



ELSEVIER

Computer Physics Communications 88 (1995) 309–326

Computer Physics
Communications

A new version of the multi-dimensional integration and event generation package BASES/SPRING

Setsuya Kawabata

National Laboratory for High Energy Physics, Tsukuba, Ibaraki 305, Japan

Received 26 August 1994

Abstract

The Monte Carlo integration and event generation package BASES/SPRING V1.0 S. Kawabata, Comput. Phys. Commun. 41 (1986) 127, has been upgraded so as to generate events with 50 independent variables, and to integrate functions with an alternating sign. Its ability to integrate real functions with indefinite sign is found to be useful in the numerical evaluation of interference effects among various amplitudes. Besides these changes, its program structure has been completely reformed.

NEW VERSION SUMMARY

Title of program: BASES/SPRING V5.1

Catalogue number: ADBJ

Program obtainable from: CPC Program Library, Queen's University of Belfast, N. Ireland (see application form in this issue)

Reference to original program: Comput. Phys. Commun. 41 (1986) 127; *Cat. number:* AAFW

Authors of original program: S. Kawabata

Does the new version supersede the original program? no

Licensing provisions: none

Computer for which the new version is designed and others on which it is operable: HP750, FACOM, HITAC, IBM, VAX and others with a FORTRAN77 compiler.

Computer: HP750, FACOM-M780; *Installation:* National Laboratory for High Energy Physics (KEK), Tsukuba, Ibaraki, Japan

Operating system: UNIX, OSIV/F4 MSP

Programming language used: FORTRAN77

Memory required to execute with typical data: 392 Kwords

No. of bits in a word: 32

Peripherals used: disc file and printer for output

No. of lines in distributed program, including test data, etc.: 6012

Key words: Multi-dimensional integration, Monte Carlo simulation, event generation and four momentum vector generation

Nature of physical problem

The previous version of the numerical integration and event generation package BASES/SPRING has been useful to obtain total cross sections and to generate events of elementary processes in high energy physics. It is applicable to processes with up to ten independent variables. In order to study, for example, e^+e^- physics at much higher energies, we often need more than ten independent variables to describe pro-

cesses of our interest. As far as the numerical integration by BASES is concerned, it is easy to extend the dimension of integral. However, the event generation requires a huge memory space with the previous generation algorithm of SPRING.

Method of solution

BASES/SPRING is suited for the integration and generation of a very singular function. The number of those independent variables which give the function a singular behavior is usually small. Then, if we divide the subspace spanned by these singular variables into hypercubes, the number of hypercubes becomes not too large. Applying the previous BASES/SPRING algorithm only to this subspace and handling the other variables as additional integral variables, we could extend the dimension of integral and event generation up to 50 variables.

Typical running time

The running time is essentially determined by the complexity of the function program which gives the differential cross section of an elementary process. If we take the process $e^+e^- \rightarrow \nu\bar{\nu}\gamma$ as an example, the two dimensional integration and the generation of 10000 events require 4.4 seconds in total on a FACOM M780 computer.

LONG WRITE-UP

1. Introduction

The numerical integration and event generation package BASES/SPRING V1.0 has been applied to many elementary processes [2]. In recent years very high energy e^+e^- colliders, JLC (KEK), NLC (SLAC), and CLIC (CERN), are being studied as future plans. At the energies of these machines, the number of final state particles in some elementary processes is so high that we cannot apply the version V1.0 directly because of its limited number of dimensions (ten at maximum). Extension of the number of dimensions is easy for the integration package BASES, but it is not for the event generation package SPRING. Before discussing this point, it is useful to describe the algorithm of the version V1.0 briefly.

BASES integrates a function by means of the importance and stratified sampling method.

In order to realize the importance sampling each variable axis is divided into N_{region} regions, each of which is subdivided into m subregions. In total each variable axis has $N_g (= N_{\text{region}} \times m)$ sub-

regions. The numbers N_{region} and m are determined from the dimension of integral N_{dim} automatically by BASES. The full integral volume is covered by a grid of $N_g^{N_{\text{dim}}}$ subregions. We call a subspace of the integral volume a *hypercube* which is spanned by the regions of each variable axis.

In each hypercube, sample points of a definite number N_{sample} are taken, and the integral and its variance are evaluated. The estimate of integral and its variance are calculated by summing up results of all hypercubes. We call this process an *iteration*. A cumulative estimate and its error are calculated from the results of all iterations statistically.

Execution of BASES consists of the grid optimization and the integration steps.

At the first iteration of the former step the grid is uniformly defined for each variable axis. After each iteration the grid is adjusted so as to make the sizes of subregions narrower at the parts with the larger function value and wider at the parts with the smaller one. In this way a suited grid to the integrand is obtained.

In the integration step, the probability to select each hypercube and the maximum value of the function in it are calculated as well as the estimate of integral with the frozen grid determined in the former step.

In SPRING, a hypercube is sampled according to its probability, calculated by BASES, and a point in the hypercube is selected by sample-and-reject method for which the maximum value of the function is used. If the grid is sufficiently optimized, a high efficiency of generating a set of values of the independent variables (i.e. event generation) can be achieved.

The difficulty in the straightforward extension of the number of dimensions in BASES/SPRING V1.0 can be easily seen by considering a 25 dimensional case (i.e. 25 independent variables). Even though we suppose to have only two regions for each variable axis, there are $N_{\text{cube}} = 2^{25}$ hypercubes in total. The calculation of such a huge number N_{cube} of probabilities requires much CPU time and their storage space is also too large. It seems to be unrealistic.

By assuming, however, an additional condition

to the integrand, this difficulty can be avoided. The algorithm of the new version BASES/SPRING V5.1 is described in the next section, and its program structure and usage are shown in Section 3. In Section 4 we will discuss its test run.

