Chapter 13
Random Sampling Methods

13.1 Introduction

Most applications require random number generators that follow a particular probability density function (pdf) which is not a uniform distribution on the interval [0, 1]. We present in this chapter methods to generate random numbers that follow some arbitrary pdf. As a source will serve uniformly distributed random numbers as they are generated with the help of the methods we discussed in Chap. 12.

The two most prominent techniques to generate random numbers from an arbitrary distribution, are the *inverse transformation method* and the *rejection method*. They will be discussed in Sects. 13.2 and 13.3, respectively. In addition, we comment in Sect. 13.4 briefly on sampling from piecewise defined pdfs and combined pdfs. It has to be emphasized that these methods are in many cases not sufficient and a more powerful approach is required. One of these is based on the idea of *importance sampling* and is referred to as the METROPOLIS method. It will be discussed briefly in Chap. 14.

Nevertheless, it is also possible to obtain quite easily random numbers for some specific pdfs by *direct sampling* [1]. For instance, suppose x_1, x_2 are two uniformly distributed random numbers. Hence, their pdf is given by

$$p_{\mathrm{u}}(x) = \begin{cases} 1 & x \in [0,1] \,, \\ 0 & \text{elsewhere.} \end{cases} \tag{13.1}$$

and the corresponding *cumulative distribution function* (cdf) follows:

$$P_{\mathrm{u}}(x) = \int_0^x \mathrm{d}x' p_{\mathrm{u}}(x') = \begin{cases} 0 & x < 0 \,, \\ x & x \in [0,1] \,, \\ 1 & x > 1 \,. \end{cases} \tag{13.2}$$

B.A. Stickler, E. Schachinger, *Basic Concepts in Computational Physics*,
DOI 10.1007/978-3-319-27265-8_13

One can prove that the new random number y

$$y = \max(x_1, x_2) \ , \tag{13.3}$$

conforms to the cdf

$$F(y) = y^2 \ , \tag{13.4}$$

and to the pdf[1]:

$$f(y) = 2y \ . \tag{13.5}$$

The consequence is an elegant method to generate random numbers z which follow the pdf

$$g(z) = kz^{k-1} \ , \tag{13.6}$$

by defining

$$z = \max(x_1, x_2, \ldots, x_k) \ . \tag{13.7}$$

Here, the random numbers x_i are uniformly distributed and can be obtained with the help of the methods introduced in Chap. 12.

Another equally elegant method can be employed to calculate random numbers which follow a normal distribution:

$$p(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \ . \tag{13.8}$$

Again, we act on the assumption that the random numbers x_i are uniformly distributed within the unit interval $[0, 1]$. We take two random numbers (x_1, x_2) and construct two random numbers (z_1, z_2) using the transformation:

$$z_1 = \cos(2\pi x_2)\sqrt{-2\ln x_1}, \qquad z_2 = \sin(2\pi x_2)\sqrt{-2\ln x_1} \ . \tag{13.9}$$

[1]This follows from the transformation of pdfs (see Chap. 14):

$$f(y) = \int_0^1 \mathrm{d}x_1 \int_0^1 \mathrm{d}x_2 \, \delta[y - \max(x_1, x_2)] = 2y.$$

It is easy to demonstrate that (z_1, z_2) follow the pdf (13.8). We introduce the joint distribution $p_u(x_1, x_2) = p_u(x_1)p_u(x_2)$ (assumption of no correlations). The transformation of probabilities [2–4] gives

$$p(z_1, z_2)\mathrm{d}z_1\mathrm{d}z_2 = p_u(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 \;, \tag{13.10}$$

or, equivalently, the JACOBIAN determinant

$$p(z_1, z_2) = \frac{\partial(x_1, x_2)}{\partial(z_1, z_2)} \;, \tag{13.11}$$

where we employed Eq. (13.1). We recognize that Eq. (13.9) is equivalent to

$$x_1 = \exp\left(-\frac{z_1^2 + z_2^2}{2}\right), \qquad x_2 = \frac{1}{2\pi}\tan^{-1}\left(\frac{z_2}{z_1}\right). \tag{13.12}$$

