Published in final edited form as:

IEEE Trans Image Process. 2009 June ; 18(6): 1228–1238. doi:10.1109/TIP.2009.2017139.

Electronic Noise Modeling in Statistical Iterative Reconstruction

Jingyan Xu[Member, IEEE] and Benjamin M. W. Tsui[Fellow, IEEE]

The Johns Hopkins University, Baltimore, MD 21287-0859 USA

Jingyan Xu: jxu@jhmi.edu

Abstract

We consider electronic noise modeling in tomographic image reconstruction when the measured signal is the sum of a Gaussian distributed electronic noise component and another random variable whose log-likelihood function satisfies a certain linearity condition. Examples of such likelihood functions include the Poisson distribution and an exponential dispersion (ED) model that can approximate the signal statistics in integration mode X-ray detectors. We formulate the image reconstruction problem as a maximum-likelihood estimation problem. Using an expectation-maximization approach, we demonstrate that a reconstruction algorithm can be obtained following a simple substitution rule from the one previously derived without electronic noise considerations. To illustrate the applicability of the substitution rule, we present examples of a fully iterative reconstruction algorithm and a sinogram smoothing algorithm both in transmission CT reconstruction when the measured signal contains additive electronic noise. Our simulation studies show the potential usefulness of accurate electronic noise modeling in low-dose CT applications.

Index Terms

Compound Poisson distribution; electronic noise; low dose X-ray CT; sinogram restoration; statistical image reconstruction

I. Introduction

Accurate electronic noise modeling is important for low-dose X-ray CT applications where the number of detected photons are limited. For clinical X-ray CT systems utilizing energy-integrating detectors, it has been shown that the detected X-ray intensity follows a compound Poisson distribution, see, e.g., [1]–[3]. Because of its complicated expression, it is not yet clear how an iterative reconstruction algorithm can be derived based on the exact form of the likelihood function. Researchers have proposed various approximations, either in the sinogram domain (after the log-transform) using a mean-variance dependent Gaussian density [4], [5], or in the transmitted intensity using a (scaled) Poisson approximation [6], [7]. In the transmitted intensity domain when the Poisson approximation is used, the electronic noise compensation problem can be dealt with by a shifted Poisson approximation [6]. This approximation was proposed in [8] and [9] to compensate for readout noise in CCD cameras and was also used in PET reconstruction [10] to model randoms precorrected transmission data.

We have previously introduced a subclass of the exponential dispersion (ED) model to approximate the exact compound Poisson distribution of the signal statistics in integration

mode X-ray detectors [11]. In this work, we further extend our model to include electronic noise when the X-ray detected signal is not necessarily Poisson distributed. Specific examples include the Poisson and the ED model. We formulate the image reconstruction problem as a maximum-likelihood (ML) estimation problem. However, the likelihood function of the sum of the ED model and a Gaussian distributed electronic noise is not analytically tractable. To circumvent this difficulty, we propose an expectation-maximization (EM) framework and combine it with the surrogate function approach and the optimization transfer principle [12]–[16] for electronic noise compensation. The result of this combination, a simple “substitution rule,” constitutes the central idea of this paper.

The organization of this paper is as follows. In Section II, we briefly summarize the properties of the compound Poisson distribution of the X-ray signal statistics as described in [1]–[3]. In Section III, we present the central idea of this paper, i.e., a simple “substitution rule,” which states that under certain linearity conditions, if $\mu = \mathcal{F}(y)$ is a reconstruction algorithm where y is the measured X-ray intensity signal without electronic noise contamination, then $\mu = \mathcal{F}(E(Y|Z=z))$ is a reconstruction algorithm where the measured signal $Z = Y + N$ is contaminated by electronic noise N . As we demonstrate in Section V, the word “reconstruction” should be understood in the general sense of parameter estimation. More specifically, in $\mu = \mathcal{F}(E(Y|Z=z))$, μ is merely the unknown parameter in a maximum-likelihood estimation problem. We describe one method for calculating the conditional expectation $E(Y|Z)$. This method can be regarded as a generalization of the shifted Poisson approximation in [8] and [9]. The substitution rule can be applied to a fully iterative reconstruction algorithm for X-ray CT or a statistical sinogram restoration algorithm. In Section V, we provide two examples of how to apply the substitution rule in these two settings. We present simulation studies in Section VI to demonstrate the effectiveness of accurate electronic noise modeling for transmission CT reconstruction. Section VII concludes this paper and outlines some possible future work. Throughout this paper, we use upper case letters, e.g., Y and Z to denote a random variable, and their lower case counterparts, e.g., y and z to denote a sample realization.

II. Background

In clinical X-ray CT systems, a simplified model of the detected signal can be described by the energy weighted photon counts [1]–[3], [6]. If we denote by Y_i the X-ray intensity measured at the i th detector bin, then Y_i can be modeled as the following:

$$\begin{aligned} Y_i &= \sum_k E_k \lambda_i(E_k) \\ &= \sum_k E_k \text{Poisson}[\bar{\lambda}_i(E_k)] \end{aligned} \quad (1)$$

where

$$\bar{\lambda}_i(E_k) = N_i(E_k) e^{-\sum_j s_{ij} \mu_j(E_k)}$$

and k is the energy-bin index, $N_i(\cdot)$ is the incident photon number spectrum (in units of # of photons per detector element), $\mu_j(E_k)$ is the energy-dependent linear attenuation coefficient at the pixel j , and s_{ij} is the intersection pathlength between the boundaries of j th pixel and the ray connecting the i th detector to the X-ray focal spot. The notation $\text{Poisson}[\cdot]$ represents a Poisson distributed random variable (r.v.) with its argument as the mean. If we consider the contribution to Y_i from each individual detected photon, denote which by $X_{i,1}$,

$X_{i,2}$, etc., then each photon generates a signal E_k from a probability distribution that is proportional to the detected spectrum. Equivalently, the detected signal can be written as

$$Y_i = X_{i,1} + X_{i,2} + \cdots + X_{i,\Lambda} \quad (2)$$

where $\Lambda = \sum_k \lambda_i(E_k)$ is the (random) total number of detected photons regardless of their energies, and

$$X_{i,\cdot} = E_k, \quad \text{with the probability of } \frac{\bar{\lambda}(E_k)}{\sum_k \bar{\lambda}(E_k)}. \quad (3)$$

As is written in (2), the random variable Y_i follows a compound Poisson model [17, p.177], which by definition is the summation of a Poisson distributed, random number of independent and identically distributed (i.i.d.) r.v.'s $X_{i,\cdot}$. For the particular compound Poisson distribution in energy-integrating X-ray detection (1), the Poisson variable Λ has a mean of $\sum_k \bar{\lambda}_i(E_k)$ [cf. (1)], and each component follows the probability distribution in (3). More rigorous derivations and discussions about the density function of Y_i can be found in [1]–[3]. Since the exact form of the likelihood function of Y is not amenable for algorithm derivation, researchers have often approximated the compound Poisson distribution by either the Poisson distribution [6] or the Gaussian distribution. We have previously proposed an exponential dispersion (ED) model to approximate the exact X-ray signal statistics, and derived a maximum-likelihood reconstruction algorithm based on the ED model [11]. The ED model is a class of probability density functions that includes a special compound Poisson distribution with i.i.d. Gamma distributed components. This work further extends our previous work by incorporating electronic noise modeling into the iterative reconstruction process.

