An Extension of Alias Sampling Method for Parametrized Probability Distributions

Lucreţiu M. Popescu

Institute of Environmental Research and Engineering
294 Spl. Independenţei, 77703 Bucharest, Romania
email: lmp@mailbox.ro

March 30, 1999
Revised November 2, 1999

Keywords: Monte Carlo, random variable, sampling, probability, Molière, multi-scatter
MSC:

Running head: Alias sampling method for parametrized PDFs

Corresponding author:

Lucreţiu M. Popescu

Environmental Radioactivity Laboratory

Institute of Environmental Research and Engineering

294 Spl. Independenţei, 77703 Bucharest, Romania

telephone: (40-1) 240 26 95, facsimile: (40-1) 221 92 04

email: lmp@mailbox.ro

Abstract

A general method for sampling random variables with parametrized probability distributions was developed. It takes advantage of modern computer architectures with large amounts of cheap memory, by using discrete representations of probability distribution functions. The sampling is done by fast interpolation techniques involving only elementary logical and arithmetical operations, allowing to keep a higher degree of accuracy as the grids spacing is controlled by the user. It is based on an extension of the alias sampling technique for distribution functions depending on a number of parameters. By this method it is possible to obtain the value of interest by a direct interpolation between the sampled values obtained with the same set of random numbers for the grid values of the parameters adjacent to the values of interest. Sampling tests carried for the case of Molière electron multi-scatter angle distribution show that this method can successfully replace the sampling arrangements commonly used in Monte Carlo codes, where, for complex probability distributions, after a careful study of functions proprieties, a combination of sampling techniques, like superposition, rejection and inverse function methods are used.

1. Introduction

In a Monte Carlo code for simulating physical processes, as a particle transport simulation system, a significant part of the running time is spent for generating random variables according to different probability density functions (PDFs). Therefore for developing such codes it is a necessity to have methods which provide fast algorithms for generating random variables according to the various PDFs given by the laws describing the physical processes involved. As the shape of probability distributions and also the form of presentation (mathematical formula, tabulated data, or algorithmic description) can be very different from case to case, for rapid development of simulation codes is very useful to have a general method for sampling random variables which can treat all these cases in an uniform manner.

2. Sampling a random variable

A sampling method suitable, theoretically, for any continuous random variable is the rejection method proposed in its original form by von Newman [1, 2]. In this method for sampling a random variable $x \in (a, b)$ with probability distribution $p(x)$, first two uniformly distributed numbers are sampled, $\hat{x} \in (a, b)$ and $\hat{y} \in (0, \max_{\hat{x}\in(a,b)} p(\hat{x}))$. The $\hat{x}$ value is accepted if $\hat{y} \leq p(\hat{x})$ and rejected otherwise. Random sampling is repeated until a point $\hat{y}$ under $p(\hat{x})$ is selected. The disadvantage of this method is that it can become extremely time consuming if the distribution is peaked and narrow.

Theoretically any random variable $x \in (a, b)$ with probability density $p(x)$, satisfying $1 = \int_a^b p(x)\mathrm{d}x$ and cumulative probability distribution:

$$F(x) = \int_a^x p(x')\mathrm{d}x' \,, \tag{1}$$

can be sampled using inverse function method or direct method, which gives

$$\hat{x} = F^{-1}(\zeta) \,, \tag{2}$$

where ζ is a random number uniformly distributed in the interval $(0,1)$. But to apply this method it is necessary to have an analytical formula for the inverse function $F^{-1}:(0,1)\to(a,b)$ which is possible only in few particular cases. This disadvantage can be overcame by using the equally probable intervals spacing method. In this method a discretization in N points of F^{-1} over a equally spaced division of interval $(0,1)$ is taken resulting in a division of the random variable domain (a,b) into equally probable intervals delimited by the points $\{x_i\}_{i=\overline{0,N}}$, where $x_i = F^{-1}(\frac{i}{N})$. For sampling, the interval i is selected first as the integer part of $[\zeta N]$, where ζ is a random number uniformly distributed in $(0,1)$, afterwards a value $\hat{x}$ in the selected interval (x_i, x_{i+1}) is sampled. As a first approximation a uniformly distributed number

$$\hat{x} = (1-t)x_i + tx_{i+1} \tag{3}$$

where t is the fractional part of $\{\zeta N\}$, can be used. A higher degree of accuracy is achieved if the linear interpolation formula of probability distribution over the selected interval is used:

$$f_i(x) = (1-u)p_i + up_{i+1} \tag{4}$$

where $u = \frac{x-x_i}{x_{i+1}-x_i}$, and $p_i = p(x_i)$. An elegant method using only elementary operations is described in [3]. First the uniformly distributed value $\hat{x}$ is selected:

$$\hat{x} = tx_i + (1-t)x_{i+1} \tag{5}$$

Then a second random number ζ uniformly distributed in the interval $(0, p_i+p_{i+1})$ is selected. If $f_i(\hat{x}) > \zeta$ the value $\hat{x}$ is accepted, otherwise the following value is chosen:

$$\hat{x} = (1-t)x_i + tx_{i+1}\,. \tag{6}$$

In fact this is a modified rejection method exploiting the symmetry properties of linear distributions.

