2005

Density resolution of proton computed tomography

R. W. Schulte  
*Loma Linda University Medical Centre, USA*

V. Bashkirov  
*Loma Linda University Medical Centre, USA*

M. C. Loss Klock  
*Loma Linda University Medical Centre, USA*

T. Li  
*State University of New York, USA*

A. J. Wroe  
*University of Wollongong*

*See next page for additional authors*

http://ro.uow.edu.au/engpapers/146

**Publication Details**

This article was originally published as Schulte, RW, Bashkirov, V, Loss Klock, MC et al, Density resolution of proton computed tomography, Medical Physics, 32(4), April 2005, 1035. Copyright American Association of Physicists in Medicine. Original journal available [here](http://ro.uow.edu.au/engpapers/146).

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au
Density resolution of proton computed tomography

Reinhard W. Schulte and Vladimir Bashkirov
Department of Radiation Medicine, Loma Linda University Medical Center, Loma Linda, California 92354

Márgio C. Loss Klock
Federal Center of Technological Education in Paraná State, Curitiba, PR 80230-901, Brazil and
Department of Radiation Medicine, Loma Linda University Medical Center, Loma Linda, California 92354

Tianfang Li
Department of Physics and Astronomy, State University of New York Stony Brook, Stony Brook, New York 11790

Andrew J. Wroe
Centre for Medical Radiation Physics, Wollongong University, Wollongong, NSW 2522, Australia

Ivan Evseev
Polytechnic Institute, Rio de Janeiro State University, Nova Friburgo, RJ 28630-050, Brazil

David C. Williams
Santa Cruz Institute of Particle Physics, UC Santa Cruz, Santa Cruz, California 95064

Todd Satogata
Collider-Accelerator Department, Brookhaven National Laboratory, Upton, New York 11973

(Received 22 September 2004; revised 14 February 2005; accepted for publication 14 February 2005; published 22 March 2005)

Conformal proton radiation therapy requires accurate prediction of the Bragg peak position. Protons may be more suitable than conventional x rays for this task since the relative electron density distribution can be measured directly with proton computed tomography (CT). However, proton CT has its own limitations, which need to be carefully studied before this technique can be introduced into routine clinical practice. In this work, we have used analytical relationships as well as the Monte Carlo simulation tool GEANT4 to study the principal resolution limits of proton CT. The noise level observed in proton CT images of a cylindrical water phantom with embedded tissue-equivalent density inhomogeneities, which were generated based on GEANT4 simulations, compared well with predictions based on Tschalar’s theory of energy loss straggling. The relationship between phantom thickness, initial energy, and the relative electron density resolution was systematically investigated to estimate the proton dose needed to obtain a given density resolution. We show that a reasonable density resolution can be achieved with a relatively small dose, which is comparable to or even lower than that of x-ray CT. © 2005 American Association of Physicists in Medicine. [DOI: 10.1118/1.1884906]

Key words: proton computed tomography, density resolution, Monte Carlo simulation

I. INTRODUCTION

Proton beams have distinct advantages compared to other radiation treatment options because they have the potential to deliver the radiation energy precisely to the tumor while leaving the tissue around the tumor mostly undamaged. This is possible due to the characteristics of the proton depth dose curve: a relatively low entrance dose followed by a high-dose Bragg peak that can be positioned within the tumor tissue. Beyond the Bragg peak the dose fall-off is very steep, i.e., from 90% to 20% of the peak dose within a few millimeters. However, successful sparing of normal tissues requires a very accurate prediction of the position of the Bragg peak within the patient’s body.

By contrast, when an x-ray beam traverses the patient, it delivers radiation energy along its entire path. If a tumor is near a critical region like the spinal cord or the optic nerve, proton beams have a clear advantage by leaving the adjacent region unexposed when it is located behind the tumor (relative to the beam direction). X-ray beams can only avoid the critical structure when it is not in the beam path. Due to the obvious advantage of proton beams, several medical proton accelerator facilities have been established over the last 15 years in the U.S. and Japan, and new proton treatment facilities are in the planning phase.

In existing proton treatment centers, the dose calculations are currently performed based on x-ray computed tomography (CT), and the patient is positioned with the help of x-ray radiographs, hence direct visualization of the actual three-dimensional patient anatomy in the treatment room is presently not performed. Using the therapeutic proton beam for CT imaging would make the proton radiation treatment more precise by defining the position of the Bragg peak more accurately and enabling verification of the actual patient and tumor position with respect to the proton beam.

Several early publications demonstrated that proton imaging is feasible. In the late 1960s, Koehler showed that with
160 MeV protons one could produce radiographic films with much greater contrast than x-ray radiographs taken under the same conditions. Since that time, a number of publications about proton and heavy particle radiography and tomography have appeared in the literature, mainly discussing proton imaging as a diagnostic tool. However, because most of the technological efforts successfully went into improving diagnostic x-ray CT, the interest in medical proton CT stagnated.

The situation changed with the development of medical proton gantries and an increasing number of patients treated with proton therapy. The need for an accurate prediction of proton gantries and an increasing number of patients treated with proton therapy demands the development of accurate imaging techniques. In this report, we focus on the inherent precision of quantitative proton CT and its limitations due to energy loss straggling. We will derive an analytical expression that relates the resolution of an ideal proton CT scanner to dose and compare it to the equivalent expression of an ideal x-ray CT scanner. The predicted performance of a proton CT scanner is tested with reconstructions of a water cylinder with different tissue inhomogeneities using GEANT4 simulated measurements of energy loss.

