{cases} 1 & (i, j) = (s, t) \\ 0 & (i, j) \neq (s, t) \end{cases}, \]

where \( (i, j) \) are the integer coordinates of the pixels in the discrete channel function; \( (s, t) \) is the location of the unit impulse; the origin of the integer coordinates \((0, 0)\) is at the lower left corner of the ROI.

The specific channel set employed for the breast CT simulation consists of fourteen channels. The first ten are the discrete Laguerre-Gauss channels, \( u_n(r|a) \), with \( n \in [0, 9] \) and \( a = 0.5 \), and the remaining for are the single-pixel channels, \{\( u(63, 63), u(63, 64), u(64, 63), u(64, 64) \}\}. Considering the channel functions as column vectors of length 128x128, where the pixel elements are in lexicographical order, the 14 channels form a channelization matrix \( U \) of size 16,384 \( \times \) 14 (16,384 = 128 \( \times \) 128).

The channelized linear classifier is computed by estimating the mean channelized signal and the channelized image covariance. To compute these quantities, the channelized images are first obtained from the reconstructed training images by

\[ [u_{\text{sig}+}]_i = U^T \left[ f_{\text{sig}+}^{(\text{recon})} \right]_i, \]

\[ [u_{\text{sig}-}]_i = U^T \left[ f_{\text{sig}-}^{(\text{recon})} \right]_i, \]

where \( i \) is the realization index, which runs from 1 to \( N_{\text{real}} = 4000 \); and \( f_{\text{sig}+}^{(\text{recon})} \) \( (f_{\text{sig}-}^{(\text{recon})}) \) is a column vector with pixels values from the central 128x128 ROI extracted from the
II. METHODS

II.C. SKE/BKE signal detection model

reconstruction from signal-present (signal-absent) data. The first \( i = 1 \) through \( N_{\text{train}} = 2000 \) realizations are assigned to the training set, and the rest of the realizations \( i = N_{\text{train}} + 1 \) through \( N_{\text{train}} + N_{\text{test}} \) are assigned to the testing set. The mean channelized signal is

\[
s_u = \frac{1}{N_{\text{train}}} \sum_{i=1}^{N_{\text{train}}} ([u_{\text{sig}+}]_i - [u_{\text{sig}-}]_i).
\]

Using the small signal approximation, the training images under both hypotheses can be combined to provide a covariance estimate

\[
K_u = \frac{1}{(2N_{\text{train}} - 1)} \left[ \sum_{i=1}^{N_{\text{train}}} ([u_{\text{sig}+}]_i - \bar{u}_{\text{sig}+})^\top ([u_{\text{sig}+}]_i - \bar{u}_{\text{sig}+}) \\
+ \sum_{i=1}^{N_{\text{train}}} ([u_{\text{sig}-}]_i - \bar{u}_{\text{sig}-})^\top ([u_{\text{sig}-}]_i - \bar{u}_{\text{sig}-}) \right],
\]

where the barred variables indicate mean over the ensemble of corresponding realizations. The channelized Hotelling template is computed as

\[
w_u = K_u^{-1} s_u,
\]

and the ROI Hotelling template can be reconstituted by matrix-vector multiplication

\[
w_{\text{image}} = U w_u.
\]

Dotting a test image with the Hotelling template \( w_{\text{image}} \) provides the test statistic, which can be compared with a set threshold to make the classification into either signal-present or signal-absent hypotheses.

The detectability metric in the image domain is estimated by running a 2-AFC experiment with the hybrid-CHO for every possible combination of signal-present and signal-absent test images

\[
PC_{\text{image}} = \frac{1}{N_{\text{test}}^2} \sum_{i=1}^{N_{\text{test}}} \sum_{j=1}^{N_{\text{test}}} c(a_i; b_j),
\]

\[
a_i = w_{\text{image}}^\top \left[f^{(\text{recon})}_{\text{sig}+} \right]_{i+N_{\text{train}}},
\]

\[
b_j = w_{\text{image}}^\top \left[f^{(\text{recon})}_{\text{sig}-} \right]_{j+N_{\text{train}}},
\]
and the two-sample kernel function $c(a; b)$ is defined

$$c(a; b) = \begin{cases} 
1 & a > b \\
0.5 & a = b \\
0 & a < b 
\end{cases}.$$ 

In the 2-AFC experiment, the Hotelling template is dotted with a pair of test images, where one is drawn from the signal-present realizations and the other is drawn from the signal-absent realizations. Whichever dot product yields the higher value, the hybrid-CHO classifies the corresponding image as a signal-present image. The summation over the two-sample kernel function essentially counts all of the times that the hybrid-CHO identified the signal-present image correctly.