2. Algorithm of version 5.1

In usual practice the function of our interest has only a limited number of variables that cause its singular behavior. We call such kind of variables *wild variables*. The rest variables, on which the function depends weakly, are called *gentle variables*. If we divide only the subspace spanned by the wild variables into hypercubes, the number of hypercubes is not too large. Thus, by applying the BASES/SPRING V1.0 algorithm only to this subspace and by handling the gentle variables as additional integral variables, we can overcome the difficulty in the event generation of higher dimensions.

In order to see the difference between the old and the new algorithms, we consider the following integral of three dimensions:

$$I = \int_0^1 f(x, y, z) dx dy dz.$$

In the old algorithm, the estimate of integral is given by

$$I \approx \sum_{j=1}^{N_{\text{cube}}} \frac{1}{N_{\text{sample}}} \sum_{i=1}^{N_{\text{sample}}} \frac{f(x_i^j, y_i^j, z_i^j)}{p(x_i^j, y_i^j, z_i^j)},$$

where x_i^j , y_i^j and z_i^j indicate i th sample point in j th hypercube, $p(x_i^j, y_i^j, z_i^j)$ is the probability density at the point for the importance sampling, and N_{sample} is the number of sample points in each hypercube.

On the other hand, if we take the new algorithm and suppose the variables x and y are wild and z is gentle, the estimate is given by

$$I \approx \sum_{j=1}^{N_{\text{cube}}} \frac{1}{N_{\text{sample}}} \sum_{i=1}^{N_{\text{sample}}} \frac{f(x_i^j, y_i^j, z_i)}{p(x_i^j, y_i^j, z_i)},$$

where z is sampled in the full range of z variable. Since the importance sampling takes place even for the gentle variable z , the numerical

integration converges fast enough. An efficient event generation is guaranteed if the function

$$F^j(z) = \int_{j\text{-th hypercubes}} f(x^j, y^j, z) dx^j dy^j$$

has similar dependences on z for all j .

The numbers of hypercubes are N_{region}^3 and N_{region}^2 in the old and the new algorithms, respectively, where N_{region} is the number of regions per variable. Generally speaking, if N_{dim} and N_{wild} are the numbers of independent variables and the wild variables ($N_{\text{wild}} \leq N_{\text{dim}}$), respectively, the number of hypercubes N_{cube} is given by

$$\begin{aligned} N_{\text{cube}} &= N_{\text{region}}^{N_{\text{dim}}} \quad (\text{old}) \\ &= N_{\text{region}}^{N_{\text{wild}}} \quad (\text{new}). \end{aligned}$$

In version 5.1, the maximum number of wild variables is equal to 15.

3. Program structure and usage

In the old version, users must prepare many subprograms, e.g. USERIN, FUNC, USROUT, SPINIT, SPEVNT, SPTERM etc. according to the specification of each subprogram. This complexity seems to give a pedestrian a high threshold to use this package. Furthermore, the program flow is completely controlled by the BASES/SPRING package, and the numerical integration and the event generation are designed to be successively performed by two separate procedures. This makes it less flexible particularly in the event generation, to pass generated events to detector simulation and data analysis. In order to dissolve these inconveniences, the program structure has been completely reformed.

3.1. Program components

When we study an elementary process, we first calculate the effective cross section by integrating the differential cross section with some kinematical cuts by BASES, generate events of four momentum vectors of final state particles by SPRING, pass through a detector simulation program to simulate the experimental data, and then

```

: C      Main Program for BASES/SPRING V5.1
: C
1  IMPLICIT REAL*8(A-H,O-Z)
2  EXTERNAL FUNC
:
:
3  PARAMETER (NXDIM = 50)
4  COMMON /BPARAM1/ XL(NXDIM),XU(NXDIM),NDIM,NWILD,IG(NXDIM),NCALL
5  COMMON /BPARAM2/ ACC1,ACC2,ITMX1,ITMX2
:
: C
6  COMMON/KINEM/W,EM,ZM,ZGAM,CZ,CV,CA,FACTOR,XK,COSTH
7  REAL*4 P(4)
8  DATA PI,ALP,GEVNB,GENER /
:      3.1415926D0, 137.036D0, 0.38927D6, 3.0D0 /
:
: *****
: * Initialization of BASES/SPRING V5.1
: *****
: ***** Initialization of BASES by calling BSINIT
:
: C      CALL BSINIT
:
: ***** Initialization of parameters for kinematics etc.
:
10  EM = 0.511E-3
11  ZM = 90.0
12  WM = 78.97
13  ZGAM = 2.6
14  ALPHA = 1./ALP
15  RAD = PI/180.
16  TWOP1 = 2.D0*PI
17  SQ = SQRT(ZM**2-WM**2)
18  CZ = ZM**2/(2.*WM*SQ)
19  CA = CZ*(1./2.)
20  CV = 2.*CA*(-1./2.+2.*SQ**2/ZM**2)
21  FACTOR = GENER*CZ**2*ALPHA**3/12.*GEVNB
22  COSMIN = 15.
23  COSMAX = 180.-COSMIN
24  W = 105.0D0
25  XKMIN = 1./W
:
: ***** Initialization of BASES parameters
:
26  NDIM = 2
27  NWILD = 1
28  NCALL = 5000
29  ITMX1 = 10
30  ACC1 = 0.1D0
31  ITMX2 = 100
32  ACC2 = 0.05D0
:
:
33  XL(1) = XKMIN
34  XU(1) = W/2.
35  XL(2) = DCOS(COSMAX*RAD)
36  XU(2) = DCOS(COSMIN*RAD)
:
: ***** Initialization of Histograms
:
37  CALL XHINIT(1,XL(1),XU(1),40,'Photon Energy (GeV)')
38  CALL XHINIT(2,XL(2),XU(2),50,'Cos(theta) of Photon')
39  CALL DHINIT(1,XL(2),XU(2),50,XL(1),XU(1),50,
:      ' x : cos(theta) -- y : Photon Energy ')
:
: *****
: * Numerical Integration by BASES V5.1
: *****
:
40  CALL BASES( FUNC, ESTIM, SIGMA, CTIME, IT1, IT2 )
:
:      LU = 6
41
:      CALL BSINFO( LU )
42
:      CALL BHPlot( LU )
43
: *****
: * Event generation by SPRING V5.1
: *****
: ***** Initialization of additional histograms
: *****
44  CALL XHINIT( 5, 0.0D0, 4.0D1, 40,
:      'Photon Transverse Energy (GeV)')
:
: ***** Event generation loop
:
:      MXTRY = 50
:      MXEVT = 10000
45
46
:      DO 100 NEVNT = 1, MXEVT
47
:          CALL SPRING( FUNC, MXTRY )
48
:          *      Compute the four vectors of generated event
:          *      from the kinematical variables
:
49          PHI = TWOP1*DRN(DUMY)
50          PXV = XK*SQRT(1.0 - COSTH+COSTH)
51          P(1) = PXV*COS(PHI)
52          P(2) = PXV*SIN(PHI)
53          P(3) = XK*COSTH
54          P(4) = XK
:
:          C      WRITE(20) P
:
55          CALL XHFILL( 5, PXV, 1.D0 )
:
:          100 CONTINUE
56
:      CALL SPINFO( LU )
57
:      CALL SHPlot( LU )
58
:      STOP
59  END
60
: *****