II. THE PHYSICAL BASIS OF PROTON CT

A. Interactions of protons with matter

When traversing matter, protons lose most of their energy via inelastic collisions with the outer atomic electrons, leading to ionizations and excitations. Furthermore, they will be deflected by multiple small-angle scattering from the nuclei of the target material [multiple Coulomb scattering (MCS)]. These two main processes, occurring a great number of times along the macroscopic path length of the proton, lead to the macroscopic effects of the interaction of protons with matter: loss of energy and a deflection from their original direction. As individual interaction processes occur randomly, this results in a statistical distribution of the principal quantities observed in proton imaging: (1) the amount of energy lost by each proton after traversing a layer of given thickness, and (2) the lateral and angular displacement of the proton from its incident direction. It will be seen that for the density resolution of proton CT, the amount of energy-loss variation (straggling) is the limiting physical process, while MCS is the principal limiting factor of the spatial resolution. Since this work emphasizes the density resolution of proton CT, it will focus on energy loss straggling; MCS and spatial resolution will not be considered in detail.

Protons in the energy range used for proton CT also undergo nonelastic nuclear interactions, leading to reduction of proton transmission in a depth-dependent manner. Protons undergoing nuclear interactions mostly deposit their energy locally and hence contribute to the dose within the patient without contributing to the image formation.

B. Characteristics of tissues and energy requirements

The protons used for proton CT must have sufficient energy to traverse the body part to be imaged. According to the NIST PSTAR database, the range in the continuously slowing-down approximation of 200 MeV protons in A150 tissue-equivalent plastic is 25.8 cm, which is sufficient to penetrate an adult human skull (nominal width of 20 cm in anterior-posterior direction). For 250 MeV protons the range is 37.7 cm, sufficient to penetrate an adult trunk (nominal width of 34 cm, excluding arms).

Human tissues are composed of atoms of relatively low atomic number (Z) and weight (A). In particular, the ratio Z/A varies little between different tissues and usually lies between 0.50 and 0.55. Density differences between various human soft tissues are of the order of a few percent, and their density value scatters around that of water. Compact bone is about 80% denser than muscle, and fat tissue is about 12% less dense than muscle.

C. Mean energy loss

For the energy range important for proton CT (10–250 MeV), the mean energy loss of protons per unit path length, also called stopping power, is well described by the Bethe–Bloch theory. For protons in the stated energy range, density effect and shell corrections are relatively small, and were therefore neglected in this work. In this case, the Bethe–Bloch formula may be written in the following form, convenient for proton CT reconstruction:

\[
- \frac{dE}{dx}(r) = \eta_e(r)S(I(r), E(r))
\]

where \(\eta_e\) is the relative electron density with respect to water, \(I(r)\) is the mean excitation potential of the material, which for water is 75 eV, \(E(r)\) the proton energy, and \(S\) is the proton stopping power in water, which can be expressed as

\[
S(I(r), E(r)) = K \frac{1}{\beta^2(E)} \left[ \ln \left( \frac{2m_e c^2}{I(r)} \frac{\beta^2(E)}{1 - \beta^2(E)} \right) - \beta^2(E) \right].
\]

The constant \(K = 170\) MeV/cm combines various physical parameters, \(m_e\) is the electron mass and \(\beta\) is the proton velocity relative to the speed of light. The dependence on the position vector \(r\) is a reminder that all quantities in Eq. (2) can vary with spatial position in the object.

Equation (2) was used in this work to determine the mean energy of protons of incident energy \(E_0\) after traversing a homogeneous water layer of thickness \(d\) (\(\eta_e = 1\)). For this, Eq. (2) was written in its integral form as
\[ d = \int_{E_{\text{out}}(d)}^{E_{\text{in}}} \frac{1}{S(I,E)} dE \]  

(3)

and was solved numerically for the outgoing energy \( E_{\text{out}}(d) \).

D. Energy loss straggling

When traversing an object of certain thickness and density, monoenergetic protons experience varying numbers of collisions. Further, the energy transferred by a proton to the atoms (mainly to their outer electrons) of the object is also subject to statistical fluctuations. In consequence, a monoenergetic beam incident on an object will have an energy distribution after traversing the object, which was first described mathematically by Bohr’s theory \(^{13}\) and later by others.

For energy losses not exceeding 20% of the initial energy, but large enough that the central limit theorem applies, the energy-loss distribution is well described by a normal (Gaussian) distribution, for which only the first two moments of the distribution are different from zero. For relativistic protons, the variance of the energy loss distribution after passing through a layer of thickness \( d \) can then be described by Bohr’s theory \(^{13}\):

\[ \sigma_E^2(d) = \eta_s K \int_0^d \frac{1}{1 - \beta^2(E(x))} dx, \]  

(4)

where \( E(x) \) is the mean energy of protons of incident energy \( E_{\text{in}} \) after traversing a path length \( x \) through the material, and \( \eta_s \) and \( K \) are defined as above.

For energy losses larger than 20% of the initial energy, Bohr’s theory becomes inaccurate and moments of higher order contribute to a skewed (non-Gaussian) energy-loss distribution that develops a tail toward higher energy losses. For this case, which is more relevant for proton CT, one has to use a different theory, for example, the theory described by Tschalar. \(^{14,15}\) The variance of the energy loss distribution in Tschalar’s theory can be expressed by the differential equation:

\[ \frac{d}{dx} \sigma_E^2(x) = \kappa_t (x) - 2 \left( \frac{d}{dE} \kappa_l (E(x)) \right) \sigma_E^2(x) + \text{higher order terms}, \]  

(5)

where

\[ \kappa_t (x) = \eta_s K \frac{1}{1 - \beta^2(E(x))} \]  

(6)

and \( \kappa_l (E(x)) \) is identical to the stopping power given by the Bethe–Bloch formula (1), and higher-order terms contain higher-order derivations of \( \kappa_l (E(x)) \) with respect to energy.

In this work, differential equation (5) was solved numerically for four different incident proton energies (160, 175, 200, 250 MeV) using the software package MATHCAD \(^{11}\) (Mathsoft Inc., Cambridge, MA) after excluding higher-order terms. The results were used to provide a more accurate analytical estimate of the energy loss spread after traversing a homogeneous water layer than that given by Bohr’s theory [Eq. (4)].