Once PC_{image} is computed it can be compared with PC_{data} to provide a measure of loss of signal detectability. The quantity PC_{data} is known analytically so the corresponding value does not have error bars. The value PC_{image}, on the other hand, is estimated from realizations, and thus it has variability due to the randomness of the testing set. There is also variability in the training of the hybrid-CHO because it is computed from the random training images. To account for both sources of variability we employ the level 2 variability estimation from Chen et al.\textsuperscript{28}, and the 95% confidence intervals are reported.

\section*{II.D. Test phantom for visual correspondence}

In order to illustrate the correspondence between visual image quality and the image quality metrics, the same simulation parameters and scan configurations are investigated using a test phantom with a structured fibro-glandular tissue model\textsuperscript{29}, shown in Fig. 2. This breast phantom is composed of a 16 cm disk containing background fat tissue, attenuation 0.194 cm$^{-1}$, skin-line and randomly generated fibro-glandular tissue at attenuation 0.233 cm$^{-1}$. These components of the phantom are defined on a 2048$\times$2048 grid of dimensions 18$\times$18 cm$^2$. The structured background allows for visualization of fine details. In order to have a more direct comparison with a signal detection task, a contrast-detail (CD) insert is included in the phantom consisting of an 8x8 grid of point-like signals. The signals are defined as analytic disks so that the line-integrals through the signals can be computed exactly and their contribution to the projection data is not subject to pixelization of the test phantom image. The disk contrast in the CD insert increases linearly from 0.01 cm$^{-1}$ to 0.05 cm$^{-1}$.
Figure 2: Computerized breast phantom with a contrast-detail (CD) insert. The displayed images are the image of the phantom (top, left), the ROI focused on the CD insert (top, right), an unregularied FBP reconstruction (bottom, left), and a regularized FBP image (bottom, right). For reference, the RMSE values of the unregularized and regularized FBP images are 0.0198 and 0.01155 cm$^{-1}$, respectively. The gray scale window for all panels is $[0.174,0.253]$ cm$^{-1}$.

Going from left to right, and the disk radius starts at 200 microns and increases linearly to 500 microns going from top to bottom. For reference, the reconstruction grid’s image pixel width is 350 microns. To appreciate the noise level of the breast CT simulation, ROIs are shown of images reconstructed by FBP using a ramp filter and FBP followed by Gaussian blurring. For the FBP reconstructions, the $N_{\text{views}} = 512$ scan configuration is used. Due to the high-level of speckle noise in the unregularized FBP image, it is difficult to see even the most conspicuous of signals in the CD insert. With regularization, the larger, higher contrast corner of the CD insert becomes visible.

III. Results

The hybrid-CHO signal detection figure-of-merit and image RMSE are computed alongside TV-LSQ reconstructed images of the breast phantom, exploring the three parameters of the CT-simulation: $N_{\text{iter}}$, $N_{\text{views}}$, and TV constraint parameter $\gamma$. The TV constraint is reported as a fraction of the TV of the ground truth image.
IIIA. Signal detectability as a function of iteration number

The first set of results focus on $N_{\text{views}} = 128$ and $\gamma = 1.0$, i.e. the TV constraint is equal to the ground truth phantom TV. A series of ROI images are shown in Fig. 3 as a function of iteration number for the TV-LSQ reconstruction of the breast phantom. From the perspective of accurate recovery of the phantom, the gray level estimation appears to improve with increasing iteration number, as a general trend, which is to be expected because the TV constraint is selected to be the TV of the test phantom. From the perspective of visualizing the fine details in the image, the trend with iteration is more complex. The structure detail in the fibro-glandular tissue and many of the signals are visible already at 20 iterations, where it is clear from the overall gray value that the image is far from the solution to the TV-LSQ problem. As the iterations progress, the larger signals of the CD insert appear
more conspicuous as the speckle noise amplitude is reduced. On the other hand, some of the more subtle features in the image appear to become distorted as the iterations progress, and the numerically converged image has a classic patchy look where it is difficult to distinguish noise from real structures.

