

# A proposed alternative to phase-space recycling using the adaptive kernel density estimator method

Neelam Tyagi

*The University of Michigan Department of Nuclear Engineering and Radiological Sciences, Ann Arbor, Michigan 48109-2104, and The University of Michigan, Department of Radiation Oncology, Ann Arbor, Michigan 48109-0010*

William R. Martin

*The University of Michigan Department of Nuclear Engineering and Radiological Sciences, Ann Arbor, Michigan 48109-2104*

J. Du

*Power Information Network, An Affiliate of J.D. Power and Associates, Troy, Michigan 48098*

A. F. Bielajew

*The University of Michigan Department of Nuclear Engineering and Radiological Sciences, Ann Arbor, Michigan 48109-2104*

Indrin J. Chetty<sup>a)</sup>

*The University of Michigan, Department of Radiation Oncology, Ann Arbor, Michigan 48109-0010*

(Received 13 June 2005; accepted for publication 8 December 2005; published 31 January 2006)

We have implemented a nonparametric density estimation technique, the adaptive kernel density estimator (AKDE), to generate additional phase space (PS) variables in the vicinity of simulated PS points in Monte Carlo linear accelerator simulation. The method involves the placement of kernels at simulated PS points that have a “window width” that depends on the density of simulated PS points. This method has been tested on known one-dimensional (1-D) and two-dimensional (2-D) probability density functions (PDFs) and has been used to sample (photons only) from PS files generated from accelerator simulations. The original simulated PS vector  $(x, y, u, v, E)$  was reduced to a rotationally invariant PS vector  $(r, \theta, \alpha, E)$  that takes advantage of the azimuthal symmetry ( $\phi$ ) above the collimating jaws. The new PS vector  $(r', \theta', \alpha', E')$  is sampled in the vicinity of the sampled PS vector  $(r, \theta, \alpha, E)$ . The first step in assessing the accuracy of the method was a correlation analysis among the AKDE generated PS variables compared with correlations among the original PS variables. “In-air” particle fluence distributions between AKDE samples and the original PS distribution showed agreement within 2% (–8.8% to 6.8%) across the entire phase space plane. Central axis energy distributions and angular distributions agreed on average to within 1.5% (range=–1.5% to 6.6%) and 0.1% (range=0 to 3.0%), respectively. Dose profiles were calculated for field sizes  $3 \times 3 \text{ cm}^2$ ,  $10 \times 10 \text{ cm}^2$ , and  $30 \times 30 \text{ cm}^2$  for AKDE and compared against calculations performed with PS recycling. AKDE calculated depth doses and profiles were within 2% and 2%/1 mm, respectively, of those computed using PS recycling. © 2006 American Association of Physicists in Medicine. [DOI: 10.1118/1.2163250]

Key words: adaptive kernel density estimator, Monte Carlo, phase space, linac simulation

## I. INTRODUCTION

Monte Carlo simulation of the linear accelerator treatment head using the BEAMnrc<sup>1</sup> code yields an output file that records the phase space (PS) particle vector  $\{x, y(\text{position}), u, v(\text{direction}), E(\text{energy}), \text{latch}, \text{wt}(\text{weight})\}$  for each particle crossing the scoring plane. The variable *latch* contains the particle charge, the number of times the particle has crossed the scoring plane, and information that allows the particle’s history to be traced. When the PS file is used as an input for patient-dependent simulation, it may have to be recycled several times in order to achieve acceptable statistics in the final dose calculations. Walters *et al.* have described in detail the effect of restarting and recycling the PS on the uncertainty of the dose calculation during simulation with BEAMnrc, and have pointed out how correla-

tions arising particularly from particle restarting can significantly underestimate the uncertainty.<sup>2</sup> In this paper we propose an alternative to PS recycling using an approach called the adaptive kernel density estimator (AKDE) originally proposed by Silverman<sup>3</sup> to estimate population densities. This method [the response kernel density estimation Monte Carlo method (RKMC)] has been previously investigated at our institution with its application to electron transport simulation in the EGS4 Monte Carlo code.<sup>4</sup> We have now applied this technique to PS treatment head simulation. The basic method can be summarized as follows: (a) a PS vector is read from the PS input file, (b) AKDE is used to sample additional “on the fly” PS vectors in the vicinity of the existing vector, and (c) the new PS distribution, now containing many more samples than the original PS file, is used for the

patient-dependent calculation. The AKDE technique is novel because it generates PS samples without the need for direct particle simulation and, in principle, has the potential to reduce the latent variance given a minimum number of simulated histories in the PS file. Our intent in this paper is to describe the AKDE method and its application in treatment head simulation. We illustrate the accuracy of the AKDE method through evaluation of fluence distributions and in-phantom dose calculations in comparison to other more well-established methods, such as PS recycling.

In what follows we will discuss the theory associated with AKDE and its application in Monte Carlo sampling, discuss preprocessing of the PS variables for input to AKDE, and finally compare AKDE with direct PS simulation and PS recycling in linac treatment head simulation and in-phantom dose calculations.

## II. THEORY

### A. Adaptive Kernel Density Estimator (AKDE)

A theoretical derivation of the AKDE method is reported in the text by Silverman.<sup>3</sup> We present necessary parts of the theory here and in the Appendix for completeness. Suppose we have a simulated dataset of  $n$  independent and identically distributed samples  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ , (each of dimension  $d$ ), from an unknown probability distribution function (PDF)  $f(\mathbf{x})$ , that we would like to estimate. Because of the complexity and multidimensionality of the data, it is very difficult to use parametric density estimation (e.g., Gaussian with unknown mean and variance) to obtain accurate results. In other fields, kernel density estimators are widely used in nonparametric density estimation and these are investigated in this report. For a fixed kernel density estimator, the (usually unknown) PDF  $f(\mathbf{x})$  is approximated by the following PDF:

$$\hat{f}(\mathbf{x}) = \frac{1}{n|\mathbf{H}|^{1/2}} \sum_{i=1}^n k[(\mathbf{x} - \mathbf{x}_i)^T \mathbf{H}^{-1}(\mathbf{x} - \mathbf{x}_i)]. \quad (2.1)$$