The method of equally probable intervals spacing is fast, as only few elementary operations are involved, but in the case of intervals where the events are less probable the spacing is larger and the accuracy decreases; even more, in the case of distributions given as tabulated data the transformations of the original grid to a grid with equally probable intervals may lead to an avoidable loss of precision. If the points spacing does not correspond to equally probable intervals then the selection of the interval requires to search a table of cumulative probabilities, operation which is very time consuming.

2.1. Alias sampling method

An alternative to the search operation is the use of alias sampling method [4, 1]. Suppose we have an arbitrary division $\{x_i\}_{i=\overline{0,N}}$ of random variable domain (a, b) and consider as elementary events the selection of one of the intervals (x_i, x_{i+1}). In alias sampling method the above elementary events are grouped in pairs of two in order to obtain a table of equally probable compound events. The selection is done in two steps, first, as in the method of equally probable intervals such a compound event is sampled and second, from the pair elementary events one of them is selected, according to their relative probabilities. For a PDF given in tabulated form $\{x_i, p_i\}_{i=\overline{0,N}}$ where p_i is the density probability in the point x_i, such a discretization divides the whole domain into N intervals (x_i, x_{i+1}), $i = \overline{0, N-1}$ each with the probability of selection

$$P_i = \frac{p_i + p_{i+1}}{2}(x_{i+1} - x_i), \text{ satisfying: } 1 = \sum_{i=0}^{N-1} P_i \,. \tag{7}$$

The compound events are stored into a special data arrangement $\{j_i, r_i\}_{i=\overline{0,N-1}}$ into which to each interval i is associated a second interval j_i and a real number r_i. For selecting the interval, first a compound event i is selected with equal probabilities, then a second random number $\zeta \in (0, 1)$ is used to decide whether the associated interval j_i (if $\zeta > r_i$) or the interval i (if $\zeta \leq r_i$) is selected. An elegant method for making such arrangement of data is the Walker algorithm [4].

Successively the array $\{P_k\}_{k=\overline{0,N-1}}$ is searched for the intervals with the minimum (P_i) and respectively maximum (P_j) probabilities; in fact is sufficient to search for pair of intervals were $P_i < P_{ave} = 1/N$ and $P_j > P_{ave}$. The compound event i is formed from interval with the same index i and probability P_i to which the interval $j_i = j$ with the probability $1/N - P_i$ is added. In this way each compound event has the probability $1/N$ and $r_i = P_i N$. The probability of interval j, P_j is updated by subtracting the quantity added to the compound event $P_j \leftarrow P_j - (1/N - P_i)$. A new search is performed, but excluding the interval i, just prepared, from the search array $\{P_k\}$, and so on, until it becomes empty. By mathematical induction it can be proved that this arrangement is always possible.

3. The case of parametrized PDFs

3.1. Direct interpolation

Let's consider first the case of a random variable x with the PDF depending on a single parameter $p(x, \alpha)$ where α is a real number taking values in the interval $[\alpha_{\min}, \alpha_{\max}]$. In this case a grid with a convenient spacing, linear or logarithmic, is created $\{\alpha_k\}_{k=\overline{0,N_\alpha}}$. For techniques like the inverse function method, or equally probable intervals spacing method, a sample at a given point $\alpha \in [\alpha_i, \alpha_{i+1})$ can be determined by interpolating samples taken from distributions in the points α_i and α_{i+1} using the same random numbers.

Briefly the method is: using the same random numbers, the values, $\hat{x}_1$, according to distribution $p(x, \alpha_k)$ and $\hat{x}_2$, according to distribution $p(x, \alpha_{k+1})$, are selected. The sampled value for the distribution $p(x, \alpha)$ is finally given by:

$$\hat{x} = (1-u)\hat{x}_1 + u\hat{x}_2\,, \text{ where: } u = \frac{\alpha - \alpha_k}{\alpha_{k+1} - \alpha_k}\,,\ u' = 1 - u\,. \tag{8}$$

By this procedure the random variabile is obtained by interpolation between

the inverse of the cumulative distribution functions

$$\tilde{F}^{-1}(y,\alpha) = u'F^{-1}(y,\alpha_k) + uF^{-1}(y,\alpha_{k+1})\,. \tag{9}$$

And if for α_k the random variabile takes values inside the interval (a_k, b_k), and for α_{k+1} it takes values inside the interval (a_{k+1}, b_{k+1}), then the interpolated random variabile will take values inside the interval $(u'a_k + ua_{k+1}, u'b_k + ub_{k+1})$, being obtained also an adjustment of the domain in which the interpolated variable lies.