The JACOBIAN determinant is readily evaluated and gives[2]:

$$\begin{aligned}
\frac{\partial(x_1, x_2)}{\partial(z_1, z_2)} &= \begin{vmatrix} \frac{\partial x_1}{\partial z_1} & \frac{\partial x_1}{\partial z_2} \\ \frac{\partial x_2}{\partial z_1} & \frac{\partial x_2}{\partial z_2} \end{vmatrix} \\
&= \begin{vmatrix} -z_1 x_1 & -z_2 x_1 \\ -\frac{z_2}{2\pi(z_1^2+z_2^2)} & \frac{z_1}{2\pi(z_1^2+z_2^2)} \end{vmatrix} \\
&= \frac{x_1}{2\pi} \\
&= \frac{1}{2\pi}\exp\left(-\frac{z_1^2 + z_2^2}{2}\right) \\
&= p(z_1)p(z_2) \;.
\end{aligned} \tag{13.13}$$

This is the product of two normal distributions and, thus, z_1 and z_2 follow indeed a normal distribution.

[2]We make use of:

$$\frac{\mathrm{d}}{\mathrm{d}x}\tan^{-1}(x) = \frac{1}{1 + x^2}.$$

13.2 Inverse Transformation Method

The inverse transformation method is one of the simplest and most useful methods to sample random variables from an arbitrary pdf [1, 5–7]. Let $p(x)$, $x \in [x_{\min}, x_{\max}]$, denote the pdf from which we want to obtain our random numbers. The corresponding cdf will be denoted by

$$P(x) = \int_{x_{\min}}^{x} \mathrm{d}\,x' p(x') \; . \tag{13.14}$$

It follows immediately from the positivity and the normalization condition of pdfs (Appendix, Sect. E.5) that $P(x)$ is monotonically increasing and, furthermore, that $P(x_{\min}) = 0$ and $P(x_{\max}) = 1$. Let ξ denote some random number uniformly distributed within the interval [0, 1]. We obtain from the conservation of probability [2–4]

$$p_{\mathrm{u}}(\xi)\mathrm{d}\xi = p(x)\mathrm{d}\,x \Longrightarrow 1 = p_{\mathrm{u}}(\xi) = p(x)\left(\frac{\mathrm{d}\xi}{\mathrm{d}\,x}\right)^{-1} \; , \tag{13.15}$$

which can be solved by the choice $\xi = P(x)$, since

$$\frac{\mathrm{d}}{\mathrm{d}\,x}P(x) = p(x) \; . \tag{13.16}$$

Hence, we arrive at

$$x = P^{-1}(\xi) \tag{13.17}$$

where P^{-1} denotes the inverse of P. It is an obvious caveat of this method that it requires the inverse $P^{-1}(\xi)$ to exist and that $P(x)$ must be calculated and inverted analytically. This is, for instance, not possible in the case of the normal distribution (13.8).

Let us illustrate this method with two simple examples:

1. Suppose we want to draw random numbers which are uniformly distributed within the interval $[a, b]$. The corresponding pdf reads

$$p(x) = \frac{1}{b-a} \; , \tag{13.18}$$

   and the cdf takes on the form

$$P(x) = \frac{x-a}{b-a} \; , \tag{13.19}$$

where we set, in this particular example, $x_{\min} = a$. Hence, we have

$$\xi = \frac{x-a}{b-a} , \tag{13.20}$$

which is uniformly distributed within [0, 1]. Consequently, we determine random numbers $x \in [a, b]$ uniformly distributed via

$$x = a + (b-a)\xi . \tag{13.21}$$

2. We are interested in random numbers x drawn from a pdf given by the exponential distribution:

$$p(x) = \frac{1}{\lambda} \exp\left(-\frac{x}{\lambda}\right) . \tag{13.22}$$

Here, $\lambda > 0$ and $x \in [0, \infty)$. These could, for instance, describe the free path x of a particle between interactions, where the mean free path $\langle x \rangle = \lambda$. From Eq. (13.17) we obtain

$$\xi = \frac{1}{\lambda} \int_0^x \mathrm{d}x' \exp\left(-\frac{x'}{\lambda}\right) = 1 - \exp\left(-\frac{x}{\lambda}\right) , \tag{13.23}$$

and consequently the relation

$$x = -\lambda \ln(1-\xi) , \tag{13.24}$$

gives random variables x which comply to the exponential distribution (13.22) if ξ follows the pdf $p_{\mathrm{u}}(\xi)$ of Eq. (13.1). Moreover, it follows from the symmetry of the uniform distribution that

$$x = -\lambda \ln(\xi) , \tag{13.25}$$

without affecting the resulting random numbers. In Fig. 13.1 we show a histogram with random numbers drawn according to (13.25).