The bandwidth matrix  $\mathbf{H}$  is a  $d$ -dimensional symmetric positive definite matrix that can be thought of as a covariance matrix for the kernel, and it typically has the following form  $\mathbf{H} = \{\text{diag}(h_1^2, h_2^2, \dots, h_d^2), h_1, h_2, h_d > 0\}$ , where  $h_i$  is the global bandwidth for variable  $x_i$  (the  $i$ th component of vector  $\mathbf{x}$ ) and is a measure of the “spread” of the variable  $x_i$ , and the kernel  $K(\mathbf{x}) = k(\mathbf{x}^T \mathbf{x})$  is a symmetric, non-negative function that is centered at zero and integrates to 1:

$$\int K(\mathbf{x}) d\mathbf{x} = 1, \quad \int \mathbf{x} K(\mathbf{x}) d\mathbf{x} = 0 \quad (2.2)$$

and

$$\int \mathbf{x}^2 K(\mathbf{x}) d\mathbf{x} = k_2 \neq 0.$$

The deviation of the density estimator  $\hat{f}$  from the true density  $f$  is usually quantified using a measure termed the mean in-

tegrated square error (MISE) as a criterion of the global accuracy:

$$\text{MISE}(\hat{f}) = E \int \{f(\mathbf{x}) - \hat{f}(\mathbf{x})\}^2 d\mathbf{x}, \quad (2.3)$$

where  $E(\circ)$  denotes the expectation value of the quantity  $\circ$ .

It is possible to choose a global bandwidth to minimize MISE. For example, in one dimension, the following expression for MISE holds:

$$\text{MISE}(\hat{f}) \approx \frac{1}{4} h^4 k_2^2 \int f''(x)^2 dx + n^{-1} h^{-1} \int K(t)^2 dt. \quad (2.4)$$

The details of the proof are shown in the Appendix. The first term in the above expression represents the bias term (truncation error) and the second term is the variance term (statistical error). The above expression also represents two fundamental difficulties of density estimation, which is true for any number of dimensions: (1) MISE cannot be evaluated unless  $f(x)$  is known and (2) a compromise has to be made between bias and variance to reduce the MISE. The bandwidth  $h$  needs to be small to reduce the bias; however, reducing the variance requires a large value of  $nh$ . Unless  $n$  is sufficiently large,  $h$  has to be large. Silverman derived the following expression for the optimal value using a Gaussian kernel for the global bandwidth such that the MISE is minimized:

$$h_{\text{opt}} = 0.9 A n^{-1/5}, \quad \text{where } A = \min\left(\sigma, \frac{\text{IQ}}{1.34}\right). \quad (2.5)$$

In this equation,  $\sigma$  is the standard deviation in the sample dataset and IQ is the interquartile range, or the distance between the 25th and 75th percentiles of the dataset. For a distribution with dimension  $d$ , Eq. (2.5) becomes  $h_{\text{opt}} \propto n^{-1/(d+4)}$  if the variables in each dimension are independent.

However, if the unknown PDF has a narrow peak and long tails, it becomes difficult to find a good global bandwidth. Either the central peak is underestimated or the noise in the long tails dominates. This becomes even more troublesome in high-dimensional density estimations (in particular, if the variables in higher dimensions are not independent), because both the peak area and the tails can be important in different regions of phase space. The AKDE overcomes this difficulty by using local bandwidths according to the density of the data points. In areas where data points are dense, small bandwidths are used whereas larger bandwidths are used in areas where the data is sparse. The algorithm requires an initial estimate of the approximate point density; this estimate yields a pattern of bandwidths corresponding to the various data points and these bandwidths are used to construct the adaptive estimator. This process is accomplished in the steps shown in Sec. 2 of the Appendix. Equation (A12) (Sec. 2 of the Appendix) illustrates the use of a Gaussian kernel for density estimation. The diagonal form of  $\mathbf{H}$  makes Monte Carlo sampling from the multidimensional Gaussian straightforward.

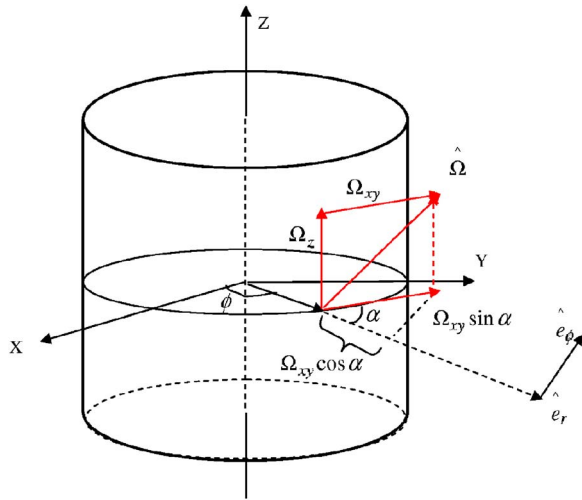


FIG. 1. A schematic showing the cylindrical coordinate geometry used for the preprocessing of the phase space variable.  $\hat{\Omega}$  is the particle direction,  $\Omega_z$  is the z component,  $\Omega_{xy}$  is the x-y (planar) component, and  $\alpha$  is the azimuthal angle of  $\Omega_{xy}$  w.r.t. the radial vector.

## B. Monte Carlo sampling using AKDE

Monte Carlo implementation of AKDE does not require calculating  $\hat{f}(\mathbf{x})$  at each observation point,  $\mathbf{x}$ . Instead, a phase space vector  $\mathbf{x}$  can be sampled using the following algorithm.

- Generate a uniform random number  $\xi$  in the interval  $[0, 1)$ .
- Sample a PS vector index  $i = \xi n + 1$ , where  $n$  is the total number of PS vectors in the original PS (note all PS vectors are equally probable by construction, for our application).
- Sample a new PS vector in the vicinity of  $\mathbf{x}$  by sampling from the multivariate Gaussian distribution as defined by Eq. (A12) (Sec. 2 of the Appendix).

