LBL-29674
Preprint

Lawrence Berkeley Laboratory

UNIVERSITY OF CALIFORNIA

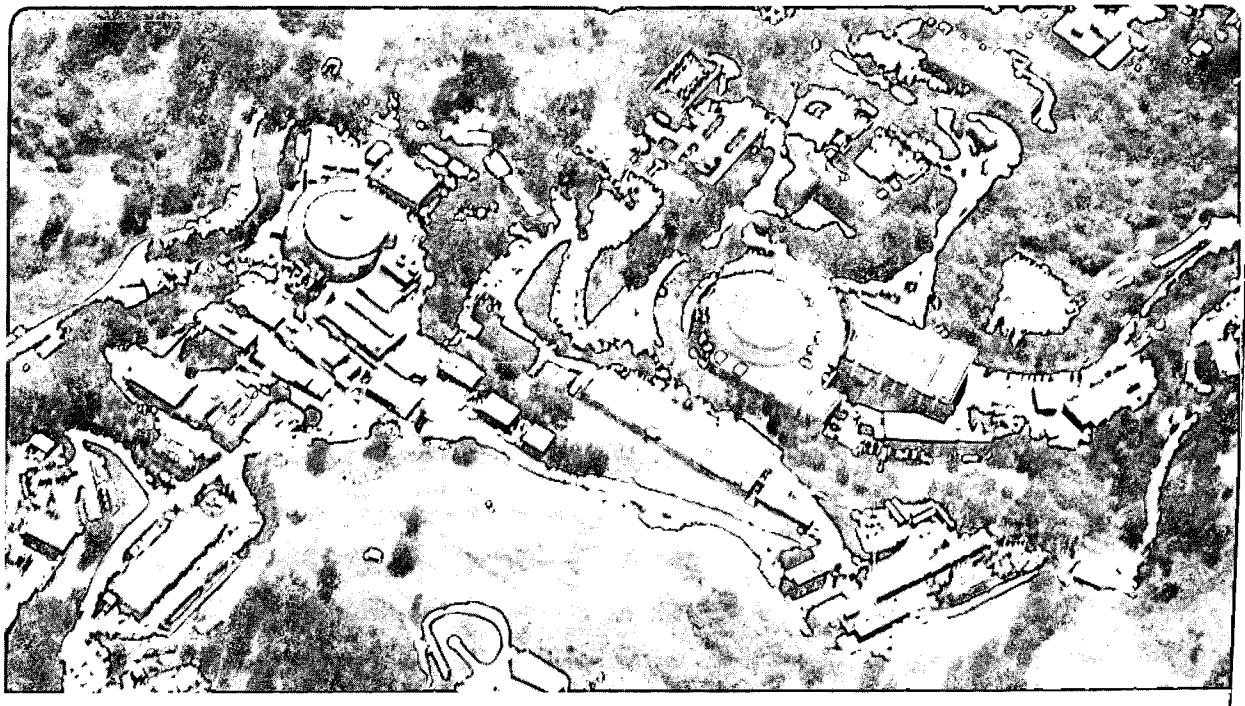
Physics Division

Submitted for publication

Example of Monte Carlo Integrals where Reference Functions are Useful

P.H. Eberhard and O.P. Schneider

October 1990



LOAN COPY
Circulates
for 2 weeks
Bldg. 50 Library.
LBL-29674
Copy 2

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–29674

Example of Monte Carlo Integrals where Reference Functions are Useful*

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.

Example of Monte Carlo Integrals where Reference Functions are Useful*

Philippe H. Eberhard
and
Olivier P. Schneider

October 4, 1990

Abstract

In the preceding paper, it has been shown that use of reference functions could decrease errors on integrals estimated by Monte Carlo. Here the statement is illustrated by a physics example. It concerns the measurement of the $\mu^+\mu^-$ forward-backward production asymmetry in e^+e^- interactions, taking radiative corrections and detector distortions into account. The parameters of the detector are chosen arbitrarily just for purpose of illustration.

The result using reference functions is about 4 times as accurate as without using them. The reference functions also make the results less sensitive to some biases in the Monte Carlo generator.

*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 The Usual Monte Carlo Integration 2

1.2 Monte Carlo Integration Using Reference Functions 3

2 Physics Considerations **4**

2.1 Muon Pair Production, Detection and Reconstruction 5

2.2 Integrals to Be Evaluated 7

2.3 Born– and Numerical Approximations 8

3 The Monte Carlo Estimates **10**

3.1 The Monte Carlo Generation 10

3.2 Choosing the Reference Functions 12

3.3 Results With and Without Use of Reference Functions 14

3.4 Systematic Errors Due to Biases in the Monte Carlo Generator 17

References **19**

Figures **20**

1 Introduction

In the preceding paper [1], it was pointed out that "reference functions" can decrease errors on Monte Carlo estimates of multidimensional integrals. Mathematics were developed to make best use of such reference functions. Here, we construct an example where these mathematics are applied and are useful. The example is a problem of the type encountered in high energy physics. It concerns the production angle θ of muon pairs produced in e^+e^- interactions, and more precisely the prediction of the average of $\cos\theta$ as measured in the detector.

The theoretical prediction of the expectation value $z = \overline{\cos\theta}$ of the cosine of θ is based on integrals of a type that has to be evaluated by Monte Carlo. In this paper, for a set of conditions chosen arbitrarily, z is estimated by the two following methods using the same number of Monte Carlo events:

1) by the usual and straightforward Monte Carlo method, i.e., applying the well known formulae of Sect. 1.1 below;

2) by the method of Ref. [1], Sect. 3.3, based on reference functions, i.e., applying the formulae recapitulated in Sect. 1.2 of this paper.

Each kind of estimation is repeated 1000 times and the results are histogramed on Fig. 1a. The averages of the estimates are the same for both methods, but the spread of their values is much narrower in the case of estimates with reference functions. This shows that, in this particular example, reference functions significantly improve the accuracy of Monte Carlo predictions for the same number of events. Alternately, for a given accuracy, the number of Monte Carlo events that have to be generated could be decreased. In addition, estimations are also made introducing a bias in the Monte Carlo generator and the results plotted on Fig. 1b. There is evidence that, in this example, estimates computed using reference functions are less affected by some types of flaw in the physics generator than the usual Monte Carlo estimates.

