

An improved physics-based approach for unfolding megavoltage bremsstrahlung spectra using transmission analysis

E. S. M. Ali^{a)} and D. W. O. Rogers^{b)}

Carleton Laboratory for Radiotherapy Physics, Department of Physics, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6, Canada

(Received 21 September 2011; revised 2 February 2012; accepted for publication 2 February 2012; published 1 March 2012)

Purpose: To develop a physics-based approach to improve the accuracy and robustness of the ill-conditioned problem of unfolding megavoltage bremsstrahlung spectra from transmission data.

Methods: Spectra are specified using a rigorously-benchmarked functional form. Since ion chambers are the typical detector used in transmission measurements, the energy response of a Farmer chamber is calculated using the EGSnrc Monte Carlo code, and the effect of approximating the energy response on the accuracy of the unfolded spectra is studied. A proposal is introduced to enhance spectral sensitivity by combining transmission data measured with multiple detectors of different energy response and by combining data from multiple attenuating materials. Monte Carlo methods are developed to correct for nonideal exponential attenuation (e.g., scatter effects and secondary attenuation). The performance of the proposed methods is evaluated for a diverse set of validated clinical spectra (3.5–25 MV) using analytical transmission data with simulated experimental noise.

Results: The approximations commonly used in previous studies for the ion-chamber energy response lead to significant errors in the unfolded spectra. Of the configurations studied, the one with best spectral sensitivity is to measure four full transmission curves using separate low-Z and high-Z attenuators in conjunction with two detectors of different energy response (the authors propose a Farmer-type ion chamber, once with a low-Z, and once with a high-Z buildup cap material), then to feed the data simultaneously to the unfolding algorithm. Deviations from ideal exponential attenuation are as much as 1.5% for the smallest transmission signals, and the proposed methods properly correct for those deviations. The transmission data with enhanced spectral sensitivity, combined with the accurate and flexible spectral functional form, lead to robust unfolding without requiring *a priori* knowledge of the spectrum. Compared with the commonly-used methods, the accuracy is improved for the unfolded spectra and for the unfolded mean incident electron kinetic energy by at least factors of three and four, respectively. With simulated experimental noise and a lowest transmission of 1%, the unfolded energy fluence spectra agree with the original spectra with a normalized root-mean-square deviation, $\% \Delta(\psi)$, of 2.3%. The unfolded mean incident electron kinetic energies agree, on average, with the original values within 1.4%. A lowest transmission of only 10% still allows unfolding with $\% \Delta(\psi)$ of 3.3%.

Conclusions: In the presence of realistic experimental noise, the proposed approach significantly improves the accuracy and robustness of the spectral unfolding problem for all therapy and MV imaging beams of clinical interest. © 2012 American Association of Physicists in Medicine. [http://dx.doi.org/10.1118/1.3687164]

Key words: transmission analysis, spectral unfolding, bremsstrahlung reconstruction, inverse problems, photon beams, functional form, ion-chamber energy response, EGSnrc

I. INTRODUCTION

In clinical photon beams, knowledge of the photon spectrum is needed for dose calculations in patients and for the calculation of spectrum-averaged dosimetric quantities and detector correction factors. Typically, the spectrum is estimated by best matching measured depth-dose and profile data. While this “self-tuning” approach is suitable for many applications, it has been shown that accurate knowledge of the spectrum is needed for accurate dose calculations around tissue interfaces and heterogeneities,^{1,2} and for accurate modeling of the energy response of detectors.³ A reliable method

to determine the true spectra would be useful for more robust beam commissioning, for stricter testing of the dose calculation engines of treatment planning systems (Monte Carlo or otherwise), and for better modeling of the energy response of different detectors in a given beam.

Transmission analysis is a clinically-viable indirect method to determine linac photon spectra. In this method, transmission signals are acquired after the beam passes through different attenuator thicknesses. In an ideal attenuation geometry, the normalized transmission signals can be expressed in terms of the unknown spectrum through a homogeneous Fredholm equation of the first kind given by

$$T_{\text{ideal}}(d, x_i) = \frac{M(d, x_i)}{M(d, 0)} = \frac{\int_{E_l}^{E_m} R(d, E) \psi(E) \exp(-\mu(E)x_i) dE}{\int_{E_l}^{E_m} R(d, E) \psi(E) dE}, \quad (1)$$

where $M(d, 0)$ and $M(d, x_i)$ are, respectively, the measured signal of detector d (most commonly an ion chamber) without an attenuator and with an attenuator of thickness x_i , $T_{\text{ideal}}(d, x_i)$ is the corresponding normalized transmission signal in good-beam geometry, $R(d, E)$ (called henceforth the energy response of detector d) is the ion-chamber air-cavity dose per unit energy fluence from monoenergetic photons of energy E , $\psi(E)$ is the unknown photon energy fluence at energy E at the center of the ion chamber when the chamber is not present (this definition is chosen for the mere convenience of making $\psi(E)$ independent of the buildup cap thickness for side-on irradiation using caps of different thicknesses), $\mu(E)$ is the linear attenuation coefficient of the attenuator material at E , and E_l and E_m are the lowest and maximum photon energies, respectively. It is well known that the weak dependence of μ on E at therapy energies makes the problem of unfolding spectra from noisy transmission data ill-conditioned, and the problem has been extensively investigated over the years, as referenced below.

Previous studies on transmission analysis in the megavoltage energy range were mainly concerned with answering the following question: given that transmission data offer little spectral differentiation, how can we devise an unfolding algorithm clever enough to still provide useful spectra? Answers to this question include the methods of Laplace transform pairs,⁴⁻⁶ direct matrix inversion,^{7,8} neural networks,⁹ and iterative unfolding (including least-squares, expectation-maximization, and simulated annealing) with: (a) *a priori* spectral knowledge and smoothing constraints,¹⁰⁻¹⁷ (b) regularization,¹⁸⁻²¹ or (c) spectrum parameterization.²²⁻³³ In this study, before approaching the unfolding issue, we first try to answer the more fundamental question: how can we make the transmission data themselves contain more (and accurate) spectral information? To this end, EGSnrc (Refs. 34 and 35) is used to accurately model the energy response of the detector used (Sec. II A); the optimum use of multiple detectors and attenuator materials is investigated (Secs. II B – II D); and, correction methods for nonideal exponential attenuation are developed (Sec. II E). The resulting improvements are independent of the unfolding details and would thus improve the relative performance of any unfolding algorithm.

Next, we try to answer the question: using the transmission data with enhanced spectral sensitivity, how can we make the unfolding truly robust and portable for all therapy and MV imaging beams? While many of the unfolding methods mentioned above demonstrated some success under custom conditions, they typically have strong caveats that limit their portability and usefulness. Laplace methods are extremely sensitive to the chosen pair and require the attenuation coefficient to be a strong function of energy. Direct inversion methods are extremely sensitive to the experimental and numerical noise. Neural network methods are limited by the scope, resolution and accuracy of the training set. Iterative methods that

require *a priori* knowledge, smoothing constraints, and/or regularization [(a) and (b) above] are sensitive to the initial estimate of the spectrum, the penalty/smoothing function, the regularization parameter, etc. Therefore, in this study we choose iterative unfolding with spectrum parameterization (Sec. II F). In addition to its compact specification of a spectrum, this method has the most potential for true robustness with standard least-squares minimization without requiring unrealistic measurement accuracy or *a priori* knowledge of the spectrum. Parameterization does not restrict the spectrum any more than the smoothing or regularization constraints do in the bin-by-bin iterative unfolding methods, provided that the functional form used in the parameterization is accurate and flexible. The functional form used in this study was constructed and rigorously benchmarked in a precursor study.³⁶ It was shown to be accurate, robust, and flexible enough to represent the complete range of relevant MV photon spectra.

The performance of the approach proposed in this study is evaluated for a diverse set of validated clinical spectra using analytical transmission data with simulated experimental noise. A companion study³⁷ presents the experimental validation of the proposed approach using a dedicated research linac whose photon spectra and incident electron beam parameters are directly and independently known.

II. METHODS

II.A. Accurate modeling of the detector energy response

Accurate modeling of the energy response of the detector used in the transmission measurements is a prerequisite for accurate unfolding of the spectra. If only one detector is used, then the quantity unfolded from Eq. (1) is the product $R(d, E)\psi(E)$. In this case, errors in $R(d, E)$ translate into type-B energy-dependent errors in $\psi(E)$ such that the product above remains unchanged. Those errors affect $\psi(E)$ *only after the minimization*, but the minimization itself is not driven in the wrong direction by the errors in $R(d, E)$. This is one of the reasons that the accuracy of $R(d, E)$ was not central in previous studies. If, however, data from more than one detector are simultaneously used in the unfolding (as proposed in this study), then the quantity unfolded is $\psi(E)$ rather than the product above; therefore, errors in $R(d, E)$ (where d runs over all the detectors used) adversely affect the minimization itself. This puts more emphasis on the importance of accurate modeling of the detector energy response.

In previous transmission analysis studies, the ion-chamber energy response was roughly approximated by one of the following three methods: (1) assuming that the material of the chamber wall and buildup cap are air-equivalent, thus the air-cavity dose, D_{cav} , equals the collision air-kerma, $K_{\text{col}}^{\text{air}}$, and consequently, $R(d, E)$ is proportional to $(\mu_{\text{en}}/\rho)^{\text{air}}$, the mass energy absorption coefficient for air;^{16,17,23,24,26,31,32} (2) measuring the response at a few known low energies (e.g., ¹³⁷Cs and ⁶⁰Co) then interpolating or extrapolating to other energies;^{4-6,11-13} or, (3) developing approximate expressions to take into account the deviation from air equivalence caused

by the wall and/or cap materials.^{8,14,27} To the authors' knowledge, no previous study on transmission analysis used Monte Carlo for detailed detector modeling. In this study, the EGSnrc usercode *cavity*³⁸ is used to simulate the detector energy response. The results are used to illustrate the effect of approximating the energy response on the accuracy of the unfolded spectra, and to propose new methods that enhance the spectral sensitivity in the measured transmission signals.

