LBL-29673
Preprint

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

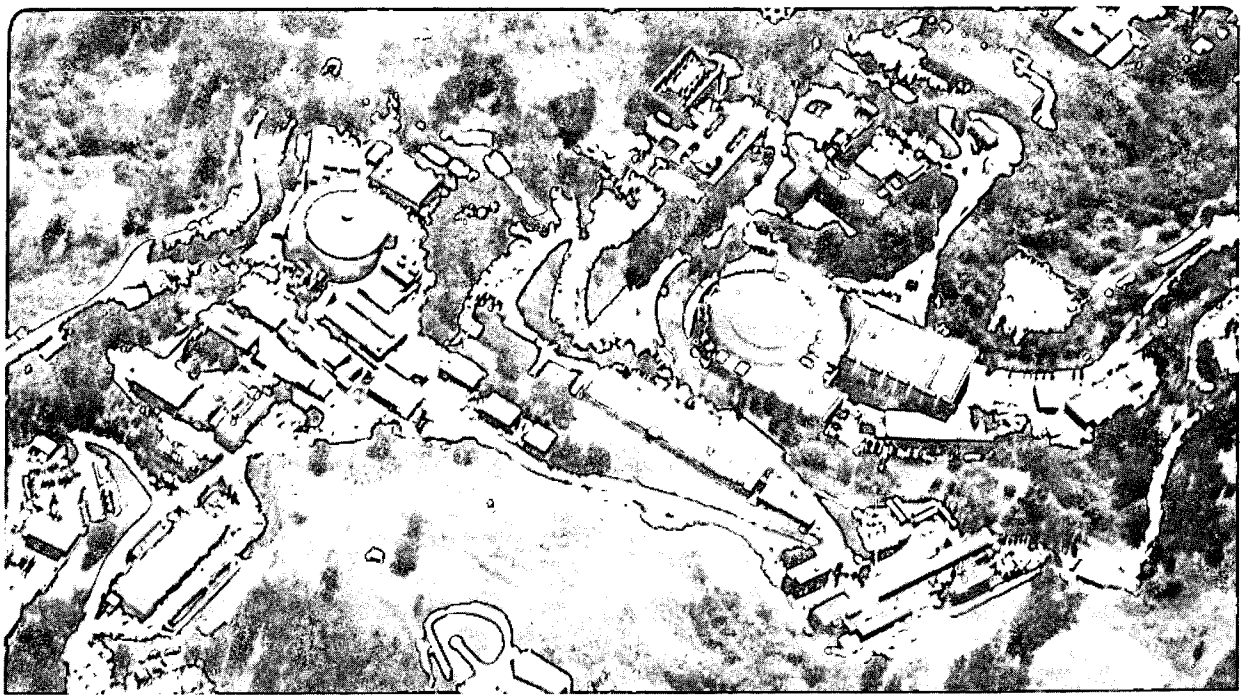
Physics Division

Submitted for publication

Reference Functions to Decrease Errors in Monte Carlo Integrals

P.H. Eberhard and O.P. Schneider

October 1990



LOAN COPY
Circulates
for 2 weeks
Bldg. 50 Library.
Copy 2
LBL-29673

Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098.

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

LBL–29673

Reference Functions to Decrease Errors in Monte Carlo Integrals*

Philippe H. Eberhard and Olivier P. Schneider

Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

October 1990

*This work was supported in part by the Director, Office of High Energy Physics of the U.S. Department of Energy and Nuclear Physics, Division of High Energy Physics of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098 and by the Swiss National Science Foundation.

Reference Functions to Decrease Errors in Monte Carlo Integrals*

Philippe H. Eberhard
and
Olivier P. Schneider

October 4, 1990

Abstract

Given a set of points generated by Monte Carlo, the accuracy of a Monte Carlo integral of functions $f(x)$ performed using these points can be improved if there are "reference functions", i.e., other functions that can be integrated both analytically and by Monte Carlo using the same points. This statement is shown to be true. Formulae are developed to make practical use of it. Guidelines are given as how to choose reference functions, recommending that there be a linear combination of these functions that approximate the functions $f(x)$ as well as possible. Not only statistical errors but some systematic errors due to some biases in the Monte Carlo generator are reduced too.

Monte Carlo integrations using reference functions may be considered as bridging the gap between numerical and usual Monte Carlo integrations.

*This work was supported in part by the Director, Office of High Energy Physics of the U.S. Department of Energy and Nuclear Physics, Division of High Energy Physics of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098 and by the Swiss National Science Foundation.

Contents

1 Introduction **1**

1.1 Generalities about Monte Carlo Integrals 1

1.2 A Simple Case of Use of a Reference Function 3

1.3 Ratio of Integrals 5

1.4 Implicit Normalization 7

2 Extracting the Maximum Information Out of the Reference Functions **9**

2.1 The Poisson Process 10

2.2 The Case of Several Reference Functions 11

2.3 The Best Estimate of F 13

2.4 The Least Square Fitted Reference Function 16

3 Generalizations and Practical Formulae **19**

3.1 Integrating Linear Combinations of Reference Functions . . . 19

3.2 Mathematics for Computers 22

3.3 Several Functions to Be Integrated 23

3.4 Weighted Distributions — General Formulae 25

References **28**

1 Introduction

Monte Carlo techniques are used to estimate quantities that are difficult to calculate by analytic means [1,2]. In general, the problem can be formulated in terms of evaluating some integrals. Reference functions are functions that can be integrated both by Monte Carlo and by analytic means. Such functions can be used to reduce the error on Monte Carlo integrals of other functions for which there is no such analytic means.

Reference functions can be useful when Monte Carlo simulations need high accuracy. In physics, they are likely to be useful in cases where the theoretical probability distribution of a physical quantity, ξ_1, has an analytic expression but the detector distortions that affect the measured value, ξ_2, of ξ_1 have to be simulated by Monte Carlo. If the expectation value, $\overline{\xi_1}$, of the undistorted quantity, ξ_1, can be calculated analytically, then the predicted expectation value, $\overline{\xi_2}$, of the measured value ξ_2 under these conditions can be computed more accurately with proper use of reference functions derived from the distribution of ξ_1. An example is given in the next paper [3].

1.1 Generalities about Monte Carlo Integrals

Basically, Monte Carlo simulation is used to evaluate integrals of the form

$$F = \int f(x) p^{MC}(x) dx \ , \tag{1}$$

where x is a multidimensional variable of dimension M_x (large), $f(x)$ is a function of the variable x that can be expressed in closed form,[1] and $p^{MC}(x)$ is a function of x that can be simulated by a distribution of points generated by Monte Carlo. The distribution $p^{MC}(x)$ may not be expressed in closed form but a Monte Carlo algorithm exists that generates points k of coordinates ξ_k with a probability $p^{MC}(x)dx$ for each point k to fall in the

[1] In this context, "closed form" means the possibility for an expression to be calculated to the computer accuracy in a few machine cycles, i.e., in a time negligible with respect to the time needed to compute integrals numerically.

volume dx around x. By definition,

$$\int p^{MC}(x)dx = 1 \ . \tag{2}$$

Note that limits of integrations have not been spelled out in Eqs. (1) and (2). These integrals are to be understood as integrals over the whole range of x. This does not restrict the generality of these equations since we can always define a function to be integrated over any finite domain as a function that is zero outside of that domain.