The purpose of this paper is not to solve a real problem of physics but only, by an example, to show the usefulness of reference functions. However

an effort has been made to make the problem resemble a real one as much as possible. For that reason, radiative corrections are reasonably well approximated in spite of their complexity. The experimental setup has distortions, finite resolution, limited acceptance, and less than 100% efficiency. Aside of that, the detector parameters and the analysis cuts are completely arbitrary. The physical problem is described in Sect. 2.

In Sect.3, details of the Monte Carlo integrals with and without reference functions are described and results are discussed.

1.1 The Usual Monte Carlo Integration

Let us suppose there are M_f integrals to be evaluated. Each one can constitute a component in a vector $|F|$ of the form

$$|F| = \int |f(x)| a(x) dx \ , \tag{1}$$

where x is a multidimensional variable, $|f(x)|$ is a vector with M_f components constituted with functions of x known in closed form, and $a(x)$ is a function of x whose shape can be simulated by a distribution of points generated by Monte Carlo. In this paper, we consider only cases where $a(x)$ is a probability distribution, normalized to unity.

$$\int a(x) dx = 1 \tag{2}$$

The Monte Carlo integration methods use the probability distribution $a(x)$ to generate the integration points. We call N the number of such points.

In the usual Monte Carlo method, estimates $|\widehat{F}|$ of the integrals $|F|$ are obtained from the formula

$$|\widehat{F}| = \frac{1}{N} \sum_k |\varphi_k| \ , \tag{3}$$

where

$$|\varphi_k| = \textit{values of } |f(x)| \textit{ at the Monte Carlo point } k \ . \tag{4}$$

An estimate $||\widehat{E}||$ of the error matrix on $|\widehat{F}|$ of Eq. (3) is given by

$$||\widehat{E}|| = \frac{1}{N^2} \, ||\epsilon_f|| - \frac{1}{N} \, |\widehat{F}| \, \widetilde{|\widehat{F}|} \ , \tag{5}$$

where

$$||\epsilon_f|| = \sum_k |\varphi_k| \, \widetilde{|\varphi_k|} \, . \tag{6}$$

1.2 Monte Carlo Integration Using Reference Functions

It has been shown in Ref. [1] that more sophisticated estimates $|\widehat{F}|$ can be defined than the ones of Eq. (3), if there are M_t linearly independent functions, which we consider as components of a vector $|t(x)|$, that can be expressed in closed form and correspond to integrals $|T|$ that are also known in closed form.

$$|T| = \int |t(x)| a(x) dx \tag{7}$$

Such functions are called "reference functions". Using reference functions, integration points are still generated by Monte Carlo as in Sect. 1.1. Then one computes

$$\begin{aligned}
|\vartheta_k| &= \textit{values of } |t(x)| \textit{ at the Monte Carlo point } k \, , && (8)\\
|\Theta| &= \sum_k |\vartheta_k| \, , && (9)\\
||\epsilon_t|| &= \sum_k |\vartheta_k| \, \widetilde{|\vartheta_k|} \, , && (10)\\
\tau &= \widetilde{|\Theta|} \, ||\epsilon_t||^{-1} |T| \, , && (11)\\
\kappa &= \frac{\widetilde{|T|} \, ||\epsilon_t||^{-1} |T|}{\tau} \, , && (12)\\
|\Delta| &= ||\epsilon_t||^{-1} \left(\kappa |\Theta| - |T| \right) \, . && (13)
\end{aligned}$$

An integration weight η_k is associated to each integration point k,

$$\eta_k = \kappa - \widetilde{|\vartheta_k|} \, |\Delta| \, , \tag{14}$$

and the integrals $|F|$ of Eq. (1) can now be estimated as

$$|\widehat{F}| = \sum_k \eta_k \, |\varphi_k| \, . \tag{15}$$

An estimate $||\widehat{E}||$ of the error matrix on $|\widehat{F}|$ of Eq. (15) is given by

$$||\widehat{E}|| = \kappa^2 \left(||\epsilon_f|| - \widetilde{||\Psi||} \, ||\epsilon_t||^{-1} ||\Psi|| \right) + \frac{\kappa}{\tau} |\Omega| \, \widetilde{|\Omega|} \, , \tag{16}$$

where

$$||\Psi|| = \sum_k |\vartheta_k| \, \widetilde{|\varphi_k|} \, , \tag{17}$$

$$|\Omega| = |\hat{F}| - \widetilde{||\Psi||} \, ||\epsilon_t||^{-1} |T| \, . \tag{18}$$

Furthermore, the quantity

$$\chi^2_{min} = \frac{\widetilde{|\Theta|} \, |\Delta|}{\kappa} \tag{19}$$

allows a test of consistency between the distribution of the Monte Carlo points and the integrals of the reference functions. It should be interpreted as a χ^2 associated with $M_t - 1$ degrees of freedom.

2 Physics Considerations

The problem concerns the study of the following reaction:

$$e^+e^- \rightarrow \mu^+\mu^- \, (\gamma) \, . \tag{20}$$

The two incoming particles have the same energy but opposite momenta. They may emit a photon (γ) before they annihilate and produce the muon pair. The detector is less than 100% efficient and distorts the measurement of the production angle θ.

At LEP energies a non-zero forward-backward asymmetry is expected in muon pair production. At any given energy, the theoretical prediction for that asymmetry depends on the Z^0 mass, M_Z, the total Z^0 width, Γ_Z, and the Z^0 coupling constants. Comparing the prediction of that asymmetry with an experimental measurement is a crucial test of the standard model [2]. The test may be done by measuring the average value $< \cos\theta >$ of $\cos\theta$ over the events detected in the experiment and accepted in the analysis, and then by comparing the measured value to the theoretical expectation value of $< \cos\theta >$, $z = \overline{\cos\theta}$, computed assuming given values for the Z^0 parameters. For non-ideal detectors, this expectation value has to be computed by Monte Carlo.