III. Expectation-Maximization Formulation for Electronic Noise Modeling

Electronic noise can be adequately modeled as a Gaussian distributed random variable with a certain mean m and a variance σ^2 . The offset mean m can be estimated accurately prior to each scan and subtracted from the measured intensity [18]. Without loss of generality, we assume the raw detected intensity Z is the summation of a random variable Y and a zero-mean Gaussian distributed r.v. N modeling electronic noise. We do not specify the distribution of Y yet, but our result requires a certain linearity condition on the log-likelihood function of Y which will become clear shortly.

We formulate the maximum-likelihood reconstruction problem in the expectation-maximization (EM) framework. The raw detected intensity Z is written as

$$Z = Y + N \quad (4)$$

where Y is the detected X-ray intensity whose distribution depends on the unknown parameter μ , $N \sim \mathcal{N}_i(0, \sigma^2)$ describes the Gaussian distributed electronic noise, and the random variables Y and N are independent. Here, we omit the detector index i for notational clarity, since it is not needed to distinguish between different detector elements at this point. Let $S = (Y, N)$ be the complete data in the EM formulation, and Z be the incomplete data. The incomplete log-likelihood can be written as [12]

$$\log P\{z;\mu'\} = E_{S|Z;\mu} \log P\{s;\mu'\} - E_{S|Z;\mu} \log P\{s|Z;\mu'\}. \quad (5)$$

Note that $\log P\{z;\mu'\}$ is the objective function in a maximum likelihood reconstruction algorithm. Here, μ' denotes the parameter to be estimated, and μ denotes the parameter estimate at the current iteration. The notation $E_{S|Y;\mu} f(s; \mu')$ denotes the conditional expectation of $f(s; \mu')$ given, represented symbolically as

$$E_{S|Y;\mu} f(s;\mu') = \int f(s;\mu') d(\text{probability measure of } S|Y_\mu)$$

where the distribution of Y is parameterized by μ . The first term on the right hand side (RHS) of (5), in the langrange of optimization transfer [16], is a surrogate function for the incomplete log-likelihood $\log P\{z; \mu'\}$. Since Y and N are independent, the first term on the right-hand-side (RHS) of (5) splits into two terms, i.e.,

$$E_{S|Z;\mu} \log P\{s;\mu'\} = E_{Y|Z;\mu} \log P\{y;\mu'\} + E_{N|Z;\mu} \log P\{n\}. \quad (6)$$

Note that in the second term of (6), we dropped the dependency of the electronic noise on the unknown parameter μ . The log-likelihood part of the first term in (6) $\log P\{y; \mu'\}$ would be the objective function in an iterative reconstruction algorithm if electronic noise were not present, whose maximization is often accomplished, at least in transmission CT reconstruction, through maximizing a surrogate function $Q(y, \mu'; \mu)$ at each iteration. Our first result is that maximizing the conditional expectation of this surrogate ensures a nondecreasing change in the incomplete log-likelihood $\log P\{z; \mu'\}$.

Theorem 1: Let $Q(y, \mu'; \mu)$ be a surrogate function for $\log P\{y; \mu'\}$ that satisfies the following conditions [15]1:

$$\log P\{y;\mu'\} \geq Q(y, \mu'; \mu), \text{ and } \log P\{y;\mu\} = Q(y, \mu; \mu). \quad (7)$$

Then if we choose μ' such that

$$E_{Y|Z;\mu} Q(y, \mu'; \mu) \geq E_{Y|Z;\mu} Q(y, \mu; \mu) \quad (8)$$

then

$$\log P\{z;\mu'\} \geq \log P\{z;\mu\}. \quad (9)$$

Proof: We replace the log-likelihood $\log P\{y; \mu'\}$ in (6) by its surrogate $Q(y, \mu'; \mu)$ and combine with (5) to define the following:

¹Here, we only consider differentiable surrogate functions.

$$\bar{Q}(z, \mu'; \mu) \triangleq E_{Y|Z, \mu} Q(y, \mu'; \mu) + E_{N|Z, \mu} \log P\{n\} - E_{S|Z, \mu} \log P\{s|Z; \mu'\}. \quad (10)$$

Using (7), we see that

$$\log P\{z; \mu'\} \geq \bar{Q}(z, \mu'; \mu) \quad (11)$$

$$\log P\{z; \mu\} = \bar{Q}(z, \mu; \mu). \quad (12)$$

$\bar{Q}(z, \mu'; \mu)$ is a legitimate surrogate function for the incomplete log-likelihood that we want to maximize. Since the function $E_{S|Z, \mu} \log P\{s|Z; \mu'\}$ reaches a maximum at $\mu' = \mu$, choosing μ' to satisfy (8) will imply

$$\bar{Q}(z, \mu'; \mu) \geq \bar{Q}(z, \mu; \mu)$$

which, in turn, forces an increase in $\log P\{z; \mu'\}$, i.e., (9).

Effectively, the EM formulation (5) provides a first surrogate (the first term on the RHS, subject to a constant difference), then the function $\bar{Q}(z, \mu'; \mu)$ defined in (10) can be also understood as a second surrogate of this EM surrogate for maximizing the incomplete log-likelihood $\log P\{z; \mu'\}$. The idea of Theorem 1 is simple. If electronic noise were not considered, we would like to maximize $\log P\{y; \mu'\}$. Due to its complicated relationship with the unknown parameter μ' , we instead maximize a surrogate function $Q(y, \mu'; \mu)$ iteratively. Theorem 1 says that if we consider electronic noise, we can keep the derived surrogate function $Q(y, \mu'; \mu)$ and calculate its conditional expectation given the raw measurement Z . The resulting $E_{Y|Z, \mu} Q(y, \mu'; \mu)$ is a surrogate function for the incomplete log-likelihood $\log P\{z; \mu'\}$. We then maximize $E_{Y|Z, \mu} Q(y, \mu'; \mu)$ iteratively with respect to μ' . Theorem 1 holds regardless of the specific distribution function of Y as long as Y is independent of N , and the distribution of N does not depend on the unknowns μ .