## III. PREPROCESSING OF THE PS VARIABLES

The PS simulation of a Varian 21EX linear accelerator was carried out using the BEAMnrc Monte Carlo code. The simulation included patient-independent structures such as target, primary collimator, flattening filter, ion chamber, and mirror. The phase space variables of all the particles were scored on the plane below the mirror and above the collimating jaws, at a distance of 25 cm downstream from the target. In this study, we focused only on photons in the PS file (i.e., electrons were excluded) in order to illustrate a proof of principle.

The original 5-D PS vector  $(x, y, u, v, E)$  was transformed to a 4-D rotationally invariant PS vector  $(r, \theta, \alpha, E)$  by taking advantage of azimuthal symmetry ( $\phi$ ), as shown in Fig. 1, for the patient-independent simulation. (This assumption would, however, not hold for an asymmetric beam.) This reduces the dimensionality of the PS variables that are sampled by AKDE. This was accomplished using the follow-

ing transformation between the initial (input) and transformed (output) variables: PS input= $(x, y, u, v, E)$ , PS output= $(r, \theta, \alpha, E)$ :

Transformation,

$$\phi = \cos^{-1}\left(\frac{x}{\sqrt{x^2 + y^2}}\right) = \cos^{-1}\left(\frac{x}{r}\right),$$

$$r = \sqrt{x^2 + y^2},$$

$$\cos \alpha = \left(\frac{u \cos \phi + v \sin \phi}{\sin \theta}\right), \quad (3.1)$$

$$\theta = \cos^{-1}(w).$$

Here,  $r$  is the radius vector;  $u, v$ , and  $w$  are the direction cosines with respect to the  $x, y$ , and  $z$  axes, respectively; and  $\alpha$  is the azimuthal angle of the planar component of the particle's direction with respect to the cylindrical radial vector  $r$ . The reason to use angles in the transformed phase space is to take advantage of symmetry. This allowed us to reduce the dimensionality of the phase space, which minimized the sampling of variables with AKDE. The  $(r, \theta, \alpha, E)$  coordinate system is rotationally invariant; hence they can be sampled independent of the azimuthal angle  $\phi$ . The direction variables  $u$  and  $v$ , on the other hand, are fixed in the Cartesian frame, hence are dependent on  $\phi$ , and their use will require more transformations and will result in a smaller population to sample from. When a PS file is used as an input for dose calculation, a data point  $(r, \theta, \alpha, E)$  is sampled uniformly (or in accordance with the particle weights) from the PS. AKDE is then used to sample the new data point  $(r', \theta', \alpha', E')$  in the vicinity of  $(r, \theta, \alpha, E)$ , as described above. The azimuthal angle  $\phi'$  for the position coordinate is sampled uniformly between 0 and  $2\pi$ . The direction cosine vectors for a particle sampled at  $(r', \phi')$  with sampled direction  $(\theta', \alpha')$  are calculated as follows:

$$u' = \sin \theta' \cos \phi' \cos \alpha' - \sin \theta' \sin \phi' \sin \alpha',$$

$$v' = \sin \theta' \sin \phi' \cos \alpha' + \sin \theta' \cos \phi' \sin \alpha',$$

$$w' = \cos \theta'. \quad (3.2)$$

## IV. PHASE SPACE ANALYSIS USING AKDE

In this section we compare the AKDE generated PS with the original PS. AKDE calculations were performed using the transformed PS variables using global bandwidths calculated from Eq. (2.5). Note that although the AKDE method assumes independent and identically distributed samples, the PS entries are not independent due to multiple PS particles from the same primary particle. AKDE can still be applied to all entries (as in recycling) when PS is used as an input. However, to estimate the uncertainty correctly, one needs to tally all entries that correspond to the same history. The bandwidths calculated for these variables are shown in Table I for  $50 \times 10^6$  and  $100 \times 10^6$  particles. Boundary conditions were placed on the boundaries of the  $r$  and  $\theta$  variables to

TABLE I. Global bandwidths for AKDE variables.

	$50 \times 10^6$	$100 \times 10^6$
$r(\text{cm})$	0.238	0.21
$\theta(\text{rad})$	0.012	0.01
$\alpha(\text{rad})$	0.448	0.312
$E(\text{keV})$	134 960	86 943

avoid negative values of these variables. Specifically,  $r$  and  $\theta$  were transformed to  $\log_{10}(r)$  and  $\log_{10}(\theta)$ , respectively, which avoided negative  $r$  and  $\theta$ .

The low energy cutoff on  $E$  and the periodic behavior of  $\alpha$  on  $[0, 2\pi)$  allowed us to avoid the log transformations for these variables. As expected, the bandwidths decrease with increasing sample size for all variables.

We performed an analysis to assess the correlations among the variables in AKDE samples taken from the transformed phase space, in comparison to the correlations among variables in the original phase space. The coefficients were calculated using the covariance  $Cov(x, y)$  and correlation coefficient  $\rho_{xy}$ ,

$$\rho_{x,y} = \frac{Cov(x,y)}{\sigma_x \sigma_y}; \quad -1 \leq \rho_{xy} \leq 1,$$

$$Cov(x,y) = \frac{1}{n} \sum_{j=1}^n (x_j - \mu_x)(y_j - \mu_y). \quad (4.1)$$

The results presented in Table II using 50 000 PS points were analyzed using the statistical analysis toolkit in Excel, and were in some instances, independently verified. AKDE sampled PS variables retain the correlations (particularly the strong correlations), even though the local correlations among the changes in the PS variables are neglected in our approach. This is justified by noting that the first step in AKDE samples the original phase space record, or the actual trajectory, implying that the first-order correlations among the PS variables are preserved. For the variables with weaker correlations, e.g.,  $(r, \alpha)$ ,  $(\alpha, E)$ ,  $(\theta, \alpha)$ , the correlations with AKDE-generated variables are higher than those from the PS. Although this is a small effect, it is being further investigated. At this stage we believe that it may be due to systematic bias introduced by the finite bandwidth used with AKDE. The overall behavior shows that  $r$  and  $\theta$  are strongly correlated due to the forward bremsstrahlung production

TABLE II. Correlation coefficients from the original PS distribution and KDE transformed PS distribution.