2.1 Muon Pair Production, Detection and Reconstruction

We assume that each real event is described, at least approximately, by the following four-step process:

a) Initial state radiation. One of the two electrons emits a photon. The now virtual e^+e^- system has a mass $\sqrt{s^*}$ and a non-zero momentum, both dependent on the photon energy E_γ. The set of variables necessary to describe the virtual system will be referred to as ξ^{CMS}. The probability distribution of ξ^{CMS} is a function $p^{rad}(x^{CMS})$ of the possible values x^{CMS} of ξ^{CMS}, normalized to unity. It is approximately described by a $1/E_\gamma$ spectrum multiplied by the Z^0 Breit-Wigner shape, between two values $E_\gamma^{(min)}$ and $E_\gamma^{(max)}$ which can be derived from Ref. [3].

$$p^{rad}(x^{CMS}) \quad \propto \quad \frac{1}{E_\gamma} \cdot \frac{1}{(s^* - M_Z^2)^2 + M_Z^2\Gamma_Z^2} \tag{21}$$

b) Muon pair production. A μ^+ and a μ^- leptons are created. Let θ^* be the production angle defined as the angle between e^- (after radiation) and μ^- in the virtual e^+e^- center of mass system described in a); θ^* is also the angle between e^+ (after radiation) and μ^+ in the same center of mass system. The μ^+ and μ^- four-momenta in the laboratory frame are totally defined by ξ^{CMS} and θ^*. For each value x^{CMS} of ξ^{CMS}, the probability distribution of $\cos\theta^*$, $p^{prod}_{x^{CMS}}(\cos\theta^*)$,[1] normalized to unity, is given by

$$p^{prod}_{x^{CMS}}(\cos\theta^*) = \frac{3}{8}(1 + \cos^2\theta^*) + A_{FB}(x^{CMS})\cos\theta^* \ , \tag{22}$$

[1] In the notation of step a) above, as it is proper in probability theory and as it is done in Ref. [1], a clear distinction is made between the random variables such as ξ^{CMS} and the non-random quantities such as the values x^{CMS} that the random variables ξ^{CMS} can take. Here, it was feared that keeping up with this convention entirely may obscure the physical content of the example. Therefore, as is generally done in physics papers, we sometimes use the same symbol for a random variable and the values that this random variable can take.

where $A_{FB}(x^{CMS})$ is the Born expression of the forward-backward asymmetry of muon pair production at an energy $\sqrt{s^*}$. The coefficient $A_{FB}(x^{CMS})$ depends on s^* and the Z^0 parameters, and it is known in closed form [3].

c) Particle detection and reconstruction. Each muon, unless it misses the detector completely, is scattered by the pieces of equipment it traverses (we assume scattering in the detector to be multiple scattering). Some components may react and register the muon as it passes through. If detection is successful, reconstruction of the event is attempted, leading to measurement errors and, sometimes, failure to be reconstructed. Let $\theta_\pm$ be the measured angle (in the laboratory frame) between the $e^\pm$ and $\mu^\pm$ trajectories. Because of initial state radiation, detector distortions, and measurement errors, these angles θ_+ and θ_- may be different from the production angle θ^* in the center of mass and different from each other. As an approximation θ to the unknown angle θ^*, one can use many expressions; here we compute θ as the angle formed by the difference between the μ^- and the μ^+ momenta with respect to the incident e^- in the laboratory.

d) Event analysis. If the event reconstruction is successful, then experimental cuts are applied in order to reduce various undesired effects. We assume here that only two such cuts have a non negligible effect on muon pair events. The first cut limits $\cos\theta$ to a range of values where the acceptance is close to 100% and where $\cos\theta$ is not too far from $\cos\theta^*$; it should reduce the importance of detector distortions. The second one is a collinearity cut which requires the two angles θ_+ and θ_- not to be too different from each other; it should reduce the background. These cuts are expressed as

$$-C_{lim} < \cos\theta < C_{lim} \ , \tag{23}$$

$$-B_{lim} < \theta_+ - \theta_- < B_{lim} \ , \tag{24}$$

where C_{lim} and B_{lim} are values chosen by the experimenter after studying his detector and the background. Trial values are given in Table 1. For each event, we define an index λ equal to 1 if the event is detected, reconstructed and passes all experimental cuts, and equal to 0 otherwise. For each value x^{CMS} of ξ^{CMS} and for each value of $\cos\theta^*$, there is a probability distribution of λ and $\cos\theta$, $p^{apparatus}_{x^{CMS},\cos\theta^*}(\lambda,\cos\theta)$,[1] which takes into account the behavior of the apparatus, the reconstruction procedure and the analysis cuts. It satisfies

$$\sum_{\lambda=0}^{1}\int_{-1}^{1} p^{apparatus}_{x^{CMS},\cos\theta^*}(\lambda,\cos\theta)\ d(\cos\theta) = 1\ , \tag{25}$$

whatever the values of x^{CMS} and $\cos\theta^*$ are.

The physical probability distribution is

$$\begin{aligned} &p^{phys}(\lambda,\cos\theta,\cos\theta^*,x^{CMS}) = \\ &\qquad p^{rad}(x^{CMS})\ p^{prod}_{x^{CMS}}(\cos\theta^*)\ p^{apparatus}_{x^{CMS},\cos\theta^*}(\lambda,\cos\theta)\ . \end{aligned} \tag{26}$$

2.2 Integrals to Be Evaluated

The expectation value z of $\cos\theta$ observed in our hypothetical experiment may be written as

$$z = \overline{\cos\theta} = \frac{F_1}{F_2}\ , \tag{27}$$

where

$$F_1 = \sum_\lambda \int \lambda\ \cos\theta\ p^{phys}(\lambda,\cos\theta,\cos\theta^*,x^{CMS})\ dx^{CMS}\ d(\cos\theta^*)\ d(\cos\theta)\ , \tag{28}$$

$$F_2 = \sum_\lambda \int \lambda\ p^{phys}(\lambda,\cos\theta,\cos\theta^*,x^{CMS})\ dx^{CMS}\ d(\cos\theta^*)\ d(\cos\theta)\ . \tag{29}$$

To compute $z = \overline{\cos\theta}$, it is thus necessary to know the integrals F_1 and F_2. As in Sects. 1.1 and 1.2, it is convenient to handle them as components of a vector

$$|F| = \begin{vmatrix} F_1 \\ F_2 \end{vmatrix}\ . \tag{30}$$

From estimates $|\widehat{F}|$ of $|F|$, one computes an estimate $\widehat{z}$ of z as

$$\widehat{z} = \frac{\widehat{F_1}}{\widehat{F_2}} \,. \tag{31}$$

From an estimate $||\widehat{E}||$ of the error matrix on $|\widehat{F}|$, one computes an estimate $\widehat{S^2}(\widehat{z})$ of the square of the error on $\widehat{z}$ as

$$\widehat{S^2}(\widehat{z}) = \widetilde{|\Gamma|}\ ||\widehat{E}||\ |\Gamma| \,, \tag{32}$$

where $|\Gamma|$ is the gradient of z at $F_1 = \widehat{F_1}$ and $F_2 = \widehat{F_2}$.