One advantage of this EM formulation becomes obvious by examining the expression of $Q(y, \mu'; \mu)$. As our example later shows, despite its complicated dependence on μ' , $Q(y, \mu'; \mu)$ can be a simple (e.g., affine) function of y . Since the conditional expectation in $E_{Y|Z, \mu} Q(y, \mu'; \mu) = Q(E(Y|Z; \mu), \mu'; \mu)$ is calculated at the old parameter μ , $Q(E(Y|Z; \mu), \mu'; \mu)$ maintains exactly the same relationship with respect to μ' as $Q(y, \mu'; \mu)$. Hence, the first derivative, second derivative of $E_{Y|Z, \mu} Q(y, \mu'; \mu)$ with respect to μ' can be written out from $Q'(*, \mu'; \mu)$ and $Q''(*, \mu'; \mu)$ directly, with $*$ replaced by $E(Y|Z; \mu)$. These derivatives are needed in deriving an update equation for μ' using, e.g., the Newton's method. Suppose $\mu^{n+1} = \mathcal{F}(y, \mu^n)$ is a "reconstruction" algorithm given the measured intensity y without electronic noise contamination, then our result implies that to account for electronic noise, we can use the update equation $\mu^{n+1} = \mathcal{F}(E(Y|Z=z), \mu^n)$; this we will refer to as the 'substitution rule' for electronic noise modeling in a maximum-likelihood iterative reconstruction algorithm.

Unlike Theorem 1, the substitution rule works only when $Q(y, \mu'; \mu)$ is affine in the measurement $Y = y$. Examples of such surrogate functions are abundant in the literature. For instance, the convex surrogate $Q(\mu | \mu^n)$ for the monochromatic transmission model [14,

(5)]; the $Q_1(x, x^k)$ function in emission tomography [12, (26)] (but not $Q_2(x, x^k)$ [12, (27)]) which is quadratic in the measurements); the quadratic surrogate function $Q(\mu; \mu^n)$ [15, (12)], using either the maximum curvature or the optimum curvature. In all these examples, follows the Poisson distribution. One example of a non-Poisson distributed can be found in [16].

IV. Calculating the Conditional Expectation

To model and compensate for electronic noise, when the surrogate function $Q(y, \mu'; \mu)$ is affine in y , we only need to replace all occurrences y by $E(Y | Z = z)$ in a previously derived algorithm totally ignoring electronic noise. The calculation of $E(Y | Z)$ would depend on the distribution of Y . We have previously introduced an exponential dispersion (ED) model [19], [20] as an accurate approximation of the exact compound Poisson density function of in energy-integrating X-ray detectors. This ED model greatly facilitates the derivation of a maximum-likelihood iterative reconstruction algorithm for X-ray CT [11]. It is a special class of compound Poisson distributions with Gamma distributed components and its log-density function satisfies the following relationship [20]:

$$\frac{\partial \log f}{\partial \bar{Y}} = \frac{y - \bar{Y}}{\varphi \bar{Y}^p} \quad (13)$$

where $\bar{Y}$ is the mean of the random variable Y , and $\varphi \bar{Y}^p$ is its variance, $1 < p < 2$ and $\varphi > 0$ are the so-called power and dispersion parameters, respectively. For $\varphi = 1$ and $p = 0, 1, 2, 3$, (13) can be shown to reduce to the familiar Gaussian, Poisson, Gamma, and inverse Gaussian distribution [19], [20].

Next, we describe how to calculate the conditional expectation $E(Y | Z)$ using (13) as the model of Y . Snyder *et al.*'s work [8], [9] on electronic readout noise compensation for CCD cameras considered calculating $E(Y | Z)$ where Z is the summation of a Poisson distributed Y and a Gaussian noise N . They proposed two methods, one based on a shifted Poisson approximation, the other by saddlepoint approximation. The following discussion can be regarded as a generalization of the shifted Poisson approximation method; applying the saddlepoint approximation to the ED model can be found in [21].

It is well known that $E(Y | Z)$ satisfies the following minimal error property:

$$E[Y - E(Y|Z)]^2 \leq E[Y - u(Z)]^2 \quad \text{for all } u(Z).$$

In other words, among all functions of Z , $E(Y | Z)$ has the minimal estimation error. When Y has a distribution described by (13), $E(Y | Z)$ is conceivably a very complicated function of Z . We can approximate $E(Y | Z)$ by an affine function $a_0 + a_1 Z$, and seek the coefficients a_0 and a_1 such that

$$E\{[Y - (a_0 + a_1 Z)]Z^i\} = 0, \quad \text{for } i=0, 1. \quad (14)$$

The relationship in (14) is obtained from the normal equations for least square estimations. Solving (14) for a_0 and a_1 , we obtain the best affine estimator for $E(Y | Z)$ as the following:

$$\begin{aligned}
E(Y|Z) &\approx \frac{\bar{Y}\sigma^2 + Z \cdot \text{Var}Y}{\text{Var}Y + \sigma^2} \\
&= \frac{\bar{Y}\sigma^2 + \varphi \bar{Y}^\rho Z}{\varphi \bar{Y}^\rho + \sigma^2} \triangleq E_{\text{aff}}(Y|Z).
\end{aligned} \tag{15}$$

The best affine estimator $E_{\text{aff}}(Y|Z)$ is a convex combination of $\bar{Y}$ and Z with the relative variances σ^2 and $\varphi \bar{Y}^\rho$ as coefficients. Note that $E_{\text{aff}}(Y|Z = \bar{Y}) = \bar{Y}$. From (15), we see that if Y had followed a Poisson distribution, which is a special case of the ED model by setting $p = \varphi = 1$, the best affine estimator would be $\bar{Y}(\sigma^2 + Z)/(\bar{Y} + \sigma^2)$. This coincides with the shifted Poisson method [8], [9]. In the shifted Poisson approximation, we first augment the measurement Z by σ^2 , and treat $Z + \sigma^2 = Y + (N + \sigma^2)$ as the summation of two independent Poisson distributions, one with mean $\bar{Y}$, the other with mean σ^2 . The conditional expectation of Y given Z (or equivalently $Z + \sigma^2$) is then just $\bar{Y}(Z + \sigma^2)/(\bar{Y} + \sigma^2)$. Applying the substitution rule, we obtain, $\mu^{n+1} = Q(E(Y|Z), \mu^n)$. If we take the expectation on both sides, $E\mu^{n+1} = E Q(E(Y|Z), \mu^n) = Q(E(E(Y|Z)), \mu^n) = Q(\bar{Y}, \mu^n)$, which means that the ensemble mean of noisy-reconstruction is equal to the noise-free reconstruction.

V. Application of the Substitution Rule in X-Ray CT Reconstruction

We present two examples of applying the substitution rule in X-ray CT reconstruction. The first example is based on the well-known convex algorithm [14] and demonstrates how a fully iterative reconstruction algorithm can be modified to account for additional electronic noise in the measurements. This example may have applications in X-ray CT systems using CCD cameras. The second example applies the substitution rule in a sinogram restoration problem. It is also more pertaining to clinical X-ray CT imaging as we use the ED model to approximate the compound Poisson distribution of energy-integrating X-ray detectors.