E. Multiple Coulomb scattering

When passing through the image object, protons in the energy range used for proton CT experience multiple small-angle deflections due to scattering at the nuclear potential of the target atoms leading to a random macroscopic deviation from the original direction by up to a few degrees and a random displacement of the exit point with respect to the entry point by up to a few millimeters. While MCS is the main limitation of the spatial resolution of proton imaging as discussed elsewhere \(^{16,17}\) its only contribution to the energy loss spread is a small, random increase of the proton path length from scattering in a layer of material of given thickness. Therefore, MCS will not be considered further in this article.

F. Nonelastic nuclear interactions

When imaging with protons, one has to account for nonelastic nuclear interactions leading to a loss of the primary proton. The probability of nonelastic nuclear interactions of protons in the energy range used for proton CT becomes important above 100 MeV \(^{18,19}\) Nuclear interactions result in a reduction of proton fluence with increasing thickness of the traversed object and, in the case of imaging based on energy loss, an unwanted contribution to patient dose. In the continuous-slowing down approximation, the attenuated fluence of protons at a depth \( x \) can be expressed as

\[ \Phi(x) = \Phi_0 \exp(-\kappa x), \]  

(7)

where \( \kappa \) is the total nonelastic macroscopic cross section for protons in water. Cross sections for water can be inferred from experimental and theoretical data for incident proton on oxygen reactions \(^{20}\). The microscopic total nonelastic cross section for protons on oxygen depends only weakly on energy for proton energies 100 MeV < \( E < 300 \) MeV, and has an approximate value of 300 mbarn, which corresponds to a macroscopic cross section of \( \kappa = 0.01 \) cm\(^{-1}\). For example, the fluence of a 200 MeV proton beam reduces to 90.5% at a depth of 10 cm and to 81.8% at a depth of 20 cm.

In nonelastic nuclear collisions, primary protons transfer an energy-dependent fraction \( \gamma(E) \) of their kinetic energy to secondary charged particles, mostly to protons and less commonly to heavier charged particles, assuming that the incident primary proton is lost in the nuclear interaction. For incident proton energies between 100 and 250 MeV, this fraction ranges from 0.62 to 0.67 for nonelastic interactions of protons with oxygen nuclei \(^{21}\). The remaining kinetic energy is transferred to neutrons and photons, which are not expected to contribute notably to the object dose. The energy transferred to particles heavier than protons is deposited locally, whereas secondary protons are further transported. Since most of the secondary protons have energies below 10 MeV \(^{20}\) secondary charged particle equilibrium can be assumed, which implies that the energy transferred to secondary charged particles equals the energy deposited locally.

The energy spectra of the secondary protons produced in nonelastic nuclear collisions peak at low energies (3–5 MeV) but have a long tail reaching up to the energy of
the primary proton involved in the collision. The secondary protons of higher energies may, therefore, leave the object and contribute to the reconstructed image. The vast majority of these protons can be easily identified because their energy will generally be below the 3σ boundary of primary protons traversing the object.

III. DESIGN PRINCIPLES AND NOISE OF PROTON CT

A. Design of a proton CT scanner

Our approach to the conceptual design of a proton CT system registering individual protons is illustrated in Fig. 1. The object is traversed by a broad (ideally but not necessarily parallel) beam of protons of known energy $E_{in}$. Using an active proton beam scanning system, one may also adjust the energy while scanning the object to optimize density resolution according to spatial variations of the object thickness. A detector system is arranged on both sides of the patient and records the exit energy $E_{out}$ of individual protons, as well as their entrance and exit points and angles with respect to the detector system. Technical considerations for these components of the detectors have been discussed elsewhere.22

Using a Gaussian approximation of MCS and a $\chi^2$ formalism, it was recently shown that it is possible to construct a closed-form expression for the most likely path of a proton in a uniform material incorporating the effect of continuous energy loss when the entrance and exit positions and angles are known with sufficient accuracy.17 This technique also provides estimates of the probability of the particle to deviate from the most likely path. Further, it was shown that the most likely path predicts the true path to better than 1 mm despite the broadening of a pencil beam to a size of several millimeters under typical proton CT conditions. Hence, in our calculations, it was assumed that the position of individual protons is known, provided the voxel size of the proton CT data set is not smaller than 1 mm.$^3$

B. Image reconstruction

A relationship between the mean energy loss of a proton and the integral of the (usually spatially dependent) relative electron density can be obtained by integrating the reciprocal stopping power, given by the Bethe–Bloch equation (1), between incident and outgoing energy along the proton’s path $L$:

$$\int_{E_{in}}^{E_{out}} \frac{dE}{S(E_{water},E)} = \int_{L} \eta_s(r)dl. \quad (8)$$

Note that the spatially dependent mean excitation potential $\nu(r)$ was replaced by the constant mean excitation potential of water (75 eV). The error implied in this approximation will be discussed in the following (see Sec. VI). Equation (8) would be in the format of the Radon transform if proton paths were straight lines confined to a two-dimensional plane: the right side is the line integral of the $\eta_s$ along the proton path $L$, and the left side is a unique function of the proton energy difference. In a first approximation, used in this work, one may replace the nonlinear path $L$ by a straight line between the known entry and exit position of the proton with respect to the object.

In this work, we have used Eq. (8) in the straight-line approximation to convert proton energy losses simulated with Geant4 into the approximate line integral of $\eta_s$, and to reconstruct the density of a phantom with density inhomogeneities by the filtered backprojection algorithm.

C. Analytical description of proton CT noise

Successful implementation of proton CT for applications in radiation treatment planning requires that the relative electron density of the target and surrounding normal tissues be determined with a high degree of accuracy (of the order of 1%), maintaining a sufficient degree of spatial resolution (of the order of 1 mm). The random noise in the energy measurement of outgoing protons will ultimately limit the ability to measure small density differences. The principal noise limit in x-ray CT is owed to counting statistics of detected photons, whereas in proton CT there is a physical limit of the accuracy due to energy straggling of protons traversing the object. As in x-ray CT, the only way to improve measurement accuracy is to increase the fluence of protons, thereby improving the statistics of the measurement. Since the proton fluence is proportional to dose, one can establish a dose-density discrimination relationship for a given density resolution and voxel size.