```

Fig. 1. Main program of the test run.

analyze the resultant simulation data to compare with real data.

If the process has a few particles in the final states, the numerical integration and the event generation do not take much computing time. It is better to carry out the simulation study in one procedure for simplicity. If the process has many final state particles, however, the numerical integration takes much computing time. It is then recommended to separate the integration part from the other parts consisting of the event generation and the detector simulation, etc. Since in the present version the numerical integration and the event generation are carried out by calling subroutines **BASES** and **SPRING**, respectively, the simulation study can be performed either in one procedure or in separate two or more procedures. In any case users are requested to prepare main and function programs. In addition to the subroutines **BASES** and **SPRING**, a set of subroutines, described below, are prepared for use. It should be noticed that all the real variables used in the system are defined to be of double precision.

BSINIT: To set the **BASES/SPRING** parameters equal to defaults.

BSDIMS: To set the numbers of dimensions and the wild variables, and the lower and the upper limits of the integral variables.

BSGRID: To set the sampling flags for the integral variables.

BSPARM: To set the parameters for integration.

BASES: To make numerical integration and to prepare the probability information for event generation.

BSINFO: To print a parameter list and convergence behavior of the integration.

BHPLLOT: To print histograms and scatter plots, taken in the integration.

SPRING: To generate an event with weight one.

SPINFO: To print statistical information about the event generation.

SHPLLOT: To print histograms and scatter plots, taken in the event generation.

BSWRIT: To store the probability information into a file.

BSREAD: To restore the probability information from a file.

An example of outputs from **BSINFO**, **BHPLLOT**, **SPINFO**, and **SHPLLOT** is shown and explained for the test run output in Section 4.

The program flow is fully controlled by the main program written by users. Its specification is described in the next subsection together with the usage of these subroutines.

3.2. Main program

The main program can be divided into the initialization part of **BASES/SPRING**, the numerical integration part, and the event generation part. A list of the main program for the test run is given in Fig. 1 as an example.

(1) Initialization of **BASES/SPRING**

At the beginning, **BSINIT** is called to set the parameters for **BASES/SPRING** equal to default values. Then the parameters for the integration and event generation are to be initialized to their proper values, which are passed to the system through the labeled commons **BPARM1** and **BPARM2**. All real parameters in these two labeled commons are to be given with double precision. The content of common **BPARM1** is as follows;

```
COMMON / BPARM1 / XL(50), XU(50),  
NDIM, NWILD, IG(50), NCALL
```

NDIM: The number of independent variables (dimension of integral) up to 50.

NWILD: The number of the wild variables up to 15.

XL(i), **XU(i)**: The lower and the upper limits of *i*th integral variable. Notice that the first **NWILD** variables must correspond to the wild variables in each array.

IG(i): The flag to switch the sampling method for *i*th variable. The default setting is the importance sampling (0/others) = (uniform/importance sampling).

NCALL: The number of sample points per iteration.

The true number of sample points $N_{\text{call}}^{\text{real}}$ differs from a given number **NCALL**, which is automatically determined by the following algorithm.

The number of regions per variable N_{region} is determined as the maximum number which satisfies the two inequalities:

$$N_{\text{region}} = \left(\frac{N_{\text{call}}}{2} \right)^{\frac{1}{N_{\text{wild}}}} \leq 25,$$

and

$$N_{\text{region}}^{N_{\text{wild}}} < N_{\text{cube}}^{\text{max}},$$

where $N_{\text{cube}}^{\text{max}}$ is the maximum number of hypercubes and is set equal to 32768. The number of hypercubes is given by $N_{\text{cube}} = N_{\text{region}}^{N_{\text{wild}}}$, then the number of sample points per cube is $N_{\text{sample}} = N_{\text{call}}/N_{\text{cube}}$. Since the number N_{sample} is an integer, the calculated number $N_{\text{call}}^{(\text{real})} = N_{\text{sample}} \times N_{\text{cube}}$ may differ from the given number $N_{\text{call}}^{(\text{given})}$ ($= \text{NCALL}$).

The content of common BPARM2 is as follows;

```
COMMON/BPARM2/ACC1, ACC2, ITMX1,
  ITMX2
```

ACC1: The desired accuracy of integration for the grid optimization step in units of percent.

ACC2: The desired accuracy of integration for the integration step in units of percent.

ITMX1: The number of iterations for the grid optimization step.

ITMX2: The number of iterations for the integration step.

For those users, who do not like to set these parameters directly on the labeled commons, a set of subprograms is prepared.

To set the dimension, the number of wild variables, the lower and upper limits of integral variables,

```
CALL BSDIMS (NDIM, NWILD, XL, XU),
```

where XL and XU are arrays with the double precision real numbers.