PS variables	Original PS	KDE transformed PS
$(r, \theta)$	0.942	0.94
$(r, \alpha)$	-0.002	-0.013
$(r, E)$	-0.252	-0.252
$(\theta, \alpha)$	-0.002	-0.038
$(\theta, E)$	-0.292	-0.290
$(\alpha, E)$	0.0002	0.025

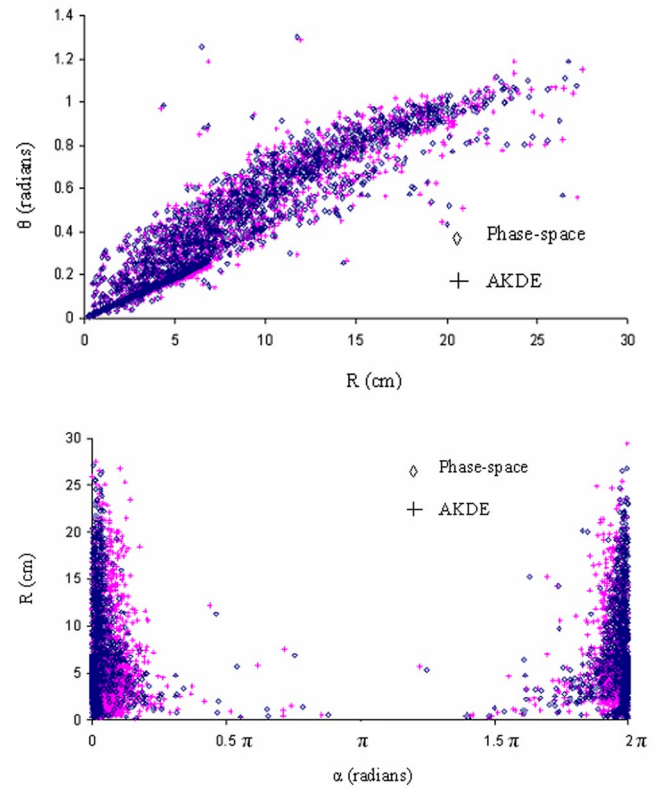


Fig. 2. Overlay of scatter plots of various PS parameters: (a) radius vector ( $r$ ) versus polar angle ( $\theta$ ) and (b) Azimuthal angle of the planar component of the particle's direction ( $\alpha$ ) versus radius vector ( $r$ ).

from the tungsten target. This correlation decreases for larger  $r$  due to increased multiple scattering. A weaker correlation is seen between  $E$  and  $\theta$  due to the inverse relation between angle of scattering and energy loss. The correlation between  $r$  and  $E$  is consistent with the observed correlations between  $r$  and  $\theta$  and  $E$  and  $\theta$ . The variable  $\alpha$  is weakly correlated to the other variables. Figures 2(a) and 2(b) show scatter plots for pairs of transformed variables  $(r, \theta)$  and  $(\alpha, r)$  for the original PS and those generated using AKDE. The above correlations are almost identical for the original PS and the AKDE generated samples, even for  $(\alpha, r)$  that shows the maximum deviation from the original PS distribution.

Photon fluence distributions as a function of radius outward from the central axis, as computed by the AKDE generated PS and the original PS distribution, are shown in Fig. 3(a). Overall fluence differences are less than 1% with local differences of up to  $\pm 9\%$  seen in the range 6–7 cm, corresponding to the edge of the primary collimator at the phase space plane. The finite width of the kernels has the greatest effect on the high gradient regions of the density distribution because AKDE convolves the data point with a finite bandwidth kernel. The angular distribution [Fig. 3(b)] and the energy distribution [Fig. 3(c)] calculated in a 1 cm radius around the central axis show an average deviation of 0.1% (range=0% to 3.0%) and 1.5% (range=-1.5% to 6.6%), respectively, in comparison to the original PS.