This kind of interpolation fails if it is applied directly with the alias sampling method. Therefore it was considered that the major difficulty with the alias sampling technique is that conventional interpolation between tabulated distribution is not possible [3] since the specially formatted distribution no longer allows direct manipulation of the probabilities.

3.2. Statistical interpolation

An interpolation technique which can be used in conjunction with alias sampling method is considered to be the statistical interpolation. It gives a way of approximating an interpolated distribution by sampling from one of the two neighboring distributions. Given two known distributions at α_k and α_{k+1} and a point $\alpha \in [\alpha_k, \alpha_{k+1})$, choosing a random number $\zeta \in (0,1)$, statistical interpolation calls for sampling from the α_k distribution if:

$$\zeta \leq \frac{\alpha_{k+1} - \alpha}{\alpha_{k+1} - \alpha_k} \tag{10}$$

otherwise the α_{k+1} distribution is used. In fact the statistical interpolation overlaps, with different weights, the PDFs at the extreme points of the interval. The PDF of the random variable by this procedure is

$$\tilde{p}(x,\alpha) = u'p(x,\alpha_k) + up(x,\alpha_{k+1})\,, \tag{11}$$

where u and u' have the same meaning as in (8). The values domain of the interpolated random variable is then $(a_k, b_k) \cup (a_{k+1}, b_{k+1})$.

Due its simplicity statistical interpolation with alias sampling is very convenient to be used with PDFs given in tabulated form. According to [3, 5] the statistical interpolation yields satisfactory results for various PDFs used for Monte Carlo simulation of particle transport, but in certain cases, as that presented in figure 1, the simple weighted superposition is not recommended. In such cases additional re-scalings or application of variable transformations must be considered in order to improve the results.

3.3. Extension of alias sampling technique to direct interpolation

The impossibility to apply the alias sampling technique with conventional (direct) interpolation is due to the attempt to use independently the arrangements for interval selection at each parameter value α_k and α_{k+1}. This inconvenience can be avoided if for both values α_k and α_{k+1} an unique arrangement for sampling the intervals is used, only the values at the extreme points of the intervals being different.

More precisely, if at the start, for each value α_k of the parameter we have a division of the random variable domain into the points $\{x_i^{(k,1)}\}$, $i = \overline{0,N}$ with corresponding probability densities $\{p_i^{(k,1)}\}$, and with the alias sampling arrangement $\{j_i^{(k)}, r_i^{(k)}\}_{i=\overline{0,N-1}}$, for each next value of the parameter α_{k+1} an additional division $\{x_i^{(k,2)}\}$ is created with the condition that the probabilities for corresponding intervals to be the same as for the grid $\{x_i^{(k,1)}\}$ at the parameter value α_k.

If for α_k the values of the cumulative distribution function are:

$$Q_i = F(x_i^{(k,1)}, \alpha_k)\,, \quad i = \overline{0,N} \tag{12}$$

Then for α_{k+1} the points $x_i^{(k,2)}$ are given by:

$$x_i^{(k,2)} = F^{-1}(Q_i, \alpha_{k+1})\,, \quad i = \overline{0,N} \tag{13}$$

Finally the values $p_i^{(k,2)} = p(x_i^{(k,2)}, \alpha_{k+1})$ are computed.

By this procedure the alias sampling arrangement obtained with the Walker algorithm for the division $\{x_i^{(k,1)}\}$ for α_k, is identical with the corresponding arrangement obtained for the division $\{x_i^{(k,2)}\}$ for α_{k+1}. A graphical representation of this kind of arrangement is given in figure 2.

For a parameter point $\alpha \in [\alpha_k, \alpha_{k+1})$ the sampling is done as follows:

- Using a first random number an interval i according to alias sampling arrangement $\{j_i^{(k)}, r_i^{(k)}\}$ prepared for α_k is selected.