Corresponding to the image series in Fig. 3, quantitative image quality metrics are also plotted, showing image RMSE and signal-detection $PC_{\text{image}}$. As expected, the RMSE trend shows improvement with iteration number, and the RMSE converges to a value well below that of the FBP reference images in Fig. 2. Again, $\gamma$ is set to the truth value and the test phantom has a high-degree of gradient sparsity; thus the solution to the TV-LSQ optimization problem is expected to yield a mathematically accurate solution and this is reflected in low RMSE values and the fact that the RMSE steadily improves as the TV-LSQ algorithm progresses toward the solution. This trend coincides with the visual gray-level accuracy seen in the series of images. It is interesting to note that the RMSE at $N_{\text{iter}} = 500$ is substantially below the value of 0.01155 corresponding to the regularized FBP image in Fig. 2.

The iteration number trend for $PC_{\text{image}}$, however, runs opposite to the image RMSE. There is a clear decline in the signal detectability at early iterations, and as convergence is achieved this metric plateaus to a value well below the data domain signal detectability. The trend in image detectability coincides with the visual appearance of the the increasing patchiness of the images shown in Fig. 3.

The main point of the $PC_{\text{image}}$ metric is that it should reflect the disappearance of small subtle details in the image, and in this example we see correspondence between this metric and the overall patchiness of the images. Thus the quantitative $PC_{\text{image}}$ metric appears to capture the desired image properties, providing a quantitative measure of over-regularization. How to use this information to determine algorithm parameters depends on the goal of the CT system design. Clearly, the results of the iteration number study indicate that $PC_{\text{image}}$ cannot be used alone to determine the optimal iteration number, because it has the largest value with one iteration. As an aside, we note that a similar behavior was observed for the maximum likelihood expectation maximization (MLEM) algorithm using a ROI-observer$^{10}$, and we take up a comparison of these experiments in Sec. IV.

Using $PC_{\text{image}}$ in concert with image RMSE, which has the opposite trend, provides
complimentary information. As an example of how it can be used, the desired image could be specified by minimizing RMSE with the constraint that the loss in signal detectability is bounded by a parameter $\epsilon$, i.e. $\text{PC}_{\text{image}}/\text{PC}_{\text{data}} \geq \epsilon$. Subjectively, the first frame at 20 iterations, in the series of images shown in Fig. 3, has the best visibility for the signals in the CD insert and image texture realism. The next image at 50 iterations already has a patchy appearance. Visualization of the intermediate frames (not shown) suggests that a value of $\epsilon = 97.0\%$ for this particular example provides a useful bound. However, the details of how $\epsilon$ is chosen and how the detection task is designed depends on the desired imaging goal. Here, we only aim to establish correspondence between $\text{PC}_{\text{image}}$ and the subjective image quality of patchiness or over-regularization with non-linear image reconstruction.

### III.B. Signal detectability as a function of $N_{\text{views}}$ and $\gamma$

For the next set of results, we fix $N_{\text{iter}} = 100$ and vary the other two parameters of the breast CT simulation. In Fig. 4, a grid of images is shown with each row and column corresponding to fixed $N_{\text{views}}$ and $\gamma$, respectively. As a general trend the lower $\gamma$ values reduce the speckle noise and streaks in the image; however, it is also clear that the heavy regularization imposed by $\gamma = 0.75$ effectively renders the borderline signals in the CD insert invisible. In terms of conspicuity of the signals in the CD insert, the images for $\gamma = 1.5$ and above appear to have similar numbers of signals visible. The trend in $N_{\text{views}}$ is more difficult to discern because the conditions of the scan are set up to be equal dose. For the larger $\gamma$-values $N_{\text{views}} = 128$ images appear to have streak artifacts in addition to the speckle noise. In general, there is a different noise texture for the various equal-dose scan configurations.