The estimation of the integral of Eq. (1) involves the following quantities:

$$\varphi_k = f(\xi_k) \quad \textit{for each Monte Carlo point } k \ , \tag{3}$$

$$N = \textit{number of Monte Carlo points} \ , \tag{4}$$

$$\Phi = \sum_k \varphi_k \ . \tag{5}$$

Let us use the symbol ^ above a quantity to designate an estimate of that quantity. An estimate of F is

$$\widehat{F} = \frac{\Phi}{N} \ . \tag{6}$$

To justify Eq. (6), it is easy to demonstrate that the expectation value of $\widehat{F}$ is F. Let us designate the expectation value of a quantity by a bar above that quantity.

$$\overline{\widehat{F}} = \frac{\overline{\Phi}}{N} = \overline{\varphi_k} = \int f(x)p^{MC}(x)dx = F \tag{7}$$

An idea about the discrepancy δF between $\widehat{F}$ and F,

$$\delta F = \widehat{F} - F \ , \tag{8}$$

can be given by the expectation value, $S^2(\widehat{F})$, of the square of δF.

$$S^2(\widehat{F}) = \overline{(\delta F)^2} = \frac{1}{N}\int [f(x) - F]^2 \, p^{MC}(x)dx \tag{9}$$

$$= \frac{1}{N}\left(\int f^2(x)p^{MC}(x)dx - F^2\right) \tag{10}$$

The quantity $S(\widehat{F})$ is the one-standard deviation error on $\widehat{F}$. For N large enough, $S^2(\widehat{F})$ given by Eq. (10) as an integral involving $f(x)$ and $f^2(x)$ can be approximated by a quantity $\widehat{S^2}(\widehat{F})$, which depends only on quantities computed at the Monte Carlo points.

$$\widehat{S^2}(\widehat{F}) = \frac{1}{N^2}\sum_k \left(\varphi_k - \frac{\Phi}{N}\right)^2 \tag{11}$$

$$\overline{\widehat{S^2}}(\widehat{F}) = S^2(\widehat{F}) \tag{12}$$

We write:

$$\epsilon_f = \sum_k \varphi_k^2 , \tag{13}$$

$$\widehat{S^2}(\widehat{F}) = \frac{1}{N^2}\left(\epsilon_f - \frac{\Phi^2}{N}\right) . \tag{14}$$

The quantity $\widehat{S^2}(\widehat{F})$ is the practical way to estimate the square of the error on $\widehat{F}$. That is the kind of estimate we will be mostly using here.

1.2 A Simple Case of Use of a Reference Function

Consider another function, $t(x)$, for which the integral

$$T = \int t(x) p^{MC}(x) dx , \tag{15}$$

unlike the one of Eq. (1), is known in closed form. The function $t(x) \cdot p^{MC}(x)$ can be integrated by both Monte Carlo and analytic means. Thus a new estimate $\widehat{F}$ of F, different from the one of Eq. (6), can be defined [4].

$$\vartheta_k = t(\xi_k) \tag{16}$$

$$\Theta = \sum_k \vartheta_k \tag{17}$$

$$\widehat{F} = T + \frac{\Phi - \Theta}{N} = T + \frac{1}{N}\sum_k (\varphi_k - \vartheta_k) \tag{18}$$

That estimate can be justified, like the one of Eq. (6), by writing an equation similar to Eq. (7).

$$\overline{\widehat{F}} - T = \frac{\overline{\Phi} - \overline{\Theta}}{N} = \overline{\varphi_k} - \overline{\vartheta_k} = \int [f(x) - t(x)] p^{MC}(x) dx = F - T \tag{19}$$

The integration procedure of Eq. (18) amounts to applying Eq. (6) to the function $f(x) - t(x)$ and then add the known value T of the integral of $t(x) \cdot p^{MC}(x)$. The integral T is not affected by statistical fluctuations. The absolute error on $\widehat{F}$ using Eq. (18) is the same as the error one would have integrating the function $[f(x) - t(x)] \cdot p^{MC}(x)$ using Eq. (6). To get an estimate of the square of that error, one has to substitute $\varphi_k - \vartheta_k$ for φ_k in Eq. (11), which then reads

$$\widehat{S^2}(\widehat{F}) = \frac{1}{N^2} \sum_k \left(\varphi_k - \vartheta_k - \frac{\Phi - \Theta}{N} \right)^2 . \tag{20}$$

We write:

$$\epsilon_t = \sum_k \vartheta_k^2 , \tag{21}$$

$$\Psi = \sum_k \varphi_k \, \vartheta_k , \tag{22}$$

$$\widehat{S^2}(\widehat{F}) = \frac{1}{N^2} \left(\epsilon_f - 2\Psi + \epsilon_t - \frac{(\Phi - \Theta)^2}{N} \right) . \tag{23}$$

Suppose the function $t(x)$ approximates $f(x)$ better than a constant, i.e., $f(x) - t(x)$ is smaller in absolute value than $f(x) - F$ in the regions where $p^{MC}(x)$ is substantial, therefore for most of the Monte Carlo points. Then, by comparing Eqs. (20) and (11), it can be seen that integrating $f(x) \cdot p^{MC}(x)$ using Eq. (18) is more accurate than using Eq. (6), as pointed out in Ref. [4].

In Eq. (18), the function $t(x)$ is used as a "reference" for the function $f(x)$. This is possible because the integral T of Eq. (15) is known. Therefore we call any function for which we know the integral of type (15) in closed form a "reference function".

As an example in physics, let us consider the simulation of events, each one characterized by a quantity ξ_1, which is then measured. In general, the measurement of ξ_1 will yield a value ξ_2 close to but different from ξ_1 because of detector distortions and measurement errors. Suppose the distribution of the physical quantity ξ_1 and its expectation value $\overline{\xi_1}$ can be expressed in closed form. Suppose the detector distortions cannot be expressed analytically and will have to be simulated by Monte Carlo. The

predicted expectation value, $\overline{\xi_2}$, of ξ_2 can be computed by generating Monte Carlo events simulating the real events. The probability distribution of the quantities ξ_1 and ξ_2 for each event is a function $p^{MC}(x_1, x_2)$.

$$\overline{\xi_2} = \int x_2 \, p^{MC}(x_1, x_2) dx_1 dx_2 \tag{24}$$

This equation is of the form of Eq. (1). Since the expectation value of ξ_1 is known and equal to

$$\overline{\xi_1} = \int x_1 \, p^{MC}(x_1, x_2) dx_1 dx_2 \ , \tag{25}$$

Eq. (18) can be used with the function x_1 plugged in as the reference function $t(x)$ of Eqs. (15) and (16). Since ξ_1 and ξ_2 are always close, the result using Eq. (18) should be much more accurate than using Eq. (6). In this particular example, the use of a reference function amounts to relying on the Monte Carlo only to determine the bias

$$\overline{\xi_2} - \overline{\xi_1} = \int (x_2 - x_1) p^{MC}(x_1, x_2) dx_1 dx_2 \tag{26}$$

due to detector distortion.

A more elaborate example, using several reference functions, is given in the next paper [3].

Because of Eq. (2), the integral of any constant is known. Therefore any constant independent of x is a reference function. Any constant introduced as $t(x)$ in Eqs. (15) and (16) makes the estimate $\widehat{F}$ of Eq. (18) identical to the one of Eq. (6). It makes also the errors of Eq. (23) equal to expression (14). Equations (6), (9), and (14) are just a special case where $t(x)$ has been chosen as a constant.

1.3 Ratio of Integrals

Reference functions can be introduced in a different context. In Sect. 1.2, we focussed on the difference between integrals involving the functions $f(x)$ and $t(x)$. Now we consider the ratio between these two integrals. Using T

of Eq. (15), Θ of Eq. (17), and Φ of Eq. (5), a new estimate for F can be derived.

$$\widehat{F} = T\,\frac{\Phi}{\Theta} \tag{27}$$

This estimate can be justified because, for N large,

$$\Phi \;\cong\; \overline{\Phi} = N\overline{\varphi_k} = N\int f(x)p^{MC}(x)dx = N\;F\;, \tag{28}$$

$$\Theta \;\cong\; \overline{\Theta} = \ldots = N\;T\;, \tag{29}$$

$$\widehat{F} \;\cong\; T\frac{\overline{\Phi}}{\overline{\Theta}} = F\;. \tag{30}$$

The difference, δF, between $\widehat{F}$ and F is related to the statistical fluctuations of Φ and Θ, $\delta\Phi$ and $\delta\Theta$, which are correlated with each other.

$$\delta F = \frac{T}{\Theta}\left(\delta\Phi - \frac{\Phi}{\Theta}\delta\Theta\right) \cong \frac{1}{N}\left(\delta\Phi - \frac{F}{T}\delta\Theta\right) \tag{31}$$

Averaging stochastic quantities, one gets

$$\begin{aligned} S^2(\widehat{F}) = \overline{(\delta F)^2} &= \frac{1}{N^2}N\int\left([f(x)-F] - \frac{F}{T}[t(x)-T]\right)^2 p^{MC}(x)dx \\ &= \frac{1}{N}\int\left(f(x) - \frac{F}{T}t(x)\right)^2 p^{MC}(x)dx\;, \end{aligned} \tag{32}$$

a quantity which can be approximated using only the quantities φ_k and ϑ_k of Eqs. (3) and (16) computed at the integration points,

$$\widehat{S^2}(\widehat{F}) = \left(\frac{T}{\Theta}\right)^2\sum_k\left(\varphi_k - \frac{\Phi}{\Theta}\vartheta_k\right)^2 \tag{33}$$

$$= \left(\frac{T}{\Theta}\right)^2\left(\epsilon_f - 2\frac{\Phi}{\Theta}\Psi + \left(\frac{\Phi}{\Theta}\right)^2\epsilon_t\right)\;, \tag{34}$$

where ϵ_f, Φ, and ϵ_t are the quantities defined in Eqs. (13), (5), and (21).