$$|\Gamma| = \left| \begin{array}{c} \partial z/\partial F_1 \\ \partial z/\partial F_2 \end{array} \right|_{\text{at } |F| = |\widehat{F}|} = \left| \begin{array}{c} \widehat{z}/\widehat{F_1} \\ -\widehat{z}/\widehat{F_2} \end{array} \right| \tag{33}$$

2.3 Born– and Numerical Approximations

Let us first consider the ideal case where radiative corrections are negligible and where, in a given range of $\cos\theta^*$, one can consider both measurement distortions to be zero and detection efficiency (i.e., all detection, reconstruction and collinearity cut efficiencies) to be 100%. Let us choose C_{lim} to be the maximum value of $\cos\theta^*$ in this range, which is then defined by the inequalities

$$-\,C_{lim} < \cos\theta^* < C_{lim} \,. \tag{34}$$

Our hypothesis implies that, in this range, $\cos\theta$ is equal to $\cos\theta^*$ and that events satisfying inequalities (34) are the same as those that satisfy inequalities (23). They are the accepted events. In that ideal case, the expectation value, z, of $\cos\theta$ over these accepted events is indistinguishable from the expectation value, $\overline{\cos\theta^*}$, of $\cos\theta^*$ over the range defined by inequalities (34). Let us define the function $\ell(\cos\theta^*)$ to be 1 when inequalities (34) are satisfied and 0 otherwise. In cases that can be approximated by our ideal case, we can use estimates $\widehat{z}$ of z such as

$$\widehat{z} = \overline{\cos\theta^*} = \frac{F_1^*}{F_2^*} \,, \tag{35}$$

where

$$|F^*| = \left| \begin{array}{c} F_1^* \\ F_2^* \end{array} \right| = \int \left| \begin{array}{c} \cos\theta^* \\ 1 \end{array} \right| \; p^{rad}(x^{CMS})\; p^{prod}_{x^{CMS}}(\cos\theta^*)\; \ell((\cos\theta^*) \; dx^{CMS}\; d(\cos\theta^*)\;. \tag{36}$$

Since radiative corrections are also negligible, the integral of Eq. (36) extends only to one set of values of the virtual e^+e^- system characteristics x^{CMS}, i.e., to the one corresponding to no photon emitted by the incident electrons at step a) of Sect. 2.1. In the integral of Eq. (36), the function $p^{prod}_{x^{CMS}}(\cos\theta^*)$ defined in Eq. (22) takes values corresponding to only one value of the parameter $A_{FB}(x^{CMS})$, A_{FB}^{Born}, i.e., the Born expression of A_{FB} corresponding to that case where the energy $\sqrt{s^*}$ of the virtual e^+e^- system after step a) is the same as the initial state energy $\sqrt{s}$. The integrals $|F^*|$ become integrals of polynomials of $\cos\theta^*$; they can be carried out in closed form.

$$|F^*| = \int \left| \begin{array}{c} \cos\theta^* \\ 1 \end{array} \right| \left[\frac{3}{8}(1+\cos^2\theta^*) + A_{FB}^{Born}\cos\theta^*\right] \ell(\cos\theta^*) d(\cos\theta^*) \tag{37}$$

$$= \left| \begin{array}{c} \frac{2}{3}C_{lim}^3 A_{FB}^{Born} \\ \frac{1}{4}C_{lim}(3+C_{lim}^2) \end{array} \right| \tag{38}$$

According to Eq. (35), this leads to the following estimate $\hat{z}^{(Born)}$ of $\overline{\cos\theta}$:

$$\hat{z}^{(Born)} = \overline{\cos\theta^*} = \frac{8}{3+9/C_{lim}^2}\; A_{FB}^{Born}\;. \tag{39}$$

We will refer to this estimate as the Born-approximation. Since the functions to be integrated in this approximation are easily integrable polynomials of $\cos\theta^*$, they give us a hint that powers of $\cos\theta^*$ may have the properties recommended in Ref. [1], Sect. 2.4, for efficient reference functions.

One may want to consider a little more sophisticated case, where radiative corrections are not negligible but the detector can still be considered as an ideal detector, with 100% efficiency and without significant measurement distortions. Equations (35) and (36) are still valid. The integration

over $\cos\theta^*$ can be carried out analytically in Eq. (36), which then reduces to an integration over x^{CMS}. One is led to the following estimate $\hat{z}^{(num)}$ of $\overline{\cos\theta}$:

$$\hat{z}^{(num)} = \overline{\cos\theta^*} = \frac{8}{3 + 9/C_{lim}^2}\,\overline{A_{FB}}\ , \tag{40}$$

where

$$\overline{A_{FB}} = \int A_{FB}(x^{CMS})\, p^{rad}(x^{CMS})\, dx^{CMS}\ . \tag{41}$$

The quantity $A_{FB}(x^{CMS})$ is a function of the single variable s^*. The distribution of s^* due to radiative corrections is easy to deduce in closed form from Eq. (21). Therefore $\overline{A_{FB}}$ given by Eq. (41), thus $\hat{z}^{(num)}$ of Eq. (40), can be obtained integrating over the single variable s^* and thus can be evaluated numerically using standard numerical procedures. We will refer to the estimate $\hat{z}^{(num)}$ as the numerical approximation.

3 The Monte Carlo Estimates

Of course, for best accuracy, the integrals of Eqs. (28) and (29) have to be evaluated making no approximation about the shortcomings of the detector. Multivariable integrations have to be performed, generally without known closed form expressions for the result. Numerical techniques become impractical because of the prohibitive number of points required in numerical multidimensional integrals. Therefore Monte Carlo integrations have to be performed.