The corresponding image RMSE and $\text{PC}_{\text{image}}$ IQ metrics are also plotted in Fig. 4. The image RMSE favors $\gamma = 1.0$, the ground truth TV value, although the RMSE for $\gamma = 0.75$ is only slightly larger. Also, the RMSE values decrease weakly with increasing $N_{\text{views}}$. The $\text{PC}_{\text{image}}$ values favor an opposite trend, where the signal detectability increases with $\gamma$. Interestingly, for the different $N_{\text{views}}$ configurations, the intermediate value $N_{\text{views}} = 256$ is slightly favored, although the values for 256 and 512 have overlapping error bars.

Again, we point out that the metrics are complimentary. Going by $\text{PC}_{\text{image}}$ alone the TV constraint would be abandoned. Going by image RMSE alone, however, can also lead to an equally pathological situation where the image is over-regularized. Using $\text{PC}_{\text{image}}$ in
RESULTS

III.B. Signal detectability as a function of $N_{\text{views}}$ and $\gamma$

Figure 4: (Top row) Images reconstructed by TV-LSQ for $N_{\text{iter}} = 100$, varying $N_{\text{views}}$ from top to bottom and varying $\gamma$ from left to right. These parameters are indicated in the figure panels. The gray scale window is [0.174,0.253] cm$^{-1}$. (Bottom, left) Plot of the corresponding image RMSE values. For reference, the RMSE of the FBP and regularized FBP images from Fig. 2 are 0.0198 and 0.01155 cm$^{-1}$, respectively, and the latter value is indicated in the plot with a dashed line. (Bottom, right) Plot of the corresponding signal detectability metric, percent correct for an ideal-observer 2-AFC experiment. The dashed line indicates the theoretical maximum PC performance inherent in the data domain. Concert with RMSE yields a more useful picture. We observe that, while it is true that PC$_{\text{image}}$ is monotonically increasing with $\gamma$, there is clearly diminishing returns for $\gamma \geq 1.5$, where this metric appears to plateau. The RMSE, on the other hand, favors lower $\gamma$ on the
III.C. Estimation of subject TV and its impact on IQ metric trends

The dependence of the simulation results on $\gamma$ are all referred to the ground truth TV value, which is object dependent. Thus applying the simulation-based IQ metrics to an actual scanning situation, where the ground truth is unknown, raises two important questions: (1) how to determine the subject TV, and (2) does the subject TV reference value yield universal IQ metric dependence on $\gamma$. Two simulations are performed to address both of these questions.

To estimate the subject TV, $\gamma_0$, we have successfully applied a validation technique\textsuperscript{30} where image reconstruction is performed with a fraction of the available data and the remaining test data are compared with the projection of the estimated image. The constraint value is estimated to be the value that yields the smallest discrepancy between the test data and the corresponding estimated data. We perform this computation in the context of the present breast CT simulation for $N_{\text{views}} = 128$ and $N_{\text{iter}} = 500$. Image reconstruction is performed with 90% of the available line-integration data, chosen from the sinogram at

---

**Figure 5:** Plot of RMSE for the data used in image reconstruction (blue) and the RMSE on the left-out testing data used for validation (red) set as a function of the TV constraint parameter. The validation RMSE has a minimum at $\gamma = 0.9$ in units scaled to the ground truth image TV. PC\textsubscript{image}-plateau. Thus a prescription that combines the two metrics could reasonably select an intermediate $\gamma$ value such as $\gamma = 1.5$, where again $\epsilon = \text{PC}_{\text{image}}/\text{PC}_{\text{data}} \geq 97\%$. At this setting, we observe that the TV-LSQ reconstructed images in Fig. 4 do not have the patchy appearance of over-regularization with TV. Also, compared with the FBP images, the image RMSE is lower and more CD insert signals are visible for TV-LSQ at $\gamma = 1.5$. 

