CompHEP 4.4 - Automatic Computations from Lagrangians to Events

E. Boos, V. Bunichev, M. Dubinin, L. Dudko, V. Ilyin, A. Kryukov, V. Edneral, V. Savrin, A. Semenov, A. Sherstnev
(the CompHEP collaboration) a

aSkobeltsyn Institute of Nuclear Physics, Moscow State University, 119992 Moscow, Russia

We present a new version of the CompHEP program (version 4.4). We describe shortly new issues implemented in this version, namely, simplification of quark flavor combinatorics for the evaluation of hadronic processes, Les Houches Accord based CompHEP-PYTHIA interface, processing the color configurations of events, implementation of MSSM, symbolical and numerical batch modes, etc. We discuss how the CompHEP program is used for preparing event generators for various physical processes. We mention a few concrete physics examples for CompHEP based generators prepared for the LHC and Tevatron.

1. INTRODUCTION

The increase of collider energies requires efficient and precise calculation of processes with multi-particle final states. The NLO, NNLO or resummation corrections to the leading order results should be incorporated when available.

At LEP1 the most interesting physical channels were 2-fermion production processes at the Z pole, while at LEP2 mainly 4-fermion processes with some photons or gluons radiated off have been considered. At the Tevatron, LHC and Next Linear Colliders a majority of interesting channels include 5, 6 and 8 or even more fermions in the final states with additional photons or gluons. For example, the top quark pair production leads to 6 fermions in the final state, single top production in various modes gives 4, 5 and 6 fermions, the channel $t\bar{t}H$ produces 8 fermions in the final state, etc. Calculations of event characteristics with multi-particle final states are needed to be done at the complete matrix element level in order to include correctly the spin correlations, non-trivial kinematics and in some cases finite width effects in a gauge invariant manner.

In practical computations, especially for hadron colliders, there are many contributing Feynman diagrams even at the tree level and a large number of subprocesses leading to the same multi-particle final state. To perform such computations a number of automatic programs have been created - CompHEP [1], GRACE [2], MadGraph [3], AlpGen [4], O’ Mega [5] with WHIZARD [6], Amegic [7], HELAC-PHEGAS [8] and other.

However, in order to get reliable predictions and reduce a dependence on scales involved, like QCD normalization and factorization scales, one should compute and simulate not only processes with many final particles but also include various corrections and perform resummations of large logs. Recently a few calculations of NLO corrections for $2 \rightarrow 3$ processes have been performed [9,10,11,12].

Precise calculations including the NLO corrections to the rate and distributions are however not sufficient for experimental studies insofar as one should also provide a corresponding NLO event generators. Since the real final states are not partons but hadrons (jets) these generators should be matched in some way to the parton shower generators, such as PYTHIA [13] and HERWIG [14], which simulate a realistic picture of hadronization. In fact the problem of matching of all the needed ingredients, multi-particle matrix element calculations with all spin correlations, NLO cor-
rections to, at least, basic production processes, hadronization taking into account initial and final state radiation is highly nontrivial. As a result one likes to have an event generator which includes all these parts without "double counting" of contributions and with smooth distributions in the whole range of the phase space. Although this problem is far from a final solution, recently interesting and promising approaches for different parts of the above mentioned matching have been proposed \[15,16,17,18,19,20\].

In this paper we discuss mainly a new aspects of the CompHEP latest version 4.4 and present some recent CompHEP applications to physics studies at colliders. We discuss shortly the SingleTop generator, created with the help of CompHEP to simulate the electroweak top quark production at the Tevatron and LHC, where proper matching of different processes with main contributions to different phase space regions is performed in the final event flow.

CompHEP is based on quantum theory of gauge fields and includes the Standard Model Lagrangian in the unitary and 't Hooft-Feynman gauges as well as several other MSSM based models. CompHEP is able to compute basically the LO cross sections and distributions but with many particles (up to 4-6) in the final state taking into account, if necessary, all the QCD and EW diagrams, masses of fermions and bosons and widths of unstable particles. Processes computed by means of CompHEP could be used as a "new" external processes for generators like PYTHIA.