3.1 The Monte Carlo Generation

For the problem at hand, two different kinds of generation of Monte Carlo points have been considered to get estimates of the integrals $|F|$ of Eqs. (28) and (29):

u) The natural and usual way: to generate events according to a probability distribution $p_{(u)}^{MC}$ that approximates the distribution of the real

events as closely as possible. Using Eq. (26),

$$\begin{aligned} p^{MC}_{(u)}(\lambda, \cos\theta, \cos\theta^*, x^{CMS}) &= p^{phys}(\lambda, \cos\theta, \cos\theta^*, x^{CMS}) \\ &= p^{rad}(x^{CMS})\; p^{prod}_{x^{CMS}}(\cos\theta^*)\; p^{apparatus}_{x^{CMS},\cos\theta^*}(\lambda, \cos\theta)\;. \end{aligned} \tag{42}$$

In that case we express the integrals $|F|$ of Eqs. (28) and (29) as

$$|F| = \sum_{\lambda} \int |f^{(u)}(\ldots)|\; p^{MC}_{(u)}(\ldots)\; d(\cos\theta)\; d(\cos\theta^*)\; dx^{CMS}\;, \tag{43}$$

where

$$|f^{(u)}(\lambda, \cos\theta, \cos\theta^*, x^{CMS})| = \lambda \cdot \begin{vmatrix} \cos\theta \\ 1 \end{vmatrix}\;. \tag{44}$$

v) The simplified distribution: Monte Carlo events are generated with a probability distribution of ξ^{CMS} and with detector distortions as expected for the real events, but the distribution of $\cos\theta^*$ is chosen uniform between -1 and $+1$ at production time, i.e., at stage b) of the process described in Sect. 2.1, for each value of s^*. Therefore the overall distribution, $p^{MC}_{(v)}$, is different from the one of the real events.

$$\begin{aligned} p^{MC}_{(v)}(\lambda, \cos\theta, \cos\theta^*, x^{CMS}) &= \frac{p^{phys}(\lambda, \cos\theta, \cos\theta^*, x^{CMS})}{2p^{prod}_{x^{CMS}}(\cos\theta^*)} \\ &= p^{rad}(x^{CMS})\; \frac{1}{2}\; p^{apparatus}_{x^{CMS},\cos\theta^*}(\lambda, \cos\theta) \end{aligned} \tag{45}$$

One advantage of the simplified distribution is that it will simplify the computations with reference functions, because the Monte Carlo integrals $|\Theta|$ of the reference functions in case v) will involve a simpler expression of the vector $|\vartheta_k|$. Another advantage is that, unlike the function $p^{MC}_{(u)}(\ldots)$ of Eq. (42), the function $p^{MC}_{(v)}(\ldots)$ of Eq. (45) does not depend on the Z^0 coupling constants. Then there is no need to generate a new sample of Monte Carlo events for each value of the Z^0 coupling constants. In this case v), we rewrite $|F|$ as

$$|F| = \sum_{\lambda} \int |f^{(v)}(\ldots)|\; p^{MC}_{(v)}(\ldots)\; d(\cos\theta)\; d(\cos\theta^*)\; dx^{CMS}\;, \tag{46}$$

where

$$|f^{(v)}(\lambda, \cos\theta, \cos\theta^*, x^{CMS})| = \lambda \cdot \left| \begin{array}{c} \cos\theta \\ 1 \end{array} \right| \cdot 2\, p^{prod}_{x^{CMS}}(\cos\theta^*)\ . \quad (47)$$

The factor $2p^{prod}_{x^{CMS}}(\cos\theta^*)$ is the factor that takes into account the difference between the distributions of Monte Carlo and the one of the real events.

To demonstrate how the integrals $|F|$ of Eqs. (43) or (46) can be computed by Monte Carlo (using the events generated according to $p^{MC}_{(u)}(\ldots)$ or $p^{MC}_{(v)}(\ldots)$ respectively), we show below the relevant correspondences between the terms of their definition and the terms in the integrals of Eq. (1).

$$\begin{array}{lcl} \int dx & \longleftrightarrow & \sum_\lambda \int d(\cos\theta)\ d(\cos\theta^*)\ dx^{CMS} \\ a(x) & \longleftrightarrow & p^{MC}_{(u)}(\ldots) \quad or \quad p^{MC}_{(v)}(\ldots) \end{array} \quad (48)$$

In both cases u) and v), estimates $|\widehat{F}|$ of $|F|$ can be obtained using the usual Monte Carlo method described in Sect. 1.1.

3.2 Choosing the Reference Functions

If a suitable set of reference functions can be found, the integration method of Sect. 1.2 may be applied too. However, one can show that this procedure applied with any one single reference function ends up with the following expressions for the estimate $\widehat{z}$ of $z = \overline{\cos\theta}$ and the estimate $\widehat{S^2}(\widehat{z})$ of the square of the error on $\widehat{z}$:

$$\widehat{z} = \frac{\sum_k \varphi_{k,1}}{\sum_k \varphi_{k,2}}\ , \quad (49)$$

$$\widehat{S^2}(\widehat{z}) = \left(\frac{T}{\Theta}\right)^2 |\widetilde{\Gamma}|\ ||\epsilon_f||\ |\Gamma| \cong \frac{1}{N^2}\ |\widetilde{\Gamma}|\ ||\epsilon_f||\ |\Gamma|\ , \quad (50)$$

where $\varphi_{k,1}$ and $\varphi_{k,2}$ are the first and second component of the vector $|\varphi_k|$ of Eq. (4), and where T and Θ are the one-dimensional vectors corresponding to

Eqs. (7) and (9). The same results are obtained when the usual integration method of Sect. 1.1 is used. This property is due to the fact that z depends only on the *ratio* of the two integrals F_1 and F_2, making the gradient vector $|\Gamma|$ of Eq. (33) orthogonal to the vector $|\widehat{F}|$. Therefore, at least two linearly independent reference functions are needed to improve the estimate of z.

