

Automatic Computation of LHC Processes

E. BOOS, V. ILYIN, K. KATO, A. PUKHOV, A. SEMENOV, A. SKATCHKOVA

Automatic computation is a new approach to HEP computing. The first such systems, GRACE [1], FeynArt/FeynCalc [2] and CompHEP [3], were reported at the 1st International Workshop AIHENP held on March, 1990 in Lyon-Villeurbanne (France). Under this terminology, *automatic computation system (ACS)*, we assume, as a distinguishing feature, the generation of the computing code for a specific collision process with the aid of another code.

ACS's are now used widely by phenomenologists for the calculation of many collision processes. For example, the GRACE and CompHEP systems were used in the LEP2 Workshop [4], and for evaluation of processes at TeV linear colliders [5]. With ACS one can calculate *all* collision processes within a given physical model, where by physical model we mean the set of Feynman rules. Recent developments with the LanHEP package [11] have opened a possibility to derive Feynman rules in the form of the ACS intrinsic physical model in a fully automatic way, starting from the Lagrangian. Now, not only the Standard Model but a number of its extensions, like SUSY models, are implemented in ACS. A general review of this new approach is given in these Proceedings by K.Kato together with discussion of main directions of the ACS development. Here, we discuss in more detail specifics of the ACS applications in LHC phenomenology, and in particular to the evaluation of QCD processes.

To close this preview we list below the main ACS options in order to provide an idea for users of what is available:

- i) selection of physical model (Lagrangian) and hard subprocess,
- ii) Feynman diagram generation,
- iii) generation of the code for matrix element,
- iv) convolution with parton distributions,
- v) generation of kinematics (phase space parameterization) with regularization of kinematical peaks,
- vi) integration over the phase space (evaluation of cross section),
- vii) generation of events at partonic level, including the interface to hadronization tools.

1 The problem of multiparticle final states: why automatic computations?

We start from the problem of the accurate evaluation of hard subprocesses in the case of multiparticle final states.

When physicists simulate HEP processes with such generators as PYTHIA [6], ISAJET [7] or HERWIG [8] they use a data base of hard subprocesses implemented in these packages. It means that a) matrix elements are stored as formulas, and b) a knowledge about the behaviour of matrix elements as phase space integrands are coded in the form of modelling functions in order to get a fast generation of the partonic events. One can note that these data bases include a rather simple variety of subprocesses, mainly of the $2 \rightarrow 2$ type.

If one tries to include a hard subprocesses with 3, 4 and more particles in the final state, large problems appear. Indeed, the size of matrix elements increases very fast. For example, in the $2 \rightarrow 4$ case, the size of the code for evaluation of helicity amplitudes for one subprocess is at the 100's of Kbyte level. However, the main problem lies elsewhere; it is impossible to construct an analytical formula for matching peaks and other structures of the rather singular behaviour of matrix elements. In the $2 \rightarrow 3$ case, phase space has 4 dimensions plus two for convolution with the PDF's; in the $2 \rightarrow 4$ case 7+2 dimensions are present, and so on. As a result, the set of kinematical singularities has, as a rule, a very complicated positioning in the multidimensional phase space. This particular problem was not solved accurately, e.g., when the Zbb final state was implemented in PYTHIA 5.7.

Let us discuss further this somewhat delicate point. It is necessary to integrate the squared matrix element over the phase space in order to obtain the cross section. Precise information about the behaviour of the integrand then is necessary for further event generation. This information can be obtained at the step of the phase space integration. The problem is that the integrand, as a rule, has a singular behaviour with sharp kinematical peaks connected with different denominators (propagators) of Feynman diagrams. This problem is caused, in particular, by the circumstance that one has to take into account nonzero masses of particles in many important cases, especially if accurate calculations are needed. The masses of elementary particles can have extremely small values, e.g. the masses of the 1st generation quarks (few MeV), and can also be zero (for the photon and gluon). At the same time, other parameters are of the order of a hundred GeV, e.g. masses of W and Z bosons and t -quark. Moreover the collision energy can also have a very large value, e.g. 14000 GeV for LHC processes, and some other important variables, like the transverse energy of jets, are at the hundred GeV scale or even greater. This huge scale interval for different parameters causes serious computational problems which result in the appearance of sharp peaks for the integrand. So, at the step of phase space parameterization, one has to include a regularization of the integration measure in order to smooth the singularities of the integrand.