The CompHEP project started in 1989 at SINP MSU. During the 90's the package was developed, and now it is a powerful tool for automatic computations of collision processes. The CompHEP program has been used in the past for many studies (see [21] and the citation to CompHEP [1] for more complete information) as shown schematically in the Fig. [1].

2. COMPHEP 4.4

In the latest CompHEP version 4.4 (http://theory.sinp.msu.ru/comphep) there are several new improvements incorporated which make the program significantly more efficient to use for event simulation, especially for the case of hadronic collisions.

2.1. QUARK FLAVOR COMBINATORICS

A serious computational problem for the hadronic collisions is a large number of partonic subprocesses with many contributing Feynman diagrams in each of them. This takes place, in particular, for the reason of numerous possible combinations of initial and final quarks with different flavors, giving different subprocesses. Many additional diagrams appear in each of these subprocess due to CKM quark mixing. In fact many of the subprocesses and many diagrams have very similar topological and Lorentz structure, so one would like to compute all the similar pieces only once.

In the paper [23] the method for a simplification of quarks combinatorics has been proposed. The method has been implemented to the CompHEP 4.4 and tested on many examples showing a significant reduction in a number of contributing subprocesses and diagrams.

The method is based on two approximations:

- mixing with the third generation is neglected
- $V_{CKM} \Rightarrow \begin{pmatrix} V & 0 \\ 0 & 1 \end{pmatrix}$
- $V = \begin{pmatrix} \cos \vartheta_e & \sin \vartheta_e \\ -\sin \vartheta_e & \cos \vartheta_e \end{pmatrix}$

Figure 1. Very incomplete list of processes computed by means of CompHEP in the past
where $\vartheta_c$ is the Cabbibo angle.

- masses of the first four quarks are neglected $M_u = M_d = M_s = M_c = 0$

One should stress that usually both approximations work very well for collider processes, although in the case when masses are needed to be taken into account one should use the regular CompHEP Standard Model.

Under the assumptions above the method is based on the two main ideas:

- rotation of down quarks, thus, transporting the mixing matrix elements from the vertices of subprocess Feynman diagrams to the parton distribution functions

- the contributing diagrams are splitted into exact gauge invariant subclasses \[24\] for each of the subprocess according to a different topology of the quark lines involved.

As a result one finds new rules for a convolution with the structure functions where each gauge invariant subclass of squared diagrams with certain topology of the quark lines is convoluted with a corresponding combination of structure functions. In general there are only three different possibilities which we refer to as a convolution rules number 1, 2 and 3.

**Scattering topology. 1st Rule**

For this topology there are two quark loops after squaring, and each loop includes only one initial quark. In such a case the structure functions are summed over the two flavors for each initial state and the dependence on Cabbibo angle disappears. The following formula gives a particular example for this case when one initial state is down quarks while another initial state is up antiquarks:

$$|D_a^{C C}|^2 = \int dx_1 dx_2 \left[ f_d(x_1) f_u(x_2) \cos^2 \vartheta_c + f_u(x_1) f_d(x_2) \cos^2 \vartheta_c + f_d(x_1) f_c(x_2) \sin^2 \vartheta_c + f_c(x_1) f_d(x_2) \sin^2 \vartheta_c \right] |\mathcal{M}|^2$$

**Annihilation NC topology. 3d Rule**

This topology is similar to the previous one leading to one quark loop in the squared diagram. But now both initial quark and antiquark are either up or down type. Generic example for the convolution with structure functions in this case is the following (note, there is no dependence on the Cabbibo angle as for the first rule):

$$|D_a^{N C}|^2 = \int dx_1 dx_2 \left[ f_d(x_1) f_d(x_2) + f_u(x_1) f_u(x_2) \right] |\mathcal{M}|^2$$

A special rule is used in the cases when the final quarks and antiquarks form a loop in a squared diagram not involving the initial quarks. In these cases each loop of this sort gives a factor of 2 after summation over flavors. Of course, this rule is valid only under the assumption that the fragmentation of the four light quarks and antiquarks leads to indistinguishable jets. If one includes nontrivial fragmentation functions, e.g. for a $c$-quark, the above rules have to be modified.