The example detector modeled is an Exradin A19 Farmer-type chamber [Fig. 1(a)]. The chamber is fitted with one of three buildup caps made of, respectively, polymethylmethacrylate (PMMA), aluminum or a tungsten-alloy with 90%

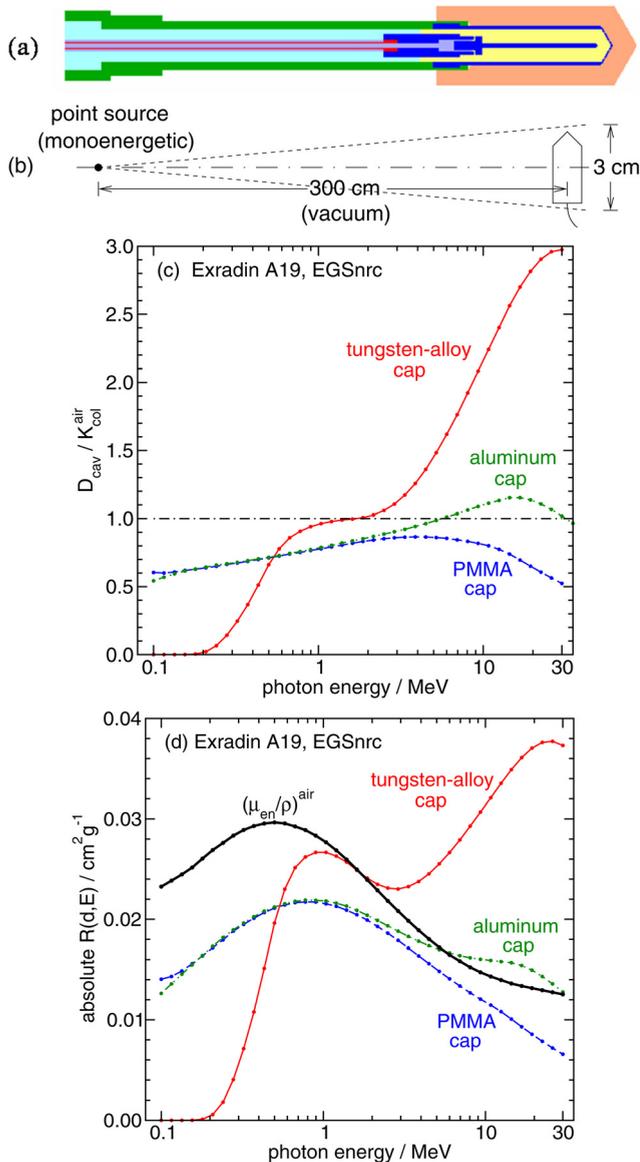


Fig. 1. (a) egs++ model (Ref. 42) of an Exradin A19 ion chamber (from blueprints) fitted with a tungsten-alloy buildup cap. (b) Irradiation geometry for the energy response calculations (not to scale). (c) Air-cavity dose, D_{cav} , per unit collision air-kerma, $K_{\text{col}}^{\text{air}}$, to examine the commonly-used assumption that $D_{\text{cav}}/K_{\text{col}}^{\text{air}}$ is flat versus energy. (d) Air-cavity dose per unit energy fluence. For comparison $(\mu_{\text{en}}/\rho)^{\text{air}}$ is also shown. The ratio of the data for $R(d, E)/(\mu_{\text{en}}/\rho)^{\text{air}}$ from panel d gives the respective data in panel c.

tungsten, 5% nickel, and 5% copper (the responses of an NE2571 chamber with the PMMA and the tungsten-alloy caps were previously validated using in-air off-axis ratio measurements³). The caps have wall thicknesses roughly equal to the CSDA range of 10 MeV electrons in their respective materials, which is sufficient in practice to provide full buildup for typical therapy beams. The irradiation geometry used for the energy response calculations is shown in Fig. 1(b). Calculations are done at 40 energies from 100 keV to 30 MeV, equispaced in $\log(E)$ to best capture the variation of the response with energy at both low and high energies. The most accurate low- and high-energy physics available in EGSnrc are used for all simulations³⁵ because the calculations are done over a large energy range. NIST cross section data are used for photons (XCOM) and for charged particles. The kinetic energy thresholds for the production and transport of charged particles and photons are 10 keV. The statistical uncertainty is kept below 0.1% so that its effect on the uncertainty of the unfolded spectra is negligible.

Figure 1(c) shows that the assumption that the ratio $D_{\text{cav}}/K_{\text{col}}^{\text{air}}$ is constant with energy, which is the basis for approximation 1 above, is not satisfied, even for low-Z caps. The variation in this ratio, relative to its mean value over the energy range shown, is $\pm 26\%$, 35% , and 120% for the PMMA, aluminum, and tungsten-alloy caps, respectively. The effect of this approximation on the accuracy of the unfolded spectra when data from one or more detectors and/or attenuator materials are used in the unfolding is discussed in the results (Sec. III A).

II.B. Combined use of detectors with different spectral sensitivity

Figure 1(d) shows the values of $R(d, E)$ that correspond to the data in panel c. The data in panel d show that the ion chamber with a high-Z tungsten-alloy cap responds significantly more to higher-energy photons than it does to lower-energy ones (mainly due to the larger pair production cross section for high-Z materials at higher energies). On the other hand, the same chamber with a low-Z PMMA cap exhibits an opposite trend. This observation suggests that, for a given attenuator material and a given total number of transmission measurements, if some of the chamber measurements are made with a high-Z cap while the rest are made with a low-Z cap, the spectral information contained in the combined data will be more than the spectral information from typical transmission measurements in previous studies which used an ion chamber with one generic cap for buildup. In other words, different $R(d, E)$ can be used as spectral weighting functions to improve energy differentiation. The large steady increase in response at high energies for the chamber with a high-Z cap is particularly appealing because it amplifies the slow variation of $\mu(E)$ in that energy range and thus improves the conditioning of the problem. Optimization of the measurement configurations with multiple detectors is discussed in Sec. II D.

The concept of using two detectors can be extended to using a combination of physically-different detectors with

different energy responses for further energy differentiation. However, using one chamber with two caps of very different Z has the advantages of experimental simplicity and consistency in the combined input data to the unfolding algorithm, without having to deal with differential detector effects.

II.C. Combined use of multiple attenuating materials

To avoid degeneracy in the solution when unfolding spectra from transmission data using one attenuating material, the μ values for the attenuator must be monotonic with energy—i.e., the minimum μ must occur at an energy larger than the maximum photon energy for the spectrum of interest. Therefore, in previous studies high- Z materials (e.g., lead with μ_{\min} at ~ 2.5 MeV) were deemed unsuitable as attenuators for most therapy beams, and, except for the work of Huang *et al.*¹⁰ (discussed below), only a single low- or medium- Z attenuator was used (e.g., aluminum, copper, or water/graphite with μ_{\min} at ~ 7 , 19, and 30 MeV, respectively). It is, however, the slow variation of $\mu(E)$ for that single attenuator that causes the problem to be ill-posed. This study proposes a different approach to the choice of the attenuator materials as follows. Figure 2 shows that past 8 MeV the mass attenuation coefficient, μ/ρ , changes with energy much more rapidly for lead than it does for graphite (by as much as a factor of 8 at 25 MeV), with the actual μ/ρ increasing for lead and decreasing for graphite. Therefore, for a fictitious spectrum with no photons below 8 MeV, a high- Z attenuator clearly provides much better energy differentiation than a low- Z one. This observation suggests that, for realistic high-MV beams (e.g., 15–25 MV) and a given total number of transmission measurements, if some of the measurements are made with a high- Z attenuator alone while the rest are made with a low- Z attenuator alone to eliminate degeneracy (proven below), the spectral information contained in the combined data will be more than the spectral information from trans-

mission measurements made with only one low- Z attenuator. Additionally, μ/ρ changes more rapidly for lead than it does for graphite below 1.5 MeV (Fig. 2). Optimization of the measurement configurations with multiple attenuators and detectors is discussed in Sec. II D.

The following are additional relevant observations. Our computational tests show that a third attenuating material does not provide additional advantage. It also does not replace the use of different- Z caps proposed in Sec. II B because it only introduces attenuation effects, whereas the responses of the ion chamber with the different caps shown before are the result of a combination of attenuation and scatter effects. For low-MV beams (e.g., 3.5–6 MV), Fig. 2 shows that making some of the measurements using a high- Z attenuator does not provide good energy differentiation past 1.5 MeV. Finally, Fig. 2 shows that copper has no advantage over graphite for low-MV beams or over lead for high-MV beams and should thus always be avoided.

In the context of unfolding a 4 MV spectrum, Huang *et al.*¹⁰ proposed to start with lead as an attenuator and then at an optimum transmission value, T_{switch} , switch to aluminum while keeping the lead in the beam. The value of T_{switch} was determined by checking after each measurement which of the two attenuators would provide more change in the average μ of the attenuated spectrum per unit transmission for the upcoming measurement. The check was achieved experimentally by making an extra measurement after each attenuator thickness using a thin ‘chopper’ material with monotonic μ (they used polystyrene), in which the average μ is determined. The approach basically uses a high- Z material when the bulk of the photon energies falls below 1.5 MeV and switches to a low/medium- Z material as the spectrum hardens. On the contrary in our study, the high- Z attenuator is used to accentuate the signal from the higher-energy photons, regardless of their relative fraction in the total energy fluence. Also in the approach of Huang *et al.*, determining T_{switch} requires extra measurements and on-the-fly calculations after each measurement until the material switch is made. Its logic also makes it applicable only to low-MV beams (e.g., 4 MV).