We pointed out already that it is certainly a caveat of this method that the cdf $P(x)$ has to be calculated and inverted analytically. Even if $P(x)$ is not analytically invertible, it is possible to employ the inverse transformation method by calculating $P(x)$ for certain grid-points x_i and then interpolating $P(x)$ piecewise between these points with the help of an invertible function. However, in many cases it is advantageous to employ the rejection method, which will be discussed next.

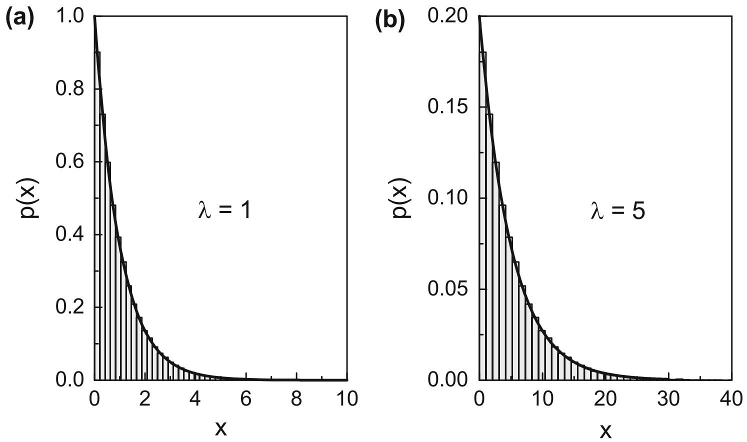


Fig. 13.1 The histogram representation of the pdf $p(x)$ vs x generated by random numbers drawn from an exponential distribution Eq. (13.22) with the help of the inverse transformation sampling method. Two different values for λ have been considered, namely **(a)** $\lambda = 1$ and **(b)** $\lambda = 5$. $N = 10^5$ random numbers have been sampled. The *solid line* corresponds to the pdf $p(x)$ according to Eq. (13.22)

13.3 Rejection Method

The rejection method is particularly suitable if the inverse transformation method fails [1, 6, 7]. One of the most prominent versions of the rejection method is the METROPOLIS algorithm. It will be introduced in Sect. 14.3.

The basic idea of the rejection method is to draw random numbers x from another, preferably analytically invertible pdf $h(x)$ and check whether or not they lie within the desired pdf $p(x)$. If this is the case the random number x is accepted, otherwise it will be rejected. This is also the basic idea of the *hit and miss* version of Monte-Carlo integration which will be discussed in Sect. 14.2.

We specify the rejection method: Let $p(x)$ denote the pdf from which we want to draw random numbers. Furthermore, let $h(x)$ be another pdf, which can easily be sampled (for instance with the help of the inverse transformation method) and which is chosen in a such a way that the inequality

$$p(x) \leq c\, h(x) \,, \tag{13.26}$$

holds for all $x \in [x_{\min}, x_{\max}]$, where $c \geq 1$ is some constant. The function $c\, h(x)$ is referred to as the *envelope* of $p(x)$ within the interval $[x_{\min}, x_{\max}]$. The strategy is clear: we sample a random variable x^t (trial state) from $h(x)$ and accept it with probability $p(x)/[c\, h(x)]$. This procedure is sketched in Fig. 13.2. Let $p(A|x)$ denote the probability that a given value x is accepted and $g(x)$ denotes the probability that

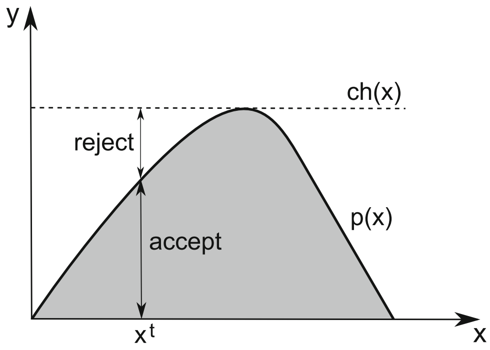


Fig. 13.2 Schematic illustration of the rejection method. The trial state x^t is accepted with probability $p(x)/[c\,h(x)]$

we produce a variable x with the help of this algorithm. Furthermore, $P(x = x^t)$ stands for the probability that a trial state x^t is generated. We have

$$\begin{aligned} g(x) &\propto P(x = x^t)p(A|x^t) \\ &= h(x^t)\frac{p(x^t)}{c\,h(x^t)} \\ &\propto p(x^t)\ . \end{aligned} \tag{13.27}$$