III.C. Estimation of subject TV and its impact on IQ metric trends
III. RESULTS

III.C. Estimation of subject TV and its impact on IQ metric trends

Figure 6: Detectability metrics using different background images. The label “uniform” and refers to the use of the background image shown in Fig. 1, and “structured” refers to using the breast phantom in Fig. 2 as the background image. Note that the data domain percent correct is lower for the structured background because it is more attenuating.

random. This leaves 10% of the data for independent testing. The resulting reconstructed image is projected and the RMSEs for the reconstruction and testing data are plotted in Fig. 5 as a function of $\gamma$. From Fig. 5, we observe that there is a monotonically decreasing trend in the reconstruction data RMSE as a function of $\gamma$, but the data RMSE of the testing set shows a minimum at $\gamma = 0.9$, which is close to the true value of $\gamma = 1.0$. This result demonstrates that this validation technique can provide an estimate of the subject TV to within 10 percent.

To address the universality question, the uniform background used in the process of estimating $PC_{\text{image}}$ is changed to the non-uniform, but known, background of the breast phantom. This modification alters $\gamma_0$ dramatically; thus it is of interest to compare the resulting $PC_{\text{image}}$ curve as a function of $\gamma$. In Fig. 6 this metric is plotted for $N_{\text{views}} = 128$ and $N_{\text{iter}} = 100$. From the graph it is clear that there is some numerical discrepancy between the numerical values of $PC_{\text{image}}$ for the same value of the scaled parameter $\gamma$; however, the trend of this metric as a function of $\gamma$ matches fairly well. That there is discrepancy in the absolute numerical values is perhaps not too surprising considering the large difference in background structure. The similarity in trends is further evidence of the potential utility of the proposed IQ metric.
IV. Discussion

The proposed signal detectability index for non-linear image reconstruction bears some similarity with the signal detectability studies on MLEM iteration number studies presented by Abbey et al.\textsuperscript{10}. In particular, the ROI observer from that investigation showed a steadily decreasing trend with iteration number. The two detectability indices, however, are different and need to be interpreted differently. The detection task considered in Abbey et al.\textsuperscript{10} was meant to have direct relevance to a clinical detection task, and furthermore the authors were seeking correspondence between model and human observers on signal detection. For the detectability metric, presented here, the signal size and amplitude are chosen so that the signal is on the edge of detectability by the ideal observer in the data space. This signal is much too small to be detected by a human observer; thus the detection task design itself is not directly relevant to a clinical detection task. The design and purpose of this detection task is meant to be a surrogate for the subjective image property of patchiness specific to over-regularization in TV-LSQ reconstructed images.

The reduction of $PC_{\text{image}}$ relative to $PC_{\text{data}}$ represents an irretrievable loss of information in distinguishing signal-present and signal-absent hypothesis. No post-processing operations can improve on $PC_{\text{image}}$. This metric, however, only captures loss of detectability due to non-invertibility of the image reconstruction algorithm. It does not necessarily reflect distortion of the signal. For example, regularizing FBP images with moderate blurring, such as what is seen in Fig. 2, is invertible and does not cause a reduction in PC even though the signal itself is broadened by the blurring operation. Reconstructing FBP images onto an image grid of large pixels, on the other hand, is a non-invertible and does cause loss in detectability\textsuperscript{14}.

The SKE/BKE detection task paradigm with a small rotationally-symmetric signal and uniform projection-angle sampling allows for the hybrid-CHO to accurately represent the ideal linear observer with a relatively small set of channels, because the detection task is well-suited to the rotationally-symmetric LG channels. Considering non-rotationally symmetric signals or scanning angular ranges less than $2\pi$ breaks this symmetry. In such cases, a different channel representation and possibly a larger channel set would need to be developed in order for the hybrid-CHO to represent the ideal linear observer. The SKE/BKE detection task also considers the signal at one location in the image. For the presented non-linear image reconstruction algorithm, this limitation does not impact the utility of the metric because
V. CONCLUSION

the TV regularization is applied isotropically over the image and results are not expected to change appreciably for different signal locations. Regularization techniques that involve spatially varying weighting need to consider either multiple SKE/BKE detection tasks with signals at different locations or a signal-known-statistically (SKS) detection task where the signal location is drawn from a spatially uniform probability distribution.