We will use reference functions only in conjunction with the generation v), in order to make computations simpler. To choose the reference functions in that case, we take a clue from the Born approximation described in Sect. 2.3. The approximation for the function to be integrated there was a linear combination of powers of $\cos\theta^*$ times the function $\ell(\cos\theta^*)$. Since powers of $\cos\theta^*$ are easily integrable analytically in the domain defined by Eq. (34), they are reference functions. If the Born approximation has any value, powers of $\cos\theta^*$ (up to $\cos^3\theta^*$) will be good reference functions to use according to the recommendations of Ref. [1]. They should be very effective in reducing the error on the Monte Carlo estimates of the integrals $|F|$ of Eq. (46), and consequently on the estimate of $z = \overline{\cos\theta}$ of Eq. (27). We introduce them as four linearly independent functions to constitute the components of a vector $|t(\cos\theta^*)|$.

$$|t(\cos\theta^*)| = \ell(\cos\theta^*) \cdot \left| \begin{array}{c} 1 \\ \cos\theta^* \\ \cos^2\theta^* \\ \cos^3\theta^* \end{array} \right| \tag{51}$$

Using the correspondence (48) with $p^{MC}_{(v)}(\ldots)$, we get the integrals $|T|$ of the functions $|t(\cos\theta^*)|$ in the sense of Eq. (7).

$$\begin{aligned} |T| &= \sum_{\lambda} \int |t(\cos\theta^*)|\, p^{MC}_{(v)}(\ldots)\, d(\cos\theta)\, d(\cos\theta^*)\, dx^{CMS} \\ &= \frac{1}{2} \int |t(\cos\theta^*)|\, d(\cos\theta^*) = \left| \begin{array}{c} C_{lim} \\ 0 \\ \frac{1}{3}C^3_{lim} \\ 0 \end{array} \right| \end{aligned} \tag{52}$$

3.3 Results With and Without Use of Reference Functions

In this section, we compare three Monte Carlo estimates of the quantity $z = \overline{\cos\theta}$, related to the integrals F_1 and F_2 by Eq. (27):

$\widehat{z}^{(u)} = \widehat{\overline{\cos\theta}}^{(u)}$, obtained from estimates $\widehat{F_1}$ and $\widehat{F_2}$ of F_1 and F_2 of Eq. (43) using the usual Monte Carlo method of Eq. (3) and points generated like the real events, according to the procedure defined above as u);

$\widehat{z}^{(v)} = \widehat{\overline{\cos\theta}}^{(v)}$, obtained from estimates $\widehat{F_1}$ and $\widehat{F_2}$ of F_1 and F_2 of Eq. (46) using also the usual Monte Carlo method of Eq. (3) but points generated flat in $\cos\theta^*$, according to the simplified procedure defined above as v);[2]

$\widehat{z}^{(ref)} = \widehat{\overline{\cos\theta}}^{(ref)}$, obtained from estimates $\widehat{F_1}$ and $\widehat{F_2}$ of Eq. (46) using the method of Eq. (15), the four reference functions of Eq. (51), and the points generated according to procedure v).

In all three cases, the parameter values used in the computation are those shown in Table 1. The values for the Z^0 parameters are averages of recent experimental results [5].

For a meaningful comparison, the same number of Monte Carlo events, namely 5000, are generated to compute $\widehat{z}^{(u)}$, $\widehat{z}^{(v)}$, and $\widehat{z}^{(ref)}$, as well as estimates of the errors on these quantities, called $\widehat{S}(\widehat{z}^{(u)})$, $\widehat{S}(\widehat{z}^{(v)})$, and $\widehat{S}(\widehat{z}^{(ref)})$. These errors are derived from the error matrix $||\widehat{E}||$ on the corresponding estimate $|\widehat{F}|$ and defined as the square root of the value given by Eq. (32). In order to get the shape of the probability distributions of the estimates of z, the whole estimation procedures is repeated one thousand times, using each time a new set of random numbers.

The arithmetic mean values, $< \widehat{z}^{(u)} >$, $< \widehat{z}^{(v)} >$, and $< \widehat{z}^{(ref)} >$, as well as the standard deviations of the distributions of $\widehat{z}^{(u)}$, $\widehat{z}^{(v)}$, and $\widehat{z}^{(ref)}$, are

[2] As stated in Sect. 3.2, it would be equivalent to define $\widehat{z}^{(u)}$ and $\widehat{z}^{(v)}$ as the estimates obtained using the method of Eq. (15) with one reference function instead of the usual Monte Carlo method of Eq. (3).

Parameter	Value
Number of events generated	$N_{gener} = 5000$
Beam energy	$E_{beam} = 47.\ GeV$
Z^0 mass	$M_Z = 91.161\ GeV/c^2$
Z^0 width	$\Gamma_Z = 2.534\ GeV/c^2$
Z^0 coupling constants	$\sin^2\theta_W = 0.218$
Minimum photon energy	$E_\gamma^{(min)} = 1.65\ MeV$
Maximum photon energy	$E_\gamma^{(max)} = 46.0\ GeV$
Number of radiation lengths	$n_{rad} = 1./\sin\theta_\pm$
Geometrical acceptance	$-0.95 < \cos\theta_\pm < 0.95$
Detection efficiency	$e_{detect} = 0.99$ (per muon
Cut on $\cos\theta$ (see Eq. (23))	$C_{lim} = 0.90$
Collinearity cut (see Eq. (24))	$B_{lim} = 100.0\ mrad$

Table 1: Parameter values used in the example.

indicated below:

$$< \hat{z}^{(u)} > = 0.0947 \quad , \quad \textit{standard deviation} = 0.0090 \ ; \tag{53}$$

$$< \hat{z}^{(v)} > = 0.0944 \quad , \quad \textit{standard deviation} = 0.0095 \ ; \tag{54}$$

$$< \hat{z}^{(ref)} > = 0.0945 \quad , \quad \textit{standard deviation} = 0.0023 \ . \tag{55}$$

The mean values $< \hat{z}^{(u)} >$, $< \hat{z}^{(v)} >$, and $< \hat{z}^{(ref)} >$ are equal to one another within statistical errors. This is expected because $\hat{z}^{(u)}$, $\hat{z}^{(v)}$, and $\hat{z}^{(ref)}$ are estimates of the same quantity z. As expected from the central limit theorem, these three distributions are Gaussian. Figure 1a shows the distribution of $\hat{z}^{(u)}$ as a non-shaded histogram and the distribution of $\hat{z}^{(ref)}$ as a shaded histogram.

We also compute the arithmetic mean values of the estimated errors; we obtain $< \hat{S}(\hat{z}^{(u)}) > = 0.0087$, $< \hat{S}(\hat{z}^{(v)}) > = 0.0094$, and $< \hat{S}(\hat{z}^{(ref)}) > = 0.0023$. As expected, these values are in satisfactory agreement with the standard deviations of Eqs. (53), (54), and (55).