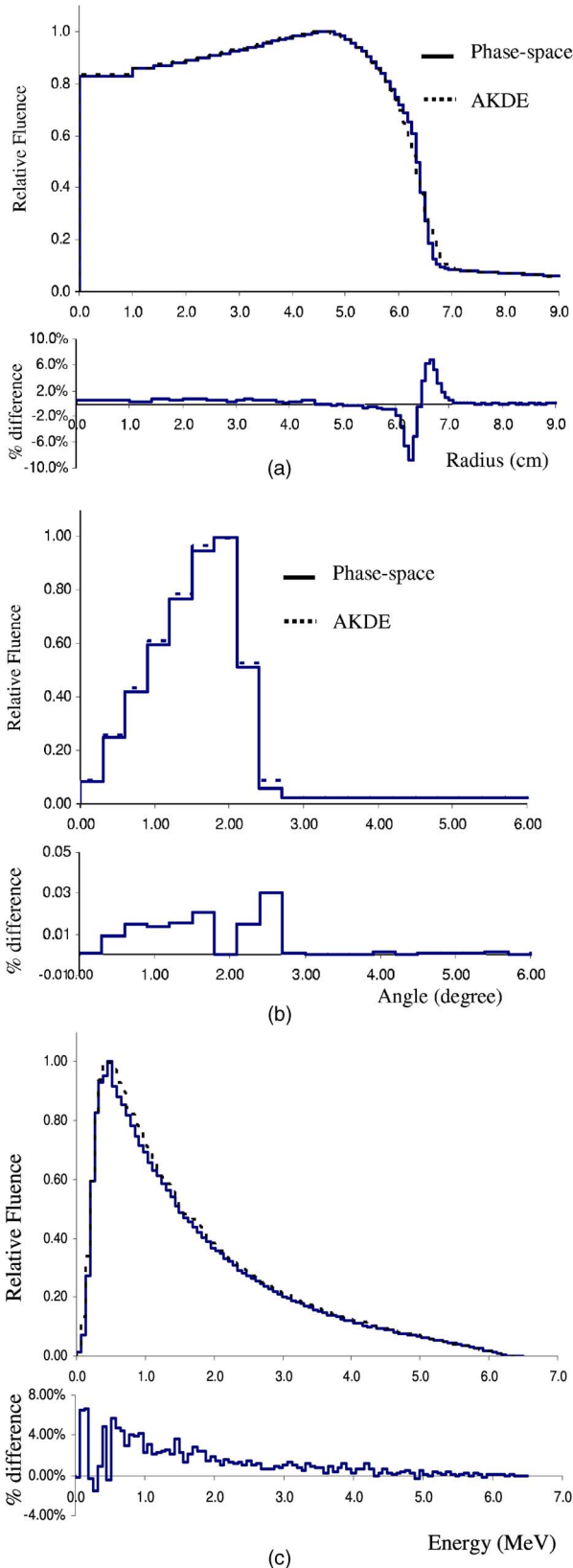


FIG. 3. In-air fluence analysis for the original PS distribution and AKDE generated distributions. Angular distributions and energy distributions were generated using a 1 cm radius around the central axis. (a) Relative fluence versus radii (cm). (b) Central axis angular distribution as a function of angle (degrees). (c) Central axis energy distribution as a function of energy (MeV).

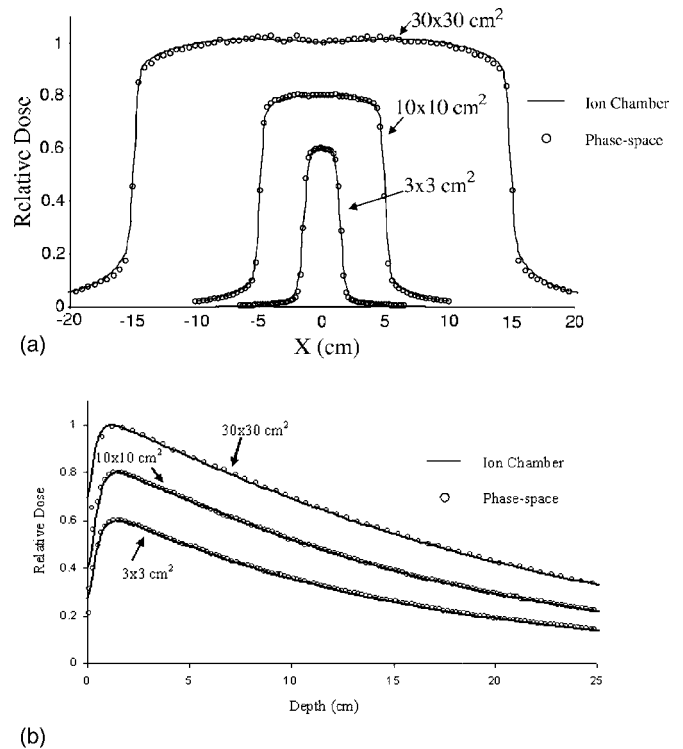


FIG. 4. A comparison between benchmark calculation and ion-chamber (CC13) data for three field sizes  $3 \times 3 \text{ cm}^2$ ,  $10 \times 10 \text{ cm}^2$ , and  $30 \times 30 \text{ cm}^2$ : (a) Transverse profiles. (b) Depth dose profiles. The relative uncertainty in the high dose region for the Monte Carlo calculation was less than 1% in each voxel.

## V. DOSE CALCULATION USING AKDE

AKDE was applied to PS distributions for in-phantom dose calculations. Calculation accuracy was assessed for the following field sizes:  $3 \times 3 \text{ cm}^2$ ,  $10 \times 10 \text{ cm}^2$ , and  $30 \times 30 \text{ cm}^2$ . Profiles and central axis depth doses were calculated in a phantom of size  $25.6 \times 30 \times 30 \text{ cm}^3$  ( $3 \times 3 \text{ cm}^2$ ),  $30 \times 30 \times 30 \text{ cm}^3$  ( $10 \times 10 \text{ cm}^2$ ), and  $40 \times 40 \times 40 \text{ cm}^3$  ( $30 \times 30 \text{ cm}^2$ ) at 10 cm depth and 90 cm SSD for a 6 MV photon beam. The voxel sizes used are 2, 3, and 5 mm along the profile directions for  $3 \times 3 \text{ cm}^2$ ,  $10 \times 10 \text{ cm}^2$ , and  $30 \times 30 \text{ cm}^2$ , respectively. A slightly larger voxel size (up to 1 cm) was used along other directions. The average uncertainty ( $1\sigma$ ) in all the voxels greater than  $\text{Dose}_{\text{max}}/2$  was less than 1.5%. All calculations of dose in phantom were carried out using the Dose Planning Method (DPM) Monte Carlo code,<sup>5,6</sup> which reads the patient-independent PS as input and includes the field-defining jaws and MLC for the phantom calculation.