![Figure 2. Scattering Topology (see the Figures for other topologies in [24])](image-url)
All the above ideas are implemented in CompHEP in the special model denominated as \( SM_{ud} \). For the processes where only QCD interactions are involved one can simplify further the light quark flavor combinatorics since QCD does not distinguish different flavors and electroweak charges. The corresponding CompHEP model is denominated as \( SM_{qQ} \).

2.2. EVENT COLOR FLOWS

As will be described below, the CompHEP has an interface to PYTHIA program. (The interface to HERWIG is in progress). For a fragmentation and hadronization PYTHIA needs an information about color configurations for the generated partonic events. For the color flows PYTHIA requires an indication of colorless pairs or, in other words, the definition of color chains. In terms of QCD the sequences of color chains correspond to the elements of the orthogonal color basis in the limit of infinite number of colors, \( N_c = \infty \). However CompHEP, as well as other matrix element programs, calculates exactly the matrix element squared with the number of colors \( N_c = 3 \).

The procedure described in this section is used in CompHEP in order to match the \( |M|^2 \) calculation with \( N_c = 3 \) and the color chains generation needed for further hadronization. Note, that originally this algorithm has been implemented in CompHEP 4.1.10 by A. Pukhov.

The matrix element is a sum of Feynman diagrams

\[
M = \sum_d L_d T_d
\]

where \( L_d \) is the Lorentz part of the diagram, and \( T_d \) is its color structure. In the limit \( N_c = \infty \) the color structure of diagrams can be decomposed over the elements of color basis (color chains) as

\[
T_d \Rightarrow T_d^{N_c=\infty} = \sum_i C_d^i t_i
\]

where the elements of the basis are orthonormal, \( t_i \otimes t_j = \delta_{ij} \). Thus, the squared matrix element can be represented in the limit \( N_c = \infty \) as a sum of positive contributions

\[
|M|_{N_c=\infty}^2 = \sum_i k_i, \quad k_i = \left| \sum_d L_d C_d^i \right|^2
\]

As a result, for each event (phase space point) one gets the set of color chains \( t_i \) with weights \( w_i = \sum k_i \).

So the event generation procedure includes two steps. First step is the standard generation of an event as a phase space point with \( |M|^2 \) calculated at \( N_c = 3 \) like in the standard event weighting procedure. At the second step color chains \( t_i \) are selected with the weighting factors \( k_i \), \( N_c = \infty \).

2.3. LANHEP AND SUSY MODELS

The LanHEP program \[25\] was developed in the framework of the CompHEP project to implement easily "new physics" models and/or anomalous effective Lagrangian terms. LanHEP input is the Lagrangian written in the compact form close to one used in textbooks. In particular, the Lagrangian terms as the LanHEP input can be written with summation over indices of the space-time or gauge symmetries and using special symbols for complicated expressions, such as the covariant derivative and the strength tensor for gauge fields. There are 2-component spinors and the superpotential formalism available. Correctness of the model can be checked for hermiticity, mass matrix diagonalization, and BRST invariance. The output of the program is a set of Feynman rules in terms of physical fields and independent parameters. This output can be written in \LaTeX{} format and in the form of CompHEP model files, which allows one to start calculations of processes in the new physical model.

We have used the description of the MSSM physical spectrum and the Lagrangian for 3 generations of quarks and leptons given in \[26\]. We also add the effective Higgs potential to the tree level MSSM Lagrangian to incorporate the radiative corrections as it is given in the FeynHiggs program \[27\] to the masses and couplings of Higgs bosons in a gauge invariant way. Details of the notation for the particles and parameters of the MSSM in CompHEP are given in \[28\]. The Feynman rules generated by LanHEP were successfully compared with the ones listed in \[26\] as well as with many results of other publications. The comparison to GRACE/SUSY \[2\] program was successfully done for many processes.
The SUGRA and GMSB SUSY models are also implemented to the CompHEP by means of linking with the ISAJET [29] library. There are also MSSM extensions available, such as R-parity violation model and the model with gravitino and goldstino fields [30].

Lagrangians for other nonstandard models, e.g. models with leptoquarks [31,32], complete two-Higgs-doublet model with conserved or broken CP invariance [33], models with anomalous top quark interactions [34] are also available for CompHEP calculations.

