IEICE TRANSACTIONS

on Fundamentals of Electronics, Communications and Computer Sciences

PAPER

Using Scattered X-Rays to Improve the Estimation Accuracy of Attenuation Coefficients: A Fundamental Analysis

Naohiro TODA[†a)], *Member*, Tetsuya NAKAGAMI[††b)], *Nonmember*, Yoichi YAMAZAKI[†††c)], *Member*, Hiroki YOSHIOKA[†d)], *Nonmember*, and Shuji KOYAMA[††††e)], *Member*

SUMMARY In X-ray computed tomography, scattered X-rays are generally removed by using a post-patient collimator located in front of the detector. In this paper, we show that the scattered X-rays have the potential to improve the estimation accuracy of the attenuation coefficient in computed tomography. In order to clarify the problem, we simplified the geometry of the computed tomography into a thin cylinder composed of a homogeneous material so that only one attenuation coefficient needs to be estimated. We then conducted a Monte Carlo numerical experiment on improving the estimation accuracy of attenuation coefficient by measuring the scattered X-rays with several dedicated toroidal detectors around the cylinder in addition to the primary X-rays. We further present a theoretical analysis to explain the experimental results. We employed a model that uses a T-junction (i.e., T-junction model) to divide the photon transport into primary and scattered components. This division is processed with respect to the attenuation coefficient. Using several T-junction models connected in series, we modeled the case of several scatter detectors. The estimation accuracy was evaluated according to the variance of the efficient estimator, i.e., the Cramer–Rao lower bound. We confirmed that the variance decreases as the number of scatter detectors increases, which implies that using scattered X-rays can reduce the irradiation dose for patients.

key words: X-ray CT, scattered X-rays, accuracy, Monte Carlo simulation, Cramer–Rao lower bound

1. Introduction

X-ray computed tomography (CT) is a medical diagnostic technique that is used to reconstruct cross-sectional images of an object from X-ray intensities measured with a rotating X-ray source-detector pair [1]. In CT, images are reconstructed from the pointwise linear attenuation coefficient (hereafter referred to as the attenuation coefficient) of the object, which is typically estimated by measuring the number of photons that pass directly through the object. Therefore,

the existence of scattered X-rays is the main cause of artifacts [1]–[3] in the reconstructed images. Many scatter-correction methods have been proposed to prevent the occurrence of artifacts. The anti-scatter grid (or post-patient collimator) [1], [4], [5] and beam-stop arrays [6]–[8] use hardware that controls the X-ray beam. Two types of scatter-correction approaches are based on software: a kernel-based method [9], [10] and a technique based on Monte Carlo simulation [10]–[13]. These scatter-correction strategies involve subtracting the scatter estimations from the original projections; that is, the scattered X-rays are discarded. Ordinary CT removes scattered X-rays because current reconstruction algorithms assume that the projection data consist of only primary X-rays owing to limited computer capabilities. Hereinafter, we refer to methods that were derived under the assumption of using only the primary X-rays as primary methods.

Because scattered X-rays are caused by objects, they provide information about them. Based on this perspective, several researchers have investigated tomographic imaging with scattered X-rays [14]–[21]. In particular, Norton's modality [17] and its improvements [18]–[21] involve a new concept. However, this is regarded as a different modality from conventional X-ray CT, and the introduction of scattered X-rays into conventional X-ray CT has rarely been discussed.

In recent years, methods using a total simulation of the CT geometric structure and interactions between X-rays and substances have been studied with advances in computer capabilities [22]–[24]. In such methods, an image is reconstructed by letting the virtually generated projection data in a computer converge to the measured projection data using an iterative optimization algorithm. We refer to these methods as simulated projection-based methods. Although this approach can also include the geometries of the grid, the introduction of scattering phenomena including multiple scattering into the simulation model gets rid of the restriction where only the primary X-rays are used. Therefore, the simulated projection-based method has the potential to estimate the attenuation coefficient by using scattered X-rays in addition to primary X-rays.

At this point, a question arises as to whether the reconstruction accuracy of the simulated projection-based method using projection data including scattered X-rays can achieve the accuracy given by a primary method with an ideal grid. Furthermore, if scattered X-rays contain information inde-

pendent from information provided by primary X-rays, the accuracy of attenuation coefficients estimated with both primary and scattered X-rays should be better than that estimated with primary X-rays only. However, so far we have found no research in the literature that addresses such issues.

In order to realize novel reconstruction schemes, our purpose in this study was to demonstrate the existence of a case where the accuracy is improved by measuring the scattered X-rays in addition to the primary X-rays. However, under the premise that the simulated projection-based method is used, the current CT geometry is too complicated for numerical experiments with currently available computers if scattered X-rays are introduced. In addition, such a situation prevents fundamental consideration.

Even if it does not lead to immediate application, once such a principle is confirmed through a simplified problem, it is expected to become a driving force for creating more practical applications in the future. In this study, we simplified the CT geometry as much as possible without losing the essence of the effect of the scattered X-rays. We observed the improved accuracy through a numerical experiment with Monte Carlo simulations using the simplified geometry. In order to confirm the fundamental concept more firmly, we present a theoretical analysis to explain the results of the numerical experiment based on mathematical statistical methods.

2. Simplification of CT Geometry

In order to clearly demonstrate that the accuracy is improved by adding the measurement of scattered X-rays, we set the following two conditions for simplification.

C-1 The estimation accuracy of the attenuation coefficient should be compared for the case of measuring the primary X-rays only and the case of measuring both primary and scattered X-rays.

C-2 The accuracy should have a variable indicating the incident X-ray photon number.

Condition **C-1** is the purpose itself. Condition **C-2** is necessary to clearly show the relation between the estimation accuracy of the attenuation coefficient—namely, the image quality of the reconstruction—and the X-ray dose. If these conditions are satisfied, the simplest geometry possible is preferred.

As shown in Fig. 1(a), in a conventional CT, cone-beam X-rays are irradiated, and the primary X-rays that pass through an object are measured with a detector array. The scattered X-rays are blocked by the grid installed in front of the detector. When the grid is removed, the scattered X-rays reach the detector. In order to capture the side-scatter and back-scatter as well, we assumed that dedicated detectors (referred to as scatter detector) are installed around the object, as shown in Fig. 1(b).

By applying a simulated projection-based method to a CT equipped with such a structure, the scattered X-rays can be introduced into estimating the attenuation coefficients.
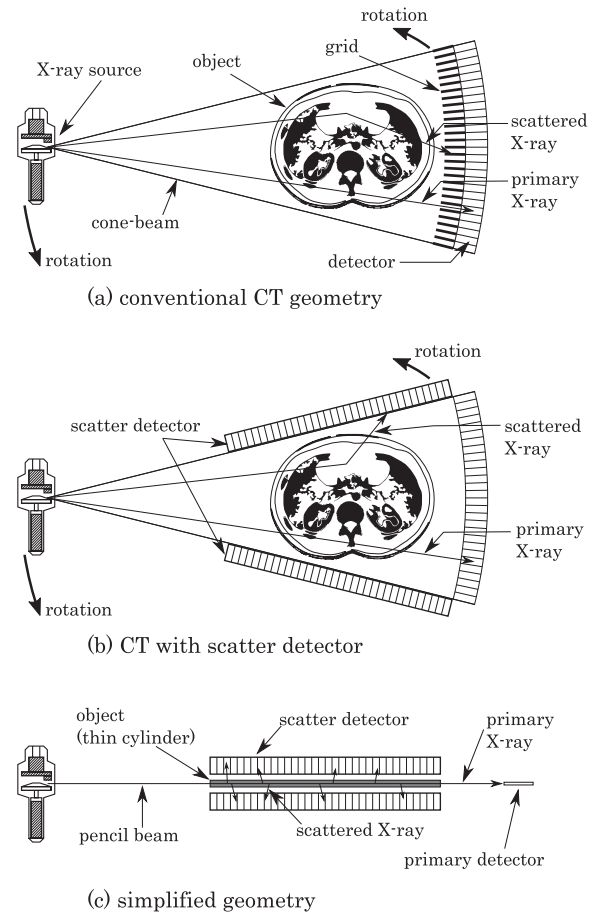

(a) conventional CT geometry

(b) CT with scatter detector

(c) simplified geometry

Fig. 1 Simplification of CT geometry.

However, as mentioned above, such a CT geometry is too complicated for a numerical experiment to be conducted. During the theoretical derivation, furthermore, the existence of other factors makes it difficult to focus only on the scattered X-rays as a control factor for the accuracy.

Here, we simplified the geometry by reducing the number of elements comprising the object to only one. The target object was a thin cylinder composed of homogeneous material, as shown in Fig. 1(c). This cylinder was exposed to a pencil-beam X-ray. Primary X-ray photons were measured by a cylindrical detector (referred to as the primary detector) with a smaller diameter than the thin cylinder. Because the scattered X-ray photons radiated around the cylinder isotropically, we used several scatter detectors with a toroidal structure for measurement, as illustrated in Fig. 2. Both scatter and primary detectors were assumed to count photons without loss. The photons were assumed to disappear instantaneously when counted by the detectors.

This geometry requires no rotation. Both the diameter of the thin cylinder and gap between the cylinder and scatter detector were set sufficiently small. Consequently, the probability of multiple scattering was kept fairly low, and any scattered photon that deviated from the pencil beam line, even if by a small angle, would be measured by the nearest scatter detector from the scattered point without crosstalk

among scatter detectors. Because almost no scattered X-rays arrived at the primary detector with this geometry, scatter correction was not required. If only the primary detector is used, this geometry is the best situation for measuring the attenuation coefficient from the viewpoint of conventional scatter correction. Our interest was whether measuring scattered X-rays would improve the estimation accuracy for the attenuation coefficient without increasing the number of incident photons. In other words, we considered whether a simulated projection-based method using scattered X-rays would improve the performance beyond that of the primary method in this case. Hereinafter, we refer to this geometry as the homogeneous thin cylinder.