The error on $\widehat{F}$ of Eq. (27) is very small if the function $t(x)$ represents a fair approximation for $f(x)$. This can be seen from Eqs. (32) and (33) if $f(x) \cong t(x)$. Then we have $F \cong T$, $\varphi_k \cong \vartheta_k$, and $\Phi \cong \Theta$. This small error is the same property as the one found in Sect. 1.2, for integrations using Eq. (18) under the same circumstances. However, here, Eqs. (27),

(32), and (33) are invariant with respect to a scaling factor of the function $t(x)$. Therefore, this time, all that is required for the small error, is that the shape of $t(x)$ approximates the shape of the function $f(x)$. The scale of $t(x)$ is of no importance as long as any change in the scale of $t(x)$ is also made on the integral T of Eq. (15).

If the function $t(x)$ is equal to a constant (which, because of Eqs. (15) and (2), must be equal to T), then the ratio Θ/T becomes equal to the N of Eq. (4). Thus $\widehat{F}$ of Eq. (27) reduces to the $\widehat{F}$ of Eq. (6), $S^2(\widehat{F})$ of Eq. (32) to the one of Eq. (9), and $\widehat{S^2}(\widehat{F})$ of Eq. (33) to the one of Eq. (11). The procedure of Sect. 1.1 and the estimate given by Eq. (6) are just a special case of the general procedure of this section characterized by Eq. (27). Of course it is not always advantageous to use a constant as the reference function $t(x)$ of Eq. (15). If there is a reference function, $t(x)$, whose shape approximates the shape of $f(x)$ better than a constant, the error using that function $t(x)$ in conjunction with Eq. (27) is smaller than the error using a constant reference function.

1.4 Implicit Normalization

The formalism of Sect. 1.3 allows us to generalize the kind of integrals that can be estimated by Monte Carlo. Let $a(x)$ be a function whose shape is the same as the expectation value $p^{MC}(x)$ of the Monte Carlo distribution but for which the normalization is not necessarily 1.

$$a(x) = A\, p^{MC}(x) \tag{35}$$

$$A = \int a(x)dx \tag{36}$$

Suppose the normalization factor A is not known in closed form but only defined implicitly as the factor that makes the integral of $t(x) \cdot a(x)$ equal to a known value T.

$$T = \int t(x)a(x)dx \tag{37}$$

Any integral of the form

$$F = \int f(x)a(x)dx \tag{38}$$

can be estimated, using Φ of Eq. (5) and Θ of Eq. (17) and introducing them in Eq. (27). This estimate is obviously as valid for F of Eq. (38) using T of Eq. (37) as it is for F of Eq. (1) using T of Eq. (15). Indeed the integral F of Eq. (38) is A times the one of Eq. (1) and the integral T of Eq. (37) is A times the one of Eq. (15). As to Φ of Eq. (5) and Θ of Eq. (17), their values are not affected by the transformation of $p^{MC}(x)$ into $a(x)$, which multiplies both hands of Eq. (27) by the factor A.

Here as before, the functions $f(x)$ and $t(x)$ are functions expressible in closed form, at least at each of the Monte Carlo points. Like $p^{MC}(x)$, the function $a(x)$ which describes the shape of the Monte Carlo distribution may not be expressed in closed form. However, here, the normalization factor A of $a(x)$ may not be equal to 1, and may not even be known. The only thing known about $a(x)$ may be the algorithm that generates the Monte Carlo points and the value T of the integral of Eq. (37).

To estimate integrals of the type of Eq. (38) involving distributions $a(x)$ not necessarily normalized to 1, we now generalize the definition of "reference function" to any function $t(x)$ for which the integral of $t(x) \cdot a(x)$, i.e., T of Eq. (37), is known in closed form.

If the normalization factor A of Eq. (36) is known, the problem can be reformulated, changing $f(x)$ into $A \cdot f(x)$, $t(x)$ into $A \cdot t(x)$, and $a(x)$ into $p^{MC}(x)$, and keeping T the same. Then the formalism of Sect. 1.3 can be applied. If A is known, there is no advantage in generalizing the integrals of the type of Eq. (1) to integrals of the type of Eq. (38). The advantage of the implicit normalization of this section becomes apparent if one has to evaluate integrals involving distributions $a(x)$ whose normalizations are given only implicitly by an equation of the type of Eq. (37). On the other hand, since Eqs. (37) and (38) reduce to Eqs. (1) and (15) for $A = 1$, there is nothing lost in developing our general formulae for integrals of the form of Eqs. (37) and (38).

As a short example in physics leading to implicit normalization, let us consider the case where the distribution $p^{MC}(x)$ of the Monte Carlo points is different from the distribution $p^{phys}(x)$ of real events, and where one wants

to estimate the expectation values of quantities associated with the physical events. These expectation values are given by integrals of the type

$$G = \int g(x) p^{phys}(x) dx \ , \tag{39}$$

where $g(x)$ is the relevant quantity expressed in closed form at the Monte Carlo points. A weighting function $w(x)$, known in closed form at least at each integration point, may be associated with the Monte Carlo distribution so that $w(x) \cdot p^{MC}(x)$ has the same shape as $p^{phys}(x)$. However the normalization of $w(x) \cdot p^{MC}(x)$ may not be known in closed form. Because $p^{phys}(x)$ is normalized to 1, one has

$$p^{phys}(x) = A w(x) p^{MC}(x) \ , \tag{40}$$

where, here, A is given by

$$A = \frac{1}{\int w(x) p^{MC}(x) dx} \ . \tag{41}$$

A function $a(x)$ can be defined according to Eq. (35). Then, Eq. (40) implies

$$\int w(x) a(x) dx = 1 \ . \tag{42}$$

Equation (42) is just a special case of Eq. (37). Thus the function $w(x)$ is a reference function. The integral of Eq. (39) is of the form Eq. (38) if we put $f(x) = w(x) g(x)$. From now on, in this example, use of the implicit-normalization procedure of this section is straightforward.[2]

2 Extracting the Maximum Information Out of the Reference Functions

In this part, our goal is to derive expressions to extract all information that reference functions can give to minimize errors on integrals. The goal is also

[2]It should be emphasized that there are cases where a weighting function $w(x)$ is used without rigorously satisfying Eq. (40). The weighting function may be just a factor in the expression of the ratio of $p^{phys}(x)$ to $p^{MC}(x)$, but not necessarily equal to a constant times this ratio. This circumstance is common if one has to compare the integrals of different physical distributions corresponding to different theories.

to give these expressions a simple enough form so that they can be used automatically. Since integrals of the type of Eqs. (37) and (38) of Sect. 1.4 allow us to use or not to use implicit normalization, they are more general than Eq. (1) and will be the types of integrals for which we derive our general formulae.