A benchmark result was calculated using  $200 \times 10^6$  particles in the patient-independent PS. The PS file was recycled 100 times (for the  $30 \times 30 \text{ cm}^2$  field) to substantially reduce the statistical fluctuation in the calculated dose. Agreement with ion-chamber (CC13, active volume  $0.13 \text{ cm}^3$ ) measurements are within 2%/2 mm, as shown in Figs. 4(a) and 4(b) for the three field sizes. Figures 5(a)–5(c) show transverse profiles for the AKDE, PS recycle, and benchmark calculation for the three field sizes along with their percentage dose

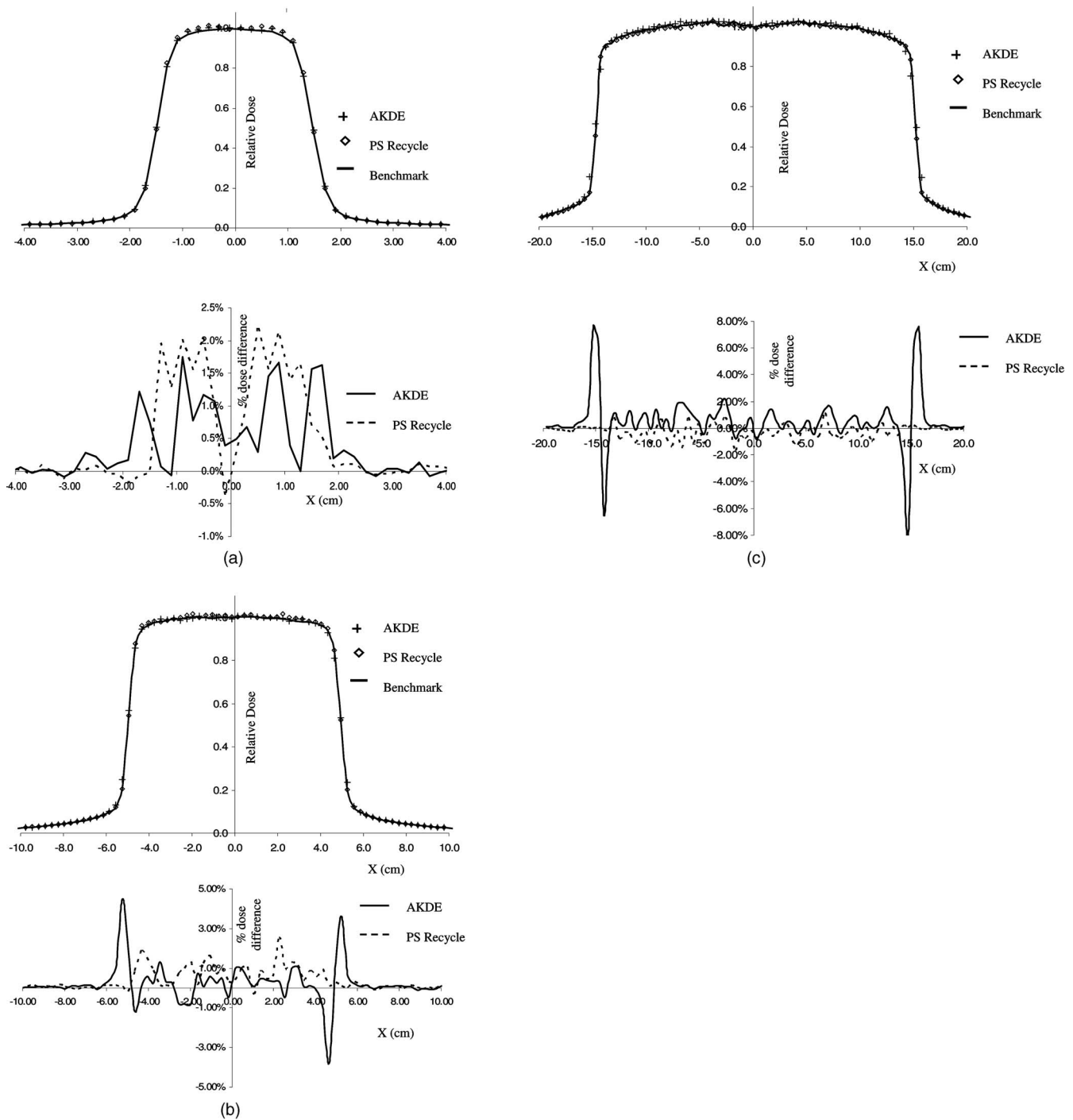


Fig. 5. Transverse profiles showing PS recycling, AKDE, and a benchmark calculation. % dose differences for PS recycling and AKDE with respect to the benchmark calculation has been shown for three field sizes: (a)  $3 \times 3 \text{ cm}^2$ , (b)  $10 \times 10 \text{ cm}^2$ , and (c)  $30 \times 30 \text{ cm}^2$ . The relative uncertainties in the high dose region for both AKDE and PS calculations were less than 1.2% ( $3 \times 3 \text{ cm}^2$ ), 0.9% ( $10 \times 10 \text{ cm}^2$ ), and 0.8% ( $30 \times 30 \text{ cm}^2$ ).

differences. The PS recycle and AKDE calculations used  $100 \times 10^6$  particles as input in the patient-independent PS. The PS calculation was recycled up to 100 times and AKDE was used to generate up to 100 times more samples (starting with  $100 \times 10^6$  samples in the original PS) for these field sizes. The agreement between AKDE and PS recycle with respect to the benchmark result is within 2% and 2.5%, respectively, in the high dose region for all the three field sizes. Larger differences are seen for AKDE in the penumbral re-

gion, where the differences are 5% and 8% for  $10 \times 10 \text{ cm}^2$  and  $30 \times 30 \text{ cm}^2$ , respectively, however, these differences are well within 1 mm distance to agreement. The percentage dose difference plot for a  $10 \times 10 \text{ cm}^2$  field size shows that PS recycle may introduce some level of systematic bias when recycled many times. This is also the case for both AKDE and PS recycle for the  $3 \times 3 \text{ cm}^2$  field size. Central axis depth dose curves were also extracted for these field sizes, as shown in Fig. 6(a). Figures 6(b) and 6(c) show