II.D. Optimizing measurement configuration with simulated noise

The following methods are used to objectively determine the best measurement configuration that employs the methods proposed above, and to compare its performance against typical configurations used in previous studies. Each configuration below is tested for eight very different point-source central-axis linac spectra, which are a subset of the validated, high-resolution Monte-Carlo spectra that were used to benchmark the functional form.³⁶ The spectra are divided into a low-MV group [Tomotherapy 3.5 MV (imaging), Varian 4 and 6 MV, and Siemens 6 MV] and a high-MV group [Varian 15 and 18 MV, Siemens 18 MV, and Elekta 25 MV]. The rationale for this division is that the optimum configuration may not be the same for both groups due to the shape of variation of μ and R with energy.

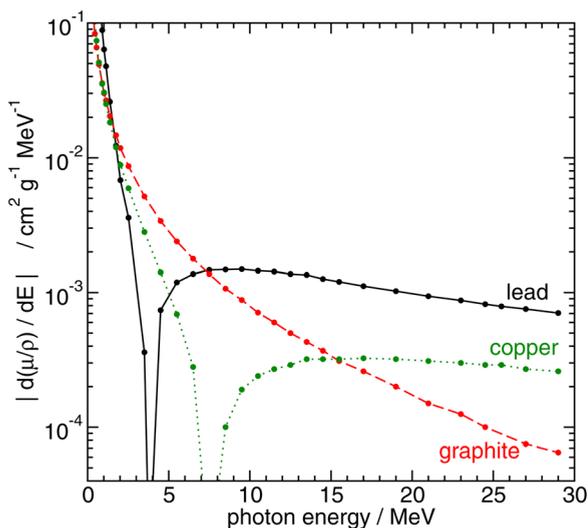


FIG. 2. Absolute value of the rate of change of the mass attenuation coefficient (μ/ρ) with energy. Lead provides better energy differentiation than graphite below 1.5 MeV and again past 8 MeV.

For all configurations below, the example low-Z/high-Z attenuators and buildup caps are graphite/lead and PMMA/tungsten-alloy, respectively. Whenever a configuration uses only one buildup cap, it is the PMMA cap because it resembles the configuration used in previous studies more closely. Unless explicitly stated, all configurations have the same *total* number of measurements ($n_m = 24$) and the same transmission cutoff ($T_{\min} = 0.01$ —i.e., 1%). For a given transmission range, the successive attenuator thicknesses are integer multiples of the smallest thickness. When measurements are made with more than one buildup cap, the thicknesses used with one cap for a given attenuator are the multiples 2, 4, 6, ... of the smallest thickness, while the multiples 1, 3, 5, ... are used with the other cap. For any given configuration, measurements using different attenuators/detectors are fed simultaneously to the unfolding algorithm.

The ten measurement configurations investigated in this study are described here, and they are summarized in the first three columns of Table II in Sec. III. In the first configuration (C1), all 24 measurements are made using a graphite attenuator and one cap. This configuration is similar to previous studies and its performance is taken as the baseline. In C2, similar to C1, only graphite is used as an attenuator but the measurements are equally split between the two caps. This evaluates the concept discussed in Sec. II B. In C3, only one cap is used, with half the measurements made with the graphite attenuator alone while the other half made with the lead attenuator alone. This evaluates the concept discussed in Sec. II C. In C4, the measurements are divided equally among the four possible combinations of the two attenuators and two caps, each forming a separate transmission curve with six data points. This configuration combines the concepts discussed in Secs. II B and II D and it is the optimum configuration proposed in this study. In C4' and C4'', configuration C4 is re-evaluated for $T_{\min} = 0.1$ instead of 0.01, and for half the total number of measurements (i.e., only three measurements per attenuator/cap combination, for a total of 12 measurements). In C5, only one cap is used and measurements are made by adding alternating thicknesses of the two attenuator materials (i.e., the first measurement is with lead, the second adds the same mass thickness of graphite, the third adds lead, etc.). The rationale for considering this configuration is that it may provide better sampling of the spectrum. In C6, the concept of alternating measurements of C5 is done for two caps, with 12 measurements each. In C7, one cap is used and lead thicknesses are used up to an arbitrary transmission value of 0.5, then additional attenuation is done with graphite. The number of measurements is 6 for lead and 18 for graphite. This configuration represents an example of nonoptimum switch of materials (compared to C9 below). In C8, the same details of C7 apply except that the starting attenuator is graphite and the switch is to lead. In C9, the approach of Huang *et al.*¹⁰ (discussed in Sec. II C) is applied to the low-MV group of spectra (not applicable to the high-MV group) using only one cap. Since it is only a computational exercise, we chose the “chopper” material, in which $\bar{\mu}/\rho$ is calculated, to be a fictitious optimal material (a mathematical construct) with strong and

monotonic variation of μ/ρ with energy. In C10, the same concept of C9 is used but for two caps, with 12 measurements each.

Evaluation of the performance of various measurement configurations is useful only if realistic experimental noise is modeled and added to the analytical data. In this study the simulated noise is sampled from Gaussian distributions with the following standard deviations, σ : $\sigma = 0.15\%T$ for $0.30 < T < 1.00$, $\sigma = 0.25\%T$ for $0.10 < T \leq 0.30$, and $\sigma = 0.40\%T$ for $0.01 \leq T \leq 0.10$. Those noise levels are based on the uncertainty budget from the companion experimental study.³⁷ They already include the uncertainties associated with the applied experimental corrections for polarity, ion recombination, room scatter, etc., as well as the uncertainties due to repeatability, linac drift, leakage, etc.

For each of the spectra and configurations above, transmission data are calculated using Eq. (1), and then smeared 1000 different times with noise sampled from the Gaussian distributions above. The noisy data are then used to unfold the spectra according to the methods of Sec. II G below. Since the unfolded spectra are specified using a functional form with a few free parameters, then the most straightforward approach to evaluate the quality of a given measurement configuration is to look at the range of variation of those free parameters (around their correct values) when the parameters are unfolded from the different noisy data. However, we found empirically that this approach is not useful for many of the configurations discussed above because of the ill-conditioned nature of the problem and the potential for some correlation among the fit parameters in the functional form. As an alternative, a set of five metrics, $\% \Delta(X)$, are devised which are based on the deviations of the unfolded spectra from the original spectra that are used to generate the input transmission data. Those metrics are defined in Table I.

II.E. Corrections for nonideal conditions

The measured transmission signals, $T_{\text{meas}}(d, x_i)$, deviate from $T_{\text{ideal}}(d, x_i)$ of Eq. (1) because Eq. (1) assumes idealized geometry in a vacuum surrounding and considers only the primary photons. Previous studies typically ignored the difference between T_{meas} and T_{ideal} . To quantify the effect of this approximation and to develop the necessary correction methods, a “best-realistic” transmission measurement setup is constructed as shown in Fig. 3. It is impractical to make the setup any closer to ideal geometry for reasons such as the following. The source-to-chamber distance is the largest possible distance with typical bunker dimensions such that the chamber is reasonably far from the attenuator (to reduce attenuator scatter) and from the room walls (to reduce backscatter). The irradiation field at the chamber is the smallest possible field that covers its active volume to avoid partial-volume irradiation uncertainties. The upstream collimator is at the closest possible distance from the virtual source located inside the linac head.

Certain deviations from ideal conditions are best corrected for computationally, while others require experimental

TABLE I. Definition of the $\% \Delta(X)$ metrics which are used to quantify the quality of spectral unfolding using different measurement configurations. The metrics are similar to those used in our earlier study which evaluates different functional forms (Ref. 36).

Symbol	Definition
n_s	Number of point-source test spectra in each linac energy group; $n_s = 4$. Low-MV group: Tomotherapy imaging 3.5 MV, ^a Varian 4 MV, 6 MV, Siemens 6 MV. High-MV group: Varian 15 MV, 18 MV, Siemens 18 MV, Elekta 25 MV. ^b
n_g	Number of times of Gaussian noise smearing; $n_g = 1000$; $\sigma = 0.15\% T$, $0.25\% T$ and $0.40\% T$ for, respectively, $0.30 < T < 1.00$, $0.10 < T \leq 0.30$, and $0.01 \leq T \leq 0.10$.
n_b	Number of energy bins in a spectrum; $n_b = 100$.
ψ_b	Differential energy fluence for energy bin b of width dE_b .
ψ_{av}	Average energy fluence; $\psi_{av} = \sum_{b=1}^{n_b} \psi_b dE_b / \sum_{b=1}^{n_b} dE_b$.
E_m, E_{mp}, E_{av}	Respectively, the maximum, most-probable, and average photon energy of a spectrum.
E_e	Mean incident electron kinetic energy in the Monte Carlo simulation that generated the photon spectrum. $E_m = E_e$ for monoenergetic electrons.
$\% \Delta^{s,g}(X)$	For spectrum s unfolded from noisy transmission set g , the per cent deviation of the unfolded, u , from truth, t , for quantity X : $= (100/\psi_{av}^{s,t}) \sqrt{(1/n_b) \sum_{b=1}^{n_b} (\psi_b^{s,g,u} - \psi_b^{s,t})^2}$ for $X = \psi$, $= 100 E_e^{s,g,u} / E_e^{s,t} - 1 $ for $X = E_e$, ^c $= 100 E_m^{s,g,u} / E_m^{s,t} - 1 $ for $X = E_m$, ^c $= 100 X^{s,g,u} / X^{s,t} - 1 $ for $X = E_{mp}$ or E_{av} .
$\% \Delta(X)$	For quantity X , the overall per cent deviation over the n_g noisy transmission sets for a given spectrum, s , then over the n_s spectra of the energy group: $= (1/n_s n_g) \sum_{s=1}^{n_s} \sum_{g=1}^{n_g} \% \Delta^{s,g}(X)$ for $X = \psi, E_e, E_m, E_{mp},$ or E_{av} .