In the next section, we set up this geometry concretely and conducted a numerical experiment.

3. Numerical Experiment on the Simplified Geometry

We performed a numerical experiment to observe the effect of using scattered X-rays on the geometry of the homogeneous thin cylinder shown in Fig. 2. The attenuation coefficient was estimated with a simulated projection-based method. We employed a Monte Carlo simulation to obtain the projection data.

3.1 Acquisition of the Measured Projection Data

Let $\mathbb{N}$ be the set of non-negative integers, i.e., $\mathbb{N} = \{0, 1, 2, \cdots\}$, and $\mathbb{R}$ be the set of real numbers. The length of the homogeneous thin cylinder was fixed to L ($L \in \mathbb{R}, 0 < L$). We used a primary detector and ℓ ($\ell \in \mathbb{N}, \ell \leq \ell_{\max}$) scatter detectors, where $\ell_{\max}$ ($\ell_{\max} \in \mathbb{N}$) denotes the maximum number of scatter detectors. We assumed that incident X-ray photons have the same energy. The attenuation coefficient of the homogeneous thin cylinder at this energy is denoted as μ^*, ($\mu^* \in \mathbb{R}, 0 < \mu^*$).

In the case $\ell = 0$, no scatter detector was used. Namely, this case corresponds to the primary method. In order to evaluate the accuracy, M_{ex} ($M_{\mathrm{ex}} \in \mathbb{N}, 1 \leq M_{\mathrm{ex}}$) independent estimation processes were performed for each ℓ (≥ 1). The length of each scatter detector was set to L/ℓ; that is, in the case of $2 \leq \ell$, the scatter detectors were produced by dividing a long scatter detector with the length L. In each estimation process, the Monte Carlo method provided by EGS5 [27] was used, and I_{p} ($I_{\mathrm{p}} \in \mathbb{N}, 1 \leq I_{\mathrm{p}}$) photons were assumed to be incident with the homogeneous thin cylinder. We defined $\ell + 1$ dimensional vectors $\boldsymbol{d}_j = (d_{1,j}, d_{2,j}, \cdots, d_{i,j}, \cdots, d_{\ell+1,j})^T$ ($i \in \mathbb{N}, 1 \leq i \leq$
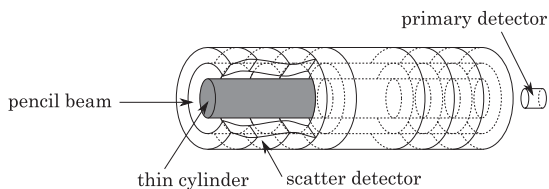

Fig. 2 Homogeneous thin cylinder.

$\ell + 1$, $j \in \mathbb{N}, 1 \leq j \leq M_{\mathrm{ex}}$, $(\cdot)^T$ means the transpose). In this notation, $d_{\ell+1,j}$ ($d_{\ell+1,j} \in \mathbb{N}, d_{\ell+1,j} \leq I_{\mathrm{p}}$) denotes the measured number of photons at the primary detector. In the cases using scatter detectors (for $1 \leq \ell$), $d_{i,j}$ ($d_{i,j} \in \mathbb{N}, d_{i,j} \leq I_{\mathrm{p}}, 1 \leq i \leq \ell$) denote the measured number of photons at the ith scatter detector.

3.2 Generation of Simulated Projection Data

When executing the simulated projection-based method, we needed to generate projection data with a virtual measurement process in a computer. We employed the Monte Carlo simulation with the same geometry used for the measurement data acquisition described in the previous Sect. 3.1. For a given attenuation coefficient μ, ($\mu \in \mathbb{R}, 0 < \mu$) and number of incident photons I_{r} ($I_{\mathrm{r}} \in \mathbb{N}, 1 \leq I_{\mathrm{r}}$), this generator outputted an $\ell + 1$ dimensional vector $\boldsymbol{v}(\mu, \ell) = (v_1^{(\mu,\ell)}, v_2^{(\mu,\ell)}, \cdots, v_i^{(\mu,\ell)}, \cdots, v_{\ell+1}^{(\mu,\ell)})^T$ ($v_i^{(\mu,\ell)} \in \mathbb{N}, v_i^{(\mu,\ell)} \leq I_{\mathrm{r}}$).

3.3 Matching Criterion

The log-likelihood is desirable as a matching criterion, or the measure of closeness between $\boldsymbol{d}_j$ and $\boldsymbol{v}(\mu, \ell)$. However, because it is difficult to derive the strict log-likelihood in actual situations, we used a pseudo log-likelihood. The number of photons entering a detector can be approximated as a Poisson random variable [1].

The probability function of each $d_{i,j}$ is given by

$$p(d_{i,j}) = \frac{\lambda_i^{d_{i,j}} \exp(-\lambda_i)}{d_{i,j}!}, \tag{1}$$

where λ_i ($\lambda_i \in \mathbb{R}, 0 < \lambda_i$) is the Poisson parameter of ith detector ($(\ell + 1)^{th}$ corresponds to the primary detector).

Here, we introduce an approximation that the measurements in all detectors are independent of each other. Then, their joint probability function is given as

$$p(\boldsymbol{d}_j) = \prod_{i=1}^{\ell+1} \frac{\lambda_i^{d_{i,j}} \exp(-\lambda_i)}{d_{i,j}!}. \tag{2}$$

Thus, the log-likelihood function $\mathrm{L_h}(\boldsymbol{\lambda} \mid \boldsymbol{d}_j)$ is derived by taking the logarithm of the above joint probability function and ignoring constant terms:

$$\mathrm{L_h}(\boldsymbol{\lambda} \mid \boldsymbol{d}_j) = \sum_{i=1}^{\ell+1} -\lambda_i + (\log \lambda_i) d_{i,j}, \tag{3}$$

where $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \cdots, \lambda_{\ell+1})^T$ is a positive real-value $\ell + 1$ dimensional parameter vector. Because an approximation has been introduced, this is a pseudo log-likelihood.

In the estimation process of the simulated projection-based method, we needed to determine the value of $\boldsymbol{\lambda}$. For a given μ, we used the notation $\widehat{\boldsymbol{\lambda}}(\mu) = (\widehat{\lambda}_1(\mu), \widehat{\lambda}_2(\mu), \cdots, \widehat{\lambda}_{\ell+1}(\mu))^T$ to estimate $\boldsymbol{\lambda}$ as

$$\widehat{\lambda}_i(\mu) = v_i^{(\mu,\ell)} \frac{I_{\mathrm{p}}}{I_{\mathrm{r}}}. \tag{4}$$

By substituting the above $\widehat{\lambda}_i$ for λ_i in Eq. (3), we could evaluate the closeness. Because I_{r} can take any number permitted by the capacity of the computer, a larger I_{r} is preferable to increase the accuracy. Likelihood definitions of the Poisson distribution similar to that presented above have been used in several iterative approximation algorithms [25], [26].

3.4 Procedure for Estimating the Attenuation Coefficient

The attenuation coefficient was estimated by the following procedure with the permissible error ε. In this problem, the unknown parameter is the attenuation coefficient μ; thus, we solved a one-dimensional optimization problem by using a golden section search [28]. When executing this procedure, we obtained an estimation of the attenuation coefficient for each j and ℓ. Thus, the estimated coefficient is denoted as $\widehat{\mu}_{j,\ell}$.

1. Initialize the search range $[\mu_{\min}, \mu_{\max}]$.
2. Calculate μ_{left} and μ_{right} given in Eq. (5).

$$\tau = (1 + \sqrt{5})/2$$
$$\mu_{\mathrm{left}} = \mu_{\min} + (\mu_{\max} - \mu_{\min})/(1 + \tau), \tag{5}$$
$$\mu_{\mathrm{right}} = \mu_{\max} - (\mu_{\max} - \mu_{\min})/(1 + \tau).$$

3. If $|\mu_{\max} - \mu_{\min}| < \varepsilon$ holds, then calculate $\widehat{\mu} \leftarrow (\mu_{\min} + \mu_{\max})/2$ and stop.
4. Otherwise, calculate $\widehat{\lambda}(\mu_{\mathrm{left}})$ and $\widehat{\lambda}(\mu_{\mathrm{right}})$ via a Monte Carlo simulation I_{r} times.
5. Calculate $\mathrm{L_h}(\widehat{\lambda}(\mu_{\mathrm{left}}) \mid \boldsymbol{d}_j)$ and $\mathrm{L_h}(\widehat{\lambda}(\mu_{\mathrm{right}}) \mid \boldsymbol{d}_j)$ with the following parameters:
6. If $\mathrm{L_h}(\widehat{\lambda}(\mu_{\mathrm{left}}) \mid \boldsymbol{d}_j) < \mathrm{L_h}(\widehat{\lambda}(\mu_{\mathrm{right}}) \mid \boldsymbol{d}_j)$ holds, then update the attenuation coefficient $\mu_{\min} \leftarrow \mu_{\mathrm{left}}$; otherwise, $\mu_{\max} \leftarrow \mu_{\mathrm{right}}$.
7. Go back to step 2.

3.5 Accuracy Evaluation

If an estimator is unbiased, evaluating the accuracy with the variance is sufficient. However, there is a possibility that a bias error is included. Thus, in this numerical experiment, we evaluated the accuracy by using the root mean square error:

$$Err(\ell) = \sqrt{\frac{1}{M_{\mathrm{ex}}} \sum_{j=1}^{M_{\mathrm{ex}}} (\mu^* - \widehat{\mu}_{j,\ell})^2}, \tag{6}$$

where μ^* denotes the true attenuation coefficient as described in Sect. 3.1.