To set the flag IG,

```
CALL BSGRID(NDIM, IG),
```

where IG is an integer array. Unless this routine is called, all integral variables are sampled by the importance sampling method as the default.

To set the number of sample points per iteration, the required accuracies and the maximum numbers of iterations both for the grid optimization and the integration steps,

```
CALL BSPARM (NCALL, ACC1, ACC2,
  ITMX1, ITMX2),
```

where ACC1 and ACC2 are to be given by the double precision real numbers.

In addition to the initialization of the BASES/SPRING parameters, parameters and constants, like center of mass energy, beam energy, kinematical cut values, the fine structure constant and the numerical value of π , etc., should be set and be passed to the function program FUNC through some labeled commons to calculate the differential cross section.

When we make histograms and scatter plots of some variables, their initialization should be done here. The system has a proper histogram package, whose description is given in Subsection 3.4.

(2) Numerical integration

In the grid optimization step, sizes of the subregions for each variable are optimized to the behavior of integrand iteration by iteration, starting from a uniform size of subregions, as briefly described in Section 1. The algorithm for the grid optimization is identical to that of the old version and VEGAS, whose more detailed description is found in Refs. [1] and [3].

In the integration step, by fixing the grid determined in the grid optimization step a cumulative estimate of the integral is calculated as well as the probability information for each hypercube spanned by the wild variables.

The both steps are terminated when the estimated accuracy of the integral reaches a required value ($\text{ACC1} / \text{ACC2}$) or the number of iterations exceeds a given number ($\text{ITMX1} / \text{ITMX2}$).

The estimate of the integral and its error etc.

of the integration step are returned to users in the arguments of the subroutine **BASES** as follows:

```
CALL BASES (FUNC, ESTIM, ERROR,
           CTIME, IT1, IT2).
```

FUNC: The name of a function program, whose specification is described in Subsection 3.3.

ESTIM: A cumulative estimate of the integral.

ERROR: The standard deviation of the estimate of the integral.

CTIME: The computing time used by the integration step in seconds.

IT1: The number of iterations made in the grid optimization step.

IT2: The number of iterations made in the integration step.

The tables of the convergency behavior of the integral both for the grid optimization and the integration steps are printed out on the standard output device. We can also keep these tables in a file by calling the subroutine **BSINFO**, for example,

```
CALL BSINFO(LU).
```

The file allocated to the logical unit **LU** should be opened in the initialization step (1). The contents of the tables are described in Subsection 4.2.

When histograms and scatter plots are made, they can be printed on a logical unit **LU** by

```
CALL BHPlot(LU).
```

Subroutines **BSINFO** and **BHPlot** do not need to be called, unless outputs from them are required.

If the numerical integration and the event generation are separately performed by different procedures, the probability information made by **BASES** should be stored on a file by calling **BSWRIT**, for instance, as follows;

```
CALL BSWRIT(LN).
```

The file allocated to the logical unit **LN** must be opened at the initialization step (1). Further-

more, at the beginning of the event generation program and just after the initialization of **BASES/SPRING** described above, subroutine **BSREAD** should be called to restore the probability information from the file as follows;

```
CALL BSREAD(LN),
```

then the event generation can be carried out.

(3) Event generation

After the integration events can be immediately generated by calling **SPRING**. For each call of subroutine **SPRING** a set of values of independent variables is generated with weight one. The calling sequence of **SPRING** is as follows;

```
CALL SPRING(FUNC, "MXTRY").
```

FUNC: The name of a function program.

MXTRY: The maximum number of trials to generate one event.

At the first call of **SPRING** a cumulative probability distribution over all the hypercubes is calculated from the differential one, prepared by **BASES**. This cumulative distribution is used to sample a hypercube according to its probability.

After a hypercube is sampled in this way, a point is selected in this hypercube by the importance sampling method and is examined whether it is accepted or not. When this point is not accepted, another point is sampled in the same hypercube and tested again. If the grid is not optimized enough, there is a possibility to get into an infinite loop of generation. To avoid this occurrence the number of trials to get an event is limited to **MXTRY** whose recommended number is 50. The case where an event could not be generated even by **MXTRY** trials is called *mis-generation*. If the number of mis-generations exceeds **MXTRY**, a warning message is printed.

The event generation by **SPRING** determines only a set of values of independent variables of the function program. Users should calculate the four-momentum vectors from this set of variables.

```

* Function program for the elementary process
* e+ e- => nu(mu) nu-bar(mu) gamma
* Two independent variables
* XK = X(1) : the energy of photon
* COSTH = X(2) : cos(theta) of photon
*
1 REAL FUNCTION FUNC=8(X)
2
3 IMPLICIT REAL*8(A-H,O-Z)
4 PARAMETER (MXDIM = 50)
5 DIMENSION X(MXDIM)
6
7 COMMON/KINEM/W,EM,ZH,ZGM,CZ,CV,CA,FACTOR,XK,COSTH
8
9 REZ(S)= S-ZM**2
10 Z2(S) = (REZ(S))**2+(ZH*ZGM)**2
11 FUNC = 0.
12 XK = X(1)
13 S = W*W
14 S1 = W*(W-2.*XK)
15 E = W/2.
16 PP = SORT(E**2-EM**2)
17 COSTH = X(2)
18 D1 = XK*(E+PP*COSTH)
19 D2 = XK*(E-PP*COSTH)
20
21 C ..... MATRIX ELEMENT SQUARE STARTS
22 C
23 C
24 C
25 C

```

Input of this package : 71 lines

```

19 DSDX = -FACTOR*YAD*XK/(E*Z2(S1))
20 FUNC = DSDX/E
21
22 *====> Fill histograms and scatter plot
23
24 CALL XHFILL(1,XK,DSDX)
25 CALL XHFILL(2,COSTH,DSDX)
26 CALL DHFILL(1,COSTH,XK,DSDX)
27
28 RETURN
29 END