The density resolution of a proton CT scanner may be defined as the $1\sigma$ spread of the reconstructed electron density value with respect to its mean value, which in proton CT is usually close to unity. It is assumed that this value is derived from $M$ energy loss values based on $N_0$ detected protons per sampling interval of length $a$, where $M$ is the number of projections distributed over $2\pi$. Detected protons with energies below the $3\sigma$ interval around the mean energy of a given bin are excluded because they are assumed to be secondary protons produced in nuclear collisions. Three main compo-
ments contribute to the spread: (1) the energy loss straggling, (2) the energy or momentum spread of the incident protons, and (3) the noise of the energy measurement detector. Because we are mainly concerned with the principal density resolution limitation of proton CT, we will assume that only energy straggling contributes to the proton CT noise, and that the other two components can be neglected.

The following relationship between the average number of protons \( N_D(0) \) detected per projection in the central bin of a cylindrical object of unit density and the standard deviation of the relative electron density reconstructed with the standard filtered backprojection algorithm was derived (see the Appendix):

\[
\sigma_{\eta_r} = \frac{\pi \sigma_{E_{\text{out}}}^2}{\sqrt{3a^2MN_D(0)S(E_{\text{out}})}},
\]

(9)

where \( S(E_{\text{out}}) \) is the stopping power in water for the outgoing energy of protons and \( \sigma_{E_{\text{out}}} \) is the standard deviation of the outgoing energy, which was determined here by solving differential equation (5) of Tscharl's theory.

Equation (9) relates the noise in the center of a reconstructed cylindrical object to the energy variance of the outgoing protons as well as to the number of detected protons per sampling interval emerging from the object center. The latter can be related to the dose \( D_c \) delivered at the object center as follows. Assuming that the exiting protons are distributed over an effective slice of height \( a \), identical to the sampling interval, the exit fluence per projection from the center of the object is \( \Phi_p(0) = N_D(0)/a^2 \) or \( N_D(0) = a^2 \Phi_p(0) \). The exit fluence is related to the incident proton fluence \( \Phi_0 \) by [see Eq. (7)]

\[
\Phi_p(0) = \Phi_0 \exp(-\kappa d),
\]

(10)

where \( \kappa \) is the macroscopic nuclear cross section for nonelastic nuclear interactions, and \( d \) the object diameter. Since the nuclear cross section depends weakly on energy for proton energies between 100 and 300 MeV, it can assumed to be independent of penetration depth. The entrance fluence can be expressed in terms of the fluence at the center of the object, \( \Phi_c \), as

\[
\Phi_0 = \Phi_c \exp(\kappa d/2).
\]

(11)

Finally, the absorbed dose \( D_c \) from \( M \) projections can be expressed as a function of fluence at the center as

\[
D_c = \frac{M \Phi_c}{\rho} (S(E_c) + \kappa \gamma E_c),
\]

(12)

where \( \rho \) is the physical density of the object and \( E_c \) is the mean proton energy at the object center. The term \( \Phi_c \kappa \gamma E_c / \rho \) describes the contribution of nuclear interactions to the dose, where \( \gamma \) is the fraction of proton energy transferred to secondary charged particles in nuclear interactions. For protons in the energy range from 100 to 250 MeV, \( \gamma \) is approximately 0.65 with weak energy dependence. Combining these relationships, Eq. (9) becomes

\[
\sigma_{\eta_r} = \frac{\pi \sigma_{E_{\text{out}}}^2}{\sqrt{3a^2MN_D(0)S(E_{\text{out}})} \rho(S(E_{\text{out}}) + \kappa \gamma E_c) \exp(-\kappa d/2)}.
\]

(13)

In this work, Eq. (13) was used to analytically evaluate the dose dependence of the density resolution for a given object diameter as well as the dependence of the resolution on object diameter and incident proton energy for a given dose. In addition, we compared our result to the equivalent relationship for an ideal x-ray CT scanner, for which the standard deviation of the reconstructed linear attenuation coefficient \( \mu \) is determined by Poisson statistics of the number of detected photons \( N_D \). At the center of a uniform circular cylindrical object and under identical scanning and reconstruction conditions as described earlier, this is given by:

\[
\sigma_{\mu_{\text{ph}}} = \frac{\pi}{\sqrt{3a^2MN_D(0)\mu(E_{\text{ph}})^2}},
\]

(14)

where \( \mu_{\text{ph}}(E_{\text{ph}}) \) is the linear attenuation coefficient of water for the photon energy \( E_{\text{ph}} \) of the x-ray scanner. Note that there are two differences between the relationship given by Gore and Tofts [their Eq. (3)] and our Eq. (14): first, they assumed a projection range of \( \pi \) rather than \( 2\pi \), which explains their factor 12 under the square root instead of 3, and second, their sigma is that of the absolute linear attenuation coefficient, whereas ours is for the relative attenuation coefficient with respect to water, which explains the additional term \( \mu_{\text{water}}(E_{\text{ph}})^2 \) in the denominator of Eq. (14).

IV. MONTE CARLO SIMULATIONS

A. Monte Carlo simulation program

The GEANT4 simulation toolkit, a computer library for the simulation of particles interacting with matter, was developed initially for high-energy physics applications. Included among the many elastic and nonelastic hadron physical processes implemented in GEANT4 are non-Gaussian models of energy straggling and MCS. In this work, we used GEANT4 to simulate the performance of an ideal proton CT scanner as shown in Fig. 1 and to compare the analytical derivations of density resolution with the results of a simulated proton CT reconstruction of a cylindrical object with tissue inhomogeneities.