3.6 Experimental Conditions

The diameter of the homogeneous thin cylinder was 0.05 cm,

its length was 20.0 cm, and its density was 1.5 times that of water. The maximum number of NaI toroidal scatter detectors $\ell_{\max}$ was set to 10. The inside diameter of the toroidal scatter detectors was 0.06 cm, and the outside diameter was 0.07 cm. Primary X-rays were measured with an ordinary cylindrical NaI detector having a diameter of 0.01 cm and length of 0.05 cm. The X-ray tube, thin cylinder, and primary detector were assumed to be located on a straight line. The distance from the X-ray tube origin to the center of the thin cylinder was 11.0 cm, and the distance from the center of the thin cylinder to the primary detector was 11.05 cm. The number of photons I_{p} used to collect the projection data at a single energy of 100 keV was set as 10^6. The number of photons I_{r} used in the estimation process with the Monte Carlo simulation was set to 10^7, and M_{ex} was 10^3. The true attenuation coefficient was 0.25919455, which was obtained by the "CALL Option" provided in EGS5, where we included Rayleigh, Compton, and coherent scattering but omitted the reduction due to bound electrons. In the estimating procedure, the permissible error was $\varepsilon = 10^{-6}$, and the initial search range was $[\mu_{\min}, \mu_{\max}] = [0.2, 0.4]$.

3.7 Results of the Numerical Experiment

The value of $Err(\ell)$ is shown in Fig. 3. Although the error with one scatter detector ($\ell = 1$) seemed to be the same as that of the primary method, it decreased extremely when several scatter detectors were used. Figure 4 shows a histogram of μ estimated with only the primary detector. Figure 5 shows a histogram of the estimation with both scatter detectors ($\ell = 10$) and the primary detector. The width of the distribution in Fig. 5 is less than that in Fig. 4. That is, the estimation accuracy of the attenuation coefficient was improved by measuring the scattered X-rays with several dedicated scatter detectors, even though the algorithm was derived by using approximations.

These results demonstrate the possible existence of a case where scattered X-rays are useful for estimating the attenuation coefficient. However, because this was a numerical experiment based on random numbers, the following questions arose. The error decreased up to $\ell = 4$, but it fluctuated when ℓ was more than 5. Was this due to random numbers?
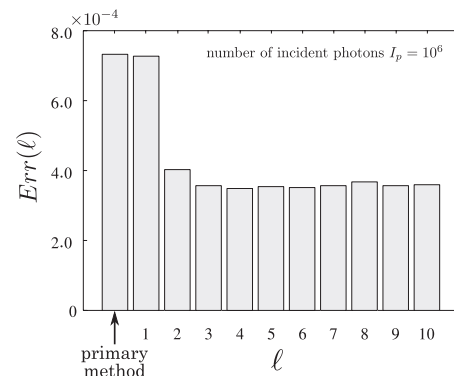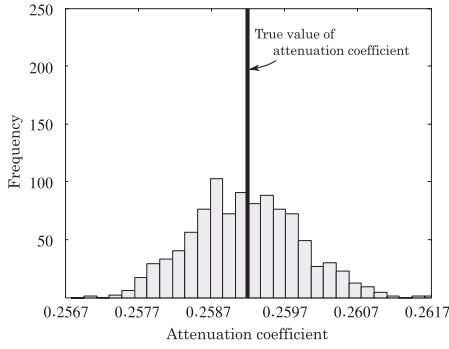

Fig. 3 Error with respect to ℓ.

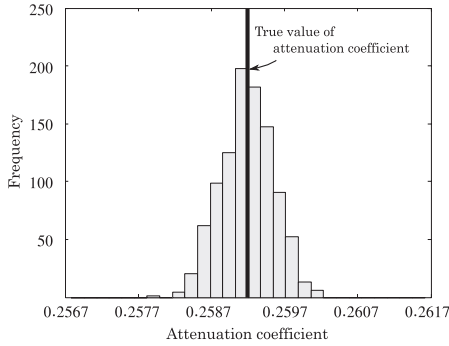**Fig. 4** Estimated distribution of μ using primary X-rays only.

Fig. 5 Estimated distribution of μ using both scattered and primary X-rays ($\ell = 10$).

Under ideal conditions (i.e., scattered photons at any angle are counted by the nearest detector from the scatter point, the number of estimation processes goes to infinity, multiple scattering does not happen, each optimization process is sufficient), is this behavior monotonically decreasing? If it is monotonic, where is the limit? Despite the fact that the scattered photons were measured, is the error of one scatter detector case exactly the same as that of the primary method? Even the exact maximum likelihood method generates a bias in cases with a small sample size, but what is the influence of the pseudo-likelihood? Although we saw a dramatic improvement in accuracy at $2 \leq \ell$, is this a reproducible and obvious phenomenon in the first place? If so, some inequality should be established.

However, we found no theoretical work that answers these questions from a statistical aspect. If a case exists such that scattered X-rays is effective, it can be a milestone for future applications, even if it is derived in an ideal situation. In the next section, to explain the results of the numerical experiment, we present a theoretical analysis based on mathematical statistical methods and introduce further simplification under conditions **C-1** and **C-2**.

4. Theoretical Analysis for a Homogeneous Thin Cylinder

In the previous numerical experiment, some approximations for the procedure or likelihood were introduced assuming
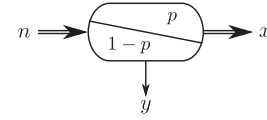

Fig. 6 T-junction model.

actual geometry. In our theoretical analysis here, we further simplify the geometry to treat strictly the probability function and accuracy evaluation.

4.1 Definitions and Evaluation Criterion for Accuracy

First, assuming a scatter detector captures all the scattered photon within its region, we can simplify the three-dimensional geometry into a one-dimensional symbolic representation by using a T-junction, as shown in Fig. 6. We ignore the multiple scattering and photoelectric absorption to derive the probability function strictly based on the binomial distribution as follows. The definitions of $\mathbb{N}$, $\mathbb{R}$, L and μ are the same as those in the previous section. For n $(n \in \mathbb{N}, 1 \leq n)$ incident photons, the number of primary photons x $(x \in \mathbb{N}, x \leq n)$ that pass directly through the T-junction is a random variable following a binomial distribution $B(x \mid n, p)$ [29]:

$$B(x \mid n, p) = \frac{n!}{x! \, (n-x)!} p^x (1-p)^{n-x},$$

where p $(p \in \mathbb{R}, 0 < p < 1)$ is the probability that a photon passes directly through the T-junction. This probability parameter is described by the length and attenuation coefficient of the homogeneous thin cylinder, i.e., $p = \exp(-\mu L)$. For simplicity, we assume that all photons that do not pass directly through the homogeneous thin cylinder are measured as scattered photons. In Fig. 6, y $(y \in \mathbb{N}, y \leq n)$ is a random variable indicating the number of scattered photons $(y = n - x)$. We refer to this model as the T-junction model.

The cases containing several scatter detectors, as shown in Fig. 2, are modeled by several T-junction models connected in series, as shown in Fig. 7. We refer to this as the k-chained T-junction model $(k \in \mathbb{N}, 1 \leq k)$. In this model, the probability parameter of each T-junction model is given by

$$p = p_k(\mu) = \exp\left(-\mu \frac{L}{k}\right). \tag{7}$$

Here, we define the finite set $\mathbb{Z}_n$ by

$$\mathbb{Z}_n = \{x \in \mathbb{N} \mid x \leq n\} \tag{8}$$

and its k direct products $\mathbb{Z}_n^k$ $(k \in \mathbb{N}, 1 \leq k)$ by

$$\mathbb{Z}_n^k = \mathbb{Z}_n \times \mathbb{Z}_n \times \cdots \times \mathbb{Z}_n \ (k \text{ times}). \tag{9}$$

With these definitions, the scattered X-ray photons are given by the measured vector ${}^k\boldsymbol{y} \in \mathbb{Z}_n^k$, which is represented as
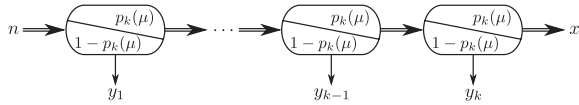
Fig. 7 k-chained T-junction model.

$$^k\boldsymbol{y} = (y_1, y_2, \cdots, y_k)^T, \quad y_i \in \mathbb{Z}_n, i \in \{1, 2, \cdots, k\}. \quad (10)$$

Because the arrival of photons at a detector is a stochastic phenomenon, the estimated attenuation coefficients fluctuate randomly. In the numerical experiment, we employed the mean squared error. As mentioned in the previous section, when the estimated parameter is unbiased, the accuracy is generally evaluated according to the variance. If the simulated projection-based method and primary method are assumed to provide satisfactory efficiency, the variances are given as the Cramer–Rao lower bound [30], which is the lower bound on the variance of the unbiased estimator.

A feature of this bound is that, if the stochastic structures (i.e., probability function) of methods differ, their accuracy can be compared analytically without explicitly constructing an estimator. Fortunately, the stochastic structures of the primary method and simulated projection-based method with several scatter detectors differ from each other, whereas the number of incident photons is identical. Consequently, we adopted this bound as a criterion for evaluating the accuracy improvement and refer to it simply as variance. The variance of the k-chained T-junction model $\mathrm{V}(k)$ is given by

$$\mathrm{V}(k) = -\frac{1}{N \cdot \mathrm{E}\left[\frac{\partial^2}{\partial \mu^2} \log\left(f_k(^k\boldsymbol{y}, x \mid n, p_k(\mu))\right)\right]}, \quad (11)$$

where $f_k(^k\boldsymbol{y}, x \mid n, p_k(\mu))$ is the probability function of the k-chained T-junction model, x and $^k\boldsymbol{y}$ are random variables, μ is the attenuation coefficient to be estimated, $\mathrm{E}[\cdot]$ denotes the expectation operator, and N ($N \in \mathbb{N}, 1 \leq N$) is the number of trials. In the following sections, we derive and analyze the properties of this variance step by step, including that of primary method.