```

```

17 ANS1=(S**2*CA**2+S**2*CV**2-2.*S*W*XK*CA**2-2.*S*W*XK*CV**2-3.*S*
. CA**2*EM**2-3.*S*CV**2*EM**2+2.*W*XK*CA**2*EM**2+2.*W*XK*CV**2*
. EM**2)/D1+S1*(-S**2*CA**2-S**2*CV**2+10.*S*CA**2*EM**2-2.*S*CV**2
. 2*EM**2-16.*CA**2*EM**4+8.*CV**2*EM**4)/(2.*D1*D2)+(-S**3*CA**2-
. S**3*CV**2+2.*S**2*W*XK*CA**2+2.*S**2*W*XK*CV**2+2.*S**2*CA**2*
. EM**2-4.*S**2*CV**2*EM**2-4.*S*W*XK*CA**2*EM**2-4.*S*W*XK*CV**2*
. EM**2+2.*W*XK*CA**2*EM**2+4.*W*XK*CV**2*EM**2)/(2.*
. D1*D2)-D2*S1*EM**2*(CA**2*CV**2)/D1+S1*D2*EM**2*(-S*CA**2-S*CV
. **2+2.*W*XK*CA**2+2.*W*XK*CV**2)/D1**2+S1*EM**2*(S*CA**2+S*CV
. **2-2.*W*XK*CA**2-2.*W*XK*CV**2)/(2.*D1**2)
18 TAU0=(-D1*S1)*(CA**2*CV**2)/D2+D1*(-S*CA**2-S*CV**2+2.*W*XK*CA**2+
. 2.*W*XK*CV**2+2.*CA**2*EM**2+2.*CV**2*EM**2)/D2-(D1*S1*EM**2)*(CA
. **2*CV**2)/D2**2+D1*EM**2*(-S*CA**2-S*CV**2+2.*W*XK*CA**2+2.*W*
. XK*CV**2)/D2**2+4.*EM**2*(CA**2*CV**2)+S1*(S*CA**2+S*CV**2-3.*CA
. **2*EM**2-3.*CV**2*EM**2)/D2+(S**2*CA**2+S**2*CV**2-2.*S*W*XK*CA
. **2-2.*S*W*XK*CV**2-3.*S*CA**2*EM**2-3.*S*CV**2*EM**2+2.*W*XK*CA
. **2*EM**2+2.*W*XK*CV**2*EM**2)/D2+S1*EM**2*(S*CA**2+S*CV**2-8.*
. CA**2*EM**2+4.*CV**2*EM**2)/(2.*D2**2)+S*EM**2*(S*CA**2+S*CV**2-
. 2.*W*XK*CA**2-2.*W*XK*CV**2)/(2.*D2**2)-(D2*S1)*(CA**2*CV**2)/D1+
. D2*(-S*CA**2-S*CV**2+2.*W*XK*CA**2+2.*W*XK*CV**2+2.*CA**2*EM**2+
. 2.*CV**2*EM**2)/D1+S1*(S*CA**2+S*CV**2-3.*CA**2*EM**2-3.*CV**2*EM
. **2)/D1+ANS1

```

*====> Calculate the value of function

Fig. 2. Function program of the test run.

A typical structure of the event generation step is as follows:

```

MXTRY=50
DO 100 NEV=1, 10000
  CALL SPRING( FUNC, MXTRY)
  ....
  Calculate four momentum vectors
  ....
100 CONTINUE

```

Since the kinematics of the process is calculated in the function program, it is better to use the result of kinematics for calculating the four-momentum vectors.

After the event generation, statistical information about the generation can be obtained by calling SPINFO as follows:

```
CALL SPINFO(LU),
```

where LU is the logical unit of an output device. Histograms and scatter plots can be also printed on the logical unit LU by

```
CALL SHPLOT(LU).
```

3.3. Function program

In the function program the value of the integrand is calculated at the sample point fed by BASES. A set of numerical values of integral variables is passed through the argument of the function program. A typical structure of function program is as follows;

```

DOUBLE PRECISION FUNCTION FUNC(X)
REAL*8 X(2)
FUNC=0.0D0
... Calculation of kinematics ...
IF(the point is outside of
  the kinematical boundary) RETURN
FUNC= is calculated from X(i) for i=1, 2
CALL XHFILL(ID, V, FUNC)
CALL DHFILL(ID, Vx, Vy, FUNC)
RETURN
END

```

A recipe for writing a function program is:

- (1) Calculate the kinematical variables by which the differential cross section is described from the integral variables, $X(i)$ for $i = 1$ to $NDIM$.
- (2) If the point fed by BASES is found to be outside the kinematical boundary, set the value of the function equal to zero and return.
- (3) If the point is inside the kinematical boundary, calculate the numerical value of the function at the point and return it as the function value.
- (4) If a histogram is required, call subroutine XHFILL once.
- (5) If a scatter plot is required, call DHFILL once.

An example of FUNC for the process $e^+e^- \rightarrow \nu_\mu \bar{\nu}_\mu \gamma$ is given in Fig. 2.

3.4. Histogram package

The system has a proper histogram package, by which histograms and scatter plots of any variables can be printed. A characteristic of this package is that there are two kinds of histograms, the one is called “original histograms” and another is “additional histograms”.

The original histogram is quite useful to check whether the frequency distribution of generated events really reproduces the distribution calculated in the integration stage. On the other hand the additional histogram prints only the frequency distribution of the generated events. A detailed description of these two kinds of histograms is found in Subsection 4.2.

The maximum numbers of histograms and scatter plots are 50 for each. If the initialization routine XHINIT (or DHINIT) is called more than 50 times, subsequent calls are neglected. Usage of this histogram package is as follows;

- (1) To initialize histograms and scatter plots the following routines are to be called:

```

CALL XHINIT(ID#,
· lower_limit ,
· upper_limit , # of bins,
· Title of histogram )

```

and

```
CALL DHINIT(ID#,
· x-lower_limit ,
· x-upper_limit , # of x bins,
· y-lower_limit ,
· y-upper_limit , # of y bins,
· 'Title of scatter plot '),
```

respectively.