Hence, we indeed generate random numbers which follow the pdf $p(x)$. We may also calculate the probability $P(A)$ that an arbitrary trial state x^t is accepted. This is done with the help of the marginalization rule (E.39):

$$\begin{aligned} P(A) &= \int \mathrm{d}\,x^t p(A \wedge x^t) \\ &= \int \mathrm{d}\,x^t p(A|x^t)P(x = x^t) \\ &= \int \mathrm{d}\,x^t\,\frac{p(x^t)}{c\,h(x^t)}h(x^t) \\ &= \frac{1}{c}\int \mathrm{d}\,x^t p(x^t) \\ &= \frac{1}{c}\ . \end{aligned} \tag{13.28}$$

More generally, the probability $P(A)$ to accept a d-dimensional random variable is given by:

$$P(A) = \frac{1}{c^d}\ . \tag{13.29}$$

We deduce that the bigger c the worse is the acceptance probability of the rejection method. It is therefore advisable to choose the envelope $h(x)$ very carefully.

As an example we aim at sampling the normal distribution (E.43) for $x \in \mathbb{R}$

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) , \tag{13.30}$$

with expectation value $\langle x \rangle \equiv x_0 = 0$ and variance σ^2. In a first step we restrict our investigation to $x \in [0, \infty)$ due to the symmetry of the pdf. The slightly modified pdf for the right-half axis reads

$$q(x) = \sqrt{\frac{2}{\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) , \qquad x \in [0, \infty) , \tag{13.31}$$

where we adjusted the normalization. The complete normal distribution (13.30) is re-obtained by sampling the sign of x in an additional step. We use as an envelope $h(x)$ the exponential distribution Eq. (13.22). Furthermore, λ and c are chosen in such a way that the acceptance probability (13.28) has a maximum under the constraint (13.26). Since this is equivalent to $c \to \min$ we have to solve the optimization problem

$$c \geq \frac{q(x)}{h(x)} \to \max . \tag{13.32}$$

The resulting $c_{\min}$ is then given by

$$c_{\min} = \frac{q(x_{\text{opt}})}{h(x_{\text{opt}})} \tag{13.33}$$

and x_{opt} is the yet unknown optimal value for x. We obtain

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\,x} \frac{q(x)}{h(x)} &= \sqrt{\frac{2\lambda^2}{\pi\sigma^2}} \frac{\mathrm{d}}{\mathrm{d}\,x} \exp\left(\frac{x}{\lambda} - \frac{x^2}{2\sigma^2}\right) \\ &= \sqrt{\frac{2\lambda^2}{\pi\sigma^2}} \exp\left(\frac{x}{\lambda} - \frac{x^2}{2\sigma^2}\right) \left[\frac{1}{\lambda} - \frac{x}{\sigma^2}\right] \\ &\stackrel{!}{=} 0 , \end{aligned} \tag{13.34}$$

and, therefore,

$$x_{\text{opt}} = \frac{\sigma^2}{\lambda} . \tag{13.35}$$

Consequently, we have

$$c_{\min} = \sqrt{\frac{2\lambda^2}{\pi\sigma^2}} \exp\left(\frac{\sigma^2}{2\lambda^2}\right) . \tag{13.36}$$

The above relation gives the minimum value of c for arbitrary values of λ. However, since $h(x)$ is our envelope, we can choose λ in such a way, that $c_{\min} \to \min$. This is achieved in a second step