The errors on $\widehat{z}^{(u)}$ and $\widehat{z}^{(v)}$ are about the same, showing that the choices u) or v) of integration points does not affect the accuracy very significantly. However the standard deviation on $\widehat{z}^{(ref)}$ is much smaller than on $\widehat{z}^{(u)}$ and $\widehat{z}^{(v)}$, by a factor 3.9, though the number of integration points is the same, 5000 in each case. This shows the effect of the reference functions. That result means that, in this particular problem, to reach a given statistical precision on our knowledge of $\overline{\cos\theta}$, we need to generate 15 times less Monte Carlo events if we use reference functions of Eq. (51) than if we do not.

The reason for the impressive reduction in statistical error lies in the fact that, at most Monte Carlo points, the function to be integrated, though a function of all the variables of the problem, is not drastically different from a linear combination of reference functions (which are functions of $\cos\theta^*$ only). However, at the limit, if that function to be integrated were such linear combination, no Monte Carlo generation would be needed because the integration could be carried out by analytic means. There is a benefit in using a Monte Carlo technique (with or without reference functions) only if the statistical error of the Monte Carlo integration is smaller than the error committed by neglecting the difference between the function to be integrated and its analytically integrable approximation.

To check that our example is one where Monte Carlo is actually needed, we carried out the computation of $\widehat{z}^{(Born)}$ according to the analytic formula (39) and of $\widehat{z}^{(num)}$ according to the numerical computation of (40) and (41) in Sect. 2.3. These computations both assume a perfect detector and negligible measurement distortions; the analytic computation also assumes that there is no radiative corrections. We found:

$$\widehat{z}^{(Born)} = 0.1548\,, \tag{56}$$

$$\widehat{z}^{(num)} = 0.0861\,. \tag{57}$$

The analytic value differs from the most accurate estimate that we have, $< \widehat{z}^{(ref)} >$ of Eq. (55), by 0.060, i.e., by 26 standard deviations of the error using reference functions and 5000 points. The numerical value differs from the best estimate by 0.0084, i.e., by 3.7 standard deviations of that same

error. It follows that Monte Carlo computation is necessary in our example as soon as an accuracy better than 0.008 is wanted. The estimate with reference functions and 5000 points is about four time more accurate than the numerical approach.[3]

With reference functions, we also computed the quantity χ^2_{min} of Eq. (19) each time an estimate $\hat{z}^{(ref)}$ was made. The value was plotted on the histogram of Fig. 2a. It is distributed according to the theoretical expectation of a χ^2 with 3 degrees of freedom (curve on Fig. 2a). This result confirms our interpretation of this quantity χ^2_{min} as a χ^2 to test the consistency between Monte Carlo generation and analytic integrals.

3.4 Systematic Errors Due to Biases in the Monte Carlo Generator

In Ref. [1], Sect. 3.1, it was stated that reference functions should make the result less sensitive to some systematic errors in the Monte Carlo generation. To test this property, in this example, a flaw was introduced in the Monte Carlo generator, discarding events randomly according to a non uniform distribution in $\cos\theta^*$. The new probability distributions of the Monte Carlo events became

$$p^{MC}_{(u)flaw}(\ldots) \propto \left(1 - Q_{flaw}(1+\cos\theta^*)\right) p^{MC}_{(u)}(\ldots) , \tag{58}$$

$$p^{MC}_{(v)flaw}(\ldots) \propto \left(1 - Q_{flaw}(1+\cos\theta^*)\right) p^{MC}_{(v)}(\ldots) , \tag{59}$$

where

$$Q_{flaw} = 0.15 . \tag{60}$$

In the estimation procedure, the same formulae as for $\hat{z}^{(u)}$ and $\hat{z}^{(ref)}$ were then applied to the remaining Monte Carlo points, as if we were not aware of the Monte Carlo imperfection (parametrized here by the constant Q_{flaw}). The values of $|T|$ given by Eq. (52) were still used. The new corresponding estimates $\hat{z}^{(u)}_{flaw}$ and $\hat{z}^{(ref)}_{flaw}$ were computed 200 times, using 5000 Monte Carlo

[3] Of course an accuracy better than 0.008 could be obtained with the usual Monte Carlo technique if more than 5000 integration points were generated.

events each time, as before. The results are shown in Figs. 1b and 2b. The estimates are still distributed according to Gaussian functions, but with the following parameters for the estimates without and with reference functions:

$$< \hat{z}^{(u)}_{flaw} >= 0.0371 \quad , \quad \textit{standard deviation} = 0.0098 \ ; \tag{61}$$

$$< \hat{z}^{(ref)}_{flaw} >= 0.0943 \quad , \quad \textit{standard deviation} = 0.0025 \ . \tag{62}$$

As can be seen by comparing these numbers to those of Eqs. (53) and (55), or by comparing Fig. 1b to Fig. 1a, the flaw in the generation has a negligible effect on the estimate $\hat{z}^{(ref)}$ which uses reference functions of Eq. (51), while it affects significantly the estimate $\hat{z}^{(u)}$ which does not use them.

Note also that, in the procedure which computed the estimate $\hat{z}^{(ref)}_{flaw}$, the flaw of the Monte Carlo generator could be detected, because the distribution of the quantity χ^2_{min} of Eq. (19) (histogram of Fig. 2b) is no longer compatible with a χ^2 distribution of 3 degrees of freedom (curve on Fig. 2b).

We conclude that, in this example, to get an accurate prediction for $\overline{\cos\theta}$ or to detect a flaw in the Monte Carlo generation, it is better to use Monte Carlo integration procedures relying on reference functions than the usual Monte Carlo method that does not. As stated in Ref. [1], the same statement should apply whenever computations involve integrating functions that can be approximated, at most of the Monte Carlo points, by a linear combination of functions whose integral can be expressed in closed form, i.e., of reference functions.

Acknowledgements. It is a pleasure to thank Dr Daniel Drijard for many fruitful discussions on subjects related to this paper.

References

1. P.H. Eberhard and O.P. Schneider, "Reference Functions to Decrease Errors in Monte Carlo Integrals", LBL–29673, Lawrence Berkeley Laboratory (1990), submitted for publication.

2. See for instance: P. Baillon, G. Barbiellini and D. Treille, in "Physics at LEP", J. Ellis and R. Peccei (eds.), CERN Report 86-02 (1986), Vol. I, p. 172.