(2) To fill the histograms and scatter plots the following routines are to be called in FUNC:

```
CALL XHFILL(ID#, V, FUNC)
```

and

```
CALL DHFILL(ID#,""Vx,""Vy,""FUNC)
```

respectively.

(3) To print the resultant histograms and scatter plots taken in the integration,

```
CALL BHPlot(LU)
```

where LU is a logical unit number.

(4) To print the resultant histograms and scatter plots taken in the event generation,

```
CALL SHPlot(LU)
```

where LU is a logical unit number.

Examples of histograms and scatter plots printed by BHPlot and SHPlot are shown in the test run output.

4. Test run

4.1. Test of algorithm

For the test of the new version we take the process $e^+e^- \rightarrow \nu\bar{\nu}\gamma$ as an example, which was also used in Ref. [1]. The integral variables are energy and polar angle of the photon. As shown in histograms of the test run output, the photon energy distribution has sharp peaks at zero and the Z^0 mass, while the angular distribution has forward and backward peaks. In order to demonstrate that the new algorithm works well, we tested our program in the following three cases:

(1) Set the integration parameters, $NDIM = 2$, $NWILD = 2$, and $NCALL = 5000$. Since $N_{dim} = N_{wild}$, the integration algorithm is exactly identical to the old one.

(2) Set the parameters, $NDIM = 2$, $NWILD = 1$, $IG(2) = 1$, and $NCALL = 5000$. Only the photon energy is considered as the wild variable and the polar angle is sampled by the importance sampling.

(3) Set the parameters, $NDIM = 2$, $NWILD = 1$, $IG(2) = 0$, and $NCALL = 5000$. The polar angle is uniformly sampled.

As summarized in Table 1, cases (1) and (2) give a consistent result and almost identical speed of convergency. The result of integration is not affected by the difference of algorithms. The efficiency of event generation by the new algorithm is slightly lower than that of the old one, but still

Table 1
The results for three cases of integration parameters

Case	Numerical integration				Event generation	
	Estimate (error)	CPU time	It-1	It-2	Efficiency	CPU time
(1)	$4.528033 (\pm 0.002230) \times 10^{-2}$	3.43	10	26	75.8	0.40
(2)	$4.527550 (\pm 0.002255) \times 10^{-2}$	3.59	10	27	69.4	0.41
(3)	$4.521658 (\pm 0.006430) \times 10^{-2}$	10.67	10	100	11.6	1.88

Case (1) is by the old algorithm. Cases (2) and (3) are by the new version with the importance and uniform samplings for the gentle variable, respectively. It-1 and It-2 are the numbers of iterations for the grid optimization and the integration steps, respectively. The generation efficiency is defined to be the ratio of the number of accepted events to that of trials. The CPU time for generation is the computing time consumed to generate 10k events and calculate their four-momentum vectors.

high. This result shows that the new algorithm works very well.

When we sample the polar angle variable uniformly (case (3)), the convergency of the integration is much worse than the other two cases and the efficiency of event generation is very low. It is very reasonable, since the sampling of the photon polar angle is never optimized.

It is, however, noted that the uniform sampling may give much faster convergence than the importance one for those variables whose distributions are known to be nearly flat. The reason is as follows. The importance sampling is done by using the distribution taken at the previous iteration. If this distribution has some bias due to less statistics of sampling, the resultant importance sampling is also biased. Although this bias is usually canceled during many iterations, it requires much computing time to reach a stable answer.

4.2. Output from BASES / SPRING

As the test run output, we take case (2). There are four sets of outputs, the outputs from the numerical integration part by calling BSINFO and BHPLLOT, and from the event generation part by calling SPINFO and SHPLLOT.

(1) Output from BSINFO

On the first page of the output, detailed information about the integration parameters and computing time is printed.

As the integration parameters,

N_{dim} : the number of integral variables (dimension),

N_{wild} : the number of wild variables,

$N_{\text{call}}^{\text{given}}$: the given number of sample points per iteration,

$N_{\text{call}}^{\text{real}}$: the true number of sample points,

N_{g} : the number of subregions per variable,

N_{region} : the number of regions per variable,

N_{cube} : the number of hypercubes,

$XL(i)$: the lower limit of i th integral variable,

$XU(i)$: the upper limit of i th integral variable,

$IG(i)$: the sampling flag for i th integral variable,

$ITMX1$: the maximum number of iterations for the grid optimization step,

$ACC1$: the desired accuracy in the grid optimization step,

$ITMX2$: the maximum number of iterations for the integration step, and

$ACC2$: the desired accuracy in the integration step

are listed. As the computing time information, the consumed CPU time (in units of hour, minute, and second) each for the grid optimization step, the integration step, and the overhead by the remainders is printed. Furthermore the expected computing time for generating 1000 events is estimated by assuming 70% of generation efficiency.

On the second page of the output, the convergency behavior for the grid optimization step is printed, which consists of the result of each iteration and the cumulative result up to the iteration. In the result of each iteration, the iteration number, the hit rate of the kinematically allowed region, the fraction of sample points resulting in negative function values, the estimate of the integral, and its accuracy are included. The cumulative result consists of the cumulative estimate of the integral, its error, and its accuracy. Since at the first iteration sampling points are uniformly distributed, the estimate of integral results in a smaller value than the final one, while its accuracy is worse than the final one. They converge to some final values due to the grid optimization.

On the third page of the output, the convergency behavior for the integration step is printed, whose content is identical to that for the grid optimization step. It is quite important to check if the estimate of the integral and its accuracy are stable in each iteration. If they fluctuate more than 5σ , a warning message is printed on a standard output device. This may happen either when the number of sampling points per iteration is too small or the choice of integral variables is inadequate for the behavior of the differential cross section.

(2) Output from BHPLLOT

BHPLLOT prints histograms and scatter plots made during the numerical integration.