2.1 The Poisson Process

To be as general as possible, one must be able to consider that, if we know the integral of the function $a(x)$ (i.e., the quantity of Eq. (36), which we call A), a constant could be used as the reference function $t(x)$ of Eq. (27). We have seen that this is possible in the final formulae but, as a concept, considering a constant as a reference function is extremely counter-intuitive. The benefit obtained from a reference function $t(x)$ used in the estimate $\widehat{F}$ of Eq. (27) comes from the correlation between the statistical fluctuations of Θ of Eq. (17) and of Φ of Eq. (5). For $t(x) = 1$, one has $\Theta = N$, i.e., a predetermined quantity to which we do not attach a statistical fluctuation. Therefore, in this approach, if we know the integral of a constant times $a(x)$, that constant has to be given a different treatment throughout the whole development than the other reference functions for which the sum (17) is a random quantity. Only at the end can it be shown that the constant plays exactly the same role in the final formulae as the other reference functions.

It is possible to take a unified approach and analyze cases where we know the integral A of Eq. (36) with exactly the same concepts as cases where we do not know A. For this, we consider that the total number N of Monte Carlo points is not predetermined but that it results from a Poisson process [5]. Now N is a random number. It has an expectation value $\overline{N}$ and a standard deviation $\sqrt{\overline{N}}$. In this context, if the integral A of the function $1 \cdot a(x)$ is known, 1 is a reference function like the others. It can be shown that the final formulae are the same if N is an outcome of a Poisson process or if N is predetermined but, with the Poisson process, the development is smoother.

In the context of a Poisson process of negligible probability for $N = 0$,

the expectation value of the estimates $\widehat{F}$ of Eqs. (6), (18), and (27) still verify Eqs. (7), (19), and (30). These estimators $\widehat{F}$ are still consistent estimators of F. However, the error $S(\Phi)$ on Φ of Eq. (5) has a different expression than if N is predetermined, because, with the Poisson process, the error on N has also to be taken into account.

$$S^2(\Phi) = \overline{N} \int f^2(x) p^{MC}(x) dx \tag{43}$$

This square of the error can be approximated by

$$\widehat{S^2}(\Phi) = \sum_k \varphi_k^2 = \epsilon_f \quad \textit{of Eq. (13)} \ . \tag{44}$$

Changing φ_k and Φ into ϑ_k and Θ, one gets the error on Θ. In the Poisson process, the error on $\widehat{F}$ of Eqs. (6), (18) or (27) is given approximately by $\widehat{S^2}(\widehat{F})$ of Eqs. (14), (23), or (34), respectively. This is true because of the statistical correlations between Φ, Θ, and N. Because of these correlations, the Poisson process is led to the same final formulae for the errors on the integrals as a predetermined N.

2.2 The Case of Several Reference Functions

If there are several reference functions, a rigorous statistical treatment has to be given to the problem of using all the available information so that the error on the estimate of F of Eq. (38) be truly minimized.

Let us consider a set of M_t reference functions $t_m(x)$, i.e., a set of functions for which the integral of the type of Eq. (37) is known. Of course any linear combination of them is also a function for which the integral is known, therefore it is another reference function. However that linear combination obviously supplies only redundant additional information. Therefore we do not lose any information in restricting our set of reference functions to a set of linearly independent functions $t_m(x)$.

Let us construct a vector $|t(x)|$ as a vector having the functions $t_m(x)$ as its components.

$$M_t = \textit{dimension of vector } |t(x)| = \textit{number of reference functions} \tag{45}$$

The known integrals T_m of these reference functions multiplied by $a(x)$ also constitute the components of a vector $|T|$ of dimension M_t.

$$|T| = \int |t(x)|a(x)dx \tag{46}$$

Along with φ_k and Φ in Eqs. (3) and (5), one can compute the following vectors:

$$|\vartheta_k| = |t(\xi_k)| , \tag{47}$$

$$|\Theta| = \sum_k |\vartheta_k| . \tag{48}$$

One can compare $|\Theta|$ to $|T|$ to get information about the statistical distribution of the Monte Carlo points. Let r be the ratio

$$r = \frac{\overline{N}}{A} . \tag{49}$$

Starting from Eqs. (35), (38), and (46), one can write equations for the Poisson process that resemble Eqs. (28) and (29) for a predetermined N. Then we obtain:

$$\overline{\Phi} = r\,F , \tag{50}$$

$$|\overline{\Theta}| = r\,|T| . \tag{51}$$

The probability distributions of Φ and $|\Theta|$ depend of r, F and $|T|$. After computing Φ and $|\Theta|$ by Eqs. (5) and (48), for any set of values r and F, we can calculate the likelihood of the outcomes being Φ and $|\Theta|$. The best estimate of F is that value $\widehat{F}$ that maximizes that likelihood. Evaluating F is similar to evaluating physical parameters from statistical distributions of real events [6].

In the Poisson process, for $\overline{N}$ large, i.e., for N large, the probability distributions for Φ and $|\Theta|$ are Gaussian. The likelihood is of the form $\exp(-\chi^2/2)$, where χ^2 is a function of the differences $\Phi - \overline{\Phi}$ and $|\Theta| - |\overline{\Theta}|$, i.e., $\Phi - r\,F$ and $|\Theta| - r\,|T|$. Let us define the vector $|\lambda_k|$ of dimension M_t+1, made of φ_k as its first component and $\vartheta_{m,k}$ as its $(m+1)^{th}$ component,

$$|\lambda_k| = \left| \begin{array}{c} \varphi_k \\ |\vartheta_k| \end{array} \right| , \tag{52}$$

and

$$|\Lambda| = \sum_k |\lambda_k| = \left| \begin{array}{c} \Phi \\ |\Theta| \end{array} \right| , \tag{53}$$

$$|L| = \left| \begin{array}{c} F \\ |T| \end{array} \right| , \tag{54}$$

$$||\epsilon|| = \sum_k |\lambda_k| \, \widetilde{|\lambda_k|} , \tag{55}$$

where the symbol ~ above a vector or a matrix indicates the transposed quantity. In the context of the Poisson process, the $(M_t + 1) \times (M_t + 1)$ matrix $||\epsilon||$ has squares of standard deviations of Φ and of the quantities Θ_m on its diagonal and correlations between them in the off diagonal terms [5]. Let us first assume that the function $f(x)$ is not a linear combination of the functions $t_m(x)$. The matrix $||\epsilon||$ is not singular. The relevant χ^2 is

$$\chi^2 = \left(\widetilde{|\Lambda|} - r\widetilde{|L|}\right) \; ||\epsilon||^{-1} \; \left(|\Lambda| - r|L|\right) . \tag{56}$$

It determines the probability distribution of $|\Lambda|$. At its minimum as a function of r and F, it yields a value $\widehat{F}$ that is the best estimate of F.

2.3 The Best Estimate of F

To minimize the χ^2 of Eq. (56), we first express the $(M_t + 1) \times (M_t + 1)$ matrices $||\epsilon||$ and $||\epsilon||^{-1}$ in a form that shows separately the elements associated with the function $f(x)$ and those associated with the reference functions $|t(x)|$. This can be done using the $M_t \times M_t$ matrix $||\epsilon_t||$ and the vector $|\Psi|$ defined as follows:

$$||\epsilon_t|| = \sum_k |\vartheta_k| \, \widetilde{|\vartheta_k|} , \tag{57}$$

$$|\Psi| = \sum_k \varphi_k |\vartheta_k| , \tag{58}$$

$$||\epsilon|| = \left\| \begin{array}{cc} \epsilon_f & \widetilde{|\Psi|} \\ |\Psi| & ||\epsilon_t|| \end{array} \right\| . \tag{59}$$

Since the functions $t_m(x)$ are linearly independent, the $M_t \times M_t$ matrix $||\epsilon_t||$ can be inverted. We define the vector $|\Pi|$ of dimension M_t,

$$|\Pi| = ||\epsilon_t||^{-1}|\Psi| \ , \tag{60}$$

and, using ϵ_f of Eq. (13),

$$\beta = \frac{1}{\epsilon_f - \widetilde{|\Pi|}\ |\Psi|} = \frac{1}{\epsilon_f - \widetilde{|\Psi|}\ ||\epsilon_t||^{-1}|\Psi|} \ , \tag{61}$$

which is > 0 and $< \infty$ because $f(x)$ is assumed not to be a linear combination of the functions $t_m(x)$.[3] Then the $(M_t + 1) \times (M_t + 1)$ inverse matrix of $||\epsilon||$ is

$$||\epsilon||^{-1} = \left\| \begin{matrix} 0 & 0 \\ 0 & ||\epsilon_t||^{-1} \end{matrix} \right\| + \beta \left| \begin{matrix} -1 \\ |\Pi| \end{matrix} \right| \left| \widetilde{\begin{matrix} -1 \\ |\Pi| \end{matrix}} \right| . \tag{62}$$