B. Simulation details

Two simulation geometries were used with GEANT4. First, to determine the fluctuation of outgoing energy due to energy-loss straggling numerically, a simple geometry was devised using right-angled parallelepipeds volumes. Upon entry into the simulation volume, protons of 250 MeV encountered a 100 cm air gap followed by a slab of water of a given thickness ranging from 1 to 30 cm. The phantom area perpendicular to the beam was chosen to be 40 cm x 40 cm. The sensitive volume in the simulation was a small slice 20 cm x 20 cm and 10 \( \mu \)m thick, located at the exit of the phantom. This sensitive volume provided the kinetic energy of the...
histories were equally distributed over $M$ by one in a sensitive volume placed at $u$ proceeding along the $u$ random vertical positions.

A straight-line approximation was used to estimate the proton path.\textsuperscript{25} First, the entry and exit points on the phantom boundary of the incident and outgoing proton tracks were mapped on the incident beam plane ($u=0$) and then connected by a straight line approximating the proton’s path through the phantom.

Equation (8) was used to convert proton energy loss into the integrated relative electron density. These values were binned into regular square intervals of the sinogram space of length $a=1$ mm and angular width $\Delta \theta=2^\circ$. The average bin values of the integrated electron density were mapped onto a regular sinogram grid $na$, $m\Delta \theta$, where $n$ and $m$ are integers and $na \pm a/2$, $m\Delta \theta \pm \Delta \theta/2$ are the boundaries of the corresponding bin interval. The phantom density reconstruction was then performed by the conventional filtered backprojection method using a Ram-Lak filter\textsuperscript{26} with $1/2$ mm cut off frequency.

C. Noise analysis

The noise in the reconstructed proton CT images was assessed quantitatively by calculating the standard deviation of reconstructed pixel values from a homogeneous circular region of interest (ROI) of 5 cm diameter (one quarter of the phantom diameter) centered on the phantom image. The influence of nonelastic nuclear interactions was assessed by comparing the noise figures for images reconstructed after including and excluding protons with energies falling outside the $3\sigma$ energy interval around the mean energy for each bin. The image noise was compared to the predicted standard deviation of reconstructed pixel values obtained with Eq. (13).

V. RESULTS

A. Energy straggling

Figure 3 shows the standard deviation of the energy loss distributions of 250 MeV protons incident on water as a function of water layer thickness obtained with the GEANT4 simulation of protons incident on water slabs of varying thickness. The figure also shows the results obtained with Bohr’s theory [Eq. (4)] and Tschalar’s theory [solution of Eq. (5)]. It is obvious that the GEANT4 simulation reproduces both Bohr’s results for thinner layers as well as the deviation from the Gaussian approximation for layers thicker than 15 cm (or >25% of the initial energy lost). Since Tschalar’s
was evaluated with Eq. 10 mGy. The relative electron density standard deviation these calculations, the central dose was chosen to be $s = 1 \text{ mm}^3$; the standard deviation of the energy loss distribution in the denominator of this formula was obtained by solving differential equation (5).

For all energies, the standard deviation at first increases with increasing phantom diameter, then goes through a maximum and for larger diameters decreases rather steeply before the protons reach their maximum range. The initial increase in standard deviation is explained by the increase in straggling with increasing proton energy (Fig. 3) whereas the stopping power term in the denominator of Eq. (13) is rather constant because the outgoing protons are in the plateau range of the Bragg peak curve. Closer to the end of the proton range, the proton stopping power increases steeply leading to the observed decline in standard deviation. At the dose level of 10 mGy to the center of the phantom, the resolution ranges from 0.4% to 1.5% with the lower proton energies having a clear resolution advantage. For example, for a phantom diameter of 16 cm the central standard deviation of a proton CT image with 160, 200, and 250 MeV protons is $0.6\%, 1.2\%$, and $1.5\%$, respectively.

Figure 4 shows the analytically predicted dose dependence of the central relative electron density resolution for a phantom diameter of 20 cm diameter imaged with 250, 200, and 175 MeV protons, respectively [Eq. (13)]. For comparison, the results for an ideal x-ray CT scanner with 75 keV photons [Eq. (14)] as well as the resolution values obtained by analyzing the simulated proton CT images (see the following) were also included.

Due to the inverse square-root dependence of the density resolution on dose for both cases, the dose-resolution relationship is described by straight lines with a slope of 0.5 in a double-log plot. The dose-resolution advantage for lower energy protons is again seen in this graph. Protons of 175 MeV, which have a range slightly higher than the phantom diameter (20.6 vs 20 cm), demonstrate the most favorable dose-resolution relationship with a resolution by a factor of 1.7 better than that of 200 MeV protons and a factor of 2.1 better than that of 250 MeV protons. The graph also shows that at the same central phantom dose, the density resolution of a proton CT image with 250 MeV protons is practically identical to that of the image obtained with an ideal x-ray CT scanner under otherwise identical conditions (beam geometry, voxel size, reconstruction algorithm). Thus, for lower proton energies there is a distinct advantage of protons versus photons with respect to density resolution.

The resolution obtained from noise analysis of the central ROI in the reconstructed images based on simulated proton CT data, shown as discrete points at three dose levels in Fig. 5, confirms the resolution is inversely proportional to the square root of dose. For the images not influenced by secondary protons, the noise level is by a factor $\sim 0.7$ smaller than that predicted for 200 MeV protons. One should remember that the noise in reconstructed CT images represents the lower limit of the noise level due to the presence of spatial correlation of the image noise. If multiple independent images were evaluated, one could expect that the noise is closer to that predicted by Eq. (13) for uncorrelated noise. As expected, the noise level of the reconstructed images in-

---

**Figure 4.** Relative electron density resolution in the center of a cylindrical phantom as a function of phantom diameter in centimeters and proton energy. The sampling interval is $a = 1 \text{ mm}$. The initial proton fluence was chosen to deliver a dose of 10 mGy at the center of the phantom.