3. S. Jadach and Z. Was, in "Workshop in Z Physics at LEP", G. Altarelli, R. Kleiss, and C. Verzegnassi (eds.), CERN Report 89-08 (1989), Vol. I, p. 235.

4. See for instance: K. Kleiss, in "Physics at LEP", J. Ellis and R. Peccei (eds.), CERN Report 86-02 (1986), Vol. I, p. 153.

5. Mark II Collaboration, G.S. Abrams et al., Phys. Rev. Lett. 63 (1989) 2173;
   L3 Collaboration, B. Adeva et al., Phys. Lett. B231 (1989) 509;
   ALEPH Collaboration, D. Decamp et al., Phys. Lett. B231 (1989) 519;
   OPAL Collaboration, M.Z. Akrawy et al., Phys. Lett. B231 (1989) 530;
   DELPHI Collaboration, P. Aarnio et al., Phys. Lett. B231 (1989) 539.

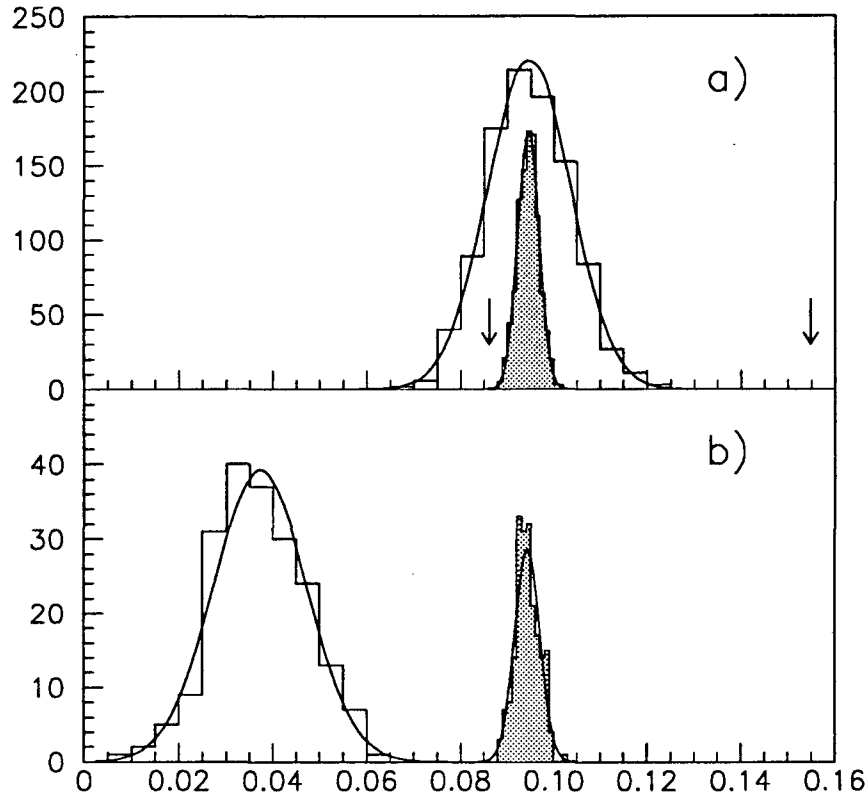


Figure 1

Distributions of $z = \overline{\cos\theta}$ estimates computed with (shaded histograms) or without (non shaded histograms) use of reference functions from Monte Carlo events produced either by a) an unbiased generator or b) a biased generator. The vertical scale indicates the number of entries per 0.001 or per 0.005 for shaded or non shaded histograms respectively. The curves represent Gaussian fits to the histograms. The arrow on the right shows the value of $\hat{z}^{(Born)}$ of Eq. (56) and the arrow in the center shows the value of $\hat{z}^{(num)}$ of Eq. (57). These distributions show how errors can be reduced by reference functions.

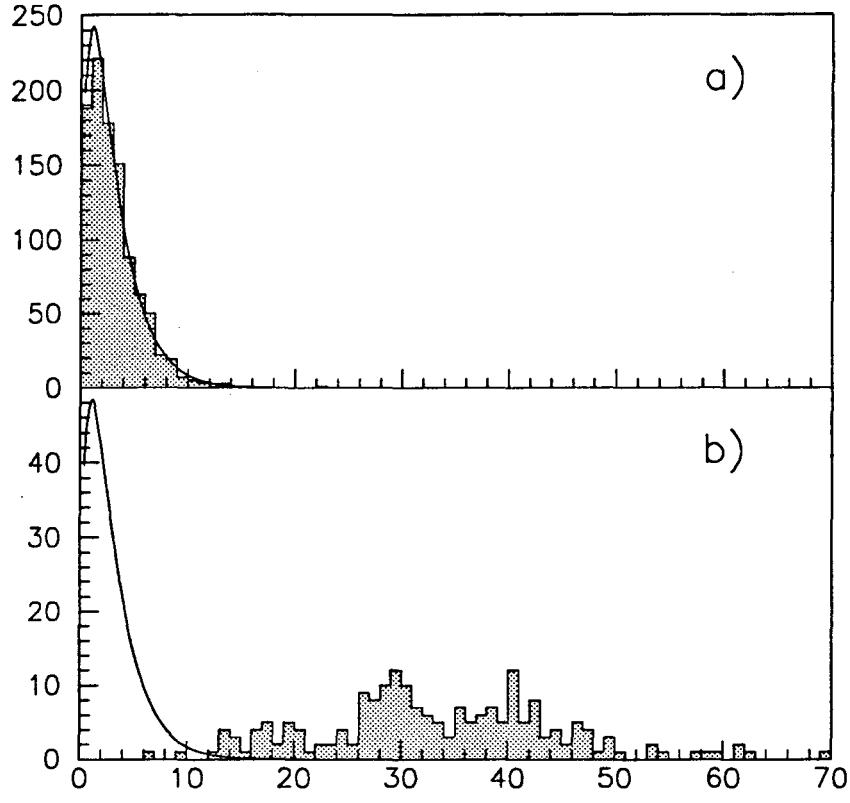


Figure 2

Distributions of the quantity χ^2_{min} computed with use of reference functions from Monte Carlo events produced either by a) an unbiased generator or b) a biased generator. The vertical scale indicates the number of entries per unit of χ^2_{min}. The curves represent the expected distribution which is a χ^2 probability distribution of 3 degrees of freedom, normalized to the number of entries in the histograms. These distributions show how some biases can be detected using reference functions.

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