V. Conclusion

We have developed and presented an image quality metric that is sensitive to the removal of subtle details in the image and that can be applied to the non-linear TV-LSQ image reconstruction algorithm. The metric is based on the detection of a small signal at the border of detectability by the ideal observer, and this metric is hypothesized to quantify the subjective visual removal of subtle image details. The design of the proposed detection task, use of the ideal observer, and connection with the 2-AFC experiment makes the metric easy to interpret. The detectability index, which is an estimate of a property of an ensemble of reconstructed images, is connected to single image realizations through the interpretation as a PC on a 2-AFC experiment. Loss of detectability through the image reconstruction process, i.e. $\text{PC}_{\text{image}} < \text{PC}_{\text{data}}$, unambiguously represents a quantitative decrease in the ability to distinguish signal-present and signal-absent images. The bounds on this metric are clear: $0.5 \leq \text{PC}_{\text{image}} \leq \text{PC}_{\text{data}}$, where the lower limit of 0.5 represents guessing on the 2-AFC experiment and the upper limit is the analytically known $\text{PC}_{\text{data}}$.

Correspondence between $\text{PC}_{\text{image}}$ and visual assessment of the reconstructed images of the breast CT simulation shows that this metric may serve to quantify TV-LSQ over-regularization. A decrease in this metric is seen to coincide with loss of borderline signals in the CD insert and with patchiness in the appearance of the images. This metric is seen to be complimentary to widely used image fidelity metrics such as image RMSE, and it may help to provide an objective means to establish useful tomographic system parameter settings. The presented methodology may also prove useful for quantifying over-regularization with other non-linear image reconstruction techniques.
A Appendix: The CPPD algorithm for TV-LSQ

The CPPD algorithm\textsuperscript{23,24} can be used to efficiently solve non-smooth convex optimization problems for CT image reconstruction\textsuperscript{25}. We provide the pseudocode for CPPD-TV-LSQ in Algorithm 1.

**Algorithm 1** Pseudocode for the CPPD-TV-LSQ algorithm. The only free algorithm parameters are $\rho$ and $N_{\text{iter}}$. Other algorithm inputs are the sinogram data $g$ and the TV constraint parameter $\gamma$. The algorithm parameters and operations are explained in the text.

1: $k \leftarrow 0$, $f^{(0)} \leftarrow 0$, $\lambda_s^{(0)} \leftarrow 0$, $\lambda_g^{(0)} \leftarrow 0$
2: while $k < N_{\text{iter}}$ do
3: $f^{(k+1)} \leftarrow f^{(k)} - \tau \left( \nu_s X^T \lambda_s^{(k)} + \nu_g D^T \lambda_g^{(k)} \right)$
4: $\tilde{f} \leftarrow 2f^{(k+1)} - f^{(k)}$
5: $\lambda_s^{(k+1)} \leftarrow \left( \lambda_s^{(k)} + \sigma(\nu_s X \tilde{f} - \nu_s g) \right) / (1 + \sigma)$
6: $\lambda^+_g \leftarrow \lambda_g^{(k)} + \sigma\nu_g D \tilde{f}$
7: $p \leftarrow |\lambda^+_g|_{\text{mag}}$
8: if $\|p\|_1 > \nu_g \gamma \sigma$ then
9: $\beta^{(k+1)} \leftarrow \text{solve}(\beta, \|\text{shrink}(p, \beta)\|_1 - \nu_g \gamma \sigma = 0)$
10: $\lambda_g^{(k+1)} \leftarrow \beta^{(k+1)} \lambda_g^+ / \max(\beta^{(k+1)}, p)$
11: else
12: $\beta^{(k+1)} \leftarrow 0$
13: $\lambda_g^{(k+1)} \leftarrow 0$
14: end if
15: $k \leftarrow k + 1$
16: end while