The first and the last bins of a histogram are underflow and overflow bins. The first column shows the lower value of each histogram bin. The second column represents the estimated differential value and its error after the characters “+–”, both of which are to be multiplied by a factor 10^x shown by “E xx”. On the right hand side of these columns a histogram of the differential values is drawn both in the linear scale with the character “*” and in the logarithmic scale with the character “0”. If negative values exist in some bins only the linear scale histogram is displayed.

A scatter plot represents only the relative height of a function. The height of the function value is described by ten characters: 1, 2, 3, ..., 8, 9, and *, while the depth (for negative function values) is displayed by the other ten characters: a, b, c, ..., h, i and #. The point which results in a negative value but larger than the threshold of level “a” is indicated by the character “–”, while the point describing a positive value less than the threshold of level “1” is given either by the character “+” (if a negative value exists somewhere) or by the character “.” (if only positive values exist).

(3) Output from SPINFO

SPINFO prints statistics of event generation, which consists of the generation efficiency, the number of generated events, the computing time each for the event generation by SPRING, for its overhead, and for the other part consumed by the calculation of four-momentum vectors and detector simulation, etc. The maximum number of trials MXTY and the number of mis-generations are also printed.

(4) Output from SHPLOT

SHPLOT prints histograms and scatter plots made during event generation. There are two kinds of histograms as mentioned in Subsection 3.4.

The original histogram is to be defined before the numerical integration by calling BASES. The contents of the histogram filled during the integration is compared with the frequency distribution made in the event generation. This comparison is made in the logarithmic scale, where the

statistical error of each bin is represented by the characters “<””. If the error is smaller than the two character space, only the frequency is shown by the character “0”. The histogram filled during the integration is represented by the character “*”.

The additional histogram is to be initialized just before starting the event generation. Since there are no corresponding data made in the integration step, only the frequency distribution of events is displayed both in the linear and the logarithmic scales.

On the first page of the output, a histogram displaying the number of trials to obtain an event is always printed in the format of an additional histogram. If this histogram has a sharp peak at the first bin, the efficiency of generation is high.

Then the original histograms and the additional histograms are successively printed if they are defined.

The scatter plots printed by SHPLOT describe only the frequency distribution in terms of the relative height with ten levels.

5. Conclusion

As the test elementary process we took a simple example: $e^+e^- \rightarrow (Z^0) \rightarrow \nu\bar{\nu}\gamma$. The physical conditions for the integration are taken to be identical to those of Ref. [4] (the center of mass energy $W = 105$ GeV and the polar angle cut for the photon $165^\circ > \theta_\gamma > 15^\circ$), except for the energy threshold of the photon $E_\gamma > 1/W$ GeV. Besides this test the new algorithm was applied to the process $e^+e^- \rightarrow t\bar{t}Z^0$, where $t \rightarrow b\bar{f}\bar{f}$ and $Z^0 \rightarrow f\bar{f}$ [5]. In this calculation, two variables were used to distinguish the different helicity combinations and another one to select decay modes of Z^0 and $W^\pm$, respectively, in addition to 20 variables for the phase space. The new version worked well even for this 23-dimensional integration and event generation.

By using the new version, we can integrate even a singular function and reproduce the behavior of the function by generating events with

weight one. There are, however, the following restrictions:

- (1) The maximum number of wild variables is limited to 15.
- (2) There should be no strong correlation among the wild variables. If it exists, singular points may run continuously along a diagonal line of the integral volume, and our algorithm cannot be applied. We should carefully choose the integral variables to avoid such correlations.

In spite of these limitations, this new version of BASES/SPRING will be useful for many applications in the field of very high energy physics.

Acknowledgements

The author greatly appreciates Prof. Y. Shimizu, Prof. T. Ishikawa, Dr. K. Fujii, Dr. K. Hagiwara, Dr. J. Fujimoto at KEK, Prof. K. Kato

at Kogakuin University, Dr. D. Perret-Gallix at LAPP in France, and other users for their valuable comments. He also thanks Prof. S. Iwata at KEK and Prof. Y. Oyanagi at the University of Tokyo for their encouragement throughout this work.

References

- [1] S. Kawabata, *Comput. Phys. Commun.* 41 (1986) 127.
- [2] For example: K. Tobimatsu and Y. Shimizu, *Prog. Theor. Phys.* 74 (1985) 567; 75 (1986) 905.
S. Kuroda et al., *Comput. Phys. Commun.* 48 (1988) 335.
J. Fujimoto and M. Igarashi, *Prog. Theor. Phys.* 74 (1985) 791.
- [3] G.P. Lepage, *J. Comput. Phys.* 27 (1978) 192.
VEGAS: Adaptive Multi-dimensional Integration Program, CLNS-80/447 (March 1980).
- [4] G. Barbiellini, B. Richter and J.L. Siegrist, *Phys. Lett. B* 106 (1981) 414.
- [5] T. Tauchi, K. Fujii, A. Miyamoto; KEK-Preprint 90-148 (November 1990).

TEST RUN OUTPUT

Date: 94/ 5/10 13:16

```

*****
*
* BBBBBB AAAA SSSSS EEEEE SS SS SS
* BB BB AA AA SS EE SS SS
* BB BB AA AA SS EE SS SS
* BBBBBB AAAAAA SSSSS EEEEE SSSSS
* BB BB AA AA SS EE SS
* BB BB AA AA SS EE SS
* BBBBBB AA SSSSS EEEEE SSSSS
*
* BASES Version 5.1
* coded by S.Kawabata KEK, March 1994
*****

```

<< Parameters for BASES >>

(1) Dimensions of integration etc.

```

# of dimensions : Ndim = 2 ( 50 at max.)
# of Wilds : Nwid = 1 ( 15 at max.)
# of sample points : Ncall = 5000(real) 5000(given)
# of subregions : Ng = 50 / variable
# of regions : Nregion = 25 / variable
# of Hypercubes : Ncube = 25