A. Convex Algorithm

The convex algorithm was proposed in [14] and has been studied and used in a wide variety of applications. Its derivation is based on the Poisson log-likelihood of the measured intensity Y and a surrogate function $Q(\mu | \mu^n)$ [14 (5)] that is affine in Y_i . To apply the substitution rule, we write

$$\mu_j^{n+1} = \mu_j^n + \frac{\mu_j^n \sum_i l_{ij} [d_i e^{-\langle l_i, \mu^n \rangle} - E(Y_i | Z_i, \mu^n)]}{\sum_i l_{ij} \langle l_i, \mu^n \rangle d_i e^{-\langle l_i, \mu^n \rangle}} \tag{16}$$

$$\approx \mu_j^n + \frac{\mu_j^n \sum_i l_{ij} \left[d_i e^{-\langle l_i, \mu^n \rangle} - \frac{\bar{Y}_i^n}{\bar{Y}_i^n + \sigma_i^2} (z_i + \sigma_i^2) \right]}{\sum_i l_{ij} \langle l_i, \mu^n \rangle d_i e^{-\langle l_i, \mu^n \rangle}} \tag{17}$$

where $\bar{Y}_i^n = d_i \exp\{-\langle l_i, \mu^n \rangle\}$. Equation (16) is obtained by substituting Y_i in [14, (7)] by its conditional expectation given the raw measurements $E(Y_i | Z_i, \mu^n)$. From (16) to (17), we used the best affine approximation (or the shifted Poisson approximation since Y is assumed Poisson) of $E(Y_i | Z_i, \mu^n)$. Equation (17) gives the explicit update equation of the convex algorithm extended to account for electronic noise.

B. Sinogram Restoration for X-Ray CT

To further illustrate the applicability of the substitution rule, we derive a penalized-likelihood sinogram restoration algorithm based on the ED model in (13). The idea of sinogram restoration is to try to recover a set of ideal line integrals, free of off-focal radiation [6], beam hardening effects, etc., from the measured X-ray transmission data that are contaminated by these physical-degrading factors. The image reconstruction is performed by analytical methods such as the filtered backprojection (FBP) from the restored 'ideal' data.

We start by assuming electronic noise does not exist, i.e., the measurements y_i are the detected X-ray intensity signals following the ED model. Assume the different detector elements are independent, the penalized log-likelihood function can be written as the following:

$$\Psi(\underline{l}, \underline{y}) \triangleq L(\underline{l}, \underline{y}) - \beta \sum_i R_i(\underline{l}) \quad (18a)$$

$$= \sum_i \frac{1}{\varphi_i} \left(y_i \frac{\bar{Y}_i^{1-p_i}}{1-p_i} - \frac{\bar{Y}_i^{2-p_i}}{2-p_i} \right) + \sum_i c(y_i, \varphi_i, p_i) - \beta \sum_i R_i(\underline{l}). \quad (18b)$$

The log-likelihood function of the ED model $L(\underline{l}, \underline{y})$ is obtained by integrating both sides of (13), $c(y_i, \varphi_i, p_i)$ is the integration constant independent of $\bar{Y}_i$. Here, $\bar{Y}_i$ denotes the mean of the detected X-ray intensity signal at the detector bin i . In this example, we use the following simplified model for $\bar{Y}_i$:

$$\begin{aligned} \bar{Y}_i &= \int E N(E) e^{-\sum_j s_{ij} m_w(E) \rho_j / m_w(E_0)} dE \\ &\triangleq \int \tilde{N}(E) e^{-m_w(E) l_i / m_w(E_0)} dE \end{aligned} \quad (19)$$

where $l_i \triangleq \sum_j s_{ij} \rho_j$. Comparing with (1), we use the model $m_w(E) \rho_j / m_w(E_0)$ to approximate the energy-dependent attenuation coefficients of the object. Here, $m_w(E)$ is the mass attenuation coefficient of water at the energy E , and E_0 is a predetermined reference energy; ρ_j is the water-equivalent attenuation at pixel j . This formulation approximates all materials in the scanned object as water with different densities. We use the notation $\tilde{N}(E) \triangleq E N(E)$ to account for the incident X-ray spectrum, the detector gain and efficiency etc. The unknowns of the sinogram smoothing problem are the line integrals l_i of the water-equivalent attenuation at the reference energy E_0 . In (18), the penalization weighting $\beta \geq 0$ controls the strength of smoothing, $R_i(\cdot)$ penalizes abrupt changes in l_i , we use the following quadratic form in [6]:

$$R_i(\underline{l}) = \sum_{j \in \mathcal{N}_i} \frac{(l_i - l_j)^2}{2}$$

where $\mathcal{N}_i$ consists of the neighborhood pixels of the pixel i . For a nonboundary pixel i , we penalize the horizontal (channel direction) and vertical (view direction) differences in the sinogram similar to [6], i.e., a neighborhood of 4 pixels with equal weights.

The objective function in (18b) is not a concave function of the unknowns. We apply the optimization transfer principle [12], [13], [15], [16] and derive in Appendix A the following separable quadratic surrogate for the log-likelihood part $L(\underline{l}, \underline{y})$ in (18a) (ignoring the penalty terms for now)

$$Q(l_i; l_i^n) \triangleq \sum_i \frac{1}{\varphi_i} \left[y_i \frac{\bar{Y}_i^{1-p_i}}{1-p_i} - \frac{\bar{Y}_i^{2-p_i}}{2-p_i} + (l_i - l_i^n) \left(\bar{Y}_i^{1-p_i} - \frac{y_i}{\bar{Y}_i^{p_i}} \right) B(l_i^n) - \frac{(l_i - l_i^n)^2}{2} C(l_i^n) \right] \quad (20)$$

where

$$\bar{Y}_i = \int \tilde{N}(E) e^{-m_w(E)l_i/m_w(E_0)} dE$$

$$B(l_i) = \int \tilde{N}(E) m_w(E) e^{-m_w(E)l_i/m_w(E_0)} / m_w(E_0) dE.$$

The expression of $C(l)$ is provided in Appendix A (32). A separable quadratic surrogate function for the penalization term in (18b) can be derived as in [13]. Combining the two parts, the update equation for l_i^n can be written as (21), shown at the bottom of the page. The constant β controls the strength of the overall smoothing, and $l_j, j \in N_i$, is the j th pixel in the neighborhood N_i of l_i , N is the total number of neighbors of pixel i ; for a non-boundary pixel i , we have $N = 4$.

$$l_i^{n+1} = \frac{\left(\bar{Y}_i^{1-p_i} - \frac{y_i}{\bar{Y}_i^{p_i}} \right) B(l_i^n) / \varphi_i + l_i^n C(l_i^n) / \varphi_i + \beta_i \sum_{j \in N_i} l_j^n}{C(l_i^n) / \varphi_i + \beta_i N} \quad (21)$$

We have up to this point completely ignored the electronic noise contribution to the raw measurements. Equation (21) is a penalized sinogram restoration algorithm assuming the ED model for the detected intensity Y_i (no electronic noise contamination). For our purpose, it suffices to point out that the curvature function $C(\cdot)$ as derived in (32) is an affine function of the measurement y_i . Then despite its complicated expression, the quadratic surrogate function in (20) is an affine function of the measurements y_i . From the results in Section III, we can write out a sinogram restoration algorithm with electronic noise compensation from (21) directly by replacing all occurrences of y_i by its conditional expectation $E(Y_i | Z_i = z_i)$. If we use the affine approximation in Section IV, then

$$E(Y_i | Z_i = z_i) \approx \widehat{E}_{\text{aff}}(Y_i | Z_i = z_i)$$

$$= \frac{\widehat{Y}_i \sigma^2 + \varphi \widehat{Y}_i^p z_i}{\varphi \widehat{Y}_i + \sigma^2} \quad (22)$$

where $\widehat{Y}_i = \bar{Y}_i^n$ is the estimated mean transmission at the current iteration.