^aThe authors call the Tomotherapy imaging spectrum a 3.5 MV beam because $E_e = 3.5$ MeV (Ref. 36).

^b E_e for the Siemens 18 MV and Elekta 25 MV beams are only 14.7 and 19.0 MeV, respectively (Ref. 36).

^c E_m can be larger than E_e because of the energy spread of the incident electron beam in typical clinical linacs. Therefore, the unfolded endpoint energy is compared with both the correct E_e and the correct E_m .

corrections. The following are the deviations that are suitable for computational corrections, and they are investigated below. (a) Signal reduction due to the additional attenuation by the column of intervening air between the isocenter and the chamber—i.e., in Fig. 3, it is the column of length $(200 - x_i - \text{cap radius})$ cm. The reduction is largest for the signal with no attenuator present. (b) Signal increase due to all sources of forward scatter [attenuator(s), collimators, and surrounding air]. (c) Signal increase due to backscatter contribution from the back wall. On the other hand, the following deviations require experimental corrections, and they are discussed in the companion experimental study.³⁷ (a) Cable leakage due to Brownian motion and natural background. (b) Radiation-induced cable leakage caused by room scatter which, in turn, is caused by the linac head leakage. (c) Increased cavity ionization due to room scatter (computational modeling of this component would require detailed knowledge of the linac head and room contents). (d) Differential ion recombination effects.

EGSnrc is used to model the setup of Fig. 3 for various attenuator/cap combinations with the point-source spectra

from Table I. Simulation results are presented in Sec. III B below and they indicate that T_{meas} can differ from T_{ideal} by as much as 1.5%, and that ignoring those differences has a non-negligible effect on the accuracy of the unfolded spectra. Therefore, the following two methods are proposed to correct for nonideal conditions, and the two methods are shown to be equivalent. The first method is iterative as shown in the flowchart of Fig. 4. It involves the following steps: (a) the measured transmission data, T_{meas} , are used without any computational correction to unfold an approximate starting spectrum, ψ' ; (b) ψ' is used to analytically calculate approximate ideal transmission data, T'_{ideal} , using Eq. (1); (c) ψ' is also used as a point-source in EGSnrc simulations of the experimental setup to calculate approximate measured transmission data, T'_{meas} ; (d) an approximate correction factor for nonideal conditions is calculated as $F'_{\text{non-ideal}} = T'_{\text{ideal}} / T'_{\text{meas}}$; (e) $F'_{\text{non-ideal}}$ is used to correct the measured experimental data, T_{meas} , such that $T''_{\text{ideal}} = F'_{\text{non-ideal}} T_{\text{meas}}$; (f) T''_{ideal} data are used to unfold a revised spectrum, ψ'' ; and (g) steps b through f are repeated until the spectrum converges. It is important to note that if minor discrepancies exist between the

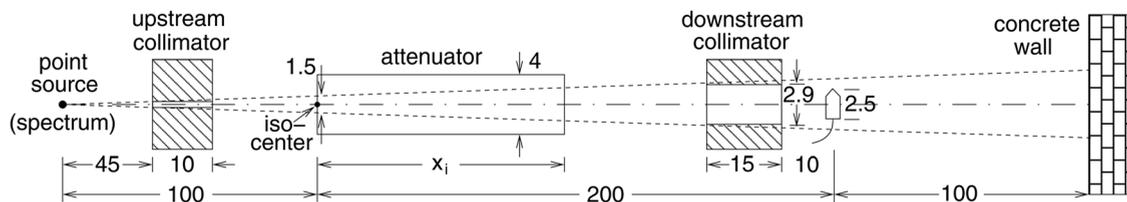


FIG. 3. A typical transmission measurement setup (dimensions in cm; not to scale). The setup is used in Sec. II E to quantify the effect of nonideal conditions and to develop the necessary correction methods.

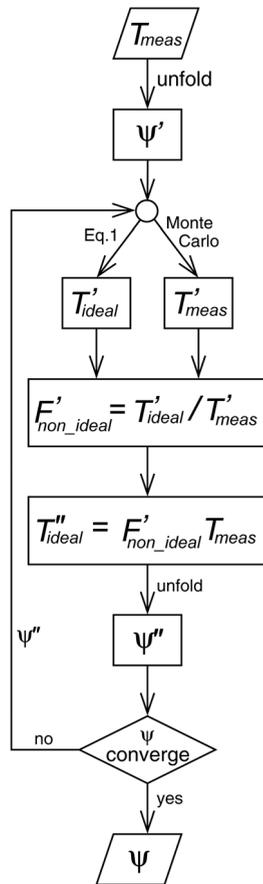


FIG. 4. A flowchart of the iterative method to correct for nonideal conditions.

μ values used in the analytical calculations of T_{ideal} and those used in the Monte Carlo calculations of T_{meas} (possibly due to interpolations using different-resolution grid points or different interpolation formulae), those discrepancies could propagate exponentially and lead to type-B errors of the same order as $|1 - F_{\text{non_ideal}}|$. For this reason, an *identical* fine-resolution grid of NIST XCOM μ values (2000 points equispaced in $\log(E)$ between 10 keV to 30 MeV) is used for both the analytical and the EGSnrc calculations.

In the second method, EGSnrc is used to generate a full system response matrix. In this matrix, an element $R(d, x_i, E)$ is the energy response per unit energy fluence for detector d to monoenergetic photons of energy E when the full experimental setup is modeled with an attenuator of thickness x_i . The equivalent of Eq. (1) is then

$$T_{\text{meas}}(d, x_i) = \frac{\int_{E_i}^{E_m} R(d, x_i, E) \psi(E) dE}{\int_{E_i}^{E_m} R(d, 0, E) \psi(E) dE}. \quad (2)$$

Since the matrix elements include the nonideal effects, no computational correction is required for the measured transmission data. In this study, the matrix elements are generated for the exact attenuator thicknesses used, and for ten energies chosen to capture the variation of the chamber response with energy. An example of the implementation of the two correction methods is given in the results (Sec. III B).

II.F. Spectral functional form

To tame the unfolding problem, the energy fluence spectra are specified using a functional form that was designed and rigorously validated in a precursor study.³⁶ The function is based on the physics of bremsstrahlung production and it has four free parameters, one of which is the mean incident electron kinetic energy, E_e . It is given by³⁶

$$\psi(E) = \left[1 + C_3 \frac{E}{E_e} + \left(\frac{E}{E_e} \right)^2 \right] \left[\ln \left(\frac{E_e(E_e - E)}{E} + 1.65 \right) - 0.5 \right] * \exp \left[- \left(\frac{\mu}{\rho} \right)_W(E) C_1^2 - \left(\frac{\mu}{\rho} \right)_{Al}(E) C_2^2 \right], \quad (3)$$

where E is the photon energy, $\left(\frac{\mu}{\rho} \right)_W(E)$ and $\left(\frac{\mu}{\rho} \right)_{Al}(E)$ are, respectively, the mass attenuation coefficients for tungsten and aluminum at energy E , and C_1, C_2, C_3 , and E_e are free parameters. The function was designed to have a clear cutoff at E_e , where $\psi(E_e) = 0$. This is useful when unfolding the higher-energy part of the spectrum from transmission data. The function was shown to outperform 11 other functional forms from the literature in terms of the combination of accuracy, flexibility, and robustness. For a diverse benchmark set of 65 high-resolution Monte Carlo spectra of typical clinical and research beams (3.5–30 MV), the function was shown to reproduce the energy fluence values in each bin with a normalized root-mean-square deviation of 1.7%. The mean incident electron kinetic energy, maximum photon energy, most-probable energy and average energy were reproduced, on average, within 1.4%, 4.3%, 3.9%, and 0.6% of their correct values, respectively. The differences between the original spectra and their functional representations lead to type-B uncertainties in the smallest transmission signals [calculated analytically using Eq. (1)] with typical and maximum values of 0.1% and 0.2%, respectively.

II.G. Spectral unfolding with incident electron energy estimation

The unfolding details presented here assume that Eq. (1) is used (in conjunction with the iterative correction method of Sec. II E). Identical methods are employed for Eq. (2). For any of the ten measurement configurations discussed in Sec. II D, the transmission data from different attenuators and/or detectors are fed simultaneously to the unfolding algorithm. Spectral parameterization and the enhanced spectral sensitivity in the transmission data make it possible to use the standard Levenberg-Marquardt least-squares minimization algorithm³⁹ without loss of robustness. Analytical first order derivatives of T_{ideal} in Eq. (1) with respect to each free parameter are calculated. The minimized objective function is the standard χ^2 , the sum of the squares of the inverse-variance-weighted difference between the input transmission data and those calculated analytically from Eq. (1) using the estimated free parameters in a given iteration. Previous studies used different methods to evaluate the integral in Eq. (1), including Simpson's rule¹⁰ and Monte Carlo.³⁰ In this study, Gauss-Legendre quadrature is used. In this quadrature

method the energies at which the weights are calculated depend on the lower and upper integration bounds [E_l and E_m in Eq. (1)], and since E_m is a free parameter (through E_e) which varies with iterations, new energies, weights, and attenuation coefficient values are needed for each iteration. To avoid this computational overhead, a fixed large upper energy bound is used to calculate the weights only once for a quadrature of order 200. This order is large enough to limit the integration errors to less than 0.01%. The lower-energy bound, E_l , is always 10 keV. The interpolation in the detector energy response is linear in $R(d, E)$ versus $\log(E)$.