```

(2) About the integration variables

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+
i XL(i) XU(i) IG(i) Wild
-----+-----+-----+-----+-----+-----+-----+
1 9.523810E-03 5.250000E+01 1 yes
2 -9.659258E-01 9.659258E-01 1 no
-----+-----+-----+-----+-----+-----+-----+

```

(3) Parameters for the grid optimization step

```

Max.# of iterations: ITM1 = 10
Expected accuracy : Acc1 = 0.1000 %

```

(4) Parameters for the integration step

```

Max.# of iterations: ITM2 = 100
Expected accuracy : Acc2 = 0.0500 %

```

<< Computing Time Information >>

```

(1) For BASES H: M: Sec
Overhead : 0: 0: 0.00
Grid Optim. Step : 0: 0: 0.98
Integration Step : 0: 0: 2.61
Go time for all : 0: 0: 3.59

```

```

(2) Expected event generation time 0.03 Sec
Expected time for 1000 events :

```

Date: 94/ 5/10 13:16

Convergence Behavior for the Grid Optimization Step

IT	Eff	R_Neg	Estimate	Acc %	Estimate(+/- Error)	Order	Acc % (H: M: Sec)
1	100	0.00	4.127E-02	3.801	4.127239(+/-0.156879)E-02	3.801	0: 0: 0.09
2	100	0.00	4.536E-02	1.992	4.434463(+/-0.078292)E-02	1.766	0: 0: 0.19
3	100	0.00	4.588E-02	0.832	4.557121(+/-0.034986)E-02	0.768	0: 0: 0.29
4	100	0.00	4.557E-02	0.333	4.557166(+/-0.013906)E-02	0.305	0: 0: 0.39
5	100	0.00	4.508E-02	0.274	4.529775(+/-0.009239)E-02	0.204	0: 0: 0.49
6	100	0.00	4.532E-02	0.255	4.530635(+/-0.007221)E-02	0.159	0: 0: 0.59
7	100	0.00	4.536E-02	0.278	4.538569(+/-0.006277)E-02	0.138	0: 0: 0.68
8	100	0.00	4.538E-02	0.238	4.538356(+/-0.005426)E-02	0.120	0: 0: 0.78
9	100	0.00	4.521E-02	0.242	4.534931(+/-0.004861)E-02	0.107	0: 0: 0.88
10	100	0.00	4.541E-02	0.261	4.535849(+/-0.004497)E-02	0.099	0: 0: 0.98

Date: 94/ 5/10 13:16

Convergence Behavior for the Integration Step

IT	Eff	R_Neg	Estimate	Acc %	Estimate(+/- Error)	Order	Acc % (H: M: Sec)
1	100	0.00	4.524E-02	0.258	4.523981(+/-0.011689)E-02	0.258	0: 0: 1.08
2	100	0.00	4.514E-02	0.253	4.519085(+/-0.008170)E-02	0.181	0: 0: 1.17
3	100	0.00	4.507E-02	0.247	4.514963(+/-0.006589)E-02	0.146	0: 0: 1.27
4	100	0.00	4.513E-02	0.250	4.514549(+/-0.005691)E-02	0.126	0: 0: 1.37
5	100	0.00	4.524E-02	0.264	4.516280(+/-0.005136)E-02	0.114	0: 0: 1.46
6	100	0.00	4.508E-02	0.233	4.514726(+/-0.004613)E-02	0.102	0: 0: 1.56
7	100	0.00	4.549E-02	0.268	4.518954(+/-0.004315)E-02	0.095	0: 0: 1.66
8	100	0.00	4.522E-02	0.260	4.519278(+/-0.004051)E-02	0.090	0: 0: 1.76
9	100	0.00	4.542E-02	0.258	4.521653(+/-0.003830)E-02	0.085	0: 0: 1.85
10	100	0.00	4.537E-02	0.267	4.523973(+/-0.003653)E-02	0.081	0: 0: 1.95
11	100	0.00	4.518E-02	0.248	4.523414(+/-0.003473)E-02	0.077	0: 0: 2.05
12	100	0.00	4.524E-02	0.261	4.523487(+/-0.003332)E-02	0.074	0: 0: 2.14
13	100	0.00	4.529E-02	0.266	4.524683(+/-0.003203)E-02	0.071	0: 0: 2.24
14	100	0.00	4.514E-02	0.247	4.523864(+/-0.003078)E-02	0.068	0: 0: 2.33
15	100	0.00	4.538E-02	0.277	4.524689(+/-0.002990)E-02	0.066	0: 0: 2.43
16	100	0.00	4.520E-02	0.264	4.524436(+/-0.002900)E-02	0.064	0: 0: 2.53
17	100	0.00	4.534E-02	0.275	4.524917(+/-0.002825)E-02	0.062	0: 0: 2.62
18	100	0.00	4.538E-02	0.257	4.525643(+/-0.002746)E-02	0.061	0: 0: 2.72
19	100	0.00	4.544E-02	0.270	4.525527(+/-0.002679)E-02	0.059	0: 0: 2.82
20	100	0.00	4.532E-02	0.261	4.526821(+/-0.002613)E-02	0.058	0: 0: 2.92
21	100	0.00	4.532E-02	0.270	4.527040(+/-0.002556)E-02	0.056	0: 0: 3.01
22	100	0.00	4.530E-02	0.266	4.527173(+/-0.002500)E-02	0.055	0: 0: 3.11
23	100	0.00	4.534E-02	0.265	4.527466(+/-0.002447)E-02	0.054	0: 0: 3.21
24	100	0.00	4.514E-02	0.240	4.526793(+/-0.002387)E-02	0.053	0: 0: 3.30
25	100	0.00	4.554E-02	0.272	4.527769(+/-0.002344)E-02	0.052	0: 0: 3.40
26	100	0.00	4.532E-02	0.259	4.527919(+/-0.002299)E-02	0.051	0: 0: 3.50
27	100	0.00	4.518E-02	0.258	4.527550(+/-0.002255)E-02	0.050	0: 0: 3.59