In the CPPD-TV-LSQ algorithm, the parameters $\nu_s$ and $\nu_g$ normalize the linear transforms $X$ and $D$:

$$\nu_s = 1/\|X\|_2, \quad \nu_g = 1/\|D\|_2,$$

where the $\ell_2$-norm of a matrix is its largest singular value. This scaling is performed so that algorithm efficiency is optimized and so that results are independent of the physical units used in implementing $X$ and $D$. Note that the sinogram $g$ in line 5 and the TV constraint parameter $\gamma$ in lines 8 and 9 must also be multiplied by $\nu_s$ and $\nu_g$, respectively. The step-size parameters $\sigma$ and $\tau$ must satisfy the inequality

$$\sigma \tau \leq 1/L^2,$$  \hspace{1cm} (7)
where $L$ is the matrix norm of $A$, which is constructed by stacking $\nu_s X$ on $\nu_g D$

$$A = \begin{pmatrix} \nu_s X \\ \nu_g D \end{pmatrix}, \quad L = \|A\|_2.$$  
Due the normalization of $X$ and $D$ and the fact that $X$ and $D$ approximately commute, $L$ should be close to 1. The step-size inequality, Eq. (7), is satisfied with equality by setting

$$\sigma = \rho/L, \quad \tau = 1/(\rho L),$$

where the step-size ratio $\rho$ is a free parameter that must be tuned because it can strongly impact CPPD convergence behavior. For all the simulations presented in this work the step-size ratio is set to $\rho = 1$.

The “solve” function at line 9, returns the value of $\beta$ that solves the equation written in its second argument, and “shrink” is defined component-wise as

$$[\text{shrink}(p, \beta)]_i = \begin{cases} p_i + \beta & p_i < -\beta \\ 0 & |p_i| \leq \beta \\ p_i - \beta & p_i > \beta \end{cases},$$

where $i$ is an index for the components of $p$. Solution of the equation at line 9 can be implemented by bisection, because $\|\text{shrink}(p, \beta)\|_1$ decreases monotonically as $\beta$ increases and the root of the equation is bracketed in the interval $[0, \max(p)]$, where max acts component-wise on $p$.

Acknowledgment

This work is supported in part by NIH Grant Nos. R01-EB026282, R01-EB023968, R01-EB020604, R01-EB028652 and The University of Chicago Women’s Board. The computational resources for this work is funded in part by the NIH S10-OD025081, S10-RR021039, and P30-CA14599 awards. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

1 I. A. Elbakri and J. A. Fessler, Statistical image reconstruction for polyenergetic X-ray computed tomography, IEEE Trans. Med. Imaging 21, 89–99 (2002).
2 C. H. McCollough, A. N. Primak, N. Braun, J. Kofler, L. Yu, and J. Christner, Strategies for reducing radiation dose in CT, Radiol. Clinics 47, 27–40 (2009).

3 E. Y. Sidky and X. Pan, Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization, Phys. Med. Biol. 53, 4777–4807 (2008).

4 G.-H. Chen, J. Tang, and S. Leng, Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic CT images from highly undersampled projection data sets, Med. Phys. 35, 660–663 (2008).

5 L. Ritschl, F. Bergner, C. Fleischmann, and M. Kachelrieß, Improved total variation-based CT image reconstruction applied to clinical data, Phys. Med. Biol. 56, 1545 (2011).

6 K. J. Batenburg and J. Sijbers, DART: a practical reconstruction algorithm for discrete tomography, IEEE Trans. Image Proc. 20, 2542–2553 (2011).

7 H. Gupta, K. H. Jin, H. Q. Nguyen, M. T. McCann, and M. Unser, CNN-based projected gradient descent for consistent CT image reconstruction, IEEE Trans. Med. Imaging 37, 1440–1453 (2018).

8 J. Adler and O. Öktem, Learned primal-dual reconstruction, IEEE Trans Med. Imaging 37, 1322–1332 (2018).

9 H. H. Barrett and K. J. Myers, Foundations of image science, Wiley, Hoboken, New Jersey, 2004.

10 C. K. Abbey, H. H. Barrett, and D. W. Wilson, Observer signal-to-noise ratios for the ML-EM algorithm, in Medical Imaging: Image Perception, edited by H. L. Kundel, volume 2712, pages 47–58, Proc. SPIE, 1996.

11 C. K. Abbey and H. H. Barrett, Human-and model-observer performance in ramp-spectrum noise: effects of regularization and object variability, JOSA A 18, 473–488 (2001).