As shown in Appendix A, the quadratic surrogate function (for the log-likelihood part) (20) is valid in the range $[l_n - \varepsilon, l_n + \varepsilon]$ for a prespecified value $\varepsilon > 0$. This constraint must be

enforced in the update values l_{n+1} calculated by (21). In our algorithm implementation, we clip l_i^{n+1} so that $l_i^{n+1} \geq 0$ and $|l_{n+1} - l_n| \leq 0.2$.

We show in Appendix B that the surrogate function (20) reduces to that in [6, (51)] for the special case where $p_i = \varphi_i = 1$ for all the detector bins. In Section IV, we have demonstrated that in this case the affine estimator coincides with the shifted-Poisson approximation. Combining these two points, it can be shown that the surrogate function in [6, (50)], and, hence, the sinogram smoothing algorithm derived using a shifted Poisson likelihood can be obtained by first ignoring the electronic noise, then applying the substitution rule using the affine estimator.

The point of the above argument is *not* that we may either apply the shifted Poisson approximation to the log-likelihood directly as done in [6] or use it at a later stage in calculating the conditional expectation as we do here. The power of the substitution rule becomes more apparent when the data is not Poisson distributed. In this case it would not be obvious how to apply an analogous “shift” and write out the approximate log-likelihood function of a non-Poisson random variable plus a Gaussian noise. However, all we need for applying the substitution rule and for compensating for the Gaussian noise is, essentially, the knowledge that the log-likelihood is affine in the measurement and a way to calculate the conditional expectation $E(Y|Z)$.

There are other examples in the literature where a simple substitution worked. In Snyder *et al.* [8], considering the update equations for nonuniform flat-field, nonzero background, but with (18) and without readout noise (15), the only difference is that we replace $r(j)$, with no readout noise, in (15) by its conditional expectation $F[r(j), \hat{u}_k, \sigma]$. Even (15) can be regarded as the result of applying the substitution rule to a nonuniform flat-field, zero-background algorithm (without the $\mu_0(j)$). However, this time the substitution rule can be applied exactly as the conditional expectation of a Poisson r.v. given the sum of i.i.d. Poisson r.v.’s has an analytical form. These examples are all based on Poisson distributions and a linear detection model. However, as we have demonstrated, this idea can be applied to non-Poisson distributed random variables and nonlinear detection models. No surrogate functions were mentioned in these examples since the algorithm derivations were based on the log-likelihood functions, which are certainly legitimate surrogates.

VI. Numerical Evaluations

We perform two simulation studies, corresponding to the sinogram restoration algorithm and the fully iterative reconstruction algorithm we discussed in Section V, to study the effectiveness of electronic noise modeling in low dose X-ray CT reconstruction.

A. Sinogram Restoration FBP Reconstruction

We simulated fan-beam projection data of polychromatic X-rays passing through a transaxial slice of a clock phantom with diameter 48 cm. The clock phantom consists of water cylinders of various densities and sizes. The uniform air scan intensity (i.e., no bowtie filter) was around 3×10^7 in units of $\text{keV} \times \# \text{ photons}$. Some other geometric simulation parameters are listed in Table I. The projection data were generated using the Siemens DRASIM© simulation package. The continuous X-ray incident spectrum was sampled at a 5-keV interval, and for each 5-keV energy bin, the mean attenuated photon number $\bar{\lambda}(E_k)$ was replaced by a Poisson random variable with mean $\bar{\lambda}(E_k)$. This noisy photon number was multiplied by the mean energy of each energy bin and the total detected signal was the sum from all the energy bins following the model in (1). We then added Gaussian noise to the transmission signal to simulate effects of electronic noise. Our incident polychromatic X-ray

spectrum has an effective energy of around 55 keV; the additive Gaussian noise has a variance of 330^2 , which corresponds to the counting statistics of about 36 X-ray photons.

The ED model parameters p_i, φ_i are assumed known *a priori*. They are obtained by fitting the ED model to the compound Poisson density of energy-integrating detectors [22]. The model parameters p_i, φ_i have the same size as the projection data, and one such set that we used in the sinogram smoothing algorithm (21) is shown in Fig. 1. These parameters together with the transmission projection data are input to the ED-model based penalized maximum-likelihood sinogram restoration algorithm (21). Since our phantom only consists of water with different densities, we chose $m_w(E)$ as the mass attenuation coefficient of water. The reference energy E_0 was the effective energy of the incident X-ray spectrum, 55 keV. We used the affine estimator to estimate $E(Y | Z)$. The stopping criterion of the sinogram smoothing algorithm was that $|l_i^{n+1} - l_i^n| \leq 10^{-5}$ for all detector bins i at two consecutive iterations. Depending on the value of the smoothing strength β , this criterion required iteration numbers ranging from 10 (small β value) to less than 200 (large β value). The final reconstruction was done by the FBP algorithm. [23]

To visualize the effect of electronic noise, we simulated one set of noise-free transmission data but added the same amount $\sigma^2 = 330^2$ of electronic noise. The reconstructed images using the FBP algorithm, without preprocessing by sinogram domain smoothing, from these two sets of data are shown in Fig. 2. Beam-hardening artifacts are obvious in both reconstructions. In addition, the reconstructed image with the electronic noise appear grainier especially toward the central region. The noise texture from electronic noise do not seem to be distinguishable from those resulting from transmission noise only.

We simulated ten independent realizations of the noisy transmission projection, and applied sinogram smoothing using different strengths β of the quadratic penalization term to sweep out the tradeoff between noise and resolution. One sample reconstructed image is shown in Fig. 3(a). To test the effectiveness of electronic noise compensation alone, for each set of data we ran the sinogram smoothing algorithm twice assuming $\sigma^2 = 0$ and $\sigma^2 = 330^2$, respectively. We also looked at the absolute baseline where electronic noise was not present in the simulation data nor was it compensated for in the sinogram smoothing process. This baseline provides a loose upper bound of the performance of an ideal electronic noise compensation method.