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\lambda} c_{\min} &= \sqrt{\frac{2}{\pi\sigma^2}} \exp\left(\frac{\sigma^2}{2\lambda^2}\right)\left(1 - \frac{\sigma^2}{\lambda^2}\right) \\ &\stackrel{!}{=} 0 . \end{aligned} \tag{13.37}$$

which results in the optimum value $\lambda_{\text{opt}} = \sigma$. This, finally, results together with Eq. (13.36) in:

$$c_{\min} = \sqrt{\frac{2e}{\pi}} . \tag{13.38}$$

The algorithm is executed in the following steps:

1. Draw a uniformly distributed random number $\xi \in [0, 1]$.
2. Calculate $x^t = -\lambda_{\text{opt}} \ln(\xi)$, where $\lambda_{\text{opt}} = \sigma$.
3. Draw a uniformly distributed random number $r \in [0, 1]$. If $r \leq q(x^t)/[c_{\min} h(x^t)]$, then $x = x^t$ is accepted and if $r > q(x^t)/[c_{\min} h(x^t)]$, x^t is rejected and we return to step 1.
4. If x^t was accepted, we draw a uniformly distributed random number $r \in [0, 1]$ and only if $r < 0.5$ we set $x = -x$ otherwise x stays as is.
5. We repeat steps 1–4 until the number N of desired random numbers has been reached.

Figure 13.3 shows random numbers obtained with the help of this method in a histogram representation. We calculated (a) $N = 10^3$, (b) $N = 10^4$, and (c) $N = 10^5$ random numbers for $\sigma = 1$. It is quite obvious that the original pdf (13.30) is the better approximated the bigger the number N of sampled random numbers becomes.

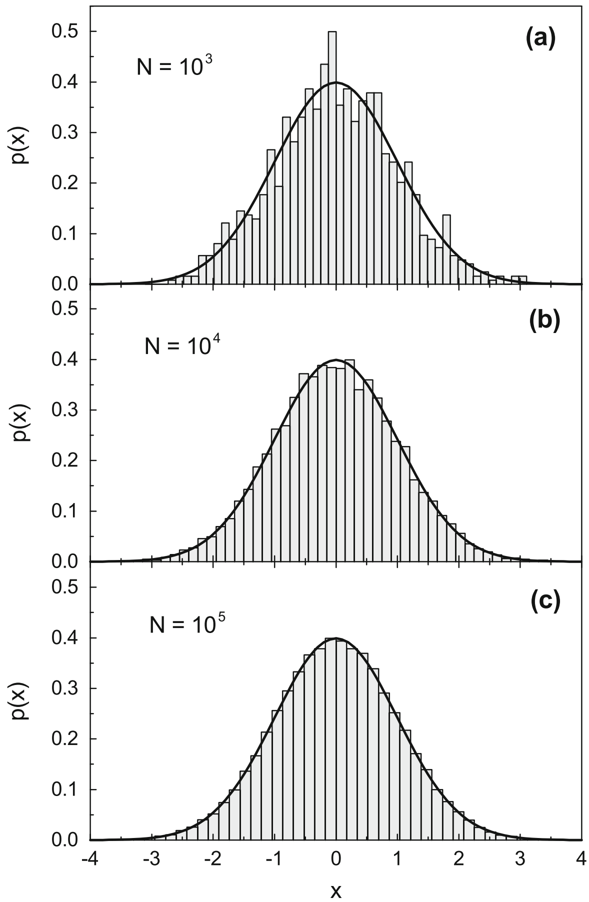


Fig. 13.3 The histogram representation of the pdf $p(x)$ vs x generated by random numbers drawn from the normal distribution Eq. (13.30) ($\sigma = 1$) with the help of the rejection method. We sampled **(a)** $N = 10^3$, **(b)** $N = 10^4$, and **(c)** $N = 10^5$ random numbers. The *solid line* represents the pdf $p(x)$ (13.30)

13.4 Probability Mixing

The method of probability mixing was developed to offer an algorithm which allows to generate random numbers by sampling piecewise defined or composite pdfs. Such a pdf is of the general form

$$p(x) = \sum_{i=1}^{N} \alpha_i f_i(x), \quad \alpha_i \neq 0 \,, \tag{13.39}$$

where the sub-pdfs $f_i(x)$ fulfill the normalization requirement

$$\int \mathrm{d}x' f_i(x') = 1 \,, \tag{13.40}$$

and are non-negative, i.e.