2.4. COMPHEP-PYTHIA INTERFACE

PYTHIA has the built-in database of matrix elements for hard subprocesses which are basically of the $2 \to 2$ type. The CompHEP-PYTHIA interface (the old version of the interface is described in [22]) allows to use the processes $2 \to 3, 4, 5, 6$ computed by means of CompHEP as the "new" processes for PYTHIA and to include in this way ISR/FSR, hadronization, and decays like it is done in PYTHIA.

CompHEP generates unweighted events and writes them to the file $events_N.txt$ where $N$ is the number of working session. The command $mixPEV$ (syntax: $./mixPEV events_{N1}.txt events_{N2}.txt ...$) mixes randomly several event files corresponding to some subprocesses in one event flow according to their relative weights, $\sigma_i/\Sigma_j \sigma_j$, where $\sigma_i$ is the cross section of the $i$-th subprocess, and writes events to the file $Mixed.PEV$. This file is ready for input to PYTHIA 6.2 and contains the data necessary for Les Houches Accord [1,15].

We provide the interface library libinterface62.a and an example of the main.f program which is very easy to be incorporated to any standard PYTHIA main.f program. Using the standard Unix "make" command a user links main.o, libinterface62.a and pythia62.o to the executable file $generator.exe$. The generator takes events from the file Mixed.PEV. The number of requested events to be generated after PYTHIA is defined by user in the file $INPARM.DAT$.

There are already many examples of the CompHEP and CompHEP-PYTHIA interface use for the real analysis at the Tevatron, LHC and LC.

Of course, as pointed out in the introduction, one has to be careful combining new higher order processes calculated with CompHEP-PYTHIA interface, and the standard PYTHIA processes. Double counting of contributions should be carefully analysed, not only when ISR/FSR are switched on, but in many other cases.

2.5. COLLIDING BEAMS IN COMPHEP

New option is implemented in CompHEP 4.4 - a possibility to enter a beam as the initial state. In CompHEP the beam is a set of partons. This is a natural terminology for hadronic collisions (proton or antiproton beams). In CompHEP 4.4 this option allows a user to enter also an electron, photon or quark itself as a beam. In this way one can set, for example, the photon as a "parton" in electron (effective photon approximation), quarks and gluons as "partons" in a photon (resolved photons), W-boson as a "parton" in electron or quark (effective W-approximation) etc. This option is specially useful for lepton and lepton-hadron colliders. Moreover, one can introduce a beam with arbitrary content for any private purposes.

The parton distribution functions are assigned now to the beam by a user before the process is entered. All predefined beams are collected into a special table which can be modified by a user if necessary. When the user sets a process (s)he indicates a particle or a beam (from the table) as an initial state. The corresponding table can be displayed on the screen by pressing the F3 key.

In the CompHEP 4.4 the latest versions of the CTEQ parton distribution functions [35] are implemented. However, some previous sets are also available since for an analysis where they could be needed.

2.6. SYMBOLIC BATCH MODE

The main goal of the symbolic batch script is to launch the symbolic calculation of squared diagrams using a command line interface (the non-GUI mode). This allows the reuse of the script, avoiding typing mistakes, and also helps to send large symbolic calculation tasks to the computer farms.

The script imitates the typing input data in
CompHEP menu (initial states, energies, final particles etc). This option was described in [36] - so-called blind mode in CompHEP 4.1.10.

All the data, necessary for further symbolic calculations, are taken from a file with a default name process.dat. Detailed instructions and explanations for the parameters are contained in this file.

After filling in the file process.dat a user launches the script. The results of the script execution will be the binary file `comphep.exe` and the file `symb_batch.log` where all details on the binary creation are stored.