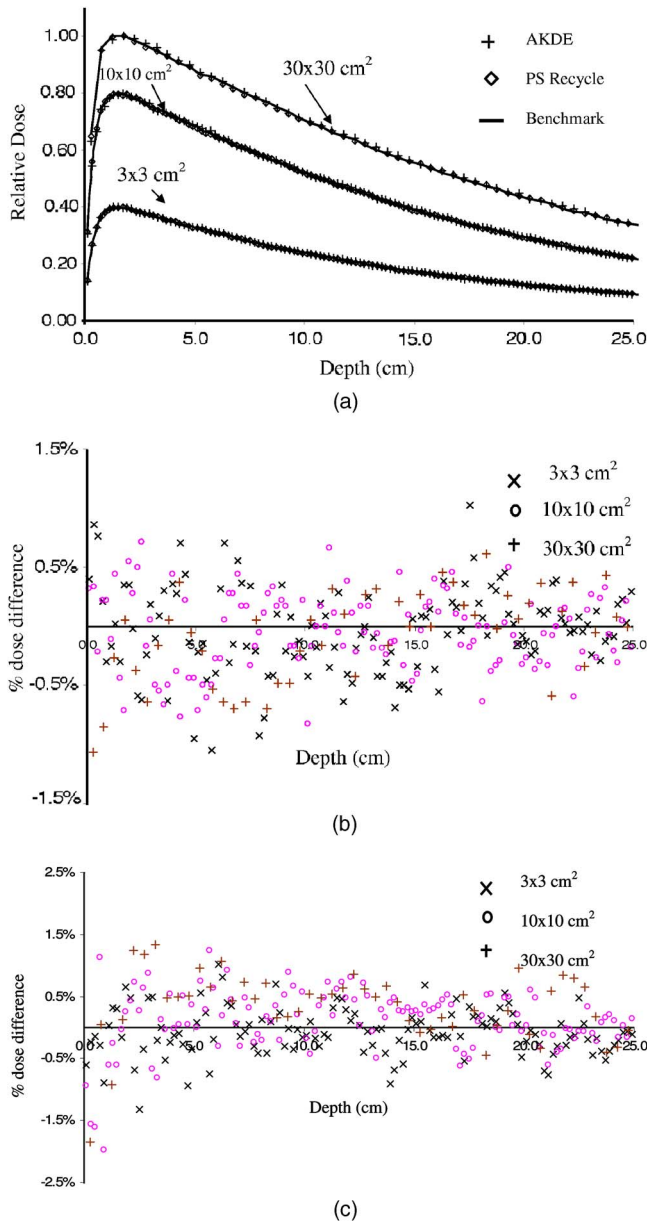


FIG. 6. Depth dose comparisons for PS recycling and AKDE with respect to the benchmark calculation. (a) The relative dose as a function of depth for a  $3 \times 3 \text{ cm}^2$ ,  $10 \times 10 \text{ cm}^2$ , and  $30 \times 30 \text{ cm}^2$  field size. (b) % dose difference between the PS recycled run and the benchmark for a  $3 \times 3 \text{ cm}^2$ ,  $10 \times 10 \text{ cm}^2$ , and  $30 \times 30 \text{ cm}^2$  field size. (c) % dose difference between AKDE and the benchmark for a  $3 \times 3 \text{ cm}^2$ ,  $10 \times 10 \text{ cm}^2$ , and  $30 \times 30 \text{ cm}^2$  field size.

percentage dose differences for PS recycle and AKDE with respect to the benchmark calculation. The agreement for both PS recycle and AKDE is at a maximum within 2%.

## VI. CONCLUSION AND DISCUSSION

In this study we have implemented a nonparametric density estimation technique, the adaptive kernel density estimator (AKDE) method, and have applied the method to Monte Carlo-based treatment head simulation. AKDE was compared to the original PS distribution by performing a correlation analysis among the PS variables. Analysis shows that although the variables have been sampled independently,

first-order correlations among the phase space variables are preserved. The in-air photon fluence distribution shows agreement to within 2% (−8.8% to 6.8%) in comparison with the original PS distribution, while central axis energy distributions and angular distributions show agreement to within 1.5% (range=−1.5% to 6.6%) and 0.1% (range=0 to 3.0%), respectively. Dose profiles were also calculated for field sizes:  $3 \times 3 \text{ cm}^2$ ,  $10 \times 10 \text{ cm}^2$  and  $30 \times 30 \text{ cm}^2$  for AKDE and compared with PS recycle and benchmark calculations. AKDE shows agreement to within 2% / 1 mm with respect to the benchmark calculation. The AKDE method generates samples “on the fly” with minimal time and memory requirements, and maintains correlations between the various PS variables. As such, this method may provide an attractive alternative to PS recycling.

AKDE may introduce some systematic bias if the initial sample size is small, just as with PS recycle. The minimum number needed in the phase space for AKDE calculations would depend on the specific application. For the purpose of this study, we have assumed that the minimum number needed for AKDE is the same as needed for conventional phase space recycling. In principle, the AKDE method may require fewer phase space particles, but this has not been examined for this paper. Similarly, there is no theoretical limit to the upper bound of the number of AKDE samples, but again we are taking a conservative stance and keeping the upper bound consistent with the number of recycled histories in order to compare the two methods. In this work all dose calculations were performed using a fixed global bandwidth for each phase space variable. The use of local bandwidths has not been explored here; however, we anticipate that local bandwidths are likely to improve further the accuracy of AKDE.

## ACKNOWLEDGMENTS

This work is supported in part by Grants No. NIH R01 CA106770 and No. P01-CA59827. The authors would like to thank the anonymous reviewers for their extremely useful comments and suggestions that have strengthened the quality of this paper.

## APPENDIX

### 1. MISE and optimum bandwidth for univariate kernel density estimators

Given a set of observations  $x_1, x_2, \dots, x_n$ , the univariate kernel density estimator is written as

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right), \quad (\text{A1})$$

and its corresponding MISE is defined by

$$\text{MISE}(\hat{f}) = E \int_{-\infty}^{+\infty} [\hat{f}(x) - f(x)]^2 dx, \quad (\text{A2})$$

where  $E(x)$  is the expectation value of  $x$ .