LHC phenomenology requires the computation of a wide spectrum of hard subprocesses with 3,4 and even more particles (partons) in the final state. This is a common need for all of the physics working groups: QCD, EW, Higgs, SUSY etc. These requirements are especially common for new physics searches. Furthermore, a major challenge results from background

analyses, where QCD subprocesses play a major role with, in many cases, multiparton final states. As a rule, for each LHC discovery reaction, one should calculate several QCD processes giving both irreducible and reducible backgrounds. The parton-shower generation of multiparticle final states is usually utilized in this situation. However, this can be too crude an approximation for many important studies leading to sometimes grossly unreliable predictions.

We emphasize that ACS can give the possibility to compute accurately a variety of LHC processes (and in particular QCD processes) with 3, 4 and more bodies in the final partonic state. Indeed, the first problem (the size of the matrix element computing code and the difficulty to obtain the exact matrix elements) is solved in ACS by the automatic generation of the corresponding code. This step is fast and pain-free from the viewpoint of the user. The second problem (the accurate integration over the multidimensional phase space) is solved in ACS by the generation of kinematics where the necessary regularizations are included. For example, in CompHEP the user has to list a set of singular propagators using the menu system. After that, the code for kinematics (with regularizations) is generated automatically. In GRACE, a library of kinematics (with regularizations) is used and the user has to make the necessary choices. Thus, the high art (mathematics and programming), needed to elaborate the sharp peaks, is enclosed in a form hidden from the user, giving him a possibility to compute complicated processes.

At the step of integration over the phase space, ACS uses adaptive Monte Carlo integrators (VEGAS [12] in CompHEP, and BASES [13] in GRACE). To match the complete set of singularities, the multichannel MC approach [10] is utilized. As a result, the phase space grid is created with an accurate mapping of the singular behaviour of the matrix element. This complex body of information (let us call it *MEgrid*) has a rather large size that rapidly increases with the number of phase space dimensions. One can consider *MEgrid* as a multidimensional analog of the modelling function used in PYTHIA and other similar packages for the effective generation of partonic events. Of course, this information can not be expressed in analytical form. It is necessary to point out also that the convolution with parton distributions should be made at the same stage as the integration over the phase space. Indeed, the contributions of different subspaces (in particular different kinematical peaks) can depend largely on the partonic collision energy, \hat{s} , resulting from the information stored in *MEgrid*.

ACS can be considered as a tool for the automatic generation of the data base of hard subprocesses for physical generators like PYTHIA, ISAJET and HERWIG. However, it is difficult to imagine that the data base created can be implemented in the code of these generators. This is due, first of all, to the size of the generated codes. Thus, we propose a *two stage* approach. At the first stage, ACS is used resulting in a cross section and *MEgrid* for the subprocess under evaluation. This can be stored in a special LHC data base. This data base can be used for the effective generation of partonic events. In GRACE, it is available with the SPRING [13] generator, and in CompHEP by a relatively straightforward procedure and an effective generator is under construction). The output is a partonic event flow that can be used as an input for physical generators like PYTHIA, ISAJET and HERWIG; this is second stage of the full simulations. At this stage partons (quarks and gluons) should be

hadronized and unstable particles decayed. We note that in PYTHIA there exists a rather flexible interface for such a *two stage* approach, the option for inclusion of external processes through the routine PYUPEV.

This is a general view on the way in which ACS (GRACE and CompHEP in particular) can be used for the simulation of LHC processes. Below we discuss some specific features of this technology with special attention to QCD aspects.

2 General Considerations about GRACE and CompHEP

With CompHEP and GRACE the user can evaluate hard subprocesses at the tree level, i.e. Feynman diagrams are generated without loops. This corresponds to the basic request for LHC phenomenology. However, it is well known that QCD next-to-leading corrections are large, as a rule, for LHC processes. In many cases these corrections can be accounted for in the form of so-called K-factors and one can include them easily in tree level calculations. Nevertheless, in many important cases an explicit evaluation of higher order corrections is necessary. At this moment it is not clear how to automate calculations of LHC processes at NLO level. The problem is connected, in particular, with the circumstance that different resummations of large logarithms should be included in order to get reliable NLO predictions. The interface between resummation techniques and event generators is under intensive discussion now, and at the present Workshop also. We note in this respect, that the GRACE package includes the code for the generation and evaluation of one-loop diagrams.