- Then using a second sequence of random numbers the following values are sampled: $\hat{x}_1$ for α_k, according to the interpolated PDF between points $(x_i^{(k,1)}, p_i^{(k,1)})$ and $(x_{i+1}^{(k,1)}, p_{i+1}^{(k,1)})$, and $\hat{x}_2$, for α_{k+1}, according to the interpolated PDF between points $(x_i^{(k,2)}, p_i^{(k,2)})$ and $(x_{i+1}^{(k,2)}, p_{i+1}^{(k,2)})$.

- Finally the value of $\hat{x}$ is obtained by interpolation between points $(\hat{x}_1, \alpha_k)$ and $(\hat{x}_2, \alpha_{k+1})$ using the equation (8).

In the case when the random variable PDF depends on two parameters $p(x, \alpha, \beta)$, then it is necessary to interpolate between the points of a two-dimensional grid $\{\alpha_k, \beta_l\}$, $k = \overline{0, N_\alpha}$, $l = \overline{0, N_\beta}$. In this case if for each grid point (α_k, β_l) with $k < N_\alpha$, $l < N_\beta$ is given a division of random variable domain $\{x_i^{(1)}\}_{i=\overline{0,N}}$ with PDF values $\{p_i^{(1)}\}_{i=\overline{0,N}}$, and with alias sampling arrangement $\{j_i, r_i\}_{i=\overline{0,N-1}}$ using analogue relations with (12), (13) three additional divisions are constructed $\{x_i^{(2)}\}$, $\{p_i^{(2)}\}$ for parameters point (α_{k+1}, β_l), $\{x_i^{(3)}\}$, $\{p_i^{(3)}\}$ for (α_k, β_{l+1}) and $\{x_i^{(4)}\}$, $\{p_i^{(4)}\}$ for $(\alpha_{k+1}, \beta_{l+1})$ in order to obtain intervals with probabilities identically with those given by the division $\{x_i^{(1)}\}$ at the parameter point (α_k, β_l). Here for a more fluent reading the k, l superscripts were skipped.

The selection method for the parameters pair $\alpha \in [\alpha_k, \alpha_{k+1})$ and $\beta \in [\beta_l, \beta_{l+1})$ is done by following steps:

- A first random number is used for sampling the interval i according to alias

sampling arrangement $\{j_i, r_i\}$ obtained for the grid $\{x_i^{(1)}\}$ at the parameter point (α_k, β_l).

- Then using the same sequence of random numbers the following values are sampled: $\hat{x}_1$ according to the interpolated PDF between points $(x_i^{(1)}, p_i^{(1)})$ and $(x_{i+1}^{(1)}, p_{i+1}^{(1)})$, $\hat{x}_2$ according to the interpolated PDF between points $(x_i^{(2)}, p_i^{(2)})$ and $(x_{i+1}^{(2)}, p_{i+1}^{(2)})$, $\hat{x}_3$ according to the interpolated PDF between points $(x_i^{(3)}, p_i^{(3)})$ and $(x_{i+1}^{(3)}, p_{i+1}^{(3)})$, and finally $\hat{x}_4$ according to the interpolated PDF between points $(x_i^{(4)}, p_i^{(4)})$ and $(x_{i+1}^{(4)}, p_{i+1}^{(4)})$.

- Finally $\hat{x}$ is obtained interpolating between the grid points $(\hat{x}_1, \alpha_k, \beta_l)$, $(\hat{x}_2, \alpha_{k+1}, \beta_l)$, $(\hat{x}_3, \alpha_k, \beta_{l+1})$ and $(\hat{x}_4, \alpha_{k+1}, \beta_{l+1})$ using the equation

$$\hat{x} = u'v'\hat{x}_1 + uv'\hat{x}_2 + u'v\hat{x}_3 + uv\hat{x}_4 \tag{14}$$

where:

$$u = \frac{\alpha - \alpha_k}{\alpha_{k+1} - \alpha_k}\,, \quad u' = 1 - u\,,$$
$$v = \frac{\beta - \beta_l}{\beta_{l+1} - \beta_l}\,, \quad v' = 1 - v\,.$$

These algorithms based on the extension of the alias sampling method for random sampling of parametrized probability distributions, were coded using the C++ object oriented programing language. Two class hierarchies were defined, the first one is introduced to provide an uniform interface to the various probability distributions given in different forms (mathematical formula, tabulated data, or algorithmic description) and the second class hierarchy is used for alias sampling arrangements preparation and runtime sampling for zero, one, or two parametric PDFs.

4. Application for sampling the Molière electron multiple scattering distribution

The multi-scatter angular distribution for swift charged particles passing through matter, usually is written in a two parametric form $p(\theta, T, s)\mathrm{d}\Omega$, where θ is the angular deflection, our random variable, and the parameters are: T the average kinetic energy during the step and s the step length. For distributions using the small angle approximation, such as the Molière theory, this is equivalent with the form $f(\theta, T, s)\theta\mathrm{d}\theta$.