4.2 Derivation of the Variance $\mathrm{V}(k)$

4.2.1 Case $k = 1$

We first consider the case $k = 1$. Because the number of primary photons x in a 1-chained T-junction model obeys a binomial distribution, its probability function $f(x \mid n, p_1(\mu))$ is given by

$$f(x \mid n, p_1(\mu)) = \frac{n!}{x!\,(n-x)!}(p_1(\mu))^x(1 - p_1(\mu))^{n-x}. \quad (12)$$

Because the number of scattered photons y_1 is

$$y_1 = n - x, \quad (13)$$

the probability function $f_1(y_1, x \mid n, p_1(\mu))$ is written as

$$f_1(y_1, x \mid n, p_1(\mu))$$

$$= \begin{cases} \frac{n!}{x!\,y_1!}(p_1(\mu))^x(1 - p_1(\mu))^{y_1}, & (y_1 = n - x) \\ 0, & (\text{otherwise}) \end{cases}. \quad (14)$$

Equation (14) represents the joint probability of y_1 and x. Generally, the expectation value of a function $\phi(y_1, x)$ is given by

$$\mathrm{E}\left[\phi(y_1, x)\right] = \sum_{(y_1, x) \in \mathbb{Z}_n^2} \phi(y_1, x) f_1(y_1, x \mid n, p_1(\mu)). \quad (15)$$

Our probability function $f_1(y_1, x \mid n, p_1(\mu))$ is zero over a large area. Consequently, the summation in Eq. (15) can be restricted to a narrow region. We can calculate the expectation value of Eq. (15) in the set $\mathbb{S}_1 = \{(y_1, x) \in \mathbb{Z}_n^2 \mid y_1 = n - x\}$. Thus, we rewrite Eq. (15) as

$$\mathrm{E}\left[\phi(y_1, x)\right] = \sum_{(y_1, x) \in \mathbb{S}_1} \phi(y_1, x) f_1(y_1, x \mid n, p_1(\mu))$$

$$= \sum_{x=0}^{n} \phi(n - x, x) f_1(n - x, x \mid n, p_1(\mu)). \quad (16)$$

Here, in the set $\mathbb{S}_1$, the second derivative of the function $\log\left(f_1(y_1, x \mid n, p_1(\mu))\right)$ with respect to μ is given by

$$\frac{\partial^2}{\partial \mu^2} \log\left(f_1(y_1, x \mid n, p_1(\mu))\right)$$

$$= \frac{\partial^2}{\partial \mu^2}\left[\log\left(\frac{n!}{x!\,(n-x)!}\right) + x \log\left(p_1(\mu)\right)\right.$$

$$\left. + (n-x) \log\left(1 - p_1(\mu)\right)\right]$$

$$= -\frac{L^2 \exp\left(-\mu L\right)}{\left(1 - \exp\left(-\mu L\right)\right)^2}(n - x). \quad (17)$$

Therefore, by applying Eq. (17) to the denominator of Eq. (11), we obtain

$$\mathrm{V}(1) = -\frac{1}{N \cdot \mathrm{E}\left[\frac{\partial^2}{\partial \mu^2} \log\left(f_1(y_1, x \mid n, p_1(\mu))\right)\right]}$$

$$= \frac{1}{\frac{N L^2 \exp\left(-\mu L\right)}{\left(1 - \exp\left(-\mu L\right)\right)^2}(n - \mathrm{E}\left[x\right])}$$

$$= \frac{1 - \exp\left(-\mu L\right)}{n N L^2 \exp\left(-\mu L\right)}. \quad (18)$$

The variance in the case where only the primary photons are measured (i.e., the primary method), which is denoted by V_p, can be derived immediately by using Eq. (12):

$$\mathrm{V}_p = -\frac{1}{N \cdot \mathrm{E}\left[\frac{\partial^2}{\partial \mu^2} \log\left(f(x \mid n, p_1(\mu))\right)\right]}$$

$$= \frac{1 - \exp\left(-\mu L\right)}{n N L^2 \exp\left(-\mu L\right)}. \quad (19)$$

We can see that $\mathrm{V}_p = \mathrm{V}(1)$. This fact, which we also observed in the numerical experiment, indicates that the additional measurement of scattered X-rays does not influence
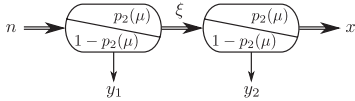
Fig. 8 Two-chained T-junction model.

the arbitrary expected value in the primary method. Consequently, in the case $k = 1$, even if scattered photons are measured, no improvement in accuracy is expected.

4.2.2 Case $k = 2$

In this section, we consider the value of measuring the scattered photons in the case $k = 2$. Let ξ be the number of primary photons of the first (left) T-junction, as shown in Fig. 8.

Here, $q(y_2, x, \xi \mid n, p_2(\mu))$ denotes the joint probability function of (y_2, x, ξ). Then, the conditional probability function $f_1(y_2, x \mid \xi, p_2(\mu))$ given ξ is

$$f_1(y_2, x \mid \xi, p_2(\mu)) = \frac{q(y_2, x, \xi \mid n, p_2(\mu))}{f_1(y_1, \xi \mid n, p_2(\mu))}. \qquad (20)$$

Thus, $q(y_2, x, \xi \mid n, p_2(\mu))$ is represented as

$$q(y_2, x, \xi \mid n, p_2(\mu))$$
$$= f_1(y_1, \xi \mid n, p_2(\mu)) \cdot f_1(y_2, x \mid \xi, p_2(\mu)). \qquad (21)$$

By substituting $n - y_1$ for ξ, we obtain

$$q(y_2, x, n - y_1 \mid n, p_2(\mu)) = q(y_2, x, \xi \mid n, p_2(\mu)). \quad (22)$$

Therefore, the probability function of the two-chained T-junction model is given by

$$f_2(^2\boldsymbol{y}, x \mid n, p_2(\mu))$$
$$= q(y_2, x, n - y_1 \mid n, p_2(\mu))$$
$$= f_1(y_1, \xi \mid n, p_2(\mu)) \cdot f_1(y_2, x \mid \xi, p_2(\mu)), \qquad (23)$$

where $^2\boldsymbol{y} = (y_1, y_2)^T$. Note that $f_2(^2\boldsymbol{y}, x \mid n, p_2(\mu))$ is zero everywhere except for the set

$$\mathbb{S}_2 = \left\{ (^2\boldsymbol{y}, x) \in \mathbb{Z}_n^3 \; \middle| \; \begin{array}{l} 0 \le y_1 \le n, \\ 0 \le y_2 \le n - y_1, \\ 0 \le x \le n - y_1 - y_2 \end{array} \right\}.$$

The probability functions $f_1(y_1, \xi \mid n, p_2(\mu))$ and $f_1(y_2, x \mid \xi, p_2(\mu))$ are given as follows:

$$f_1(y_1, \xi \mid n, p_2(\mu))$$
$$= \frac{n!}{y_1! \, \xi!} (p_2(\mu))^\xi (1 - p_2(\mu))^{y_1}$$
$$= \frac{n!}{y_1! \, (n - y_1)!} (p_2(\mu))^{n-y_1} (1 - p_2(\mu))^{y_1}, \qquad (24)$$

$$f_1(y_2, x \mid \xi, p_2(\mu))$$
$$= \frac{\xi!}{x! \, y_2!} (p_2(\mu))^x (1 - p_2(\mu))^{y_2}$$
$$= \frac{\xi!}{x! \, (\xi - x)!} (p_2(\mu))^x (1 - p_2(\mu))^{\xi - x}$$

$$= \frac{(n - y_1)!}{x! \, (n - y_1 - x)!} (p_2(\mu))^x (1 - p_2(\mu))^{n-y_1-x}. \qquad (25)$$

In the set $\mathbb{S}_2$, $f_2(^2\boldsymbol{y}, x \mid n, p_2(\mu))$ is represented as

$$f_2(^2\boldsymbol{y}, x \mid n, p_2(\mu))$$
$$= f_1(y_1, \xi \mid n, p_2(\mu)) \cdot f_1(y_2, x \mid \xi, p_2(\mu))$$
$$= \frac{n!}{x! \, (n - y_1 - x)! \, y_1!} (p_2(\mu))^{x+n-y_1} (1 - p_2(\mu))^{n-x}. \quad (26)$$

Otherwise (in the complementary set of $\mathbb{S}_2$, i.e., $\mathbb{S}_2^c = \mathbb{Z}_n^3 \backslash \mathbb{S}_2$), $f_2(^2\boldsymbol{y}, x \mid n, p_2(\mu)) = 0$.