This formula is easy to check if one multiplies it with $||\epsilon||$ of Eq. (59). The result will be the identity matrix. The χ^2 of Eq. (56) can be rewritten, taking Eqs. (53), (54), and (62) into account.

$$\chi^2 = \left(\widetilde{|\Theta|} - r\widetilde{|T|}\right)\ ||\epsilon_t||^{-1}\left(|\Theta| - r|T|\right) + \beta\left(\Phi - rF - \widetilde{|\Pi|}\ \left(|\Theta| - r|T|\right)\right)^2 \tag{63}$$

This χ^2 is obviously minimum for the following values of r and F:

$$\hat{r} \ = \ \frac{\widetilde{|\Theta|}\ ||\epsilon_t||^{-1}|T|}{\widetilde{|T|}\ ||\epsilon_t||^{-1}|T|} \ , \tag{64}$$

$$\hat{F} \ = \ \frac{\Phi}{\hat{r}} - \widetilde{|\Pi|}\ \left(\frac{|\Theta|}{\hat{r}} - |T|\right) = \widetilde{|\Pi|}\ |T| + \frac{\Phi - \widetilde{|\Pi|}\ |\Theta|}{\hat{r}} \ . \tag{65}$$

Therefore Eq. (65) gives the best possible estimate $\hat{F}$ of F.

At the minimum, the value of the χ^2 is

$$\chi^2_{min} = \widetilde{|\Theta|}\ ||\epsilon_t||^{-1}|\Theta| - \hat{r}^2\ \widetilde{|T|}\ ||\epsilon_t||^{-1}|T| \ . \tag{66}$$

[3]It is indeed easy to demonstrate that

$$\tfrac{1}{\beta} = \epsilon_f - \widetilde{|\Pi|}\ |\Psi| = \sum_k \left(\varphi_k - \widetilde{|\Pi|}\ |\vartheta_k|\right)^2 > 0 \ .$$

This quantity, χ^2_{min}, allows a test of consistency between the distribution of the Monte Carlo points and the integrals of the reference functions. The confidence level that one could derive from this test is given by this minimum value χ^2_{min} interpreted as a χ^2 associated with $M_t - 1$ degrees of freedom.

The error on $\hat{r}$ and $\hat{F}$ can be obtained from the second derivative matrix of χ^2 of Eq. (63) with respect to r and F, evaluated at the minimum.

$$\frac{\partial^2\chi^2}{\partial r^2} = 2\,|\widetilde{T}|\;||\epsilon_t||^{-1}|T| + 2\beta\left(\hat{F} - |\widetilde{\Pi}|\;|T|\right)^2 \tag{67}$$

$$\frac{\partial^2\chi^2}{\partial r\partial F} = 2\beta\hat{r}\left(\hat{F} - |\widetilde{\Pi}|\;|T|\right) \tag{68}$$

$$\frac{\partial^2\chi^2}{\partial F^2} = 2\beta\hat{r}^2 \tag{69}$$

Inverting that second derivative matrix one gets diagonal elements that are approximations $\widehat{S^2}(\hat{r})$ and $\widehat{S^2}(\hat{F})$ for the squares of the one-standard deviation errors on $\hat{r}$ and $\hat{F}$. Replacing β by its expression (61) and using the definition (60) of $|\Pi|$,

$$\widehat{S^2}(\hat{r}) = \frac{1}{|\widetilde{T}|\;||\epsilon_t||^{-1}|T|}\,, \tag{70}$$

$$\widehat{S^2}(\hat{F}) = \frac{1}{\hat{r}^2}\left(\epsilon_f - |\widetilde{\Pi}|\;|\Psi| + \frac{\left(\hat{F} - |\widetilde{\Pi}|\;|T|\right)^2}{|\widetilde{T}|\;||\epsilon_t||^{-1}|T|}\right) \tag{71}$$

$$= \frac{1}{\hat{r}^2}\left(\epsilon_f - |\widetilde{\Psi}|\;||\epsilon_t||^{-1}|\Psi| + \widehat{S^2}(\hat{r})\left(\frac{\Phi - |\widetilde{\Psi}|\;||\epsilon_t||^{-1}|\Theta|}{\hat{r}}\right)^2\right) \tag{72}$$

These formulae give a practical way to estimate errors.[4]

[4]Averaging over the possible outcomes, one gets the standard deviations for these errors. For N large, they are very close to the values obtained if $\hat{r}$, $\hat{F}$, ϵ_f, $|\Psi|$, and $||\epsilon_t||$ are replaced by their expectation values r, F, $\overline{\epsilon_f}$, $|\overline{\Psi}|$, and $||\overline{\epsilon_t}||$. In particular,

$$S^2(\hat{F}) \cong \frac{1}{r^2}\left(\overline{\epsilon_f} - |\widetilde{\overline{\Psi}}|\;||\overline{\epsilon_t}||^{-1}|\overline{\Psi}| + \frac{\left(F - |\widetilde{\overline{\Psi}}|\;||\overline{\epsilon_t}||^{-1}|T|\right)^2}{|\widetilde{T}|\;||\overline{\epsilon_t}||^{-1}|T|}\right).$$

Note that, if we have only one reference function, then $||\epsilon_t||^{-1}$ has just one element and Eq. (64) reduces to

$$\hat{r} = \frac{\Theta}{T} \ . \tag{73}$$

Then, it is easy to verify that Eq. (65) reduces to Eq. (27) and Eq. (72) reduces to Eq. (33). The procedure and formulae developed in Sect. 1.3 are just a particular case of the general procedure described here. No information is conveyed by the value of χ^2_{min}, which then is zero, for $M_t - 1 = 0$ degrees of freedom. If the single reference function is a constant and if $A = 1$, the procedure and formulae further reduce to the ones of Sect. 1.1.

2.4 The Least Square Fitted Reference Function

To understand the role played by the reference functions and to generate guidelines as how to choose them, we introduce another function, $t^{(0)}(x)$, defined as that linear combination of reference functions $t_m(x)$ that minimizes the quantity

$$H = \int \left[f(x) - t^{(0)}(x)\right]^2 a(x)dx \ . \tag{74}$$

Defining a vector $|P|$ of dimension M_t having components P_m equal to the coefficients of a linear combination of reference functions, we can express $t^{(0)}(x)$ in the form

$$t^{(0)}(x) = \widetilde{|P|}\ |t(x)| \ . \tag{75}$$

The coefficients $|P|$ of $t^{(0)}(x)$ are those minimizing H of Eq. (74). The function $t^{(0)}(x)$ is a least square approximation of the function $f(x)$.

The integral H of Eq. (74) can be approximated by the sum

$$\begin{aligned} \hat{H} = \frac{1}{\hat{r}} \sum_k \left[\varphi_k - t^{(0)}(\xi_k)\right]^2 &= \frac{1}{\hat{r}} \sum_k \left(\varphi_k^2 - 2\varphi_k \widetilde{|P|}\ |\vartheta_k| + \widetilde{|P|}\ |\vartheta_k|\ \widetilde{|\vartheta_k|}\ |P|\right) \\ &= \frac{1}{\hat{r}} \left(\epsilon_f - 2\widetilde{|P|}\ |\Psi| + \widetilde{|P|}\ ||\epsilon_t||\ |P|\right) \ , \end{aligned} \tag{76}$$

using Eqs. (13), (58), and (57). The coefficients $|P|$ are approximately equal to the coefficients $|\hat{P}|$ which make $\hat{H}$ minimum, i.e.,

$$|\hat{P}| = ||\epsilon_t||^{-1}|\Psi| = |\Pi| \ , \tag{77}$$

using the definition of Eq. (60) for $|\Pi|$. It follows that the function $t^{(0)}(x)$ that minimizes H of Eq. (74) is close to

$$\widehat{t^{(0)}}(x) = |\widetilde{\Pi}|\ |t(x)|\ . \tag{78}$$

The quantity $\widehat{H}$ is a sum of squares of differences between $f(x)$ and $\widehat{t^{(0)}}(x)$. Minimizing $\widehat{H}$ amounts also to some least square fit of $\widehat{t^{(0)}}(x)$ to $f(x)$. Like $t^{(0)}$, the function $\widehat{t^{(0)}}(x)$ could be considered as one of the linear combinations of reference functions $t_m(x)$ that approximates the function $f(x)$ best.