In the Molière [6, 7, 8] theory, by a change of variable, the above two parametric form is reduced to a single parameter one

$$f(\theta, T, s)\theta\mathrm{d}\theta = f_M(\vartheta, B)\vartheta\mathrm{d}\vartheta \tag{15}$$

where

$$\vartheta = \frac{\theta}{\chi_c\sqrt{B}}\ , \tag{16}$$

is called the reduced angle, and B and χ_c are defined by the equations:

$$B - \ln B = \ln \Omega_0 \tag{17}$$

$$\chi_c = \chi_{cc}\frac{\sqrt{s}}{\beta^2 E} \tag{18}$$

where b_c and χ_{cc} are constants which depend only on the medium in which the transport takes place. Ω_0 can be interpreted as the number of atomic collisions that contribute to the scattering and is related to the step length by the equation

$$\Omega_0 = \frac{b_c s}{\beta^2}\,, \tag{19}$$

For elements:

$$b_c = 6702.33\frac{\rho Z^{\frac{1}{3}}\,(Z+\xi)}{A}\cdot\frac{1}{1+0.000178\cdot Z^2}\ [\mathrm{cm}^{-1}] \tag{20}$$

$$\chi_{cc} = 0.39612\sqrt{\frac{\rho Z(Z+\xi)}{A}}\ [\mathrm{MeV}\cdot\mathrm{cm}^{-1/2}] \tag{21}$$

where ρ is the density in g/cm^3, s is the step length in cm, Z is the atomic number of the element, and A is the atomic mass in atomic mass units. E is the electron total energy in MeV, and ξ is a correction term introduced to take into account the inelastic scatterings with atomic electrons and it is usually taken equal to unity; an additional correction to χ_{cc} is given by Fano [9].

The Molière function is given as a series expansion:

$$f_M(\vartheta, B) = f^{(0)}(\vartheta) + \frac{1}{B} f^{(1)}(\vartheta) + \frac{1}{B^2} f^{(2)}(\vartheta) + \ldots \tag{22}$$

where:

$$f^{(n)}(\vartheta) = \frac{1}{n!} \int_0^\infty J_0(\eta\vartheta) e^{-\frac{\eta^2}{4}} \left(\frac{\eta^2}{4} \ln \frac{\eta^2}{4} \right)^n \eta \cdot \mathrm{d}\eta \ , \tag{23}$$

J_0 is the zeroth order Bessel function. In Eq. (23) for $n = 0$ we have a Gaussian curve:

$$f^{(0)}(\vartheta) = 2e^{-\vartheta^2} \ , \tag{24}$$

but for $n \geq 1$ the Molière functions must be computed using numerical methods. Terms until of order two proved to be sufficiently accurate.

Molière considered his theory valid for $\Omega_0 > 20$. Originally derived as a small-angle theory, Molière multiple-scattering theory has been modified by Bethe to predict accurately large-angle scattering by multiplying the right side of Eq. (15) with a correction term $\sqrt{\theta / \sin\theta}$. The minimum step size $s_{\min}$ required to ensure sufficient number of scatterings is given by:

$$s_{\min} = \frac{\Omega_0 \beta^2}{b_c} \ [\mathrm{cm}], \tag{25}$$

with $\Omega_0 = 20$, while the maximum step size derived by Bethe is

$$s_{\max} = \frac{E^2 \beta^4}{\chi_{cc} \ln(b_c E^2 \beta^2 / \chi_{cc})} \ [\mathrm{cm}], \tag{26}$$

For test was taken both cases of Molière distributions, the two parametric form $f(\theta, T, s)$ and the single parameter form $f_M(\vartheta, B)$.

In the case of the single parameter form, due to the requirement of preparing only material independent sampling arrangements for a unidimensional grid, the

memory usage is kept low, but the sampling speed suffer because additional calculation of the deflection angle as function of the reduced angle are required and, also, a rejection loop for the Bethe correction factor $\sqrt{\theta/\sin\theta}$, must be considered. A graphical comparison between a sampled distribution and the corresponding theoretical curve $f_M(\vartheta, B)$ is shown in figure 3.

The use of the two parametric form has the advantage of direct sampling of angular deflection, but with a large amount of memory usage because, for each material considered in the simulation, sampling arrangements for a two-dimensional grid must be prepared. An example of using this form for sampling of Molière multi-scatter angle distribution is given in figure 4.