Similarly to the case $k = 1$, the expectation value of the function $\phi(^2\boldsymbol{y}, x)$ in the set $\mathbb{S}_2$ is given by

$$\mathrm{E}\left[\phi(^2\boldsymbol{y}, x)\right] = \sum_{(^2\boldsymbol{y}, x) \in \mathbb{S}_2} \phi(^2\boldsymbol{y}, x) f_2(^2\boldsymbol{y}, x \mid n, p_2(\mu))$$
$$= \sum_{y_1=0}^{n} \sum_{x=0}^{n-y_1} \phi(^2\boldsymbol{y}, x) f_2(^2\boldsymbol{y}, x \mid n, p_2(\mu)). \quad (27)$$

Then, $\mathrm{V}(2)$ is derived as

$$\mathrm{V}(2) = \frac{4 \, (1 - \exp(-\mu L/2))}{nNL^2 \, (1 + \exp(-\mu L/2)) \exp(-\mu L/2)}. \quad (28)$$

We have

$$\mathrm{V}(1) - \mathrm{V}(2) = \frac{1 - \exp(-\mu L)}{nNL^2 \exp(-\mu L)}$$
$$- \frac{4 \, (1 - \exp(-\mu L/2))}{nNL^2 \, (1 + \exp(-\mu L/2)) \exp(-\mu L/2)}$$
$$= \frac{(1 - \exp(-\mu L/2))^3}{nNL^2 \, (1 + \exp(-\mu L/2)) \exp(-\mu L)} > 0. \quad (29)$$

Therefore, the following statement holds:

$$\mathrm{V}(1) > \mathrm{V}(2). \qquad (30)$$

Equation (30) indicates that the accuracy is improved if the scattered photons y_1 are measured along with the primary photons x. Consequently, we predict that the accuracy is improved by increasing the number of scatter detectors.

4.2.3 General Case $k \ge 2$

Here, we present the general case $k \ge 2$. Let $n \in \mathbb{N}, k \in \mathbb{N}, {}^k\boldsymbol{y} \in \mathbb{Z}_n^k, x \in \mathbb{Z}_n$,

$$\mathbb{S}_k = \left\{ ({}^k\boldsymbol{y}, x) \in \mathbb{Z}_n^{k+1} \; \middle| \; \begin{array}{l} 0 \le y_1 \le n, 0 \le y_2 \le n - y_1, \\ 0 \le y_3 \le n - y_1 - y_2, \cdots, \\ 0 \le y_k \le n - \sum_{\ell=1}^{k-1} y_\ell, \\ 0 \le x \le n - \sum_{\ell=1}^{k} y_\ell \end{array} \right\},$$

and $\mathbb{S}_k^c = \mathbb{Z}_n^{k+1} \backslash \mathbb{S}_k$.

Lemma 1. *The probability function $f_k({}^k\boldsymbol{y}, x \mid n, p_k(\mu))$ ($k \ge 2$) of the k-chained T-junction model is given by*

$$f_k({}^k\boldsymbol{y}, x \mid n, p_k(\mu))$$

$$= \begin{cases} \dfrac{n!}{x! \left(n - \sum\limits_{i=1}^{k-1} y_i - x\right)! \prod\limits_{i=1}^{k-1} y_i!} \\ \quad \cdot (p_k(\mu))^{x + (k-1)n - \sum\limits_{i=1}^{k-1}(k-i)y_i}, \quad ({}^k\boldsymbol{y}, x) \in \mathbb{S}_k \\ \quad \cdot (1 - p_k(\mu))^{n-x} \\ 0, \qquad\qquad\qquad\qquad\qquad ({}^k\boldsymbol{y}, x) \in \mathbb{S}_k^c \end{cases} \tag{31}$$

The proof of Lemma 1 is given in Appendix A.

Using Lemma 1, we derive the variance of the k-chained T-junction model in the following lemma.

Lemma 2. *The variance* $V(k)$ $(k \geq 1)$ *of the estimated parameter μ is given by*

$$V(k) = \frac{k^2 \left(1 - \exp\left(-\mu L/k\right)\right)^2}{nNL^2 \left(1 - \exp\left(-\mu L\right)\right) \exp\left(-\mu L/k\right)}. \tag{32}$$

The proof of Lemma 2 is given in Appendix B.

4.3 Properties of $V(k)$

Here, we present some properties of $V(k)$.

4.3.1 Monotonic Decreasing

Theorem . *For all $k \geq 1$, the following inequality holds:*

$$V(k) > V(k + 1). \tag{33}$$

To prove this theorem, we need Lemma 3.

Lemma 3. *For all t $(t \in \mathbb{R}, 1 \leq t)$ and u $(u \in \mathbb{R}, 0 < u)$, the following function is always positive:*

$$g(t) = u + u \exp\left(\frac{u}{t}\right) + 2t - 2t \exp\left(\frac{u}{t}\right). \tag{34}$$

The proof of Lemma 3 is given in Appendix C.

Proof of Theorem

Proof. Let $V_{\text{real}}(t)$ be a function of the real variable t $(t \in \mathbb{R}, 1 \leq t)$. This is found by replacing k with t in the expression for $V(k)$ given in Lemma 2, i.e.,

$$V_{\text{real}}(t) = \frac{t^2 \left(1 - \exp\left(-\mu L/t\right)\right)^2}{nNL^2 \left(1 - \exp\left(-\mu L\right)\right) \exp\left(-\mu L/t\right)}. \tag{35}$$

Then, the first derivative of the function $V_{\text{real}}(t)$ with respect to t is given by

$$\frac{d}{dt} V_{\text{real}}(t) = -\frac{1 - \exp\left(-\mu L/t\right)}{nNL^2 \left(1 - \exp\left(-\mu L\right)\right)}$$
$$\cdot \left[\mu L + \mu L \exp\left(\mu \frac{L}{t}\right) + 2t - 2t \exp\left(\mu \frac{L}{t}\right)\right]. \tag{36}$$

The fraction on the right side of Eq. (36) is obviously positive. Furthermore, the expression inside the square brackets is also positive because it is exactly the function in Lemma 3, with u replaced by μL. Consequently, we have

$$\frac{d}{dt} V_{\text{real}}(t) < 0. \tag{37}$$

Thus, the function $V_{\text{real}}(t)$ is a strictly decreasing function. Because k is in the set $\{t \in \mathbb{R} \mid 1 \leq t\}$, we arrive at our goal:

$$V(k) > V(k + 1). \tag{38}$$

∎

Our theorem states that the accuracy of the estimated attenuation coefficient is improved by increasing the number of scatter detectors.

4.3.2 Reduction of the Number of Incident Photons

Because $V(k)$ decreases monotonically with respect to k and is positive, it converges to a limit. By putting $u = \mu L/k$, the limit of $V(k)$ as k approaches infinity is given as follows:

$$V_\infty = \lim_{k \to \infty} V(k)$$
$$= \lim_{u \to 0} \frac{\mu^2 L^2 \left(\frac{\exp(u) - 1}{u}\right)^2}{nNL^2 \left(1 - \exp\left(-\mu L\right)\right) \exp(u)}$$
$$= \frac{\mu^2}{nN \left(1 - \exp\left(-\mu L\right)\right)}. \tag{39}$$

In this derivation, we used the following formula for the real number η:

$$\lim_{\eta \to +0} \frac{\exp(\eta) - 1}{\eta} = 1. \tag{40}$$

Achieving a variance less than the limit V_∞ with only the primary detector requires that we use a greater number of incident photons. That is, if we use a sufficient number of scatter detectors, the number of incident photons can be reduced to get the same variance as in the primary method. Let n' be the number of incident photons in the primary method for which the variance is that given by Eq. (39).

From $V_p = V_\infty$,

$$\frac{1 - \exp\left(-\mu L\right)}{n'NL^2 \exp\left(-\mu L\right)} = \frac{\mu^2}{nN \left(1 - \exp\left(-\mu L\right)\right)}. \tag{41}$$

We then have

$$\frac{n}{n'} = \frac{\mu^2 L^2 \exp\left(-\mu L\right)}{(1 - \exp\left(-\mu L\right))^2}. \tag{42}$$

To evaluate Eq. (42), we define a function $r(u)$ by subtracting the numerator from the denominator and replacing μL with u $(0 < u)$ as follows:

$$r(u) = (1 - \exp\left(-u\right))^2 - u^2 \exp\left(-u\right). \tag{43}$$

The derivative of $r(u)$ is greater than 0:

$$\frac{d}{du} r(u) = \left((u - 1)^2 \exp(u) + \exp(u) - 2\right) \exp\left(-2u\right). \tag{44}$$

This is because the expression multiplying $\exp(-2u)$ satisfies

$$\frac{d}{du}\left((u-1)^2\exp(u)+\exp(u)-2\right)=u^2\exp(u)>0 \quad (45)$$

and

$$\lim_{u\to+0}\left((u-1)^2\exp(u)+\exp(u)-2\right)=0. \quad (46)$$

Thus, $r(u)$ is strictly increasing. Furthermore, because we have

$$\lim_{u\to+0} r(u)=0, \quad (47)$$

$r(u)$ is always positive. Consequently, the value of Eq. (42) is less than 1. This means that the number of incident photons can be reduced. In Sect. 4.4, we show a numerical example of this reduction ratio. This property can be used as a principle for dose reduction.

In addition, V_∞ is thought to be related to the amount of information that can be obtained with a finite number of incident photons. Further investigation of this is required.

4.4 Numerical Example of the Theorem Based on the Maximum Likelihood Method

In the previous sections a theory was constructed. In order to make the theory more understandable, here, we explain with numerical examples.

Because a strict maximum likelihood estimator has asymptotic efficiency [30], we can verify statistical properties such as the variance by using numerical estimates with a large number of samples. Exact derivation of the maximum likelihood estimator of the attenuation coefficient of a homogeneous thin cylinder is difficult, but that of the k-chained T-junction model is derived as follows.