Using the integral of $\widehat{t^{(0)}}(x) \cdot a(x)$,

$$\widehat{T^{(0)}} = |\widetilde{\Pi}|\ |T|\ , \tag{79}$$

and the sum over the events,

$$\Theta^{(0)} = \sum_k \widehat{t^{(0)}}(\xi_k) = \sum_k |\widetilde{\Pi}|\ |\vartheta_k| = |\widetilde{\Pi}|\ |\Theta|\ , \tag{80}$$

the estimate $\widehat{F}$ of Eq. (65) can be written as the sum of two terms:

$$\widehat{F} = \widehat{T^{(0)}} + \frac{\Phi - \Theta^{(0)}}{\widehat{r}}\ . \tag{81}$$

At its minimum, $\widehat{H}$ takes the value obtained if $|\widehat{P}|$ of Eq. (77) is plugged instead of $|P|$ into Eq. (76).

$$\widehat{H} = \frac{1}{\widehat{r}}\ \left(\epsilon_f - |\widetilde{\Pi}|\ |\Psi|\right) \tag{82}$$

It can be used in an expression of the square of the error on $\widehat{F}$ obtained from Eq. (72),

$$\widehat{S^2}(\widehat{F}) = \frac{1}{\widehat{r}}\ \widehat{H} + \frac{\widehat{S^2}(\widehat{r})}{\widehat{r}^2}\ \left(\frac{\Phi - \Theta^{(0)}}{\widehat{r}}\right)^2 \tag{83}$$

$$\cong \frac{H}{r} + \frac{S^2(\widehat{r})}{r^2}\ \left(F - \widehat{T^{(0)}}\right)^2\ . \tag{84}$$

In Sect. 1.2, we considered integrals of the type of Eqs. (37) and (38) where the function $a(x)$ was the same as the Monte Carlo probability distribution $p^{MC}(x)$. Thus A of Eq. (36) was equal to 1. When estimates of the type described in Sect. 2.3 are made in cases where $A = 1$, Eq. (49) predicts

$$\widehat{r} \cong N\ . \tag{85}$$

Then Eq. (81) is equivalent to Eq. (18) if we use the least square fitted reference function $t^{(0)}(x)$ of Eq. (75) as the sole reference function in Eq. (18). The procedure leading to the estimate of Eq. (65) in Sect. 2.3 amounts to a least square fit of a linear combination of $t_m(x)$ to get $t^{(0)}(x)$, followed by an estimation of F according to the procedure of Sect. 1.2 involving only the reference function $t^{(0)}(x)$. However, in Sect. 1.2, the estimated error of Eq. (20) is smaller than $\widehat{H}$ of Eq. (76) divided by N, while, in Sect. 2.3, the estimated error is larger than $\widehat{H}/N$, as can be seen from Eq. (83). This discrepancy is not surprising because of the contribution of the fit to the error on $\widehat{F}$, which was implicitly taken into account in the procedure of Sect. 2.3. Note that the discrepancy between the two errors vanishes if, in the set of reference functions considered in Sect. 2.3, a constant is one of the reference functions used.[5]

In any case, there are two terms in the general expression of the error given by Eq. (84). The first term is equal to the integral of Eq. (74) (which expresses how far $f(x)$ and $\widehat{t^{(0)}}(x)$ are from each other) divided by the coefficient r (which is proportional to $\overline{N}$). The second of the two terms results from the uncertainty about the scaling factor r. This uncertainty multiplies the difference between the two integrals of $f(x)$ and of $\widehat{t^{(0)}}(x)$ times $a(x)$. Both terms decrease when there is a possibility to define a function $\widehat{t^{(0)}}(x)$ that fit $f(x)$ better. Therefore it is advantageous to use reference functions that define a linear combination that, at most of the Monte Carlo points, approximate the function $f(x)$ as well as possible. This property provides a guideline as how to choose reference functions. Note that reference functions

[5]Let C be a constant used as one of the reference functions. In the matrix $||\epsilon_\ell||$, the column and the row corresponding to the constant reference function is a vector equal to $C|\Theta|$. Thus the vector $||\epsilon_\ell||^{-1}|\Theta|$ has all components zero but the one corresponding to the constant reference function. That component is $1/C$ in the vector $||\epsilon_\ell||^{-1}|\Theta|$ and $C\Phi$ in $|\Psi|$. It follows that the quantity $\Theta^{(0)}$ of Eq. (80) is equal to Φ. Equation (83) of Sect. 2.3 reduces to

$$\widehat{S^2}(\widehat{F}) = \widehat{H}/\widehat{r} \cong (A/N)^2 \sum_k \left[\varphi_k - t^{(0)}(\xi_k)\right]^2 .$$

with little correlation to $f(x)$ will not be so useful but, as long as we deal with many more Monte Carlo points than reference functions, i.e., as long as

$$N \gg M_t \ , \tag{86}$$

use of irrelevant reference functions will not hurt.

There is another consequence to the fact that estimates $\widehat{F}$ of Eq. (65) and its error, Eq. (72), can be expressed as in Eqs. (81) and (83) respectively, i.e., as functions of only $f(x)$ and $\widehat{t^{(0)}}(x)$. The function $t^{(0)}(x)$ is defined as the linear combination of $t_m(x)$ that minimizes H of Eq. (74), i.e., approximately, $\widehat{H}$ of Eq. (76). It does not change if a set of functions $t_m(x)$ is replaced by a set of linear combinations of $t_m(x)$. Therefore our procedure is invariant with respect to a linear transformation of the set of reference functions. In our choice of reference functions, we should be guided only by the desire to have a set of functions $t_m(x)$ that could define linear combinations that fit $f(x)$ best for most Monte Carlo points. The actual representation of the set has no impact on the result or its error. The simplest representation will work as well as a complicated one. For instance, powers of any relevant variable will have the same effect as Legendre– or more sophisticated polynomials.

3 Generalizations and Practical Formulae

The formulae of Sect. 2 can be generalized and written in a compact form to make handling of the actual computation easier. Spelling out and justifying such general and compact formulae are the subjects of this last part of the paper.

3.1 Integrating Linear Combinations of Reference Functions

Since it is advantageous to have reference functions $t_m(x)$ such that a linear combination of them fit the function $f(x)$ well, it is important that one considers the limit where $f(x)$ *actually is* a linear combination of $t_m(x)$. That case was excluded from our general development because, to establish

the general formulae, we considered a χ^2 as in Eq. (56). This χ^2 could be formulated only if the matrix $||\epsilon||$ of Eq. (55) could be inverted and this was possible only if $f(x)$ was not a linear combination of the functions $t_m(x)$. However, if $f(x)$ is such a linear combination, the final formulae, $\widehat{F}$ of Eq. (65) and $\widehat{S^2}(\widehat{F})$ of Eq. (72), can still be computed. The linear combination $\widehat{t^{(0)}}(x)$ that minimizes $\widehat{H}$ of Eq. (76) is the one that makes $\widehat{H}$ zero, i.e., the one for which

$$f(x) = \widehat{t^{(0)}}(x) \; . \tag{87}$$

Therefore, Φ and $\Theta^{(0)}$ are equal. Thus $\widehat{F}$ gives the exact result:

$$\widehat{F} = \widehat{T^{(0)}} = F \; . \tag{88}$$

This fact is confirmed by Eq. (83), which gives

$$\widehat{S^2}(\widehat{F}) = 0 \; . \tag{89}$$

If $f(x)$ is a linear combination of the $t_m(x)$, the above procedure is equivalent to recognizing the coefficients of that linear combination by a minimum square fit to the function $f(x)$ at each Monte Carlo point, and then to computing the exact integral corresponding to that linear combination. The error is zero. This procedure gives certainly the most accurate result. Thus the estimate of F by Eq. (65) and of the error by Eq. (72) can be generalized to cases where $f(x)$ is a linear combination of the $t_m(x)$. There is no reason to limit the search toward getting reference functions with a linear combination that fits $f(x)$ best. If it fits perfectly, the same procedure applies and that special circumstance can be recognized because the estimated error will turn out to be zero.