During the minimization, E_e is treated as a free parameter. For the proposed optimum configuration (C4 in Sec. II D), E_e can be searched for concurrently with the other free parameters using the least-squares algorithm. However, for consistency with other configurations which are not as robust when noise is present, a grid search is used for E_e for all configurations (including C4) while minimizing the objective function with respect to all other free parameters. In this case the confidence limits on E_e are estimated using the graphical $\chi^2_{\min} + 1$ criterion,⁴⁰ and the covariance matrix for the free parameters is missing its off-diagonal elements for E_e . We found empirically that using such a matrix provides conservative confidence bounds on the unfolded spectra. For configuration C4 and comparable ones, the unfolding is truly robust against initial estimates of the free parameters, indicating that the global minimum is reached. For other configurations (e.g., C1 in Sec. II D, which was commonly used in previous studies), the unfolding is robust enough to yield reasonable spectra, which indicates that just using the proposed functional form improves the unfolding robustness. Typical deviations between the input and fitted transmission data are of the order of the corresponding Gaussian noise, and the reduced χ^2_{\min} with 20 degrees of freedom (24 measurements minus 4 free parameters) is of the order of 1. Minimization is virtually instantaneous for most configurations, particularly for the one proposed in this study, even with a reasonable-resolution grid search of E_e .

III. RESULTS AND DISCUSSION

III.A. Effect of the energy response of the detectors

Figure 5 shows the effect of approximating $R(E)$ on the accuracy of the unfolded spectra. The original spectra shown in Fig. 5 are generated from EGSnrc simulations of the full linac heads.³⁶ In panel a, the original 6 MV spectrum is used as a point-source to generate two sets of noise-free analytical transmission data using Eq. (1) with a graphite attenuator. Set 1 is generated using the correct $R(E)$ of the low-Z cap, while set 2 is generated using the correct $R(E)$ of the high-Z cap. The generated data are then used in spectral unfolding. When the correct $R(E)$ is employed during the unfolding, and the input transmission data are set 1 alone or set 2 alone or the combined data from the two sets, the unfolded spectra in the three cases are identical and are all given by curve *a* in Fig. 5(a). The reason that the different data yield an identical spectrum is that the data are noise-free and the minimization is completely robust. Curve *a* is almost identical to the origi-

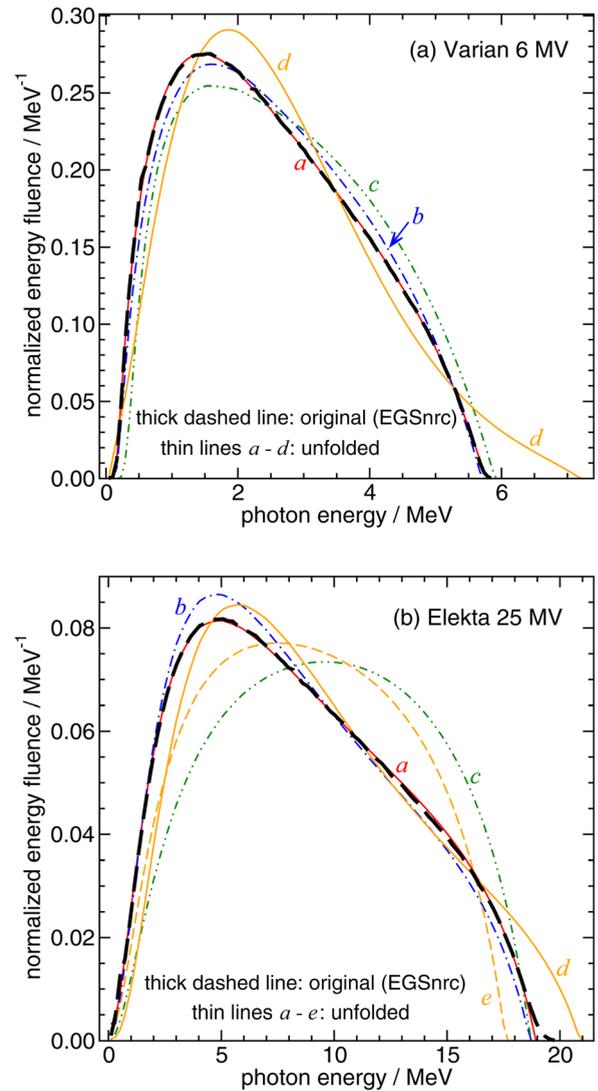


Fig. 5. The effect of approximating the detector energy response on the accuracy of the unfolded spectra for point-source spectra of (a) 6 MV and (b) 25 MV beams. Transmission data are free of Gaussian noise to isolate the effect. Spectra are normalized to unit energy fluence. See Sec. III A for the details of the different unfolded spectra.

nal spectrum, which validates the unfolding mechanics and demonstrates the flexibility of the functional form. When $R(E)$ is approximated as $(\mu_{\text{en}}/\rho)^{\text{air}}$ during the unfolding, and the input transmission data used are set 1 alone or set 2 alone or the combined data from the two sets, the unfolded spectra for the three cases are curves *b*, *c* and *d*, respectively. Since $R(E) = (\mu_{\text{en}}/\rho)^{\text{air}}$ is a worse approximation for the high-Z cap than it is for the low-Z cap [see Fig. (1)], curve *c* deviates more from the original spectrum than curve *b* does. When each bin value in curves *b* or *c* is scaled by the corresponding ratio $(\mu_{\text{en}}/\rho)^{\text{air}}/R(E)$ for the respective cap, the resulting spectrum is identical to curve *a* in both cases. On the other hand, no known scaling can be applied to curve *d* to relate it to the correct spectrum because when multiple detectors are used, the incorrect energy response adversely affects the minimization itself, as discussed in Sec. II A.

For the 25 MV spectrum [Fig. 5(b)], curves *a*–*d* are generated as done for the 6 MV spectrum above. The additional

curve e is obtained when noise-free transmission data from the four possible combinations of the attenuators and caps are simultaneously fed to the unfolding algorithm, and $R(E)$ is approximated as $(\mu_{en}/\rho)^{air}$ for both caps during the unfolding. For both curves d and e which employ more than one detector, the incorrect $R(E)$ s affect the minimization itself (not a simple scaling after the unfolding).

III.B. Effect of nonideal conditions

For the measurement setup of Fig. 3 with point-sources of Varian 6 MV and Elekta 25 MV spectra, Fig. 6 shows the EGSnrc results for the energy fluence spectra of the primary, forward-scattered and backscattered photons seen by the chamber for the smallest transmission signals. The figure shows that the spectra of scattered photons are distinctly

different from the primary ones. The forward-scattered photons collectively contribute $\sim 1\%$ to the energy fluence at the chamber location. The spectra of backscattered photons are dominated by two peaks: one is at 511 keV and it is due to annihilation events, and the other is roughly at half of 511 keV and it is due to the kinematics of Compton backscatter [cf., Eq. 6–10b in Ref. 41]. Comparing panels a and b, it can be seen that the spectrum of backscattered photons is largely independent of the incident beam, and it is thus also independent of the attenuator length. The contribution of backscattered photons to the energy fluence at the chamber location is found to be negligible (0.01% or less) and can safely be ignored.

Figure 7 shows the magnitude of the correction factor for nonideal conditions, $F_{non_ideal} = T_{ideal}/T_{meas}$, where T_{ideal} is calculated using Eq. (1) and T_{meas} is calculated using EGSnrc for the best-realistic measurement setup of Fig. 3. The deviation of T_{meas} from T_{ideal} is nonlinear with signal and can be as much as 1.5% of the smallest signals. The deviation depends on the beam energy and on the attenuators and detectors used. The magnitude of the deviation is determined by the interplay between the second-order attenuation and scatter effects in the signals with and without the attenuator.