For Eq. (31), when the jth ($j\in\mathbb{Z}_N$) set of measured values $(x^{(j)}, {}^k\boldsymbol{y}^{(j)})^T$ is given from N independent trials, the log-likelihood function L_h becomes

$$L_h(\mu|k,n)=\log\left[\prod_{j=1}^N f_k({}^k\boldsymbol{y}^{(j)}, x^{(j)} \mid n, p_k(\mu))\right]. \quad (48)$$

In the case $k=1$, according to Eq. (14), the log-likelihood function of Eq. (48) is given by

$$L_h(\mu|1,n)=\sum_{j=1}^N\log\left[\frac{n!}{x^{(j)}!\,(n-x^{(j)})!}\right]-\mu L\sum_{j=1}^N x^{(j)}$$
$$+\sum_{j=1}^N(n-x^{(j)})\log\left(1-\exp(-\mu L)\right). \quad (49)$$

When $k\geq 2$, Eq. (26) and Lemma 1 indicate that this becomes

$$L_h(\mu|k,n)=\sum_{j=1}^N\log\left[\frac{n!}{x^{(j)}!\left(n-\sum_{i=1}^{k-1}y_i^{(j)}-x^{(j)}\right)!\prod_{i=1}^{k-1}y_i^{(j)}!}\right.$$

$$-\mu\frac{L}{k}\sum_{j=1}^N\left(x^{(j)}+(k-1)n-\sum_{i=1}^{k-1}(k-i)\,y_i^{(j)}\right)$$

$$+\sum_{j=1}^N(n-x^{(j)})\log\left[1-\exp\left(-\mu\frac{L}{k}\right)\right]. \quad (50)$$

When $\widehat{\mu}_k$ is set as μ such that the log-likelihood is maximized, then for $k=1$ we can solve

$$\left.\frac{\partial}{\partial\mu}L_h(\mu|1,n)\right|_{\mu=\widehat{\mu}_1}$$

$$=-L\sum_{j=1}^N x^{(j)}+\frac{L\exp(-\widehat{\mu}L)}{1-\exp(-\widehat{\mu}L)}\sum_{j=1}^N(n-x^{(j)})=0.$$

We thus obtain

$$\widehat{\mu}_1=-\frac{1}{L}\log\left[\frac{1}{nN}\sum_{j=1}^N x^{(j)}\right]. \quad (51)$$

For $k\geq 2$, from

$$\left.\frac{\partial}{\partial\mu}L_h(\mu|k,n)\right|_{\mu=\widehat{\mu}_k}=\frac{L}{k}\cdot\frac{\exp(-\widehat{\mu}L/k)}{1-\exp(-\widehat{\mu}L/k)}\sum_{j=1}^N(n-x^{(j)})$$

$$-\frac{L}{k}\sum_{j=1}^N\left(x^{(j)}+(k-1)n-\sum_{i=1}^{k-1}(k-i)\,y_i^{(j)}\right)=0,$$

we have

$$\widehat{\mu}_k=\frac{k}{L}\log\left[\frac{knN-\sum\limits_{i=1}^{k-1}\left((k-i)\sum\limits_{j=1}^N y_i^{(j)}\right)}{(k-1)nN-\sum\limits_{i=1}^{k-1}\left((k-i)\sum\limits_{j=1}^N y_i^{(j)}\right)+\sum\limits_{j=1}^N x^{(j)}}\right]. \quad (52)$$

Thus, we derive the maximum likelihood estimator of μ for the k-chained T-junction model. However, it is difficult to analytically derive the variance. Therefore, we can obtain M_{model} estimated attenuation coefficients with random numbers by using Eq. (52). The variance $\widehat{V}(k)$ is calculated as

$$\widehat{V}(k)=\frac{1}{M_{\text{model}}}\sum_{i=1}^{M_{\text{model}}}\left(\widehat{\mu}_k^{(i)}-\overline{\widehat{\mu}_k}\right)^2, \quad (53)$$

where $\overline{\widehat{\mu}_k}$ is the mean given by

$$\overline{\widehat{\mu}_k}=\frac{1}{M_{\text{model}}}\sum_{i=1}^{M_{\text{model}}}\widehat{\mu}_k^{(i)}, \quad (54)$$

and $\widehat{\mu}_k^{(i)}$ denotes the ith estimated value.

Here, we show a numerical example. We used the specific values $n=10^5$, $N=10$, $M_{\text{model}}=10^6$, $\mu=0.17279637$, and $L=20\,\text{cm}$. Figure 9 shows the theoretical variance given by Eq. (32). The bar at the left end of
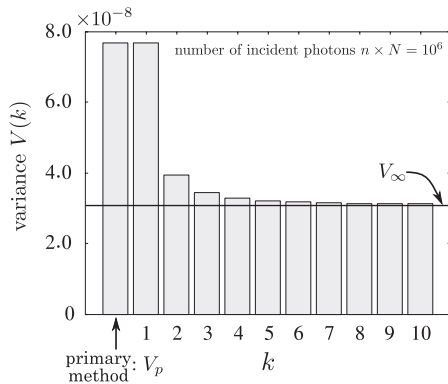
Fig. 9 Theoretically derived variance $V(k)$.
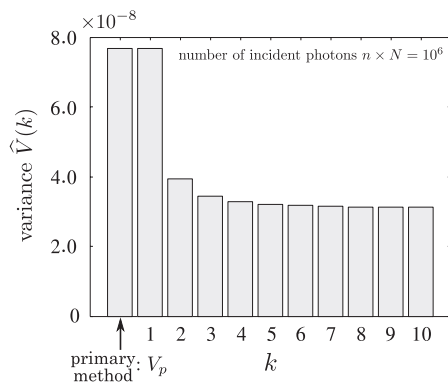

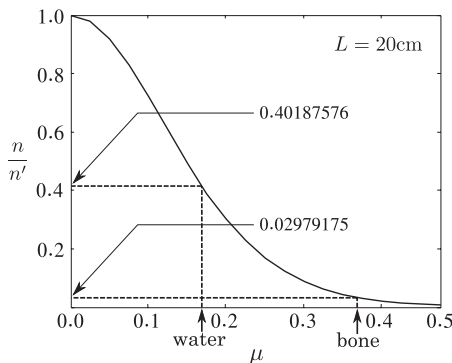**Fig. 10** Variance $\widehat{V}(k)$ estimated numerically with the maximum likelihood method.

Fig. 11 Dose reduction ratio.

the diagram indicates the variance for the primary method. Figure 10 shows the numerically estimated variance $\widehat{V}(k)$ given by Eq. (53). Because both of these quantities exhibit remarkable agreement, the theoretical analysis (Lemma 2, Lemma 3, and the Theorem) in the previous section was verified from numerical aspects.

We also calculated the dose-reduction ratio given by Eq. (42) for this specific condition. Figure 11 shows this ratio with respect to the attenuation coefficient at a length of $L = 20$ cm. According to this example, if the material is water ($\mu = 0.17279637$) or bone ($\mu = 0.37802092$), the

ratio is 0.40187576 or 0.02979175, respectively. In an ideal environment, a significant dose reduction can be expected.

5. Conclusions and Discussion

By means of a Monte Carlo simulation, we conducted a numerical experiment on simplified CT geometry with a simulated projection-based method and found that the accuracy improves when several dedicated scatter detectors are used. To explain this result, we presented a theoretical analysis with further simplification and idealization. Our results, especially the theorem in the theoretical analysis, should provide a strong motivation for verifying the effects of using scattered X-rays in actual CT geometry, as shown in Fig. 1(b).

The decrease in variance means a decrease in the ambiguity of the information about the object. This implies that scattered X-rays contain information independent from that provided by the primary X-rays.

We also clarified the existence of a limit in the accuracy achieved by using both primary and scattered X-rays. Because this limit exceeds the accuracy that can be achieved with primary X-rays only, the degree of approximation with this limit may provide a new performance criterion for future medical X-ray diagnostic devices.

Our theoretical model (k-chained T-junction model) represents a situation in which many detectors are available to extract information from a small portion of interest of an object. Thus, our results provide a basis for using a sufficient number of detectors and calculation resources to reduce the patient dose.

In future work, to put our principle into practice, various factors such as multiple scattering, photoelectric absorption, and energy characteristics have to be considered. Furthermore, other principles for utilizing scattered X-rays need to be investigated.

Acknowledgments

References

[1] J. Hsieh, Computed Tomography: Principles, Design, Artifacts and Recent Advances, SPIE Press, Bellingham, 2003.

[2] H. Kanamori, N. Nakamori, K. Inoue, and E. Takenaka, "Effects of scattered x-rays on CT images," Phys. Med. Biol., vol.30, no.3, pp.239–249, 1985.

[3] K. Tokumoto, Y. Yamazaki, N. Toda, and S. Koyama, "Evaluation of scattered radiation in cone-beam dual-energy X-ray CT with Monte Carlo simulation," The Radiological Society of North America (RSNA) 98th Scientific Assembly and Annual Meeting, pp.SSG16–02, 2012.

[4] Y. Kyriakou and W. Kalender, "Efficiency of antiscatter grids for flat-detector CT," Phys. Med. Biol., vol.52, no.20, pp.6275–6293, 2007.

[5] H. Gao, J. Boudry, A. Cohen, Y. Imai, and R. Melnyk, "Revolution CT: Overcoming wide coverage IQ challenges," White Paper, 2014.

[6] L. Zhu, Y. Xie, J. Wang, and L. Xing, "Scatter correction for cone-beam CT in radiation therapy," Med. Phys., vol.36, no.6, pp.2258–2268, 2009.

[7] A. Akbarzadeh, M.R. Ay, H. Ghadiri, S. Sarkar, and H. Zaidi, "Measurement of scattered radiation in a volumetric 64-slice CT scanner using three experimental techniques," Phys. Med. Biol., vol.55, no.8, pp.2269–2280, 2010.

[8] H. Yan, X. Mou, S. Tang, Q. Xu, and M. Zankl, "Projection correlation based view interpolation for cone beam CT: Primary fluence restoration in scatter measurement with a moving beam stop array," Phys. Med. Biol., vol.55, no.21, pp.6353–6375, 2010.