We measured the tangential (vertical) profile of the cylinder insert at 3 o'clock position [Fig. 3(b)] and fit that by a Gaussian kernel edge-spread function [24]. The full-width-at-half-maximum (FWHM) of the fitted Gaussian kernel was used as a resolution measure at that point. We averaged this FWHM from ten independent realizations to reduce effects of noise. We also calculated the coefficient of variation (COV) in a small uniform ROI near the 3 o'clock cylinder to quantify noise. In Fig. 3(c), we plot the noise-resolution tradeoff from the three different sinogram processing methods, 1) no electronic noise simulated, 2) with electronic noise but no compensation, and 3) with electronic noise and with compensation. The curves were swept out by varying the quadratic smoothing strengths β from 0.1 to 200. Fig. 3(d) is the noise-resolution tradeoff for the inner cylinder at 12 o'clock position. In both cases, we can see that with electronic noise compensation we obtain a better noise-resolution tradeoff than without.

Because of the use of the quadratic smoothing priors, the sinogram smoothing algorithm in (21) may introduce bias in the final FBP-reconstructed images, especially at larger β values. We calculated the mean reconstructed images from the ten sets of noisy data, and compared the mean reconstruction near the background region [ROI-1, Fig. 3(b) 12 o'clock position] with water's attenuation coefficient 0.22 cm^{-1} at the effective energy of 55 keV. This

comparison was repeated for the various smoothing strength β that we used for sinogram smoothing. The normalized mean bias and its standard deviation are listed in Table II. As we can see, the normalized bias is around 1 HU for all the β values that we used in the sinogram smoothing step.

B. Fully Iterative Reconstruction

We performed a second simulation study to test the effectiveness of electronic noise modeling and compensation in the fully iterative reconstruction algorithm (17). As the convex algorithm is based on a monochromatic X-ray source spectrum with a photon counting X-ray detector, we used the DRASIM© simulation package to generate noise-free counting mode data using the same scanning geometry as in Section VI.A. The monochromatic X-ray source was set at 70 keV. The same phantom composition but scaled-down to 30-cm diameter was used. The DRASIM© package by default assumes a uniform air scan intensity. To emulate the effect of a bowtie filter, from the noise-free transmission data generated by DRASIM© we adjusted the detected intensity according to a bowtie filter profile that can normalize a 20-cm water cylinder [green dashed line Fig. 4(b)], after which Poisson distributed photon noise and electronic noise ($\sigma^2 = 36$) were added. One sample projection data and its profile at view index 255 are shown in Fig. 4.

We used the original convex algorithm and its extended version (17) with electronic noise modeling for reconstruction. Fig. 5 shows two example reconstructed images using the two methods at matched resolution for the cylindrical insert at 3 o'clock position. The qualitative difference between the two reconstruction methods is clearly visible. From Fig. 4, we see that the bowtie intensity profile is not wide enough for the 30-cm phantom. As a consequence, the peripheral region receives very little radiation, and electronic noise over these region plays a relatively more significant role than the interior region. With electronic noise compensation the algorithm (17) can achieve lower noise at matched resolution.

VII. Conclusions and Discussions

We previously proposed an exponential dispersion model (ED) for approximating the compound Poisson statistics of the energy-integrating X-ray detectors, and demonstrated how such a model can be utilized in deriving an iterative X-ray CT reconstruction algorithm [11]. In this work, we extended our model to include electronic noise. We presented an expectation-maximization formulation and explained how it can be combined with the surrogate function approach and the optimization transfer principle to account for electronic noise in an ML estimation problem. This formulation is applicable to a general class of distribution functions; the only requirement is that the electronic noise component is independent of the “useful” signal, and that the distribution of the electronic noise does not depend on the unknown parameters to be estimated. If in addition the log-likelihood of the ‘useful’ signal is affine in the data, then a reconstruction algorithm with electronic noise compensation can be obtained by a minimal modification to one previously derived for the electronic-noise-free case.

The algorithm implementation involves calculating the conditional mean $E(Y|Z)$ of the nonelectronic-noise contaminated signal Y given the raw measurements $Z = Y + N$. When Y is modeled by the exponential dispersion model, we described an affine approximation for calculating $E(Y|Z)$; a saddlepoint approximation method was proposed in [21]. These two methods can be regarded as generalization of similar methods in [8] and [9] from the Poisson to the ED model.

We used the convex algorithm as an example to demonstrate how the substitution rule can be applied. In addition, we also derived a sinogram smoothing algorithm with electronic

noise modeling for low-dose CT applications where the X-ray intensity signal follows the ED model. For its simplicity, we only used the best affine estimator in combination with the sinogram smoothing algorithm and the fully iterative reconstruction algorithm. It was found in [9] that using a more accurate estimator (i.e., saddlepoint) introduced a negligible to modest improvement over a less accurate (Poisson) estimator for readout noise compensation in CCD cameras. Further studies may also investigate the same question when we have a nonlinear model such as in X-ray CT and a wider dynamic range of the signal statistics.

Acknowledgments

The authors would like to thank K. Stierstorfer at Siemens Medical Solutions for his help with the Siemens DRASIM simulation package.

This work was supported by a research contract from Siemens Medical Solutions. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Til Aach.

References

1. Whiting BR. Signal statistics of X-ray computed tomography. *Proc SPIE Physics of Medical Imaging*. 2002; 4682:53–60.
2. Whiting BR, Massouzaeh P, Earl OA. Properties of preprocessed sinogram data in X-ray computed tomography. *Med Phys*. 2006; 33:3290–3303. [PubMed: 17022224]
3. Elbakri IA, Fessler JA. Efficient and accurate likelihood for iterative image reconstruction in X-ray computed tomography. *Proc SPIE Image Processing*. 2003; 5032:1839–1850.
4. Wang J, Li T, Lu H, Liang Z. Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low-dose X-ray computed tomography. *IEEE Trans Med Imag*. Oct; 2006 25(10):1272–1283.
5. Thibault JB, Sauer KD, Bouman CA, Hsieh J. A three-dimensional statistical approach to improved image quality for multislice CT. *Med Phys*. Nov; 2007 34(11):4526–4544. [PubMed: 18072519]
6. Riviere PJJ, Bian J, Vargas PA. Penalized-likelihood sinogram restoration for computed tomography. *IEEE Trans Med Imag*. Aug; 2006 25(8):1022–1036.
7. Lasio GM, Whiting BR, Williamson JF. Statistical reconstruction for X-ray computed tomography using energy-integrating detectors. *Phys Med Biol*. 2007; 52:2247–2266. [PubMed: 17404467]
8. Snyder DL, Hammoud AM, White RL. Image recovery from data acquired with a charge-coupled-device camera. *J Opt Soc Amer*. May; 1993 10(5):1014–1023.
9. Snyder DL, Helstrom CW, Lanterman AD, Faisal M, White RL. Compensation for readout noise in CCD images. *J Opt Soc Amer*. Feb; 1995 12(2):272–283.
10. Yavuz M, Fessler JA. Penalized-likelihood estimators and noise analysis for randoms-precorrected PET transmission scans. *IEEE Trans Med Imag*. Aug; 1999 18(8):665–674.
11. Xu, J.; Tsui, BMW. A compound poisson maximum-likelihood iterative reconstruction algorithm for X-ray CT. *Proc. 9th Int. Meet. Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*; 2007. p. 108–111.
12. Pierro ARD. On the relation between the ISRA and the EM algorithm for positron emission tomography. *IEEE Trans Med Imag*. Jun; 1993 12(2):328–333.
13. Pierro ARD. A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography. *IEEE Trans Med Imag*. Mar; 1995 14(1):132–137.
14. Lange K, Fessler JA. Globally convergent algorithms for maximum a posteriori transmission tomography. *IEEE Trans Med Imag*. Oct; 1995 14(10):1430–1438.
15. Erdogan H, Fessler JA. Monotonic algorithms for transmission tomography. *IEEE Trans Med Imag*. Sep; 1999 18(9):801–814.
16. Lange K, Hunter DR, Yang I. Optimization transfer using surrogate objective functions. *J Comput Graph Statist*. Mar; 2000 9(1):1–20.
17. Feller, W. *An Introduction to Probability Theory and Its Applications*. 1. New York: Wiley; 1966.