The user interface should provide the possibility to calculate complicated processes for users not experienced in programming. CompHEP has a (graphical) menu driven system where the user proceeds through all steps of the calculation without any programming. In GRACE, the user needs to write a few simple interface routines.

The information on the GRACE system and its products can be found at
<http://www-sc.kek.jp/minami/>

The code of CompHEP is free for users and one can take it from the following Web page
<http://theory.npi.msu.su/comphep>

where the user's manual is available in PS format (see also hep-ph/9908288). The CompHEP package, adapted for LHC processes (see next section) is installed on the SUN platform

</afs/cern.ch/cms/physics/COMPHEP/v33-SUN>

and on the PC/Linux platform

</afs/cern.ch/cms/physics/COMPHEP/v33-Linux>

The interface between CompHEP and PYTHIA has been created with the corresponding code available from the address:

</afs/cern.ch/cms/physics/comp-pyth>

where one can find a short description in the file README. With this interface, the partonic event flow for any processes calculated with CompHEP can be sent to PYTHIA to generate physical events.

3 QCD aspects in automatic computations

In this section we discuss the treatment of QCD effects in the case of automatic computations, and consider CompHEP options as an example. As has been discussed above, CompHEP calculates only at tree level, and so at leading order (LO). Thus, the main problems concerning an accurate accounting of QCD effects are outside the discussion. Nevertheless, some important QCD dependencies can not be avoided even at tree level and the corresponding options are available for users. These aspects are: a) parton distributions, b) QCD scale, and c) running strong coupling constant.

Parton distributions.

In CompHEP the specification of initial states in the collision process under evaluation can include the convolution with structure function. So, in the case of hadron collisions, the cross section is evaluated as an integral

$$\sigma(s) = \int_0^1 dx_1 dx_2 f_i(x_1, Q) f_j(x_2, Q) \hat{\sigma}_{ij}(x_1 x_2 s)$$

where f_i are the corresponding parton distributions, $\hat{\sigma}$ is the partonic cross section and Q is the QCD scale.

In CompHEP v.33, installed at CERN (see address above), parton distributions from two pdf families are implemented, MRS and CTEQ, and in particular the following versions: 1) MRS(A') and MRS(G) [14], 2) CTEQ4l and CTEQ4m [15]. Note that CTEQ4l is a LO parametrization, while in all others the evolution of parton distributions is realized in the next-to-leading (NLO) approximation.

In addition, a special interface is available to include a user's defined parton distribution. By this way one can implement the most recent parametrizations (at this moment CTEQ5 and MSRT). See the CompHEP user's manual for the corresponding procedure (section 3.6.2).

Choice of QCD scale.

The factorization theorem states that parton distributions depend not only on the Bjorken variable x but also on some parameter Q which characterizes the energy (or momentum) scale at which the QCD effects give the main contribution to the hard subprocess. This parameter is set by the user for each specific QCD process. It is possible to set a fixed scale or a running scale. In the later case, Q^2 can be a squared linear combination of any set of initial and outgoing particles momenta, for example, $(p_1 - p_3)^2$, $(p_1 - p_3 - p_4)^2$, $(p_3 + p_4)^2$ and so on (initial and outgoing momenta enter with opposite signs). The corresponding settings are made through the option QCD SCALE of the numerical menu.

Running α_s .

It is the nature of strong interactions that there is no absolute normalization of the corresponding coupling constant. This is in contrast to the value $1/137$ for the electromagnetic constant known with high accuracy from classical electrodynamical experiments. Instead, we have a function for α_s rather than a constant. Even in the leading order approximation, $\alpha_s^{LO} = 6\pi / [(33 - 2n_f) \log Q / \Lambda^{(n_f)}]$, where Q is the QCD scale of the hard subprocess under

the evaluation with Λ the so-called QCD fundamental parameter. Then, n_f is the number of parton flavours with masses lower than Q . The n_f dependence in QCD parameter Λ matches the quark mass threshold effects.