**Figure 5.** Relative electron density resolution as a function of dose at the center of a 20-cm-diam cylindrical phantom for protons of 175, 200, and 250 MeV predicted by Eq. (13). For comparison, the dose-resolution relationship for an ideal x-ray CT scanner with 75 keV photons is also shown. The sampling interval is $a = 1 \text{ mm}$ for both cases. The discrete points represent the noise levels of the reconstructed 200 MeV proton CT images shown in Fig. 6 at three different dose levels. Circular symbols correspond to images reconstructed excluding secondary protons, and triangular symbols to reconstructed images including secondary protons.
including secondary protons is considerably larger (by a factor \( \sim 2 \)) than that of the images produced by primary protons (also see Fig. 6).

C. Proton CT simulation

Figure 6 depicts the reconstructed images of a cylindrical phantom with nine cylindrical inhomogeneities, corresponding to a combination of three different tissues types and three different sizes (see Fig. 2 for comparison). The reconstruction was based on GEANT4-simulated histories of up to 50,000 protons per projection. To show the influence of dose on resolution, reconstructions were performed at three different central dose levels of 3.14 mGy (208 protons per mm\(^2\)), 1.57 mGy (104 protons per mm\(^2\)), and 0.16 mGy (10 protons per mm\(^2\)), respectively. To demonstrate the influence of secondary protons produced in nonelastic nuclear interactions, reconstructions were done either including all protons registered behind the phantom (left side of Fig. 6) or after excluding registered protons with energies falling outside the 3\( \sigma \) energy interval around the mean energy for each bin (right side of Fig. 6).

The left side of Fig. 6 demonstrates that secondary protons can deteriorate the density resolution considerably, which is best seen at the lowest dose level. Since the energy loss in nuclear interactions is typically large, the reconstruction algorithm produces a high-density line crossing the image for each such event. The likelihood of a secondary proton to be registered is larger if it is produced in the phantom periphery. Therefore, the image quality suffers most from these protons in the outer part of the reconstructed phantom image. One should note that the 3\( \sigma \) exclusion of protons effectively eliminates these artifacts.

As expected, the density resolution improves with increasing dose. The diagram in Fig. 5 predicts density resolutions of 2.0\%, 2.8\%, and 8.7\% for the three dose levels used. Consistent with these numbers, the compact bone cylinders with a density of 1.85 g cm\(^{-1}\) (85\% density difference with respect to water) are well distinguished from the background water density even at the lowest dose level. The muscle-tissue-equivalent cylinders with a density of 1.04 g cm\(^{-1}\) (4\% density difference with respect to water), on the other hand, are at the limit of visibility even at the higher dose levels. The adipose-tissue-equivalent cylinder with a density of 0.92 g cm\(^{-1}\) (8\% density difference with respect to water) are quite well visible for the two higher doses, and appears at the limit of visibility for the lowest dose.

VI. DISCUSSION AND CONCLUSIONS

Proton radiation therapy is one of the most precise forms of noninvasive image-guided cancer therapy. At present, the potentials of proton therapy cannot be fully exploited because the conversion of Hounsfield values, measured with x-ray CT, to relative electron density values is not always accurate. The resulting range uncertainty is usually quoted to be between 3 and 10 mm or 3\% of the proton range in tissue, depending on the anatomical region treated and the penetration depth of the proton beam. In this study, we have explored the principal limitations of proton CT density resolution due to energy loss straggling in the body, bearing in mind that by using proton CT one may be able to reduce the range error in proton treatment planning and delivery to less than 3 mm without exceeding practical dose limits. This requires that the relative electron density resolution should be of the order of 1\%. Our results indicate that with doses \( \leq 10 \) mGy, which is clinically acceptable, the desired relative density resolution of 1\% in the center of a cylindrical phantom of 20 cm can, in principle, be achieved. The resolution at a given dose level can be optimized by adjusting the proton energy according to phantom diameter. With the development of proton beam scanning systems with variable energy, this may be technically feasible. Our results demonstrate that at the same level of dose and when using an optimized proton energy, an ideal proton CT scanner exceeds the resolution of an ideal x-ray scanner by a factor close to 2. With the availability of rotating proton gantries now in sev-
eral hospitals, these systems may be utilized also for diagnostic CT studies, in particular, when lower doses or high density resolution are desirable.

The use of cone beam CT scanners utilizing the rotating gantry of a linear accelerator in the treatment room has become a major innovation for alignment verification and image-guided photon radiation therapy in recent years. Patient exposure is an important consideration in this new technology. Both therapeutic megavoltage (MV) radiation and kilovoltage (kV) x-ray tubes mounted on the accelerator gantry in combination with flat-panel imaging systems have been tested. Doses required to distinguish soft-tissue structures in cylindrical phantoms are of the order of 1 cGy (10 mGy) for kV systems but are of one order of magnitude higher for MV systems (due to the low detection efficiency of photon detectors in the MeV energy range). According to our results, the doses required for soft-tissue density resolution with proton CT may be less than those required for kV CT. Proton CT would have the additional advantage of using the same radiation modality for treatment and imaging.

Some limitations of our study should be kept in mind. We have considered energy loss straggling as the only uncertainty component contributing to the density resolution, thus neglecting external noise sources such as the momentum spread of the proton accelerator and the uncertainty of the energy detector. Measurements at our own medical proton accelerator indicate that the energy spread is of the order of 40 keV for 250 MeV protons and of 0.016% for kV systems but are of one order of magnitude higher for MV systems (due to the low detection efficiency of photon detectors in the MeV energy range). According to our results, the doses required for soft-tissue density resolution with proton CT may be less than those required for kV CT. Proton CT would have the additional advantage of using the same radiation modality for treatment and imaging.

Further, in our reconstruction of proton CT images we have neglected the influence of MCS by assuming that protons proceed along straight lines when reconstructing the integrated density along the proton path. This may lead to a systematic underestimation of the true proton path length and, consequently, to a systematic overestimation of relative electron density, which will increase with phantom diameter. The use of reconstruction algorithms that use more realistic proton trajectories, such as algebraic reconstruction methods, may alleviate this problem.