18. Hsieh, J. Computer Tomography, Principles, Design, Artifacts, and Recent Advances. 1. Bellingham, WA: SPIE; 2003.
19. Jorgensen B. Exponential dispersion models. J Roy Statist Soc B. 1987; 49:127–162.
20. Smyth, GK. Regression analysis of quantity data with exact zeroes. Proc. 2nd Australia-Japan Workshop on Stochastic Models in Engineering, Technology and Management; 1996. p. 572-580.
21. Xu J, Tsui BMW. Electronic noise compensation in iterative X-ray CT reconstruction. Proc SPIE Physics of Medical Imaging. 2008; 6913:69 132H–69 132H-12.
22. Xu, J.; Tsui, BMW. A maximum likelihood iterative reconstruction method for X-ray CT. presented at the IEEE Medical Imaging Conf; Nov. 2006;
23. Huesman, RH.; Gullberg, GT.; Greenberg, WL.; Budinger, TF. RECLBL library users manual. Lawrence Berkeley Lab., Univ; California: 1977.
24. Riviere PJJ, Billmire DM. Reduction of noise-induced streak artifacts in X-ray computed tomography through spline-based penalized-likelihood sinogram smoothing. IEEE Trans Med Imag. Jan; 2005 24(1):105–111.
25. Ciarlet, P. Introduction to Numerical Linear Algebra and Optimisation. 1. Cambridge, U.K: Cambridge Univ. Press; 1986.

Biographies

Jingyan Xu (M'03) received the B.S. degree from Tsinghua University, China, and the M.S. and Ph.D. degrees from Stanford University, Stanford, CA, all in electrical engineering.

She was a part-time research engineer at SC Solutions, Inc., Sunnyvale, CA, during her Ph.D. study. She worked as a postdoctoral intern at Robert Bosch Corporation, Research and Technology Center, Palo Alto, CA, before joining the Division of Medical Imaging Physics, Department of Radiology, Johns Hopkins University, Baltimore, MD, in 2004, where she is currently a Research Associate. Her research interests include design, implementation, performance evaluation of image reconstruction algorithms, and the modeling and compensation for the image-degrading factors in SPECT and X-ray CT systems.

Benjamin M. W. Tsui (F'77) received the B.S. degree in physics from the Chinese University of Hong Kong in 1970, the M.A. degree in physics from Dartmouth College, Hanover, NH, in 1972, and the Ph.D. degree in medical physics from the University of Chicago, Chicago, IL, in 1977.

After graduation, he continued to work at the University of Chicago as a Postdoctoral Fellow and became an Assistant Professor of radiology in 1979. He joined the University of North Carolina, Chapel Hill, in 1982 as a Research Associate Professor of radiology and biomedical engineering (BME) and was promoted to tenured Professor and became the Director of the Medical Imaging Research Laboratory and the Associate Chair of the Department of BME in 1991. He joined the Johns Hopkins University, Baltimore, MD, in 2002, as a Professor of radiology, electrical and computer engineering, environmental health sciences, and biomedical engineering and as the Director of the Division of Medical Imaging Physics in the Department of Radiology. His research interests include imaging physics of SPECT, PET and CT, 4-D computer generated phantoms that realistically mimic human anatomy and physiology, computer simulation techniques including the use of Monte Carlo methods, statistical and quantitative image reconstruction methods, 4-D cardiac and respiratory motion compensation, and image quality evaluation using model and human observers, as well as preclinical small animal imaging instrumentation and techniques.

Appendix A Surrogate Function Derivation

We derive a quadratic surrogate function for the log-likelihood

$$\Phi(l) = y \frac{\bar{Y}^{1-p}}{1-p} - \frac{\bar{Y}^{2-p}}{2-p}, \quad 1 < p < 2. \quad (23)$$

where $\bar{Y}$ is defined as in (19). Note that we omitted the detector index i since we can look at each i separately. The nonseparability of the objective function (18) is only involved in the penalization term. We look for a quadratic surrogate function $S(l; l_n)$ satisfying the following conditions:

$$\begin{aligned} \Phi(l) &\geq S(l; l_n) = \Phi(l_n) + \Phi'(l_n)(l - l_n) - C(l_n) \frac{(l - l_n)^2}{2}, \quad C(\cdot) > 0 \\ \Phi(l_n) &= S(l_n; l_n). \end{aligned} \quad (24)$$

As a preparation for finding $C(l_n)$, we list some quantities that we will need shortly

$$\dot{\Phi}(l) = \left(\frac{y}{\bar{Y}^p} - \frac{\bar{Y}}{\bar{Y}^p} \right) \frac{\partial \bar{Y}}{\partial l} \quad (25)$$

$$\ddot{\Phi}(l) = -\frac{y}{\bar{Y}^{p+1}} \left(\frac{\partial \bar{Y}}{\partial l} \right)^2 + (p-1) \left(\frac{1}{\bar{Y}^p} - \frac{y}{\bar{Y}^{p+1}} \right) \left(\frac{\partial \bar{Y}}{\partial l} \right)^2 + \left(\frac{y}{\bar{Y}^p} - \frac{\bar{Y}}{\bar{Y}^p} \right) \frac{\partial^2 \bar{Y}}{\partial l^2} \quad (26)$$

$$\frac{\partial \bar{Y}}{\partial l} = - \int \tilde{N}(E) m'(E) e^{-m'(E)l} dE \quad (27)$$

$$\frac{\partial^2 \bar{Y}}{\partial l^2} = - \int \tilde{N}(E) [m'(E)]^2 e^{-m'(E)l} dE. \quad (28)$$