Since the result of integration is exact in the case of linear combinations, there is another interesting consequence. In this case, the result does not depend on the distribution of integration points in the integration domain. All that is required for the procedure to be valid is that the set of points generated be adequate for determining the coefficients of the development of $f(x)$ in terms of a linear combination of $t_m(x)$. That requirement means

that the set of points should give linearly independent values $\vartheta_{m,k}$, i.e., such that the matrix $||\epsilon_t||$ of Eq. (57) be not singular. In this context, at least M_t but in general not more than M_t points of integration are needed. If $||\epsilon_t||$ is not singular, any set of integration points generated by any mean is adequate. Each set of points corresponds to an integration method, which gives an exact result for linear combinations of reference functions.

Of course the case of $f(x)$ being a linear combination of the functions $t_m(x)$ is just an extreme case. There is a continuity in the behavior of the properties of integrals evaluated using Eq. (65), as one consider functions $f(x)$ that are approximated better and better by a linear combination of reference functions. Since the integrals are insensitive to the distribution of Monte Carlo points when $f(x)$ *is* a linear combination of $t_m(x)$, they will be only slightly sensitive to that distribution if $f(x)$ *is almost* a linear combination. Increasing the number of reference functions will make it possible to approximate $f(x)$ by a linear combination better. It will make the result less sensitive to the distribution of Monte Carlo points, thus to statistical fluctuations *and also* to some types of biases in the generation of these points. Reference functions may reduce not only the statistical errors on the integrals, but also some of the systematic errors due to some of possible flaws in the Monte Carlo generation. That property is illustrated in the next paper [3].

Note that the property of giving an exact result for *some* class of integrands is also a property of all numerical integrations: polynomials for the Gaussian method, trapezoids for the trapezoidal rule, etc. For the procedure advocated here, it occurs when $f(x)$ is a linear combination of $t_m(x)$. On the contrary, if reference functions have little to do with the shape of $f(x)$, this procedure reduces to the usual Monte Carlo technique. In this sense, one can say that reference functions bridge the gap between numerical and ordinary Monte Carlo integration techniques.

3.2 Mathematics for Computers

Expressions (65) and (72) of the estimates $\widehat{F}$ of F and $\widehat{S^2}(\widehat{F})$ of the square of its error were convenient for discussing properties of reference functions. There are more suitable forms to perform computations in computers [7]. Assuming one has to evaluate several integrals with the same Monte Carlo points, it is advantageous to separate terms that do not depend on the function $f(x)$ to be integrated from the terms that do depend on $f(x)$. The former can be computed once for all integrals and not every time, if computer time is to be saved.

One can reformulate Eqs. (65) and (72), replacing the vector $|\Pi|$ by its expression from Eq. (60). After computing $|\Theta|$ by the sum (48), $||\epsilon_t||$ by the sum (57), and $|T|$ by the closed form that expresses the integrals of the reference functions, we can calculate the following quantities:

$$\tau = \widetilde{|\Theta|}\ ||\epsilon_t||^{-1}|T| \ , \tag{90}$$

$$\kappa = \frac{1}{\widehat{r}} = \frac{\widetilde{|T|}\ ||\epsilon_t||^{-1}|T|}{\tau} \ , \tag{91}$$

$$|\Delta| = ||\epsilon_t||^{-1}\left(\kappa|\Theta| - |T|\right) \ , \tag{92}$$

$$\chi^2_{min} = \frac{\widetilde{|\Theta|}\ |\Delta|}{\kappa} \ . \tag{93}$$

The terms $|\Theta|$, $||\epsilon_t||$, and $|T|$ can be computed without the knowledge of $f(x)$. They could be computed during the generation of Monte Carlo points. Then Eqs. (90), (91), (92), and (93), which do not depend on $f(x)$ either, can be computed just after the generation is over. They are terms that do not have to be computed again every time a new integral involving a new function $f(x)$ has to be evaluated. Furthermore the value of χ^2_{min} provides a test of consistency of the distribution of Monte Carlo points and of the integrals involving the reference functions, if interpreted as a χ^2 corresponding to $M_t - 1$ degrees of freedom. That test is best performed just after the Monte Carlo generation.

When the integrals of functions $f(x)$ have actually to be evaluated, then it is time to compute Φ of Eq. (5), $|\Psi|$ of Eq. (58) and, if an error on

the integral is desired, ϵ_f of Eq. (13). The computation of these quantities requires one pass over the events. At the end, $\widehat{F}$ of Eq. (65) can be computed. Using the quantities defined in this section, Eq. (65) reduces to

$$\widehat{F} = \kappa\ \Phi - \widetilde{|\Psi|}\ |\Delta|\ . \tag{94}$$

If an estimate of the error is also needed, $\widehat{S^2}(\widehat{F})$ can be calculated. Equation (72) reduces to

$$\begin{aligned} \widehat{S^2}(\widehat{F}) &= \kappa^2 \left(\epsilon_f - \widetilde{|\Psi|}\ ||\epsilon_t||^{-1}|\Psi|\right) + \frac{\kappa^3}{\tau}\left(\Phi - \widetilde{|\Psi|}\ ||\epsilon_t||^{-1}|\Theta|\right)^2 \quad (95) \\ &= \kappa^2 \left(\epsilon_f - \widetilde{|\Psi|}\ ||\epsilon_t||^{-1}|\Psi|\right) + \frac{\kappa}{\tau}\left(\widehat{F} - \widetilde{|\Psi|}\ ||\epsilon_t||^{-1}|T|\right)^2 . \quad (96) \end{aligned}$$

An interesting case is one where a constant can be and *actually is* used as one of the reference functions. The general formulae of this section still apply but can be simplified.[5,6] However the simplification may not warrant using a special procedure just for this case.

3.3 Several Functions to Be Integrated

A rather common situation is one where integrals of several functions have to be computed at the same time. As examples of such circumstances, let us mention fits, where integrals of a function and of its derivatives may have to be computed at each iteration. One can construct a vector $|f(x)|$ of M_f components equal to the functions involved in the integrals in question. One defines the M_f-dimensional vectors

$$\begin{aligned} |F| &= \int |f(x)|a(x)dx = \textit{integrals to be evaluated}\,, && (97) \\ |\varphi_k| &= |f(\xi_k)|\,, && (98) \\ |\Phi| &= \sum_k |\varphi_k|\,, && (99) \end{aligned}$$

[6]In that case, Eqs. (94) and (96) reduce to the following expressions:

$$\widehat{F} = \widehat{T^{(0)}} = \widetilde{|\Psi|}\ ||\epsilon_t||^{-1}|T|\ ,\ \ \widehat{S^2}(\widehat{F}) = \kappa^2 \left(\epsilon_f - \widetilde{|\Psi|}\ ||\epsilon_t||^{-1}|\Psi|\right)\ .$$

Note that, in that particular case, the best estimate of F is equal to the integral of the least square fitted reference function $\widehat{t^{(0)}(x)}$ without additional terms.

and the $M_t \times M_f$ matrix

$$||\Psi|| = \sum_k |\vartheta_k| \, \widetilde{|\varphi_k|} \, . \tag{100}$$

Then one can obtain an estimate $|\widehat{F}|$ of the vector $|F|$ of Eq. (97), using an equation similar to Eq. (94),

$$|\widehat{F}| = \kappa \, |\Phi| - \widetilde{||\Psi||} \, |\Delta| \, . \tag{101}$$