Figure 8 shows the effect of nonideal conditions on the accuracy of the unfolded spectra and illustrates the two correction methods proposed in Sec. II F. When T_{meas} data are fed to the unfolding algorithm without correction, the unfolded spectrum is curve a . For the iterative correction method, curve a is the approximate starting spectrum. After only one iteration, the estimated F_{non_ideal} is found to be within 0.15% of its correct value because it is a ratio of a ratio, and thus not very sensitive to the exact spectrum used to calculate it. The unfolded spectrum after the first iteration is curve b , and it is already very close to the original spectrum. This suggests that empirically, only one iteration is needed. For the correction method which employs an EGSnrc system response matrix, curve c is the unfolded spectrum. The figure shows that ignoring the corrections for nonideal conditions

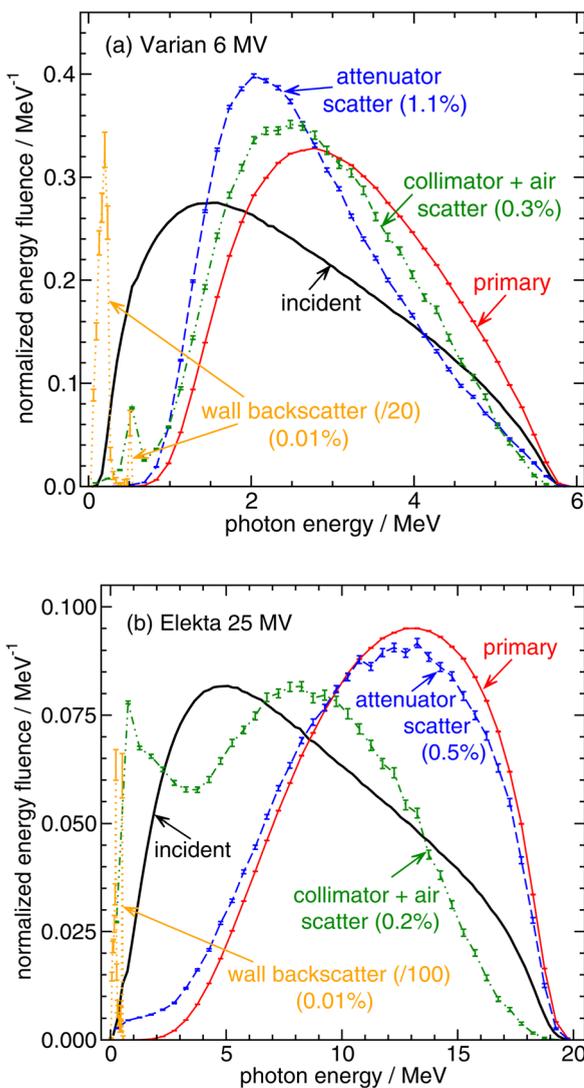


FIG. 6. For point-source spectra in the setup of Fig. 3, this figure shows the EGSnrc-generated spectra at the detector location after the beam passes through an amount of material that reduces the collision air-kerma to 1% of its original value: 8.5 cm lead for the 6 MV spectrum and 125 cm graphite for the 25 MV spectrum. Spectra are normalized to unit energy fluence. The spectra of wall backscatter are scaled down for graph clarity. The ratios of the energy fluence from different scatter sources relative to the primary are shown in brackets.

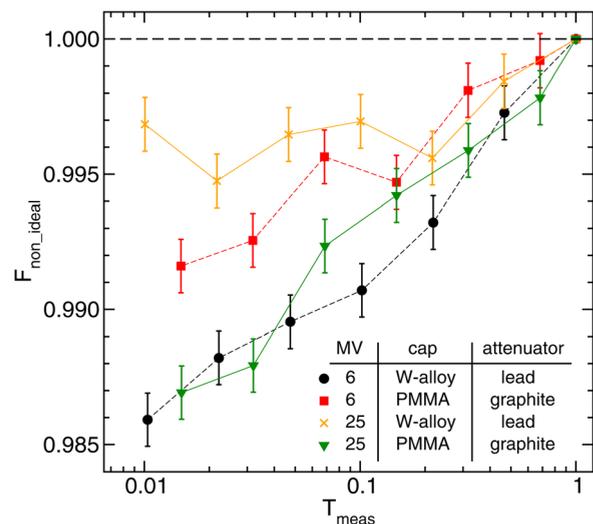


FIG. 7. The correction factor for nonideal conditions, $F_{non_ideal} = T_{ideal}/T_{meas}$.

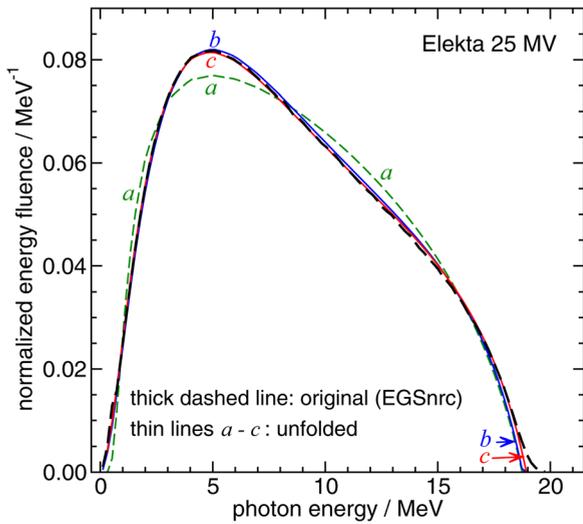


FIG. 8. The effect of nonideal conditions on the accuracy of the unfolded spectra for a point-source Elekta 25 MV spectrum in the setup of Fig. 3. Curve *a* is the unfolded spectrum when ignoring the corrections for nonideal conditions. For the iterative correction method, curve *b* compared to curve *a* shows the progression of the unfolded spectrum with improved estimate of the correction factor after only one iteration. For the correction method which employs an EGSnrc system response matrix, curve *c* is the unfolded spectrum. Spectra are normalized to unit energy fluence.

reduce the accuracy of the unfolded spectrum, and that the two proposed correction methods are equivalent.

III.C. Effect of the enhanced spectral sensitivity

Table II shows the results of evaluating the performance of the ten different measurement configurations described in Sec. II D using the eight spectra of Table I. The following observations can be made. No extremely large deviations are observed between the unfolded spectra and the original ones for all configurations. This can be attributed to the accuracy and flexibility of the functional form (without being over-

parameterized). Using multiple detectors of different energy response (as done in configuration C2), or using multiple attenuating materials (as in C3) clearly improves the unfolding accuracy for both MV groups compared with C1 (the base configuration). The unfolding accuracy of C2 and C3 are comparable to each other. When both multiple attenuators and multiple detectors are used (C4), their combined effect further improves the unfolding accuracy compared to using either of them alone (C2 or C3). In C4, the excellent ability to unfold the mean incident electron kinetic energy, E_e , (within 1.4% of its correct value) is due to both the improved sensitivity to the higher energy portion of the spectrum and the design of the functional form which has a clear cutoff at E_e . The larger deviations when the unfolded E_e is compared with the correct maximum photon energy, E_m , reflect the inherent ill-definition of E_m for typical clinical linac spectra.³⁶ Using the average $\% \Delta(X)$ values from the two MV groups in Table II, it can be said that for configuration C4, the unfolded energy fluence spectra agree with the original ones with a normalized root-mean-square deviation of 2.3%, and that the unfolded E_e, E_m, E_{mp} , and E_{av} agree, on average, with their correct values within 1.4%, 4.8%, 4.4%, and 0.7%, respectively.

Comparing C4 with C1, the enhanced spectral sensitivity alone improves the accuracy of the unfolded spectra by a factor of $(7.0 + 6.2)/(2.3 + 2.3) \approx 3$, and the accuracy of the unfolded E_e by a factor of $(7.1 + 4.8)/(1.4 + 1.4) \approx 4$. Those improvements are *in addition to* the accuracy improvements from using the functional form of Eq. (3) (as opposed to other forms in the literature), and the improvements from properly accounting for the detector energy response and the nonideal measurement conditions.

In C5 and C6, the interleaving of the attenuators does not provide an advantage over C1 and C4, respectively. The performance of C7 is better than that of C8 for both MV groups because in C8 switching to lead as the spectrum hardens

TABLE II. Overall performance of the ten measurement configurations described in Sec. II D. Analytical transmission data are calculated using Eq. (1) for point-source spectra and smeared 1000 different times with the same level of realistic Gaussian noise. Table I contains the definition of the $\% \Delta(X)$ metrics, the levels of the simulated noise and the list of spectra included in the low-MV and the high-MV groups. Unless explicitly stated, the total number of measurements, n_m , is 24 and the transmission cutoff, T_{min} , is 0.01 (i.e., 1%).

Measurement configuration			Group of low-MV spectra					Group of high-MV spectra				
Index	Attenuator materials and how they are used	# caps	$\% \Delta(X), X$ is:					$\% \Delta(X), X$ is:				
			ψ	E_e	E_m	E_{mp}	E_{av}	ψ	E_e	E_m	E_{mp}	E_{av}
C1	C full curve	1	7.0	7.1	8.5	10.9	1.6	6.2	4.8	7.1	10.1	1.4
C2		2	4.1	3.5	6.4	5.1	0.6	3.4	1.5	4.2	8.5	1.5
C3	C and Pb full curves	1	3.0	2.2	6.0	4.4	0.6	3.4	2.3	4.9	5.2	0.8
C4		2	2.3	1.4	5.2	3.2	0.4	2.3	1.4	4.3	5.6	0.9
C5	Pb + C alternating	1	7.6	9.8	11.0	7.5	3.1	6.4	4.2	5.8	10.2	1.9
C6		2	3.6	2.7	5.9	3.7	0.5	4.2	2.6	5.5	7.3	1.1
C7	$T_{switch} = 0.5$: Pb to C	1	4.9	4.4	7.1	6.4	1.8	7.1	5.9	8.1	9.4	1.1
C8	C to Pb	1	7.7	8.6	9.7	7.6	2.7	10.6	8.1	8.3	15.8	3.2
C9	optimum T_{switch} : Pb to C	1	3.8	3.9	7.2	5.1	0.6	—	—	—	—	—
C10		2	2.5	1.6	5.5	3.0	0.4	—	—	—	—	—
C4'	C4 but $T_{min} = 0.1$ not 0.01	2	3.2	2.6	5.9	3.3	0.4	3.3	2.0	4.7	6.8	0.9
C4''	C4 but $n_m = 12$ not 24	2	2.8	1.9	5.5	4.1	0.6	2.9	1.6	4.5	6.4	1.0

introduces degeneracy into the problem that is not eliminated by any other measurements, and the effect worsens for the high-MV spectra. The performance of C9 (not applicable to high-MV spectra—see Sec. II C) is better than that of C7, indicating that optimizing the transmission value for switching between materials improves the unfolding accuracy. The performance of C10 is almost the same as that for C4, which is expected for low-MV spectra. However, the necessary experimental overhead in C10 (see Sec. II C) makes configuration C4 still more favorable.