Using the Cauchy's inequality and the relationships in (27) and (28), it can be shown that

$$\left(\frac{\partial \bar{Y}}{\partial l} \right)^2 \leq \bar{Y} \frac{\partial^2 \bar{Y}}{\partial l^2}. \quad (29)$$

Since we have a maximization problem, we require $C(l_n) > 0$. One choice of $C(l_n)$ that satisfies the condition in (24) was given in [15] and [25]

$$C(l_n) = - \inf_{l \geq 0} \ddot{\Phi}(l). \quad (30)$$

However, $\Phi(l)$ may not be bounded from below for $l > 0$. This can be seen by using the special case, a monotonic model, for $\bar{Y}$. Instead of finding a global condition as in (30), we

consider a local region $[l_n - \varepsilon, l_n + \varepsilon]$ around l_n for some $\varepsilon > 0$, and look for a curvature such that

$$C(l_n) \geq - \min_{l \in [l_n - \varepsilon, l_n + \varepsilon]} \ddot{\Phi}(l). \quad (31)$$

This curvature will define the quadratic surrogate as specified in (24). In case $l_n = 0$, an asymmetric region can be used such as $[0, l_n + \varepsilon]$.

If we replace the first term in (26) by the inequality (29), we will have

$$\begin{aligned} \ddot{\Phi}(l) &\geq \frac{(p-1)}{\bar{Y}^p} \left(1 - \frac{y}{\bar{Y}}\right) \left(\frac{\partial \bar{Y}}{\partial l}\right)^2 + \left(-\frac{1}{\bar{Y}^{p-1}}\right) \frac{\partial^2 \bar{Y}}{\partial l^2} \\ &\geq -\frac{(p-1)}{\bar{Y}^p} \frac{y}{\bar{Y}} \left(\frac{\partial \bar{Y}}{\partial l}\right)^2 + \left(-\frac{1}{\bar{Y}^{p-1}}\right) \frac{\partial^2 \bar{Y}}{\partial l^2}, \quad \text{since } 1 < p < 2. \end{aligned}$$

Let $l_1 = l_n - \varepsilon$, and $l_2 = l_n + \varepsilon$. Since the quantities $\bar{Y}$, $-\partial \bar{Y} / \partial l$, and $\partial^2 \bar{Y} / \partial l^2$ are all nonincreasing functions of l , the curvature $C(l_n)$ can be obtained by

$$\begin{aligned} \min_{l \in [l_1, l_2]} \ddot{\Phi}(l) &\geq -\frac{(p-1)y}{\bar{Y}^{p+1}} \Big|_{l_2} \left(\frac{\partial \bar{Y}}{\partial l} \Big|_{l_1}\right)^2 + \left(-\frac{1}{\bar{Y}^{p-1}} \Big|_{l_2}\right) \frac{\partial^2 \bar{Y}}{\partial l^2} \Big|_{l_1} \\ &\triangleq -C(l_n). \end{aligned} \quad (32)$$

Appendix B Special Case $p = \varphi = 1$

In this special case, the exponential dispersion log-likelihood in (23) degenerates to the Poisson log-likelihood, since

$$\lim_{t \rightarrow 0} \frac{x^t - 1}{t} = \frac{e^{t \log x} - 1}{t} = \log x, \quad x > 0.$$

The second order derivative in (26) can be simplified as

$$\begin{aligned} \ddot{\Phi}(l) &= -\frac{y}{\bar{Y}^2} \left(\frac{\partial \bar{Y}}{\partial l}\right)^2 + \left(\frac{y}{\bar{Y}} - 1\right) \frac{\partial^2 \bar{Y}}{\partial l^2} \\ &\geq -\frac{y}{\bar{Y}} \frac{\partial^2 \bar{Y}}{\partial l^2} + \left(\frac{y}{\bar{Y}} - 1\right) \frac{\partial^2 \bar{Y}}{\partial l^2} \\ &= -\frac{\partial^2 \bar{Y}}{\partial l^2} \geq -\frac{\partial^2 \bar{Y}}{\partial l^2} \Big|_{l=0} \\ &= -\int \tilde{N}(E) [m'(E)]^2 dE \\ &\triangleq -C_1(l^n) = \text{const.} \end{aligned}$$

As a side note, we point out that our derivation of the quadratic surrogate function is slightly different than those used in [6] and [15], where the quadratic surrogate is obtained, not for the original objective function, but for a separable, first surrogate function of the original objective function derived using convexity arguments. In the special case where $p = \varphi = 1$, these two approaches can result in the same surrogate. If we identify $I_i b_{ij}$ in [6] with $\tilde{N}(E)$ in

our model, and the linear model $f_i(l^{\text{mono}}) \triangleq m'(E) l^{\text{mono}}$ instead of the identity function used in ([6], Approach 2), then it can be seen that $C_1(l_n)$ corresponds to the maximum curvature (slowest convergence) in [6(24)].

Fig. 1.

(a) Power parameter p_i and (b) the dispersion parameter ϕ_i in the ED-modeled based sinogram smoothing algorithm.

Fig. 2.

(a) Direct FBP reconstruction from noise-free data, and no electronic noise was added to projection data. (b) Direct FBP reconstruction with noise-free transmission data and electronic noise with variance $\sigma^2 = 330^2$. The display window for both images is $[0.18, 0.25]$ cm^{-1} .

Fig. 3. (a) FBP reconstruction from the smoothed sinogram data with electronic noise compensation. The display window is $[0, 0.4] \text{ cm}^{-1}$. (b) The definition of ROIs and line profiles for measuring resolution. (c) Noise-resolution trade-off of different reconstruction procedures for the cylinder at 3 o'clock position and (d) for the smaller cylinder at 12 o'clock position.

Fig. 4.

(a) One sample of the noisy projection data used in the second simulation study. (b) The bowtie intensity profile (dashed line) and the projection profile at view index 255 (solid line). The units of both vertical axes are photon counts per bin.

Fig. 5. Reconstructed images at matched resolution without (a) and with electronic noise compensation (b). The display window for both images is $[0, 0.4] \text{ cm}^{-1}$.

TABLE I

Parameters for Fan-Beam Simulation

# channels	# views/360°	Focus-ISO	ISO-detector
672	512	595.0 mm	490.6 mm
Fan angle	Detector size	Recon image size	Recon pixel size
52°	1.47 mm	512×512	0.96 mm

TABLE II

Bias Calculation for Sinogram Smoothing FBP Recon

β	1.0	2.0	5.0	10.0	20.0	50.0
normalized bias ($\times 10^{-3}$)	1.291	1.292	1.295	1.297	1.290	1.242
stdv bias ($\times 10^{-3}$)	0.2933	0.283	0.2657	0.2483	0.2257	0.187