Since the integrand is non-negative, the order of integration and expectation can be reversed. The optimum bandwidth is determined by minimizing the MISE. Using the definition of mean and variance, Eq. (A2) reduces to

$$\begin{aligned} \text{MISE}(\hat{f}) &= \int_{-\infty}^{+\infty} [E\hat{f}(x) - f(x)]^2 dx + \int_{-\infty}^{+\infty} \text{var} \hat{f}(x) dx \\ &= \int_{-\infty}^{+\infty} \text{bias}^2(x) dx + \int_{-\infty}^{+\infty} \text{var} \hat{f}(x) dx. \end{aligned} \quad (\text{A3})$$

Using the expression of  $\hat{f}(x)$ , the expectation and variance can be written as

$$E\hat{f}(x) = \int_{-\infty}^{+\infty} \frac{1}{h} K\left(\frac{x-y}{h}\right) f(y) dy$$

and

$$\begin{aligned} \text{var} \hat{f}(x) &= \frac{1}{n} \int_{-\infty}^{+\infty} \frac{1}{h^2} K\left(\frac{x-y}{h}\right)^2 f(y) dy \\ &\quad - \frac{1}{n} \left[ \int_{-\infty}^{+\infty} \frac{1}{h} K\left(\frac{x-y}{h}\right) f(y) dy \right]^2, \end{aligned} \quad (\text{A4})$$

Substituting the above expressions in Eq. (A3) results in an expression for MISE that still cannot be minimized since  $h$  is inside the kernel function. However, assuming that the unknown density  $f(x)$  has continuous derivatives of all orders and the kernel  $K(x)$  is narrow enough, then along with Eq. (2.2) the optimal  $h$  can be evaluated by making a linear transformation  $y=x-h$  and expanding  $f(x-h)$  using Taylor series. The expressions for bias and variance reduce to

$$\text{bias}(x) = \frac{1}{2} h^2 f''(x) k_2 + O(h^3) \approx \frac{1}{2} h^2 f''(x) k_2 \quad (\text{A5})$$

and

$$\begin{aligned} \text{var} \hat{f}(x) &= \frac{1}{nh} f(x) \int_{-\infty}^{+\infty} K(t)^2 dt + O(n^{-1}) \\ &\approx \frac{1}{nh} f(x) \int_{-\infty}^{+\infty} K(t)^2 dt. \end{aligned} \quad (\text{A6})$$

Thus the expression for MISE becomes

$$\text{MISE}(\hat{f}) \approx \frac{1}{4} h^4 k_2^2 \int_{-\infty}^{+\infty} f''(x)^2 dx + \frac{1}{nh} \int_{-\infty}^{+\infty} K(t)^2 dt. \quad (\text{A7})$$

Minimizing the above equation gives an expression for optimum bandwidth as

$$h_{\text{opt}} = k_2^{-2/5} \left[ \int_{-\infty}^{+\infty} K(t)^2 dt \right]^{1/5} \left[ \int_{-\infty}^{+\infty} f''(x)^2 dx \right]^{-1/5} n^{-1/5}. \quad (\text{A8})$$

## 2. Formulation of local bandwidths for the adaptive kernel estimator

(1) Find a pilot estimate  $\tilde{f}(\mathbf{t})$  that satisfies  $\hat{f}(\mathbf{x}_i) > 0$  for all  $i$ :

$$\tilde{f}(\mathbf{t}) = \frac{1}{n} \sum_{i=1}^n K_{\mathbf{H}}(\mathbf{t} - \mathbf{x}_i). \quad (\text{A9})$$

(2) Define local bandwidth factors  $\lambda_i$  by

$$\lambda_i = \left\{ \frac{\hat{f}(\mathbf{x}_i)}{g} \right\}^{-\alpha}, \quad (\text{A10})$$

where  $g$  is the geometric mean of  $\hat{f}(\mathbf{x}_i)$ ,

$$\log g = n^{-1} \sum_{i=1}^n \log \hat{f}(\mathbf{x}_i) \quad (\text{A11})$$

and  $\alpha$  is the sensitivity parameter that is  $1/d$  for  $d > 1$  and  $1/2$  for  $d = 1$ , where  $d$  describes the dimension. It allows the local bandwidth factors to depend on a power of the pilot density and hence gives more flexibility to the method.

Then one can define the adaptive kernel estimator as follows:

$$\hat{f}(\mathbf{t}) = \frac{1}{n} \sum_{i=1}^n K_{\mathbf{H}}(\mathbf{t} - \mathbf{x}_i),$$

$$K_{\mathbf{H}}(\mathbf{x}_i) = \frac{1}{(2\pi)^{d/2} |\mathbf{H}|^{1/2}} \exp\left(-\frac{1}{2} \mathbf{x}_i^T \mathbf{H}^{-1} \mathbf{x}_i\right),$$

$$\mathbf{H} = \{\lambda \text{diag}(h_1^2, h_2^2, \dots, h_d^2), h_1, h_2, \dots, h_d > 0\}. \quad (\text{A12})$$

<sup>a)</sup>Electronic mail: indrin@med.umich.edu

<sup>1</sup>D. W. O. Rogers, B. Walters, and I. Kawrakow, BEAMnrc Users Manual: NRC Report PIRS 509(a) revH, 2004.

<sup>2</sup>B. R. B. Walters, I. Kawrakow, and D. W. O. Rogers, "History by history statistical estimators in the BEAM code system," Med. Phys. **29**, 2745–2752 (2002).

<sup>3</sup>B. W. Silverman, in *Density Estimation for Statistics and Data Analysis* (Chapman and Hall, New York, 1986).

<sup>4</sup>J. Du and W. R. Martin, "Adaptive Kernel Density Estimation and Monte Carlo sampling," Trans. Am. Nucl. Soc. **74**, 171–172 (1996).

<sup>5</sup>J. Sempau, S. J. Wilderman, and A. F. Bielajew, "DPM, a fast, accurate Monte Carlo code optimized for photon and electron radiotherapy treatment planning dose calculations," Phys. Med. Biol. **45**, 2263–2291 (2000).

<sup>6</sup>I. J. Chetty, P. M. Charland, N. Tyagi, D. L. McShan, B. A. Frass, and A. F. Bielajew, "Photon beam relative dose validation of the DPM Monte Carlo code in lung-equivalent media," Med. Phys. **30**, 563–573 (2003).