[9] W. Zhao, S. Brunner, K. Niu, S. Schafer, K. Royalty, and G.H. Chen, "Patient-specific scatter correction for flat-panel detector-based cone-beam CT imaging," Phys. Med. Biol., vol.60, no.3, pp.1339–1365, 2015.

[10] M. Baer and M. Kachelrieß, "Hybrid scatter correction for CT imaging," Phys. Med. Biol., vol.57, no.21, pp.6849–6867, 2012.

[11] G. Jarry, S.A. Graham, D.J. Moseley, D.J. Jaffray, J.H. Siewerdsen, and F. Verhaegen, "Characterization of scattered radiation in kV CBCT images using Monte Carlo simulations," Med. Phys., vol.33, no.11, pp.4320–4329, 2006.

[12] Y. Xu, T. Bai, H. Yan, L. Ouyang, A. Pompos, J. Wang, L. Zhou, S.B. Jiang, and X. Jia, "A practical cone-beam CT scatter correction method with optimized Monte Carlo simulations for image-guided radiation therapy," Phys. Med. Biol., vol.60, no.9, pp.3567–3587, 2015.

[13] G. Bootsma, F. Verhaegen, and D. Jaffray, "Efficient scatter distribution estimation and correction in CBCT using concurrent Monte Carlo fitting," Med. Phys., vol.42, no.1, pp.54–68, 2015.

[14] T. Truong and M.K. Nguyen, Recent Developments in Compton Scatter Tomography: Theory and Numerical Simulations (Numerical Simulation: From Theory to Industry), pp.101–128, INTECH Open Access Publisher, 2012.

[15] P. Lale, "The examination of internal tissues, using Gamma-ray scatter with a possible extension to megavoltage radiography," Phys. Med. Biol., vol.4, no.2, pp.159–167, 1959.

[16] J. Battista and M. Bronskill, "Compton scatter imaging of transverse sections: An overall appraisal and evaluation for radiotherapy planning," Phys. Med. Biol., vol.26, no.1, pp.81–99, 1981.

[17] S.J. Norton, "Compton scattering tomography," J. Appl. Phys., vol.76, no.4, pp.2007–2015, 1994.

[18] M.K. Nguyen, T.T. Truong, M. Morvidone, and H. Zaidi, "Scattered radiation emission imaging: Principles and applications," Journal of Biomedical Imaging, vol.2011, pp.13–28, 2011.

[19] F. Zhao, J.C. Schotland, and V.A. Markel, "Inversion of the star transform," Inverse Probl., vol.30, no.10, pp.105001–105032, 2014.

[20] R. Krylov and A. Katsevich, "Inversion of the broken ray transform in the case of energy-dependent attenuation," Phys. Med. Biol., vol.60, no.11, pp.4313–4334, 2015.

[21] T. Truong and M. Nguyen, "New properties of the V-line Radon transform and their imaging applications," J. Phys. A: Math. Theor., vol.48, no.40, pp.405204–405232, 2015.

[22] J.B. Thaibault, K.D. Sauer, C.A. Bouman, and J. Hsieh, "A three-dimensional statistical approach to improved image quality for multislice helical CT," Med. Phys., vol.34, no.11, pp.4526–4544, 2007.

[23] Z. Yu, J.B. Thaibault, C.A. Bouman, K.D. Sauer, and J. Hsieh, "Fast model-based X-ray CT reconstruction using spatially nonhomogeneous ICD optimization," IEEE Trans. Image Process., vol.20, no.1, pp.161–175, 2011.

[24] K. Yasaka, M. Katsura, M. Akahane, J. Sato, I. Matsuda, and K. Ohtomo, "Model-based iterative reconstruction for reduction of radiation dose in abdominopelvic CT: Comparison to adaptive statistical iterative reconstruction," SpringerPlus, vol.2, no.209, 2013.

[25] J.A. Fessler, I.A. Elbakri, P. Sukovic, and N.H. Clinthorne, "Maximum-likelihood dual-energy tomographic image reconstruction," Medical Imaging, pp.38–49, 2002.

[26] J.A. O'Sullivan and J. Benac, "Alternating minimization algorithms for transmission tomography," IEEE Trans. Med. Imag., vol.26, no.3, pp.283–297, 2007.

[27] H. Hirayama, Y. Namito, and A. Bielajew, "The EGS5 code system, report 2005-8," High Energy Accelerator Research Organization, Tsukuba, Japan, 2009.

[28] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, Numerical Recipes in C, pp.397–402, Cambridge university press Cambridge, New York, NY, 1996.

[29] P. Billingsley, Probability and Measure, 3rd ed., p.256, Wiley, New York, NY, 1995.

[30] E.L. Lehmann and G. Casella, Theory of Point Estimation, 2nd ed., pp.113–123, Springer, Germany, 2003.

Appendix A: Proof of Lemma 1

Proof. We can prove the statement by mathematical induction.

Basis: Show that the statement holds for $k = 2$. From 4.2.2, in the set $\mathbb{S}_2$, $f_2(^2\boldsymbol{y}, x \mid n, p_2(\mu))$ is given by Eq. (26).

Inductive step: We assume that Eq. (31) holds when $k = m$, i.e.,

$$f_m(^m\boldsymbol{y}, x \mid n, p_m(\mu))$$
$$= \frac{n!}{x! \left(n - \sum_{i=1}^{m-1} y_i - x\right)! \prod_{i=1}^{m-1} y_i!}$$
$$\cdot (p_m(\mu))^{x+(m-1)n - \sum_{i=1}^{m-1}(m-i)y_i} (1 - p_m(\mu))^{n-x}. \quad (A\cdot 1)$$

The probability function $f_{m+1}(^{m+1}\boldsymbol{y}, x \mid n, p_{m+1}(\mu))$ in the set $\mathbb{S}_{m+1}$ of the model shown in Fig. A· 1 is given by

$$f_{m+1}(^{m+1}\boldsymbol{y}, x \mid n, p_{m+1}(\mu))$$
$$= f_m(^m\boldsymbol{y}, \xi \mid n, p_{m+1}(\mu)) \cdot f_1(y_{m+1}, x \mid \xi, p_{m+1}(\mu)). \quad (A\cdot 2)$$

By using the inductive hypothesis and Eq. (12), we get

$$f_{m+1}(^{m+1}\boldsymbol{y}, x \mid n, p_{m+1}(\mu))$$
$$= \frac{n!}{x! \, (\xi - x)! \left(n - \sum_{i=1}^{m-1} y_i - \xi\right)! \prod_{i=1}^{m-1} y_i!}$$
$$\cdot (p_{m+1}(\mu))^{\xi+x+(m-1)n - \sum_{i=1}^{m-1}(m-i)y_i} (1 - p_{m+1}(\mu))^{n-x}. \quad (A\cdot 3)$$

Because ξ is given by

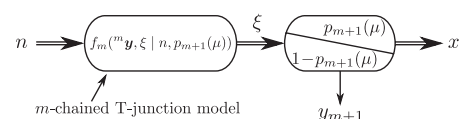$$\xi = n - \sum_{i=1}^{m} y_i, \quad (A\cdot 4)$$

Equation (A· 3) becomes

Fig. A· 1 The $(m + 1)$-chained T-junction model.

$$f_{m+1}({}^{m+1}\boldsymbol{y}, x \mid n, p_{m+1}(\mu))$$

$$= \frac{n!}{x! \left(n - \sum_{i=1}^{m} y_i - x\right)!}$$

$$\cdot \frac{1}{\left(n - \sum_{i=1}^{m-1} y_i - \left(n - \sum_{i=1}^{m} y_i\right)\right)! \prod_{i=1}^{m-1} y_i!}$$

$$\cdot (p_{m+1}(\mu))^{n - \sum_{i=1}^{m} y_i + x + (m-1)n - \sum_{i=1}^{m-1}(m-i)y_i}$$

$$\cdot (1 - p_{m+1}(\mu))^{n-x}$$

$$= \frac{n!}{x! \left(n - \sum_{i=1}^{m} y_i - x\right)! \prod_{i=1}^{m} y_i!}$$

$$\cdot (p_{m+1}(\mu))^{x + mn - \sum_{i=1}^{m}(m+1-i)y_i} (1 - p_{m+1}(\mu))^{n-x}. \quad (\text{A} \cdot 5)$$

Thus, Eq. (31) holds for $k = m + 1$, and Eq. (31) holds for all natural numbers $k \geq 2$. ∎

Appendix B: Proof of Lemma 2

Proof. The expected value of the function $\phi({}^{k}\boldsymbol{y}, x)$ ($k \geq 2$) is given by

$$E\left[\phi({}^{k}\boldsymbol{y}, x)\right] = \sum_{({}^{k}\boldsymbol{y}, x) \in \mathbb{S}_k} \phi({}^{k}\boldsymbol{y}, x) f_k({}^{k}\boldsymbol{y}, x \mid n, p_k(\mu)). \quad (\text{A} \cdot 6)$$

We can rewrite Eq. (A·6) in the same manner as for the case $k = 1$:

$$E\left[\phi({}^{k}\boldsymbol{y}, x)\right]$$

$$= \sum_{y_1=0}^{n} \sum_{y_2=0}^{n-y_1} \cdots \sum_{x=0}^{n-\sum_{i=1}^{k-1} y_i} \phi({}^{k}\boldsymbol{y}, x) f_k({}^{k}\boldsymbol{y}, x \mid n, p_k(\mu)). \quad (\text{A} \cdot 7)$$