This article has focused on the density resolution aspect of proton CT. Equally if not more important for proton treatment planning are systematic errors in the estimation of relative electron density. Systematic errors may be introduced by using approximations in the theory employed for converting energy loss into integrated density values. For example, in Eq. (8), which is central to the reconstruction of proton CT images, we have replaced the mean excitation potential $I(r)$ of the material traversed with that of water as a tissue substitute. This simplification may lead to systematic errors in the estimation of the line integral of the relative electronic density when the object contains materials with mean excitation potentials very different from that of water. Mean excitation potentials of various tissues have been published by the ICRU and are available from the NIST material database. For example, for adipose tissue, muscle tissue, and compact bone the mean excitation potentials are 63.2, 74.7, and 91.9 eV, respectively, while the value for water is 75 eV. Corresponding errors in the stopping power resulting from using the water value instead of the true tissue specific value can be calculated to be $-2\%$ for adipose tissue and $+2\%$ for compact bone, while for muscle tissue the error is practically zero. This means that if the tissue evaluated with proton CT consisted entirely of fat or bone, an additional density error of $\pm2\%$ would be introduced. However, as most body sections contain a mixture of these tissues, the actual systematic error due to the use of the mean excitation potential for water will be smaller than $2\%$. Other small systematic errors may be introduced by neglecting density corrections (at high proton energies) and shell corrections (at low proton energies) in the Bethe–Bloch equation. The amount of these systematic errors has to be investigated experimentally using standard electron density phantoms before proton CT is used clinically. Some of these effects could also possibly be corrected in the reconstruction algorithms by introducing systematic correction factors and/or through calibration of the proton CT system.

Our results highlight the importance of secondary protons produced in nonelastic nuclear interactions for the quality of the proton CT image reconstruction. Treating these protons in the same way as primary protons leads to high-density artifacts, which due to their random nature increase the images noise. On the other hand, as shown here, these protons can be easily excluded from the reconstruction by introducing a $3\sigma$ exclusion rule in the analysis.

In conclusion, we have demonstrated with analytical calculations and Monte Carlo simulations that proton CT has the potential to reduce the range uncertainty in proton therapy due to uncertainties in the relative electron densities of the tissues traversed. A density resolution of $1\%–2\%$ and a range uncertainty of the same order of magnitude seem possible and would be an improvement over current uncertainties in proton therapy. This can be achieved with proton doses of $10\;\text{mGy}$ or less, which is comparable to or better than doses stated for cone beam CT with kV x-ray beams and an order of magnitude less than the dose required to resolve soft tissues with megavoltage photon radiation.

**ACKNOWLEDGMENTS**

The authors are grateful to S. Peggs, Ph.D., from the Brookhaven National Laboratory and H. Sadrozinski, Ph.D., from the Santa Cruz Institute of Particle physics for providing the conceptual basis for this work. We would like to thank W. Preston for carefully reviewing the manuscript and making stylistic improvements. R.W.S. was supported by the National Medical Technology Testbed Inc. (NMTB) under the U.S. Department of the Army Medical Research Acquisition Activity, Cooperative Agreement No. DAMD17-97-2-7016. T.F.L. was supported by the National Institute of Health (Grant No. HL51466). M.C.L.K. received financial support from the Brazilian agency CAPES.
APPENDIX: DERIVATION OF EQUATION (9)

The following derivation of the relationship between density resolution of proton CT and the initial proton fluence follows mostly that outlined by Gore and Tofts. The x-ray CT scanner is modeled as an idealized x-ray CT scanner but is modified to take into account the particular circumstances of proton CT. Consider a homogeneous cylindrical object, which is scanned in a 2D section by an infinitely thin parallel beam of monoenergetic protons perpendicular to the principal axis of the cylinder. The direction of the beam is changed in discrete intervals \( m \Delta \theta \), where \( m \) is an integer; a total of \( M \) directions is used to scan the object. Due to multiple Coulomb scattering, the lines between entry and exit point of each proton have a distribution around \( m \Delta \theta \), which is near-Gaussian. The protons traversing the object are binned into regular square intervals of the sinogram space of length \( a \) and angular width \( \Delta \theta = 2 \pi / M \). We assume that on average \( N_D \) protons are registered per bin. As discussed earlier, the energy of each proton emerging from the object is measured and converted to the integral of the relative electron density along the straight line connecting entry and exit point. The average bin value of the integrated proton density \( \bar{p} \) is mapped onto a regular grid \( n a, m \Delta \theta \), where \( n \) is an integer and \( na \pm a/2, m \Delta \theta \pm \Delta \theta / 2 \) are the boundaries of the corresponding bin interval:

\[
\bar{p}(na, m \Delta \theta) = \sum_{i=1}^{N_D(na, m \Delta \theta)} p_i / N_D(na, m \Delta \theta). \tag{A1}
\]

In an ideal proton CT scanner, the statistical fluctuation of \( \bar{p} \) is determined by the proton energy fluctuations due to straggling. Additional variations due to the statistical variation in the number \( N_D \) of registered protons per bin can be neglected if \( N_D \gg 1 \). The variance of each projection, \( \sigma_p^2 \), is given by

\[
\sigma_p^2(na, m \Delta \theta) = \frac{\sigma_p^2(na, m \Delta \theta)}{N_D}, \tag{A2}
\]

where \( \sigma_p^2 \) is the variance of the individual projection values. The latter can be expressed in terms of the variance of the residual proton energy \( E_{\text{out}} \) using the first-order error propagation formula

\[
\sigma_p^2(E_{\text{out}}) = \sigma_{E_{\text{out}}}^2 \left( \frac{\partial p(E)}{\partial E} \right)^2_{E=E_{\text{out}}}^2 = \sigma_{E_{\text{out}}}^2 S(E_{\text{out}})^2, \tag{A3}
\]

where \( \sigma_{E_{\text{out}}}^2 \) is the variance of the outgoing proton energy due to energy straggling, which in this work was calculated based on Tschalar’s theory by solving Eq. (5). Note that we have assumed here that the incoming proton energy is constant and that the variation of the outgoing energy is the only source of fluctuation in \( p \) and \( \bar{p} \).