$$f_i(x) \geq 0 \; , \tag{13.41}$$

for all $i = 1, \ldots, N$. It follows that

$$\sum_{i=1}^{N} \alpha_i = 1 \; , \tag{13.42}$$

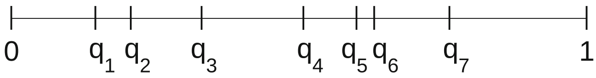
which ensures that

$$\int \mathrm{d}x' p(x') = 1 \; , \tag{13.43}$$

is fulfilled. The question is how to sample random numbers from such a pdf, since in most cases it might be hard to invert the sum (inverse transformation method) or find a suitable envelope (rejection method). However, this question can easily be answered: We define

$$q_i = \sum_{\ell=1}^{i} \alpha_\ell \; . \tag{13.44}$$

Thus, $q_N = 1$ and the interval [0, 1] has been divided according to:



The index i of the relevant pdf is determined by the condition

$$q_{i-1} < r < q_i \; , \tag{13.45}$$

where $r \in [0, 1]$ is a uniformly distributed random number. We then draw the required random number x from the sub-pdf $f_i(x)$ with any of the methods discussed above.

This procedure is quite plausible, since the coefficients α_i give the relative weight of the sub-pdfs $f_i(x)$. In particular, α_i determines the importance of the sub-pdf $f_i(x)$. It is, therefore, a natural approach to use α_i as a measure of the probability that a random variable is to be sampled from the particular sub-pdf $f_i(x)$.

Summary

To generate random numbers is essential for many application in Computational Physics. This chapter concentrated on basic methods to generate the desired random numbers: (a) the *direct sampling method* used transformations of the uniform distribution to generate the required random numbers; (b) the *inverse transformation method* was based on the availability of an inverse cdf which was in most cases required to be calculated analytically; finally, (c) the *rejection method* which was basically a hit or miss method. It used an easily invertible pdf $h(x)$ which enveloped the desired pdf $p(x)$ completely within some interval $x \in [x_{\min}, x_{\max}]$. The effectiveness of this method depended on how 'well' the envelope $h(x)$ enclosed the original pdf $p(x)$. In a last step the method of *probability mixing* was discussed. It was an easily verifiable method which allowed to sample random numbers from composite pdfs.

Problems

Draw random numbers from the following pdfs:

1. *Direct Sampling*:
   Sample the normal distribution with $\langle x \rangle = 0$ and $\sigma = 1$ with the help of the method discussed in Sect. 13.1. Check the result by plotting the random numbers against the pdf $p(x)$ in a histogram.
2. *Inverse Transformation Method*:
   Write a function which samples random numbers from the exponential distribution with the help of the inverse transformation method as discussed in Sect. 13.2. Compare the generated random numbers to the pdf in a histogram.
3. *Rejection Method*:
   Sample the normal distribution with $\langle x \rangle = 0$ and $\sigma = 1$ with the help of the exponential distribution as discussed in Sect. 13.3. Compare the generated random numbers with the pdf in a histogram. Determine the acceptance probability numerically.
4. *Probability Mixing*:
   We choose an alternative envelope for the normal distribution with $\langle x \rangle = 0$ and $\sigma = 1$. This envelope is chosen to be constant for all $|x| < x_0$, and decays exponentially for $|x| \geq x_0$. (x_0 is a parameter of your choice.) The parameters do not need to optimize the acceptance probability. Again, plot the generated random numbers in a histogram and compare the acceptance probability with the acceptance probability of point 3.

References

1. Devroye, L.: Non-uniform Random Variate Generation. Springer, Berlin/Heidelberg (1986)
2. Chow, Y.S., Teicher, H.: Probability Theory, 3rd edn. Springer Texts in Statistics. Springer, Berlin/Heidelberg (1997)
3. Kienke, A.: Probability Theory. Universitext. Springer, Berlin/Heidelberg (2008)
4. Stroock, D.W.: Probability Theory. Cambridge University Press, Cambridge (2011)
5. Bratley, P., Fox, B.L., Schrage, L.E.: A Guide to Simulation. Springer, Berlin/Heidelberg (1987)
6. Knuth, D.: The Art of Computer Programming, vol. II, 3rd edn. Addison Wesley, Menlo Park (1998)
7. Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: Numerical Recipes in C++, 2nd edn. Cambridge University Press, Cambridge (2002)