Here, in the set $\mathbb{S}_k$, the second derivative of the function $\log(f_k({}^{k}\boldsymbol{y}, x \mid n, p_k(\mu)))$ is given by

$$\frac{\partial^2}{\partial \mu^2} \log(f_k({}^{k}\boldsymbol{y}, x \mid n, p_k(\mu)))$$

$$= \frac{\partial^2}{\partial \mu^2} \left[\log\left(\frac{n!}{x! \left(n - \sum_{i=1}^{k-1} y_i - x\right)! \prod_{i=1}^{k-1} y_i!} \right) \right.$$

$$+ \left(x + (k-1)n - \sum_{i=1}^{k-1}(k-i) y_i\right) \log(p_k(\mu))$$

$$\left. + (n-x) \log(1 - p_k(\mu)) \right]$$

$$= \frac{L}{k}(n-x) \cdot \frac{1}{(1 - \exp(-\mu L/k))^2}$$

$$\cdot \left[-\frac{L}{k} \exp(-\mu L/k)(1 - \exp(-\mu L/k)) \right.$$

$$\left. - \frac{L}{k} \exp(-2\mu L/k) \right]$$

$$= -\frac{L^2}{k^2} \cdot \frac{\exp(-\mu L/k)}{(1 - \exp(-\mu L/k))^2}(n - x). \quad (\text{A} \cdot 8)$$

By replacing $\phi({}^{k}\boldsymbol{y}, x)$ in Eq. (A·6) with $\frac{\partial^2}{\partial \mu^2} \log(f_k({}^{k}\boldsymbol{y}, x \mid n, p_k(\mu)))$ and using Lemma 1, the expected value of the function $\phi({}^{k}\boldsymbol{y}, x)$ is given by

$$E\left[\frac{\partial^2}{\partial \mu^2} \log(f_k({}^{k}\boldsymbol{y}, x \mid n, p_k(\mu)))\right]$$

$$= -\frac{L^2}{k^2} \cdot \frac{p_k(\mu)}{(1 - p_k(\mu))^2} E[n - x]$$

$$= -\frac{L^2 p_k(\mu)}{k^2 (1 - p_k(\mu))^2} (n - E[x]). \quad (\text{A} \cdot 9)$$

The expected value is given as follows:

$$E[x]$$

$$= \sum_{y_1=0}^{n} \sum_{y_2=0}^{n-y_1} \cdots \sum_{x=0}^{n-\sum_{i=1}^{k-1} y_i} x \cdot f_k({}^{k}\boldsymbol{y}, x \mid n, p_k(\mu))$$

$$= \sum_{y_1=0}^{n-1} \frac{n!}{y_1! (n - y_1)!}(p_k(\mu))^{n-y_1}(1 - p_k(\mu))^{y_1}$$

$$\sum_{y_2=0}^{n-1-y_1} \frac{(n - y_1)!}{y_2! (n - y_1 - y_2)!}(p_k(\mu))^{n-y_1-y_2}(1 - p_k(\mu))^{y_2}$$

$$\cdots \sum_{x=1}^{n-\sum_{i=1}^{k-1} y_i} \frac{\left(n - \sum_{i=1}^{k-1} y_i\right)!}{(x-1)! \left(n - \sum_{i=1}^{k-1} y_i - x\right)!}$$

$$\cdot (p_k(\mu))^x (1 - p_k(\mu))^{n-\sum_{i=1}^{k-1} y_i - x}. \quad (\text{A} \cdot 10)$$

Here, by putting $\xi = x - 1$, $\left(0 \leq \xi \leq n - 1 - \sum_{i=1}^{k-1} y_i\right)$, we get

$$E[x]$$

$$= \sum_{y_1=0}^{n-1} \frac{n!}{y_1! (n - y_1)!}(p_k(\mu))^{n-y_1}(1 - p_k(\mu))^{y_1}$$

$$\sum_{y_2=0}^{n-1-y_1} \frac{(n - y_1)!}{y_2! (n - y_1 - y_2)!}(p_k(\mu))^{n-y_1-y_2}(1 - p_k(\mu))^{y_2}$$

$$\cdots \sum_{\xi=0}^{n-1-\sum_{i=1}^{k-1} y_i} \frac{\left(n - \sum_{i=1}^{k-1} y_i\right)!}{\xi! \left(n - \sum_{i=1}^{k-1} y_i - (\xi+1)\right)!}$$

$$\cdot (p_k(\mu))^{\xi+1}(1 - p_k(\mu))^{n-\sum_{i=1}^{k-1} y_i - (\xi+1)}$$

$$= n(p_k(\mu))^k \left[\sum_{y_1=0}^{n-1} \frac{(n-1)!}{y_1! (n-1-y_1)!}(p_k(\mu))^{n-1-y_1} \right.$$

$$\cdot (1 - p_k(\mu))^{y_1} \sum_{y_2=0}^{n-1-y_1} \frac{(n-1-y_1)!}{y_2! \, (n-1-y_1-y_2)!}$$

$$\cdot (p_k(\mu))^{n-1-y_1-y_2} (1 - p_k(\mu))^{y_2}$$

$$\cdots \sum_{\xi=0}^{n-1-\sum_{i=1}^{k-1} y_i} \frac{\left(n-1-\sum_{i=1}^{k-1} y_i\right)!}{\xi! \left(n-1-\sum_{i=1}^{k-1} y_i - \xi\right)!}$$

$$\left. \cdot (p_k(\mu))^{\xi} (1 - p_k(\mu))^{n-1-\sum_{i=1}^{k-1} y_i - \xi} \right]$$

$$= n(p_k(\mu))^k = n \exp(-\mu L). \tag{A·11}$$

Therefore, the expected value of the function $\phi(^k\boldsymbol{y}, x)$ is given by

$$E\left[\phi(^k\boldsymbol{y}, x)\right] = -\frac{nL^2 \left(1 - \exp(-\mu L)\right) \exp(-\mu L/k)}{k^2 \left(1 - \exp(-\mu L/k)\right)^2}. \tag{A·12}$$

Consequently, the variance $V(k)$ of the estimated parameter μ is given by Eq. (32).

When we replace $V(k)$ in Eq. (32) with $k = 1$, the result is Eq. (18). Therefore, for all $k \geq 1$, the variance is given by Eq. (32). ∎

Appendix C: Proof of Lemma 3

Proof. The first derivative of the function $g(t)$ $(1 \leq t)$ is strictly increasing:

$$\frac{d}{dt} g(t) = 2 - \frac{2t^2 - 2ut + u^2}{t^2} \exp\left(\frac{u}{t}\right). \tag{A·13}$$

This is because, for all $t \geq 1$ and $u > 0$,

$$\frac{d^2}{dt^2} g(t) = \frac{u^3}{t^4} \exp\left(\frac{u}{t}\right) > 0. \tag{A·14}$$

Moreover, the limit of Eqs. (34) and (A·13) as t approaches infinity is given as follows:

$$\lim_{t \to \infty} g(t) = 0, \tag{A·15}$$

$$\lim_{t \to \infty} \frac{d}{dt} g(t) = 0. \tag{A·16}$$

Therefore, Eq. (A·13) is always negative, i.e., $\frac{d}{dt} g(t) < 0$. Consequently, because $g(t)$ is a strictly decreasing function, the function $g(t)$ in Eq. (A·15) is always positive, i.e., $g(t) > 0$. ∎

Naohiro Toda received a doctor of engineering degree from the Graduate School of Toyohashi University of Technology in 1990. He became an Assistant Professor in Information Engineering at the same university. In 1992, he became a Lecturer on Electrical Engineering at Maizuru National College of Technology. In 1994, he became a Lecturer on Information Engineering at Toyohashi University of Technology. He also served as a Visiting Researcher at the University of Southern California in 1995. In 1999, he became an Associate Professor in the Department of Information Science and Technology of Aichi Prefectural University, and in 2002 he became a Professor. His research interests are medical imaging, signal processing, and nonlinear time series analysis. He is a full member of the IEEE and Society of Instrument and Control Engineers.

Tetsuya Nakagami received his master's degree in information science from the Graduate School of Aichi Prefectural University in 2017. He is currently working at DENSO CREATE INC. Technical Resources Department. His main research interest is theoretical analysis for estimating the attenuation coefficient of X-ray computed tomography.

Yoichi Yamazaki completed the doctoral program without a PhD in information science and technology. In 2011, he became a postdoctoral researcher at Aichi Science and Technology Foundation and served as a Visiting Researcher at Aichi Prefectural University. In 2012, he received his PhD degree from Aichi Prefectural University. In 2016, he became a fixed-term research assistant in the School of Science Engineering at Kwansei Gakuin University. His research interests are medical imaging, computational system biology, and kansei engineering. He is a full member of the IEEE, Society of Instrument and Control Engineers, and IEICE.

Hiroki Yoshioka is a Professor in the Department of Information Science and Technology at Aichi Prefectural University, Aichi, Japan. He received his MS and PhD degrees in nuclear engineering from the University of Arizona, Tucson, in 1993 and 1999, respectively. His research interests include cross-calibration of satellite data products for continuity and compatibility of similar datasets and the development of particle radiative transfer models and their inversion techniques.

Shuji Koyama became an Assistant Professor in the Department of Radiological Technology at Nagoya University College of Medical Technology in 1992. In 2001, he received the Doctor of Medicine degree from the Graduate School of Nagoya University School of Medicine. In 2005, he became a Lecturer in the Department of Radiological Technology at Nagoya University School of Health Sciences. In 2012, he became a Lecturer in the Department of Radiological and Medical Laboratory Sciences at Nagoya University Graduate School of Medicine. In 2013, he became an Associate Professor at the Brain and Mind Research Center of Nagoya University. His research interests are radiation measurement engineering and Monte Carlo simulation. He is a full member of the Japanese Society of Radiological Technology and American Association of Physicists in Medicine.