In the version of CompHEP installed at CERN, the running α_s is realized in LO, NLO and NNLO. All of the corresponding formulas are based on the choice of $\Lambda^{(6)}$ (see *Review of Particle Physics* [16] p.81. The user can find the corresponding switch in the option QCD SCALE in the numerical menu.

Therefore, to evaluate QCD processes with CompHEP, one has, first of all, to fix the normalization of α_s . The popular normalization point is the mass of Z boson, $Q = M_Z$. By changing the parameter $\Lambda^{(6)}$, the user should set the strong coupling at the appropriate value, say $\alpha_s(M_Z) = 0.118$. Then, the user has to choose the order for the running α_s (LO, NLO or NNLO). Finally, the user has to define the QCD scale Q , which will be used both for the evaluation of α_s at this scale and in the parton distributions.

Thus, the complete LO calculations of LHC processes are available, with the matrix element, parton distributions and running strong coupling constant calculated in the lowest order of the perturbation theory. This is a self-consistent starting point in the phenomenological analysis; when/where higher order corrections are available, all elements of the calculation can be calculated at the higher order and then compared to the leading order result.

However, it is also common for phenomenologists use a mixed approach, with the matrix element evaluated at LO but the parton distributions and running α_s taken in NLO approximation. Surely, only a part of the NLO corrections is accounted for in this case. We note that this option is also available for users in CompHEP calculations.

4 Partonic Subprocesses

When hadronic collision processes are evaluated, especially in the case of a large number of final state particles, one serious problem is the large number of contributing partonic subprocesses. This occurs because of the quark and gluon content of the initial hadrons and CKM quark mixing. For example, at LHC energies, 180 subprocesses contribute to the $W+2jets$ and 292 subprocesses to the $W+3jets$ production (taking into account only quarks of the first two generations). During this workshop a new method has been proposed to avoid a multiplication of channels due to the mixture of quark states [17]. The method leads to a simple modification of the rules for the evaluation of the cross sections and distributions. It is based on the unitary rotation of down quarks, thus providing the transportation of mixing matrix elements from vertices of Feynman diagrams to the parton distribution functions. As a result, one can calculate cross sections with a significantly smaller number of subprocesses contributing. For the examples mentioned above, one needs to evaluate (with the new rules) only 21 and 33 subprocesses, respectively, in order to compute the cross sections for the $W+2jets$ and $W+3jets$ processes. The matrix elements of the subprocesses are calculated without quark mixing, but with a modified PDF convolution which now depends on the quark

mixing angle and the topologies of the gauge invariant classes of diagrams contributing to the subprocesses. The method proposed has been incorporated into the CompHEP program and checked with many examples.

5 PEVLIB - library of LHC processes

Now the library of CompHEP based event generators for LHC processes has been started at the address:

`/afs/cern.ch/cms/physics/PEVLIB`

The following QCD processes are stored already in this library: $Zb\bar{b}$, $Wb\bar{b}$, $t\bar{t}b\bar{b}$ and some others. In the corresponding directories (with the names literally corresponding to the final states) unweighted events are stored (see the files `README` in these directories for details about evaluation of the corresponding samples of events).

Together with the CompHEP-PYTHIA interface code (see discussion above) these event files can be used for full LHC simulations with the help of PYTHIA package and detector simulation software in the standard way.

Let us discuss the process $Zb\bar{b}$ in order to supply more details. In the directory

`/afs/cern.ch/cms/physics/PEVLIB/Z_b_b`

the file `__pevZbb` includes about 200000 unweighted events with the final state $Zb\bar{b}$. Each event includes the Lorentz momenta of all particles in the initial and final states. In the present version of this library, there is no information about the color flow in the event. Thus, only the **Independent Fragmentation Model** can be used for the hadronization. Of course, the user can use the Lund model; for this one has to define the corresponding color flows by hand in the routine `PYUPEV`. The same remark is valid also for FSR (final state radiation), what is switched off by default in CompHEP-PYTHIA interface. In the same time ISR (initial state radiation) is switched on automatically.