A test program was written and runned on UNIX systems (Linux on Intel platforms and Digital UNIX on a Digital Alpha workstation). Several cases were tested. First were taken the single parameter form of the Molière distribution (AS1), and the two parameter form (AS2), in which for the PDFs interpolation inside the selected intervals the method given by equations (5),(6) was used. Second, were considered two simplified versions (AS1s, AS2s) which use the equation (3) for the PDFs interpolation; in this case is no more necessary to store the probability density values ($p_i^{(k,1)}$, $p_i^{(k,2)}$, etc.) and the loss of precision can be compensated by taking more refined grids. Finally, was considered a classical approach (C-EGS) based on the method used in EGS system, as it was described in [10].

The running speed results are tabulated in the table 1. As was mentioned before, in the case of the single parameter form (AS1, AS1s) the sampling speed suffers due the additional calculations made. An important factor for improving the scores is the speed of retrieving data from memory. This fact is emphasized by comparing the scores obtained on the 386 system with those obtained on the enhanced systems based on Pentium or Digital Alpha. During the generation of a random variable repeated readings of data spread on a relatively large amount of computer memory are made. Therefore a special care must be taken during the

implementation, in order to use efficient the CPU internal or external memory cache. The use of structured types of data and dynamical memory allocation to group together data needed for sampling of a certain random variable value (alias sampling arrangements and grids with interpolation points) can lead to a significant increase of speed, up to doubling it, according to our experience. Additional optimizations, which wasn't considered here, like direct sampling of memory addresses, rather than operate with indices when sampling of intervals, can also speed-up the algorithm.

5. Conclusion

A method of direct interpolation for sampling of random variables with parametrized probability distributions that allows the use of the alias sampling technique is presented. It takes advantage of modern computer architectures with large amounts of cheap and fast memories, by using discrete representations of probability distribution functions. The sampling is done by fast interpolation techniques involving only elementary logical and arithmetical operations, allowing to keep a higher degree of accuracy as the grids spacing is controlled by the user.

This method can successfully replace the sampling arrangements commonly used in Monte Carlo codes where, for complex probability distributions, from case to case, after a careful study of functions proprieties, a combination of sampling techniques, like superposition, rejection and inverse function methods are used. Unlike the methods involving discrete representation of PDFs and interpolation, this approach makes the resultant code strongly dependent on the theory used.

Acknowledgments

The author wishes to express his gratitude to professor Octavian Sima from Physics Department of Bucharest University for his support and suggestions in revising the manuscript.

References

[1] Donald E. Knuth. *Seminumerical Algorithms*, volume 2 of *The Art of Computer Programming*. Addison-Wesley, Reading, Massachusetts, second edition, 1981.

[2] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Wetterling. *Numerical Recipes — The Art of Scientific Computing*. Cambridge University Press, 1986.

[3] A. L. Edwards, J. A. Rathkopf, and R. K. Smidt. Extending the alias Monte Carlo sampling method to general distributions. In *Proc. Inter. Topical Mtg. AdvanceS in Mathematics, Computations, and Reactor Physics*. ANS, 1991. 2.1, 3–1.

[4] A. J. Walker. New fast method for generating discrete random numbers with arbitrary frequency distributions. *Electronics Letters*, 10(8), 1974.

[5] Clinton T. Ballinger. The response history Monte Carlo method for electron transport. Technical Report UCRL-ID-108040, University of California, Lawrence Livermore National Laboratory, 1991.

[6] G. Molière. Theorie der streuung scheneller geladener teilchen I. Einzelstreuung am abgeschirmten Coulomb-field. *Z. Naturforsch*, 2a:133, 1947.

[7] G. Molière. Theorie der streuung scheneller geladener teilchen II. Mehrfach- und vielfachstreuung. *Z. Naturforsch*, 3a:78, 1948.

[8] H. A. Bethe. Molière's theory of multiple scatering. *Phys. Rev.* , 89(6):1256, 1953.

[9] U. Fano. Inelastic collisions and the Molière theory of multiple scattering. *Phys. Rev.* , 93(1):117, 1954.

[10] W. R. Nelson, H. Hiroyama, and D. W. O. Rogers. The EGS4 code system. SLAC Report 256, Stanford Linear Accelerator Center, Stanford University, 1985.

List of Figures

Table 1: Comparison of running speed obtained with different versions of alias sampling technique and with the classical approach used in EGS, for the Molière electron multiple-scattering distribution*.

Method	Running Speed (generated values per second)		
	386DX 40 MHz	Pentium 133 MHz	Digital Alpha 233 MHz
C-EGS	4.12×10^3	4.91×10^4	8.05×10^4
AS1	3.32×10^3	4.34×10^4	7.20×10^4
AS1s	4.40×10^3	5.04×10^4	8.43×10^4
AS2	3.54×10^3	5.96×10^4	8.68×10^4
AS2s	6.39×10^3	7.99×10^4	14.35×10^4

* In all cases was used the same random number generator routine.