Based on the observations above, configuration C4 outperforms all others for both MV groups, and it is thus recommended in this study for *all* MV beams. With this configuration and T_{\min} down to only 0.1 instead of 0.01 (C4' in Table II), $\% \Delta(\psi)$ is 3.3%. This indicates that the enhanced spectral sensitivity, combined with the robustness of the functional form, are enough to extract accurate spectra with a more compact setup and without having to deal with the issues associated with small transmission signals. Also, with only half the total number of measurements (C4'' in Table II), $\% \Delta(\psi)$ marginally worsens to 2.85% compared with C4 (2.3%). This can be useful if data acquisition time is an issue.

Figure 9 shows the variation of the objective function during a grid search of E_e for configurations C1 and C4. The data are shown for a 25 MV spectrum where accurate spectral unfolding is most challenging. The objective function has a clear minimum in C4 compared with C1. Since the noise level and the functional form used are the same in both configurations, the increase in sensitivity to the maximum energy is exclusively from the additional spectral information in the transmission data. In C4, the objective function minimum is not identical with different noise. This leads to

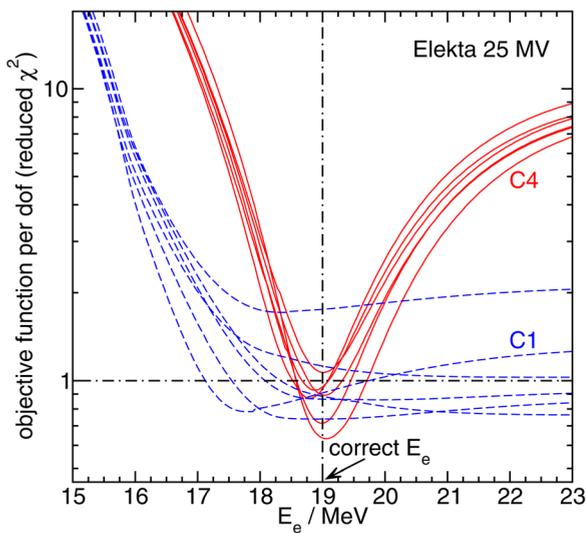


Fig. 9. For a point-source Elekta 25 MV spectrum, the figure shows the variation of the objective function per degree of freedom (the reduced χ^2) during a grid search of the mean incident electron kinetic energy, E_e , while minimizing the objective function with respect to all other free parameters. Data are shown for the optimum measurement configuration proposed in this study (C4 in Table II) and for the typical configuration in previous studies (C1 in Table II). Different lines represent smearing of the same analytical transmission data with Gaussian noise (noise levels are defined in Table I). Note that the correct E_e is only 19.0 MeV (Ref. 36).

very close (but not identical) unfolded spectra, which reflects the inherent ill-conditioned nature of the problem. The value of the reduced χ^2_{\min} is always close to unity, indicating that Eq. (1) is able to model the input transmission data in accord with the variance which, in turn, implies that the functional form used in the model is not over-parameterized or under-parameterized.

To graphically illustrate the quality of the unfolded spectra when noise is present, Fig. 10 shows the 95% confidence bounds on the unfolded spectra for configurations C1 and C4. The bounds are estimated by evaluating the average and standard deviation of the 1000 values of the unfolded energy fluence at a given energy; the bounds are twice the standard deviation above and below the average unfolded value. The following observations can be made from the figure. The bounds are much tighter with C4 than they are with C1, particularly at the higher-energy part of the spectra (which is the part that benefitted the most from the enhanced spectral

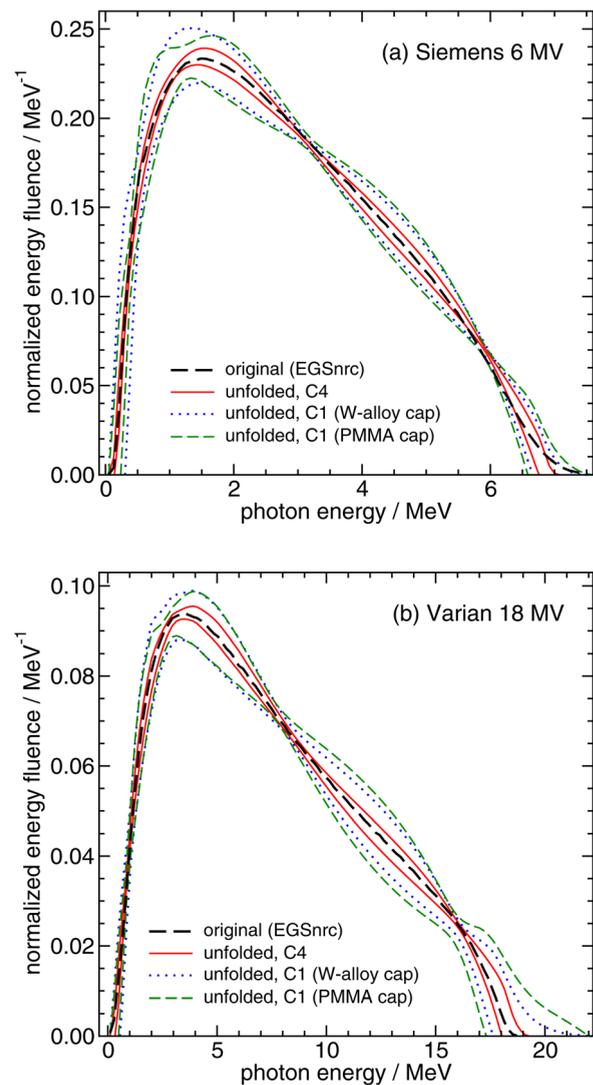


Fig. 10. The 95% confidence bounds on the unfolded spectra for: (a) 6 MV and (b) 18 MV spectra. The bounds are shown for the optimum measurement configuration proposed in this study (C4 in Table II) and for the typical configuration in previous studies (C1 in Table II). See Sec. III C for how the bounds are estimated.

sensitivity). The tight bounds for C4 indicate that making some of the measurements with a low- Z attenuator does in fact eliminate the degeneracy introduced by the high- Z attenuator measurements while maintaining the benefit of better energy differentiation. In any given configuration the bounds are tighter at the lower-energy part of the spectra because of the stronger variation of μ with energy. For C1, the bounds with either cap alone are comparable; however, it is interesting that the bounds are tighter for the PMMA cap at the lower-energy part of the spectra and tighter for the tungsten-alloy cap at the higher-energy part. This is a direct reflection of the shape of the energy response with the two caps [Fig. 1(d)]. The irregular shape of the bounds is a result of the shape of the functional form where the unfolded spectra with different noise cross each other more often at certain energies than they do at others, and thus the bounds at those energies are tighter than the bounds at neighboring energies.

For C4, the confidence bounds can also be determined directly from a single noisy transmission set using the full covariance matrix. This is possible because E_e can be searched for as a free parameter concurrently with the other free parameters in the least-squares algorithm (as opposed to a grid search). For this direct method of confidence bounds estimation, one must take into account that the scaling used to normalize ψ to unit energy fluence is in itself a function of the unfolded parameters. The bounds calculated directly are found to be almost identical to the ones in Fig. 10 for C4 (including the two bottlenecks). This observation is an independent validation for the methods used in the companion experimental study where only one set of measured transmission data is available.³⁷ It is also an indirect validation of the bounds for C1 in Fig. 10.

Finally, we used EGSnrc to quantify the effect on depth-dose curves for the difference between the spectra unfolded within the tight bounds of C4 and the original spectra. The effect on the depth of maximum dose is within 1 mm, and on the per cent depth dose at 10 and 20 cm is within 0.2% (relative to the maximum dose of 100). This indicates that the spectra unfolded using the approach proposed in this study are within the measurement uncertainties of dosimetric quantities of clinical interest.

IV. CONCLUSIONS

In this study, a physics-based approach is developed which improves the accuracy and robustness of the ill-posed problem of unfolding megavoltage bremsstrahlung spectra from noisy transmission data. The simple but surprisingly unexploited idea of using multiple detectors of different energy response, in conjunction with multiple attenuating materials, significantly improves the spectral sensitivity of the transmission data, particularly at higher energies where spectral sensitivity is least. The proposed spectral functional form plays a central role in taming the unfolding problem. Inaccurate detector energy response modeling and ignoring corrections for nonideal conditions both have a non-negligible effect on the unfolding accuracy. Compared with

the commonly-used approach in previous studies, the improvements in the accuracy of the unfolded spectra and in the estimated mean incident electron kinetic energy are *at least* factors of three and four, respectively. The proposed approach does not require knowledge of the linac head details or *a priori* knowledge of the spectrum shape or end-point energy. The approach is shown to be valid for all therapy and MV imaging beams (3.5–25 MV).

ACKNOWLEDGMENTS

The authors thank Malcolm McEwen of the National Research Council Canada for many insightful discussions and Bryan Muir of Carleton University for the Exradin A19 egs++ model. Elsayed Ali acknowledges the funding from a Vanier CGS. Dave Rogers acknowledges the funding from the CRC program, NSERC, CFI, and OIT.

^{a)}Author to whom correspondence should be addressed. Electronic mail: eali@physics.carleton.ca

^{b)}Electronic mail: drogers@physics.carleton.ca

¹B. L. Werner, I. J. Das, F. M. Khan, and A. S. Meigooni, "Dose perturbations at interfaces in photon beams," *Med. Phys.* **14**, 585–595 (1987).