One can set another name for the data file or the "results" directory by the -n and -d options respectively. If the CompHEP directory "results" is not empty before the script is launched, it is renamed to results_`old_0` and the script creates a new "results" directory. The symbolic batch script has some extra options:

- **recovery** if the `s_comphep.exe` program crashes during calculations of squared diagrams, one can launch the script with the recovery option, and the `s_comphep` resumes computations from the last calculated diagram;

- **relink** if the user has changed the `userFun.c` file and wants to relink the `n_comphep` program with a new file, (s)he can launch the script with the relink option. All details of the relinking are stored in the `symb_batch.log` file;

- **show diagrams** this option is applied if the user wants to exclude by hand some diagrams from the calculation. The script launches `s_comphep` in GUI mode in the Feynman diagrams menu. After reviewing the diagrams the user has to finish the GUI session and the script will go on.

### 2.7. NUMERICAL BATCH MODE

The CompHEP was originally created as a program with an interactive GUI interface. However in practice for many cases a command line interface (non-GUI) is very useful and it was implemented in the CompHEP 4.4 numerical batch mode. The interface allows to perform large time scale calculations using, if necessary, the computer farms (PBS and LSF systems), to compute processes with a large number of subprocesses, in particular, in parallel if computer resources are enough. The interface is very useful for repeating numerical calculations by non-expert users without making trivial mistakes.

In this mode numerical computations start after finishing the symbolic computations and creation of the executable file `n_comphep.exe` in the directory "results". At the first step a user has to create the numerical batch data file batch.dat by launching the command `./num_batch.pl` from the user working directory, where CompHEP has been started. The created file batch.dat contains the default parameters (model parameters, cuts, kinematics, etc.) listed in the file session.dat in the directory "results". If necessary the user can change any of these parameters by editing the file batch.dat or using GUI interface to modify initial session.dat file. In the latter case one can start the command `./num_batch.pl` with the option — `--add_sess2bat` to include parameters from the new session.dat file to the batch.dat.

The command (script) `./num_batch.pl` has many other useful options, for example:

- **-help** - to list the complete set of options with an explanation

- **-d dir_name** - to change the default directory results to the directory `dir_name`

- **-add_sess2bat [file]** - to add parameters from existing file session.dat in the directory results or from `[file]`, if indicated, to the batch.dat file

- **-run [vegas,max,evnt]** - to start numerical calculations including all steps by default or only the steps as indicated by the keys [vegas,max,evnt], where [vegas] means the cross section calculation, [max] - search for maxima in each of the phase space sub-cubes, [evnt] - unweighted event generation

- **-run [cleanstat,cleangrid,clean]** - to clear statistics or grid or both

- **-proc n1,n2-n3,n4...** - to run only the subprocesses with corresponding numbers


- **-pbs ["pbs prefix"]** - to start calculations in parallel with PBS batch system. The default PBS prefix is: “qsub -I ”
- **-lsf ["lsf prefix"]** - to start calculations in parallel with LSF batch system. The default LSF prefix is: “bsub -I ”

For example, the command

```
./num_batch.pl -run vegas -pbs
```

creates temporary subdirectories for each of the subprocess and run the PBS batch job for each of subprocess with PBS command: qsub -I run.sh, where the script run.sh is created automatically and starts the n_complep.exe. These temporary subdirectories will be automatically removed and the results for all subprocess will be saved in one directory.

A user can also calculate subprocesses in parallel without the PBS (LSF) batch system by starting the same command with empty PBS (LSF) prefix:

```
./num_batch.pl -run vegas -lsf
```

### 2.8. FUTURE PLANS

There are several directions for the future development of CompHEP. One important direction is a further development of distributed Monte Carlo calculation and event generation on computer clusters. Another important direction is an implementation of the FORM computer algebra program for symbolic calculations of matrix elements (some preliminary realization was described in [38]). This option will allow to introduce new complicated structures in the vertices (e.g. form-factors), implement new algorithms to increase the efficiency of symbolic calculations, perform calculations in theories with extra dimensions, use the dimensional regularization, perform polarized calculations by introducing the corresponding density matrices for external lines of squared diagrams. We plan to create the Les Houches accord 1 based interface to HERWIG and the Les Houches accord 2 based interface to PDFs as well as the SUSY Les Houches accord based interface to various codes calculating the SUSY mass spectrum.

The long term CompHEP collaboration projects which are under discussion include symbolic and numerical amplitude calculations with extension to the 1-loop case using various methods, incorporation the gauge invariant classes of diagrams, etc.