12 A. Wunderlich and F. Noo, Image covariance and lesion detectability in direct fan-beam x-ray computed tomography, Phys. Med. Biol. 53, 2471–2493 (2008).
M. Das, H. C. Gifford, J. M. O’Connor, and S. J. Glick, Penalized maximum likelihood reconstruction for improved microcalcification detection in breast tomosynthesis, IEEE Transactions on Medical Imaging 30, 904–914 (2010).

A. A. Sanchez, E. Y. Sidky, I. Reiser, and X. Pan, Comparison of human and Hotelling observer performance for a fan-beam CT signal detection task, Med. Phys. 40, 031104 (2013).

G. J. Gang, J. W. Stayman, W. Zbijewski, and J. H. Siewerdsen, Task-based detectability in CT image reconstruction by filtered backprojection and penalized likelihood estimation, Med. Phys. 41, 081902 (2014).

A. A. Sanchez, E. Y. Sidky, and X. Pan, Task-based optimization of dedicated breast CT via Hotelling observer metrics, Med. Phys. 41, 101917 (2014).

J. Xu, M. K. Fuld, G. S. K. Fung, and B. M. W. Tsui, Task-based image quality evaluation of iterative reconstruction methods for low dose CT using computer simulations, Phys. Med. Biol. 60, 2881–2901 (2015).

S. D. Rose, A. A. Sanchez, E. Y. Sidky, and X. Pan, Investigating simulation-based metrics for characterizing linear iterative reconstruction in digital breast tomosynthesis, Med. Phys. 44, e279–e296 (2017).

Y. Liu, Z. Liang, J. Ma, H. Lu, K. Wang, H. Zhang, and W. Moore, Total variation-stokes strategy for sparse-view X-ray CT image reconstruction, IEEE Trans. Med. Imaging 33, 749–763 (2013).

S. Niu, Y. Gao, Z. Bian, J. Huang, W. Chen, G. Yu, Z. Liang, and J. Ma, Sparse-view x-ray CT reconstruction via total generalized variation regularization, Phys. Med. Biol. 59, 2997–3017 (2014).

J. M. Boone, A. L. C. Kwan, J. A. Seibert, N. Shah, K. K. Lindfors, and T. R. Nelson, Technique factors and their relationship to radiation dose in pendant geometry breast CT, Med. Phys. 32, 3767–3776 (2005).

J. S. Jørgensen and E. Y. Sidky, How little data is enough? Phase-diagram analysis of sparsity-regularized X-ray computed tomography, Philos. Trans. Royal Soc. A 373, 20140387 (2015).
A. Chambolle and T. Pock, A first-order primal-dual algorithm for convex problems with applications to imaging, J. Math. Imag. Vis. 40, 120–145 (2011).

T. Pock and A. Chambolle, Diagonal preconditioning for first order primal-dual algorithms in convex optimization, in International Conference on Computer Vision (ICCV 2011), pages 1762–1769, 2011.

E. Y. Sidky, J. H. Jørgensen, and X. Pan, Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm, Phys. Med. Biol. 57, 3065–3091 (2012).

B. D. Gallas and H. H. Barrett, Validating the use of channels to estimate the ideal linear observer, JOSA A 20, 1725–1738 (2003).

W. Zhou, H. Li, and M. A. Anastasio, Approximating the Ideal Observer and Hotelling Observer for binary signal detection tasks by use of supervised learning methods, IEEE Trans. Med. Imaging 38, 2456–2468 (2019).

W. Chen, B. D. Gallas, and W. A. Yousef, Classifier variability: accounting for training and testing, Pattern Recognition 45, 2661–2671 (2012).

I. Reiser and R. M. Nishikawa, Task-based assessment of breast tomosynthesis: effect of acquisition parameters and quantum noise, Med. Phys. 37, 1591–1600 (2010).

T. G. Schmidt, R. F. Barber, and E. Y. Sidky, A spectral CT method to directly estimate basis material maps from experimental photon-counting data, IEEE Trans. Med. Imaging 36, 1808–1819 (2017).