²P. M. Charland, I. J. Chetty, L. D. Paniak, B. P. Bednarz, and B. A. Fraass, "Enhanced spectral discrimination through the exploitation of interface effects in photon dose data," *Med. Phys.* **31**, 264–276 (2004).

³E. Tonkopi, M. McEwen, B. Walters, and I. Kawrakow, "Influence of ion chamber response on in-air profile measurements in megavoltage photon beams," *Med. Phys.* **32**, 2918–2927 (2005).

⁴P.-H. Huang, K. R. Kase, and B. E. Bjärngard, "Spectral characterization of 4 MV bremsstrahlung by attenuation analysis," *Med. Phys.* **8**, 368–374 (1981).

⁵B. R. Archer, P. R. Almond, and L. K. Wagner, "Application of a Laplace transform pair model for high energy x-ray spectral reconstruction," *Med. Phys.* **12**, 630–633 (1985).

⁶S. D. Ahuja, P. G. Steward, T. S. Roy, and E. D. Slessinger, "Estimated spectrum of a 4-MV therapeutic beam," *Med. Phys.* **13**, 368–373 (1986).

⁷P. Francois, A. Catala, and C. Scouarnec, "Simulation of x-ray spectral reconstruction from transmission data by direct resolution of the numeric system $AF = T$," *Med. Phys.* **20**, 1695–1703 (1993).

⁸A. Catala, P. Francois, J. Bonnet, and C. Scouarnec, "Reconstruction of 12 MV bremsstrahlung spectra from measured transmission data by direct resolution of the numeric system $AF = T$," *Med. Phys.* **22**, 3–10 (1995).

⁹S. Hussain, "Artificial neural network model for spectral construction of a linear accelerator megavoltage photon beam," *UKSim/AMSS IEEE 1st International Conference on Intelligent Systems, Modelling and Simulation*, Liverpool, England (2010), pp. 86–91.

¹⁰P.-H. Huang, K. R. Kase, and B. E. Bjärngard, "Simulation studies of 4-MV x-ray spectral reconstruction by numerical analysis of transmission data," *Med. Phys.* **9**, 695–702 (1982).

¹¹P.-H. Huang, K. R. Kase, and B. E. Bjärngard, "Reconstruction of 4-MV bremsstrahlung spectra from measured transmission data," *Med. Phys.* **10**, 778–785 (1983).

¹²A. Piermattei, G. Arcovito, F. A. Bassi, and L. Azario, "Reconstruction of 9 MV bremsstrahlung spectrum by numerical analysis of measured transmission data," *Phys. Med.* **1**, 43–57 (1987).

¹³A. Piermattei, G. Arcovito, L. Azario, C. Bacci, L. Bianciardi, E. De Sapio, and C. Giacco, "A study of quality of bremsstrahlung spectra reconstructed from transmission measurements," *Med. Phys.* **17**, 227–233 (1990).

¹⁴A. Nisbet, H. Weatherburn, J. D. Fenwick, and G. McVey, "Spectral reconstruction of clinical megavoltage photon beams and the implications of spectral determination on the dosimetry of such beams," *Phys. Med. Biol.* **43**, 1507–1521 (1998).

¹⁵R. G. Waggner, M. M. Blough, J. A. Terry, D. Chen, N. E. Lee, S. Zhang, and W. D. McDavid, "X-ray spectra estimation using attenuation measurements from 25 kVp to 18 MV," *Med. Phys.* **26**, 1269–1278 (1999).

¹⁶A. Iwasaki, M. Kubota, J. Hirota, A. Fujimori, K. Suzuki, M. Aoki, and Y. Abe, "Characteristic features of a high-energy x-ray spectra estimation

- method based on the Waggener iterative perturbation principle," *Med. Phys.* **33**, 4056–4063 (2006).
- ¹⁷T. Shimozato, K. Tabushi, S. Kitoh, Y. Shiota, C. Hirayama, and S. Suzuki, "Calculation of 10 MV x-ray spectra emitted by a medical linear accelerator using the BFGS quasi-Newton method," *Phys. Med. Biol.* **52**, 515–523 (2007).
- ¹⁸G. Wahba, "Practical approximate solutions to linear operator equations when the data are noisy," *SIAM (Soc. Ind. Appl. Math.) J. Numer. Anal.* **14**, 651–667 (1977).
- ¹⁹A. Höcker and V. Kartvelishvili, "SVD approach to data unfolding," *Nucl. Inst. Meth. A* **372**, 469–481 (1996).
- ²⁰B. Armbruster, R. J. Hamilton, and A. K. Kuehl, "Spectrum reconstruction from dose measurements as a linear inverse problem," *Phys. Med. Biol.* **49**, 5087–5099 (2004).
- ²¹M. Manciu, F. Manciu, T. Vulcan, E. Nes, and R. Waggener, "Robust megavoltage x-ray spectra estimation from transmission measurements," *J. X-Ray Sci. Technol.* **17**, 85–99 (2009).
- ²²C. R. Baker, "Reconstruction of clinical bremsstrahlung spectra in the range 4 to 30 MeV," Ph.D. thesis, University of Surrey, United Kingdom, 1993.
- ²³R. M. Harrison, G. D. Lambert, and C. L. Chapple, "Spectral estimation and contrast calculation in the design of contrast-detail test objects for radiotherapy portal imaging," *Phys. Med. Biol.* **38**, 545–556 (1993).
- ²⁴M. Krmar, J. Slivka, I. Bikit, M. Vesković, L. Conkić, M. Bistović, and A. Rudić, "A new method for the measurement of bremsstrahlung spectra," *Phys. Med. Biol.* **38**, 533–544 (1993).
- ²⁵C. R. Baker, B. Ama'ee, and N. M. Spyrou, "Reconstruction of megavoltage photon spectra by attenuation analysis," *Phys. Med. Biol.* **40**, 529–542 (1995).
- ²⁶M. Krmar, J. Slivka, I. Bikit, M. Vesković, and L. Conkić, "Evaluation of bremsstrahlung spectra generated by a 4-MeV linear accelerator," *Med. Phys.* **23**, 651–654 (1996).
- ²⁷C. R. Baker and K. K. Peck, "Reconstruction of 6 MV photon spectra from measured transmission including maximum energy estimation," *Phys. Med. Biol.* **42**, 2041–2051 (1997).
- ²⁸P. Bloch and J. McDonough, "Extraction of the photon spectra from measured beam parameters," *Med. Phys.* **25**, 752–757 (1998).
- ²⁹M. Partridge, "Reconstruction of megavoltage photon spectra from electronic portal imager derived transmission measurements," *Phys. Med. Biol.* **45**, N115–N131 (2000).
- ³⁰S. Sawchuk, "Monte Carlo estimate to improve photon energy spectrum reconstruction," in *Advanced Monte Carlo for Radiation Physics, Particle Transport Simulation and Applications, Proceedings of Monte Carlo 2000 Conference, Lisbon*, edited by A. Kling, F. Barão, M. Nakagawa, L. Távora, and P. Vaz (Springer, Berlin, 2001), pp. 383–388.
- ³¹W. H. Hinson and J. D. Bourland, "Spectral reconstruction of high energy photon beams for kernel based dose calculations," *Med. Phys.* **29**, 1789–1796 (2002).
- ³²W. H. Hinson, W. T. Kearns, A. F. deGuzman, and J. D. Bourland, "Photon spectral characteristics of dissimilar 6 MV linear accelerators," *Med. Phys.* **35**, 1698–1702 (2008).
- ³³H. M. Garnica-Garza, "Determination of radiotherapy x-ray spectra using a screen-film system," *Med. Biol. Eng. Comput.* **46**, 1029–1037 (2008).
- ³⁴I. Kawrakow, "Accurate condensed history Monte Carlo simulation of electron transport. I. EGSnrc, the new EGS4 version," *Med. Phys.* **27**, 485–498 (2000).
- ³⁵I. Kawrakow, E. Mainegra-Hing, D. W. O. Rogers, F. Tessier, and B. R. Walters, "The EGSnrc Code System: Monte Carlo simulation of electron and photon transport," NRC Technical Report No. PIRS-701 v4-2-3-2, National Research Council Canada, Ottawa, Canada, <http://www.irs.inms.nrc.ca/inms/irs/EGSnrc/EGSnrc.html>, 2011.
- ³⁶E. S. M. Ali and D. W. O. Rogers, "Functional forms for photon spectra of clinical linacs," *Phys. Med. Biol.* **55**, 31–50 (2012).
- ³⁷E. S. M. Ali, M. R. McEwen, and D. W. O. Rogers, "Experimental methods for unfolding linac photon spectra from transmission measurements, with direct independent spectral validation," *Med. Phys.* (submitted).
- ³⁸I. Kawrakow, E. Mainegra-Hing, F. Tessier, and B. Walters, "The EGSnrc C++ class library," Technical Report No. PIRS-898 (rev A), National Research Council of Canada, Ottawa, Canada, <http://www.irs.inms.nrc.ca/EGSnrc/PIRS898/>, 2009.
- ³⁹W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran*, 2nd ed., Cambridge University Press, New York, NY (1992).
- ⁴⁰D. W. O. Rogers, "Analytic and graphical methods for assigning errors to parameters in non-linear least squares fitting," *Nucl Instrum Methods* **127**, 253–260 (1975).
- ⁴¹H. E. Johns and J. R. Cunningham, *The Physics of Radiology*, 4th ed., Charles C. Thomas, ed. (Springfield, Illinois, 1983).
- ⁴²B. R. Muir and D. W. O. Rogers, "Monte Carlo calculations of k_Q , the beam quality conversion factor," *Med. Phys.* **37**, 5939–5950 (2010).