### 3. THE COMPHEP BASED GENERATOR SINGLETOP

The CompHEP has been used to prepare a special event generator SingleTop to simulate the electroweak single top quark production with its subsequent decays at the Tevatron and LHC. Single top is expected to be discovered at the Tevatron Run II and will a very interesting subject of detail studies at the LHC (see the reviews [41][42]).

There are three main processes of singe top production at hadron colliders characterizing by the virtuality $Q_W^2$ of the participating at the process $W$-boson: t-channel, s-channel and associated $tW$ mechanisms respectively.

The generator SingleTop includes all the three processes and provides Monte-Carlo unweighted events at the NLO QCD level. We discuss shortly here only the main process with the largest rate, the t-channel production. The representative NLO diagrams are shown in Fig. 3. The top decay is not shown, however it is included with all the spin correlations.

![Figure 3. LO order and representative loop and tree NLO diagrams to the t-channel single top production](image)

We compute by means of the CompHEP the LO order process $2 \rightarrow 2$ with the b-quark in the initial state and top spin correlated $1 \rightarrow 3$ subsequent decay, put it into PYTHIA using the interface and switch on ISR/FSR. Then with CompHEP we compute the NLO tree level corrections $- 2 \rightarrow 3$ processes with additional b- and light
quarks or gluons in the final state including also the top decay with spin correlations. We split the phase space region in "soft" and "hard" parts on $p_t$ of those additional $b$ and light jets being from PYTHIA radiation in the "soft" and from the CompHEP matrix element calculation in the "hard" regions. The soft part is normalized in such a way that all parts being taken together give known from calculations the NLO cross section $[43,44]$. The splitting parameters are turned such that all the distribution become smooth after the normalization. The performed cross checks show that the computed NLO distributions $[44]$ are correctly reproduced. Therefore, prepared in that way generator does not have a double counting, gives the NLO rate and distributions, and include all the spin correlations.

The first release of the generator $[15]$ did not include the hard radiation of the light jets, while the latest version $[46]$ currently used in the analysis by the Fermilab DO and the LHC CMS collaborations includes all the mentioned properties.

There are several recent examples of the CompHEP use for various simulations, namely, the generator for MSSM Higgs bosons in the intense coupling regime at the LHC $[47]$, the generator for sfermion pair production with their subsequent decays to polarized fermions $[48]$ and the generator for measurements of the $HZ\gamma$ coupling of $[49]$ at Linear Colliders.

4. CONCLUSIONS

The CompHEP package is a powerful tool for a simulation of different physical processes at hadron and lepton colliders. The most important CompHEP 4.4 improvements include the build-in MSSM, simplification of quark flavor combinatorics, generation of unweighted events with color chains defined for the following showering and hadronization procedure, Les Houches accord based interface with PYTHIA, symbolic and numerical batch modes, the latest PDF sets implementation and a new treatment of the initial beams.

The CompHEP is basically the LO program. However it allows to include partly the NLO corrections. NLO tree level $2 \rightarrow N$ process can be computed. One can include the NLO structure functions and loop relations between the parameters like in the MSSM, K-factors when available, and loop contributions from the existing publications as form-factors. However a correct matching to complete NLO is obviously a nontrivial problem which has to be carefully considered in each particular physical case. The CompHEP based generator Single-Top is an example of a reasonable matching of the different-order contributions.

5. ACKNOWLEDGMENTS

The authors thank the Russian Ministry of Industry, Science and Technology for the support of the CompHEP project. We acknowledge the support of the RFBR grant 04-02-17448 and the grant of the program "University of Russia". We are grateful to our colleagues for using the CompHEP, for reporting us on problems and bugs.

We specially thank the Organizing Committee of the ACAT2003 Conference for a support.

REFERENCES

37. A. Kryukov, L. Shamardin. In this proceedings.
40. V. Bunichev, A. Kryukov and A. Vologdin,
42. LHC/LC Study Group Working Document.
   In preparation
45. E. Boos, L. Dudko and V. Savrin, CMS Note 2000/065
46. E. Boos, L. Dudko, V. Savrin, and A. Sherstnev, in preparation