We assume that the image reconstruction process is performed with the filtered backprojection algorithm. One first computes the Fourier transform of each projection into the spatial frequency \((k)\) domain, multiplies the Fourier transform by a bandwidth limited ramp filter \(|k| A(k)\), and calculates the inverse Fourier transform to obtain the filtered projection \( \tilde{p}(na, m \Delta \theta) \). In case of the Ram-Lak filter, introduced by Ramachandran and Lakshminarayanan, the apodizing window \( A(k) \) is given by \( \text{rect}(k/2k_m) \), where the rectangle function \( \text{rect}(x) \) is defined as \( \text{rect}(x) = 1 \) if \( |x| \leq \frac{1}{2} \) and 0 otherwise, and \( k_m = 1/2a \) is the band-limiting cutoff frequency. Instead of filtering in the frequency domain, one can perform a convolution with a one-dimensional convolution kernel, derived from the inverse Fourier transform of the filter function, in the spatial domain. In discrete terms, this may be written as

\[
\bar{p}(na, m \Delta \theta) = a \sum_{n'=a}^{\infty} p(n' a, m \Delta \theta) h((n - n') a), \tag{A4}
\]

where \( h \) is the convolution kernel of the Ram-Lak filter. To determine the corresponding convolution kernel, we express the filter function \(|k| A(k)\) as

\[
|k| \text{rect}\left(\frac{k}{2k_m}\right) = k_m \text{rect}\left(\frac{k}{k_m}\right) - k_m A\left(\frac{k}{k_m}\right), \tag{A5}
\]

where the triangle \( A(x) \) function is defined as \( A(x) = 1 - |x| \text{rect}(x/2) \). Since the inverse Fourier transform of the rectangle function is \( \text{sinc}(\pi x) \), and the inverse Fourier transform of the triangle function is \( \text{sinc}^2(\pi x) \), the convolution kernel of the Ram-Lak filter is

\[
h(t) = k_m^2 \left( 2 \text{sinc}(2\pi k_m t) - \text{sinc}^2(\pi k_m t) \right) = \frac{1}{4a^2} \left[ 2 \text{sinc} \left( \frac{\pi t}{a} \right) - \text{sinc}^2 \left( \frac{\pi t}{2a} \right) \right]. \tag{A6}
\]

For discrete arguments, \( na \), this becomes

\[
h(na) = \begin{cases} 1/4a^2 & \text{if } n = 0 \\ -1/\pi^2n^2a^2 & \text{if } n \text{ is odd} \\ 0 & \text{if } n \text{ is even}. \end{cases} \tag{A7}
\]

Since the convolution is a linear operation, the variance in the filtered projections is given by

\[
\sigma_p^2(na, m \Delta \theta) = \sum_{n'=a}^{\infty} \sigma_p^2(n' a, m \Delta \theta) \cdot |h((n - n') a)|^2. \tag{A8}
\]

The final step in the reconstruction process is the backprojection of the \( M \) filtered projections, and thereby reconstructing the object relative electron density \( \tilde{\eta} \), i.e.,

\[
\tilde{\eta}(ia, ja) = \sum_{m=1}^{M} \tilde{p}(ia \cos(m \Delta \theta) + ja \sin(m \Delta \theta), m \Delta \theta) \cdot \Delta \theta, \tag{A9}
\]

where the integers \( i \) and \( j \) index the \( x \) and \( y \) direction in the reconstructed image, respectively. Note that the discrete values at which the filtered projections are needed for Eq. (A9) do generally not match integer multiples of the interval \( a \) and, therefore, the filtered projections have to be interpolated. The additional variation in \( \tilde{\eta} \) this may cause is not considered here.
The variance in the reconstructed image, $\sigma_{\eta}(ia, ja)$, is the result of the superposition of the $M$ independent projections with individual variances given by Eq. (A8), i.e.,

$$
\sigma_{\eta}(ia, ja) = \sum_{m=1}^{M} \sigma_{p}(ia \cos(m \Delta \theta) + ja \sin(m \Delta \theta), m \Delta \theta) \cdot \Delta \theta^2.
$$

(A10)

We now evaluate the variance at the center of the object, which is given by

$$
\sigma_{\eta}(0,0) = \sum_{m=1}^{M} \sigma_{p}(0, m \Delta \theta) \Delta \theta^2 = \sigma_{p}^2(0) \frac{4 \pi^2}{M},
$$

(A11)

where we have made use of the radial symmetry of the object and assumed that the projections cover a range of $2 \pi$, thus $\Delta \theta = 2 \pi / M$. Using Eq. (A8) to express the variance of the filtered projection, and taking into account that the variance of the projections changes very little around the center of the homogenous object, we get

$$
\sigma_{\eta}(0,0) = \frac{4a^2 \pi^2}{M} \sum_{n=-\infty}^{\infty} \alpha_{u}(na)|h(na)|^2 = \frac{4a^2 \pi^2}{M} \sigma_{p}^2(0) \sum_{n=-\infty}^{\infty} |h(na)|^2.
$$

(A12)

Finally, we can express the sum of convolution kernel coefficients as

$$
\sum_{n=-\infty}^{\infty} |h(na)|^2 = \frac{1}{16a^4} + \frac{1}{\pi^2 n^4 a^4} - \frac{1}{12a^4}
$$

(A13)

which, using Eqs. (A2) and (A3), leads to

$$
\sigma_{\eta}^2(0,0) = \frac{\pi^2 \sigma_{p}^2(0)}{3a^2 M} = \frac{\pi^2 \sigma_{\text{out}}^2}{a^2 M N_D(0) S(E_{\text{out}})^2},
$$

(A14)

where $E_{\text{out}}$ is the mean outgoing energy of the $N_D(0)$ the protons traversing the center of the object.

---

Electronic mail: rschulte@dominion.llumc.edu