Note that the user can generate more events than stored in the library. In the corresponding subdirectories (indicated in the file `README`) the generators are stored in the form of the executable code (at this moment for SUN platform only). These generators are the corresponding CompHEP codes for the process with the proper set of kinematical regularizations.

The library PEVLIB is under construction now. New processes will be added. The structure and user's interface will be developed.

6 Acknowledgements

The work of V.I., A.P., A.S, E.B. and A.S. was partially supported by the CERN-INTAS grant 377 and RFBR-DFG grant 99-02-04011.

References

- [1] *New Computing Techniques in Physics Research*, ed. D.Perret-Gallix and W.Wojcik, Ed. CNRS, Paris, 1990, p.555.
 T.Tanaka, T.Kaneko and Y.Shimizu, *Comp. Phys. Commun.* **64**, 149 (1991).
 T.Ishikawa et al., *GRACE manual*, KEK Report 92-19, 1993.
- [2] *New Computing Techniques in Physics Research*, ed. D.Perret-Gallix and W.Wojcik, Ed. CNRS, Paris, 1990, p.565.
- [3] *New Computing Techniques in Physics Research*, ed. D.Perret-Gallix and W.Wojcik, Ed. CNRS, Paris, 1990, p.573.
 P.A.Baikov et al., In: Proc.of X Workshop on High Energy Physics and Quantum Field Theory (QFTHEP-95), ed.by B.Levtchenko and V.Savrin, Moscow, 1996, p.101, hep-ph/9701412.
 E.E.Boos, M.N.Dubinin, V.A.Ilyin, A.E.Pukhov and V.I.Savrin, hep-ph/9503280.
 A.E.Pukhov et al., *CompHEP user's manual, v.3.3*. Preprint INP MSU 98-41/542, 1998, hep-ph/9908288.
- [4] "Physics at LEP2", ed. by G.Altarelli, T.Sjöstrand and F.Zwirner, CERN report 96-01, vol. 1, 2.
- [5] E.Accomando et al., *Physics with e^+e^- linear colliders*, *Phys. Rep.* **C299**, 1 (1998).
- [6] T.Sjostrand, *Comp. Phys. Commun.* **82**, 74 (1994).
- [7] F.E.Paige et al., *ISAJET 7.40: a Monte Carlo event generator for pp , $\bar{p}p$, and e^+e^- reactions*, BNL-HET-98-39, Oct 1998, hep-ph/9810440.
- [8] G.Marchesini et al., *HERWIG VERSION 5.9.*, hep-ph/9607393.
- [9] H.Tanaka, *Comput. Phys. Commun.* **58**, 153 (1990).
- [10] F.A.Berends. R.Pittau and R.Kleiss, *Comp. Phys. Commun.* **85**, 437 (1995).
 V.A.Ilyin, D.N.Kovalenko and A.E.Pukhov, *Int.J.Mod.Phys.* **C7**, 761 (1996).
 A.E.Pukhov and D.N.Kovalenko, Talk on 5th Int. Workshop AIHENP'96 (September 1996, Lauzanne). *Nucl.Inst. and Meth.* **A393**, 299 (1997).
- [11] A.Semenov, Talk on 5th Int. Workshop AIHENP'96 (September 1996, Lauzanne). *Nucl.Inst. and Meth.* **A393**, 293 (1997).
 A.Semenov, *Comp. Phys. Commun.* **115**, 124 (1998).
- [12] G.P.Lepage, *J. Comp. Phys.* **27**, 192 (1978).
- [13] S.Kawabata, *Comp. Phys. Commun.* **41**, 127 (1986).
 S.Kawabata, *Comp. Phys. Commun.* **88**, 309 (1995).

- [14] A.D.Martin, R.G.Roberts and W.J.Stirling, *Phys. Lett.* **B354**, 155 (1995).
- [15] H.L.Lai et al., *Phys. Rev.* **D55**, 1280 (1997).
- [16] Particle Data Book, *Europ. Phys. J.* **C3** (1998).
- [17] E. E. Boos, V. A. Ilyin and A. N. Skachkova, [hep-ph/0004194](#) .