Alternately, a computer-time saving procedure may be designed to minimize the computation to be performed for each function. Instead of computing the sums $|\Phi|$ and $||\Psi||$, the two sums can be combined to get the result $|\widehat{F}|$ directly. After the Monte Carlo points have been generated and the vector $|\Delta|$ has been calculated via Eq. (92), a weight η_k, called integration weight,[7] can be computed for each Monte Carlo point k during the next pass over the events. Knowing $|\vartheta_k|$ from Eq. (47), one obtains the integration weight

$$\eta_k = \kappa - \widetilde{|\vartheta_k|} \, |\Delta| \, . \tag{102}$$

No matter what the functions $|f(x)|$ are, the same estimate $|\widehat{F}|$ as the one of Eq. (101) can be obtained from the following equation:

$$|\widehat{F}| = \sum_k \eta_k \, |\varphi_k| \, . \tag{103}$$

If several iterations are contemplated with different integrals to be evaluated each time, the weight η_k can be stored for future use. An interesting property of the sum of the weights η_k is that it satisfies

$$\sum_k \eta_k = \kappa(N - \chi^2_{min}) \cong \kappa N \cong A \textit{ of Eq. (36)} \; . \tag{104}$$

Equation (103) shows the analogy with numerical integrations. The weight η_k associated with each point of integration k is independent of the functions to be integrated but it depends on the location of the other points.

[7] These integration weights, η_k, should not be confused with Monte Carlo weights, such as those we talk about in Sect. 3.4.

It is also interesting to note that Eq. (6) of Sect. 1.1 is of the same form as Eq. (103), but where η_k is equal to $1/N$ for all points regardless of their locations.

Eventually one may want to estimate an error matrix for the integrals $|\widehat{F}|$. Such an estimate, $||\widehat{E}||$, is a $M_f \times M_f$ matrix expressing errors and correlations between the components of the vector $|\widehat{F}|$. It can be obtained using the $M_t \times M_f$ matrix $||\Psi||$ of Eq. (100), as well as the $M_f \times M_f$ matrix

$$||\epsilon_f|| = \sum_k |\varphi_k| \, \widetilde{|\varphi_k|} \tag{105}$$

which generalizes the quantity ϵ_f of Eq. (13). We write:

$$|\Omega| = |\widehat{F}| - \widetilde{||\Psi||} \, ||\epsilon_t||^{-1} |T| \, , \tag{106}$$

$$||\widehat{E}|| = \kappa^2 \left(||\epsilon_f|| - \widetilde{||\Psi||} \, ||\epsilon_t||^{-1} ||\Psi|| \right) + \frac{\kappa}{\tau} \, |\Omega| \, \widetilde{|\Omega|} \, . \tag{107}$$

It is easy to check Eq. (107) by proving that the integral of any linear combination $f^{(1)}(x)$ of components of the vector $|f(x)|$ will have the same error if computed from Eq. (96) or using the matrix $||\widehat{E}||$ of Eq. (107). If the M_f coefficients of the linear combination $f^{(1)}(x)$ are written in a vector $|P_f|$ with M_f components,

$$f^{(1)}(x) = \widetilde{|P_f|} \, |f(x)| \, , \tag{108}$$

then Eq. (96), applied to the integral of $f^{(1)}(x) \cdot a(x)$, leads to

$$\widehat{S^2}(\widehat{F^{(1)}}) = \widetilde{|P_f|} \, ||\widehat{E}|| \, |P_f| \, . \tag{109}$$

Therefore the matrix $||\widehat{E}||$ of Eq. (107) is indeed the error matrix for the set of functions constituting the components of the vector $|f(x)|$.

3.4 Weighted Distributions — General Formulae

Consider now the case where the integrals $|F|$ to be evaluated are expressed in the form of Eq. (97) with the distribution $a(x)$ represented by a weighted Monte Carlo distribution of the type

$$a(x) = A \, w(x) \, p^{MC}(x) \, . \tag{110}$$

Here, $w(x)$ is a weighting function expressible in closed form at each integration point, $p^{MC}(x)$ is the distribution of the integration points, and A a normalization constant which may be defined either explicitly in closed form or implicitly by integrals of reference functions. We assume that the known integrals $|T|$ of the reference functions $|t(x)|$ are also expressed in the form of Eq. (46) with the distribution $a(x)$ of Eq. (110).

In Sect. 1.4, Eq. (40), we used an example of weighted distribution where, for purpose of illustration, it was convenient to incorporate the weighting function $w(x)$ into the function $f(x)$ and, implicitly, into the function $t(x)$ too. It is often more convenient to incorporate $w(x)$ into the definition of $a(x)$ instead, as it is done in Eq. (110).

To adjust our formulae to this case, all that is needed is, at the integration points, to multiply $|f(x)|$ and $|t(x)|$ by the factor $w(x)$, since it is no longer incorporated in the definition of these functions. At each point k, we define a quantity ω_k, called Monte Carlo weight,

$$\omega_k = w(\xi_k) \ , \tag{111}$$

and replace the previous definitions of $|\varphi_k|$ and $|\vartheta_k|$, i.e., Eqs. (98) and (47), by

$$|\varphi_k| = \omega_k \ |f(\xi_k)| \ , \tag{112}$$

$$|\vartheta_k| = \omega_k \ |t(\xi_k)| \ . \tag{113}$$

Then the general procedure of Sect. 3.3 can be used, first computing quantities independent of $f(x)$, i.e, $|\Theta|$, $||\epsilon_t||$, τ, κ, $|\Delta|$, and χ^2_{min} using Eq. (48), (57), (90), (91), (92), and (93), respectively. Secondly, one can compute either the integration weights η_k from Eq. (102) and the estimate $|\widehat{F}|$ of $|F|$ from Eq. (103), or $||\Psi||$ from Eq. (100) and $|\widehat{F}|$ from Eq. (101). For an estimate $||\widehat{E}||$ of the error matrix $||E||$, one needs $||\epsilon_f||$ of Eq. (105), as well as $||\Psi||$ of Eq. (100), to introduce in Eq. (107).

The procedure described here is the most general one. If the Monte Carlo distributions are not weighted, all Monte Carlo weights ω_k have to be set to 1. If there is only one function $f(x)$ to be integrated, or if there is only

one reference function $t(x)$, the relevant vectors reduce to a one-component vector. This general procedure is the one in the program described in Ref. [7].

As seen in Sect. 3.1, for integrands that are linear combinations of reference functions, the estimates of the integrals come out exact and the estimated errors are zero.

An example of use of reference functions in a Monte Carlo simulation for a physics problem is given in the next paper [3].

Acknowledgements. The authors are indebted to Dr Daniel Drijard for useful discussions on the topics of this paper and to Lisa Goddard for reviewing the manuscript.

References

1. F. James, "Monte Carlo Theory and Practice", Rep. Prog. Phys. 43 (1980) 1145; reprinted in "Experimental Techniques in High Energy Physics", T. Ferbel (ed.), Addison-Wesley, Menlo Park CA (1987).

2. Particle Data Group, "Review of Particle Properties", Phys. Lett. B239 (1990) p. III.36.

3. P.H. Eberhard and O.P. Schneider, "Example of Monte Carlo Integrals where Reference Functions are Useful", LBL–29674, Lawrence Berkeley Laboratory (1990), submitted for publication.

4. Ref. 1, pp. 1157 ff. Also:
   P.H. Eberhard, IPNL 89–1, Université de Lausanne (1989), unpublished;
   O. Schneider, D. Drijard, and P.H. Eberhard, IPNL 89–7, Université de Lausanne (1989), unpublished.

5. See for instance: W.T. Eadie, D. Drijard, F.E. James, M. Roos and B. Sadoulet, "Statistical Methods in Experimental Physics", North-Holland, Amsterdam (1971), pp. 52 ff.

6. Ref. 5, pp. 115 ff.

7. P. Eberhard and W. Koellner, "The Optime System for Fitting Theoretical Expressions", Comp. Phys. Comm. 3 (1972) 296;
   P. Eberhard and W. Koellner, "User Manual for the Optime System", Comp. Phys. Comm. 5 (1973) 163.

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
INFORMATION RESOURCES DEPARTMENT
BERKELEY, CALIFORNIA 94720