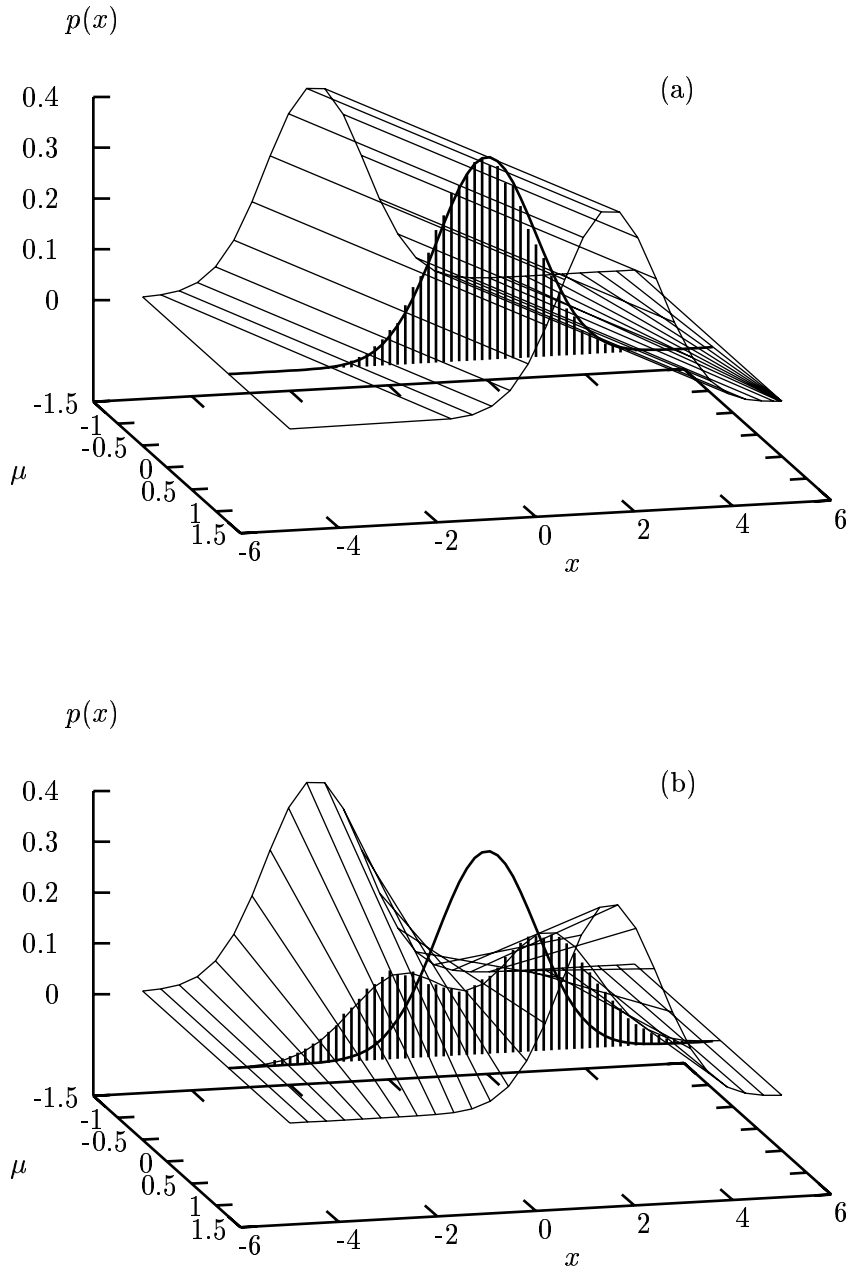
p(x)
(a)
0.4
0.3
0.2
0.1
0
-1.5
-1
-0.5
0
0.5
1
1.5
μ
-6
-4
-2
0
2
4
6
x
p(x)
(b)
0.4
0.3
0.2
0.1
0
-1.5
-1
-0.5
0
0.5
1
1.5
μ
-6
-4
-2
0
2
4
6
x

Figure 1:

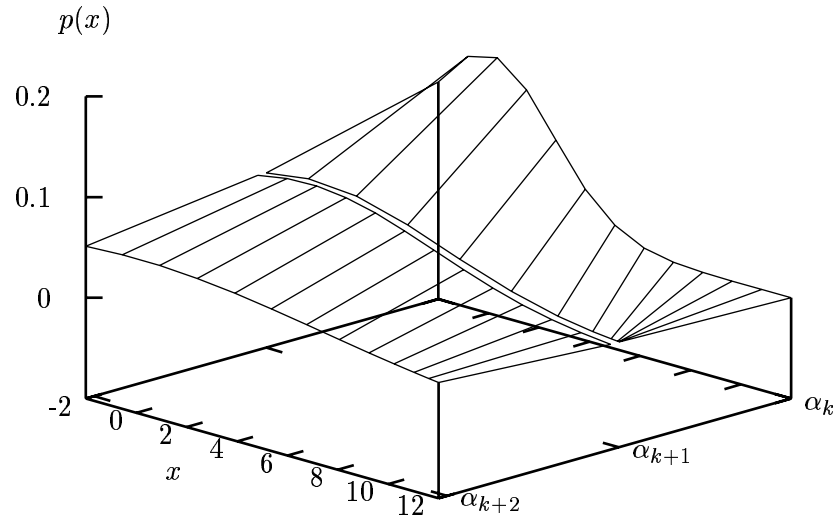

Figure 2:

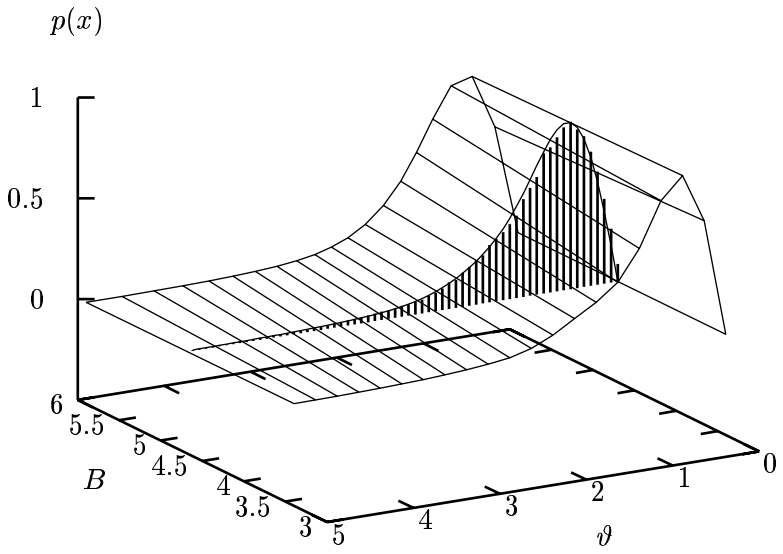

Figure 3:

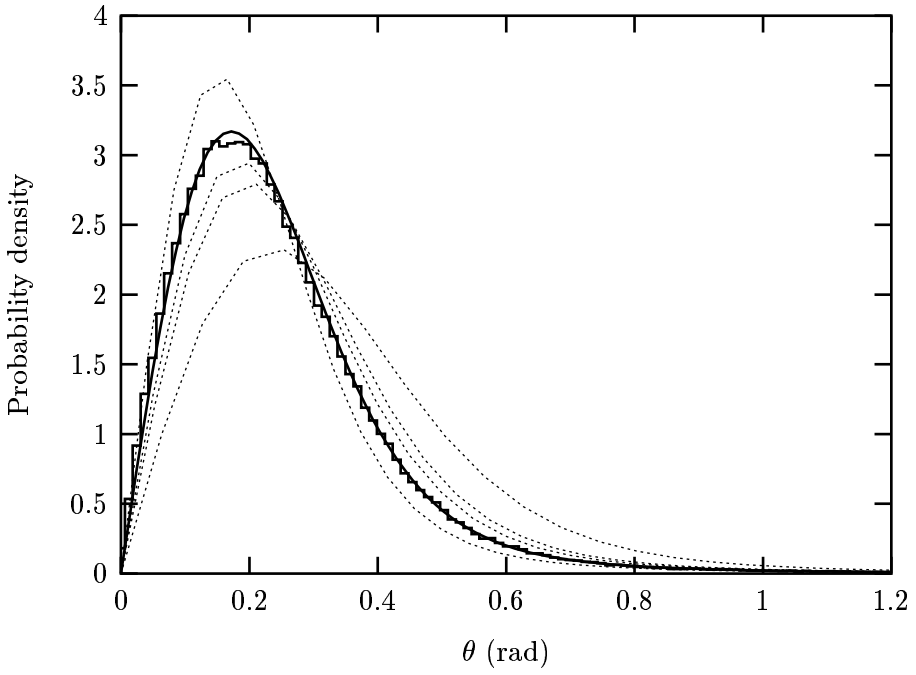

Figure 4: