

# MadGraph/MadEvent v4: The new web generation

---

**Johan Alwall**

*Stanford Linear Accelerator Center, Stanford University,  
Stanford, CA 94309, USA*

*E-mail:* alwall@slac.stanford.edu

**Pavel Demin, Simon de Visscher, Rikkert Frederix, Michel Herquet, Fabio Maltoni**

*Centre for Particle Physics and Phenomenology (CP3),  
Université Catholique de Louvain,  
Chemin du Cyclotron 2, 1348 Louvain-la-Neuve, Belgium*

*E-mails:* simon.de.visscher@fynu.ucl.ac.be, pavel.demin@fynu.ucl.ac.be,  
frederix@fyma.ucl.ac.be, mherquet@fyma.ucl.ac.be, maltoni@fyma.ucl.ac.be

**Tilman Plehn**

*School of Physics, University of Edinburgh,  
James Clerk Maxwell Building, The King's Buildings, Edinburgh EH9 3JZ, Scotland, UK*

*E-mail:* tilman.plehn@cern.ch

**David L. Rainwater**

*Department of Physics and Astronomy, University of Rochester,  
Rochester, NY 14627, USA*

*E-mail:* rain@pas.rochester.edu

**Tim Stelzer**

*Department of Physics, University of Illinois at Urbana-Champaign,  
1110 West Green Street, Urbana, IL 61801, USA*

*E-mail:* tstelzer@uiuc.edu

ABSTRACT: We present the latest developments of the MadGraph/MadEvent Monte Carlo event generator and several applications to hadron collider physics. In the current version events at the parton, hadron and detector level can be generated directly from a web interface, for arbitrary processes in the Standard Model and in several physics scenarios beyond it (HEFT, MSSM, 2HDM). The most important additions are: a new framework for implementing user-defined new physics models; a standalone running mode for creating and testing matrix elements; generation of events corresponding to different processes, such as signal(s) and backgrounds, in the same run; two platforms for data analysis, where events are accessible at the parton, hadron and detector level; and the generation of inclusive multi-jet samples by combining parton-level events with parton showers. To illustrate the new capabilities of the package some applications to hadron collider physics are presented:

- I. Higgs search in  $pp \rightarrow H \rightarrow W^+W^-$ : signal and backgrounds.
- II. Higgs CP properties:  $pp \rightarrow Hjj$  in the HEFT.
- III. Spin of a new resonance from lepton angular distributions.
- IV. Single-top and Higgs associated production in a generic 2HDM.
- V. Comparison of strong SUSY pair production at the SPS points.
- VI. Inclusive  $W$ +jets matched samples: comparison with the Tevatron data.

KEYWORDS: Event Generation, Collider physics, New Physics.

---

## Contents

<b>1. Introduction</b>	<b>2</b>
<b>2. The path to event generation</b>	<b>3</b>
<b>3. The Standard Model implementation</b>	<b>7</b>
3.1 Standard Model interactions	7
3.2 Higgs Effective Theory	7
<b>4. Going beyond the Standard Model</b>	<b>8</b>
4.1 The 2HDM implementation	9
4.2 The MSSM implementation	11
4.3 The User Model	12
<b>5. Matching of jet production by parton showers and matrix elements</b>	<b>15</b>
<b>6. Tools</b>	<b>16</b>
6.1 Analysis platforms: MadAnalysis and ExRootAnalysis	16
6.2 Decaying unstable particles in the SM and beyond	17
6.3 Hadronization and detector simulation using Pythia and PGS	17
<b>7. Applications to hadron collider physics</b>	<b>18</b>
7.1 Higgs search in $pp \rightarrow h \rightarrow W^+W^-$ : signal and backgrounds	18
7.2 Higgs CP properties: $pp \rightarrow hjj$ in the HEFT	21
7.3 Spin of a new resonance from lepton angular distributions	21
7.4 Single-top associated Higgs production in a generic 2HDM	24
7.5 Comparison of strong SUSY pair production at the SPS points	26
7.5.1 Setup and generation	26
7.5.2 Comparison of the results for the SPS points	27
7.6 Inclusive $W$ +jets matched samples: comparison with the Tevatron data	30
<b>8. Conclusions and Outlook</b>	<b>32</b>

---

## 1. Introduction

Accurate simulation of both signal and background will play a key role in making discoveries at the LHC. A well-known example is given by supersymmetric models, one of the most studied and cleanest templates for physics beyond the Standard Model. Characteristic signatures, from large rate multi-jet plus missing  $E_T$  events coming from squark pair production, to extra  $b$ -jets and  $\tau$ 's from Higgs production, have in general large Standard Model backgrounds that need to be measured from the data and/or well described by Monte Carlo's. At the same time, the most important distinctive features of the signal, such as, for instance, mass distributions, kinematic edges and angular correlations will be exploited not only to improve the signal over background ratio but also to identify the quantum numbers (*e.g.*, spin, color) of the intermediate heavy states which decay.

The need for better simulation tools has spurred an intense activity over the last five years, that has resulted in several important advances in our ability to accurately simulate hard interactions. At the matrix element level, these include the development of general purpose event generators, such as MadGraph [1] and MadEvent [2], ComPHEP/CalcHEP [3, 4], SHERPA [5] and WHIZARD [6], high efficiency multiparton generators which go beyond the usual Feynman diagram techniques, such as ALPGEN [7] and HELAC [8], as well as Monte Carlo's that include NLO corrections, such as MCFM [9] and MC@NLO [10].

An accurate simulation of a hadronic collision requires a careful integration of the matrix element hard process with the full parton showering and hadronization infrastructure [11, 12]. Here too significant advances have been made in the development of matching algorithms such that by Catani, Krauss, Kuhn and Webber (CKKW) [13, 14, 15], by Mangano (MLM) [16] and by Lavesson and Lonnblad [17] and in their comparison [18, 19]. A breakthrough was also achieved by Frixione, Webber and Nason [10, 20] who showed how to correctly interface an NLO calculation with a parton shower to avoid double counting, releasing the first event generator at NLO, MC@NLO. Further developments and refinements in these directions are ongoing [21, 22].

While automatization has not yet been achieved at NLO, at tree-level we are now in a position where the matrix elements can be calculated by several different tools, and integrated in multiple ways into different parton-shower/hadronization codes to provide multiple checks of the accuracy of the simulation.

Each of the tools mentioned above was developed with a unique approach optimized to the meet the authors intentions. This diversity of approaches, that include a wide range of overlap in physics reach is a key element to strengthening the programs and providing confidence in their results. It is important to understand the philosophy/intentions of the authors to understand and utilize their code.

The single underlying principle in the development of MadGraph and MadEvent was to develop a tool that would maximize the amount of time the physicist could concentrate on physics, and minimize the obstacles between having an inspired idea, and being able to compare it to experimental data. This originated first with MadGraph and then with MadEvent. At a time when many phenomenologists were spending enormous amounts

of time and energy performing important tree level calculations, MadGraph/MadEvent was able to automate the whole process, from the calculation of matrix elements to the generation of unweighted parton-level events, allowing physicists to concentrate on other pressing issues.

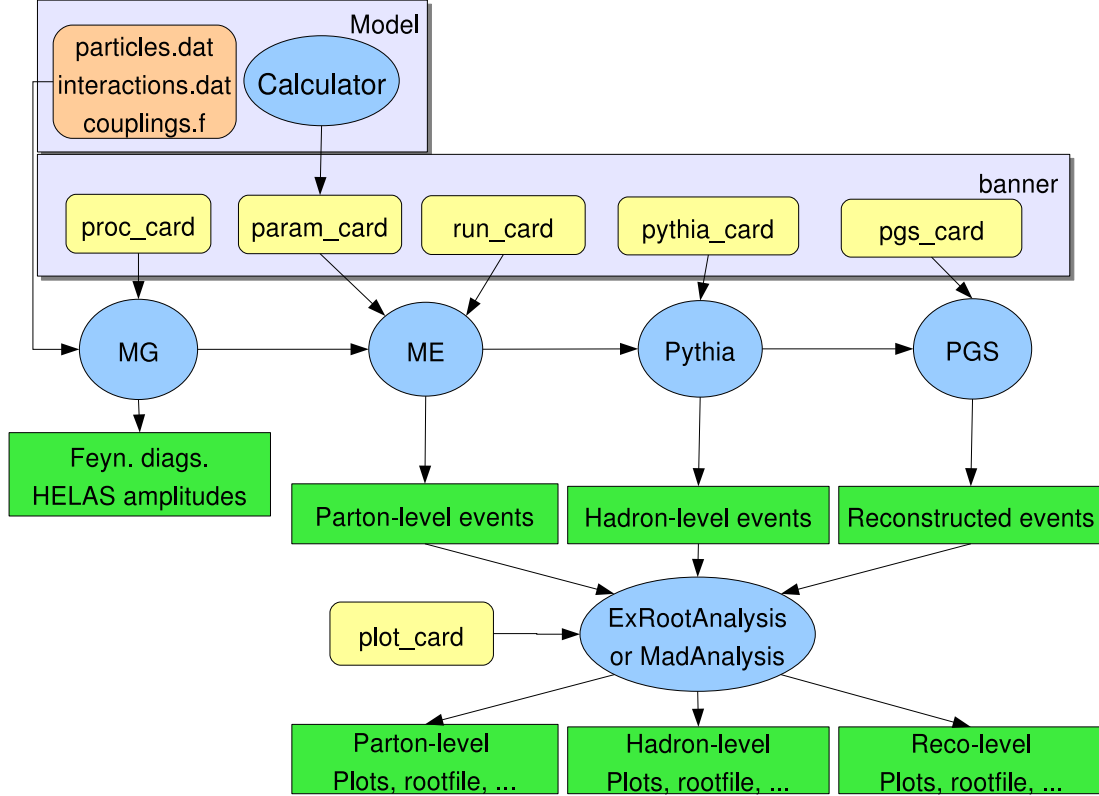
This current release of the MadGraph/MadEvent package is the natural extension of this project both in the theoretical, and the experimental directions. While the original MadGraph had the standard model “hard-coded” this version includes several new models (MSSM, 2HDM, HEFT,...), as well as the capability for user defined models. It also includes the ability to seamlessly pass the events through a full hadronic simulation so they can be subjected to a complete detector simulation. Utilizing the strengths of the web, a physicist can now literally go from an inspired concept to a full event simulation with detector reconstruction with just a few clicks of the mouse. In addition, in order to facilitate the communication between theorists and experimentalists, the complete process from diagram generation to event analysis can be recreated from a simple text file which collects all the input cards used in the various phases of the simulation.

In this work we briefly illustrate the new features of the code, focusing more on the structure and on general aspects than on the technical details. In Sections 2 and 3 we give an overview of the package. In Section 4 we discuss the models for which a dedicated implementation exists (SM+HEFT, MSSM, 2HDM) as well as a framework where new models can be implemented starting from the SM. In Section 6 our approach to generate multi-jet samples through matching is described. Section 7 collects the descriptions of some of the tools available that allow the users to handle the parton-level events in the Les Houches format for different purposes: from analysis to further simulation, to the decay of unstable particles maintaining the leading spin-correlations. In Section 8 we provide several examples, from SM measurements to the search for SUSY in inclusive signatures, of studies that can be performed with our package. We draw our conclusions and present our current line of research and development in the final section.

## 2. The path to event generation

The new structure of the MadGraph/MadEvent package is shown in Fig. 1. Each of the four steps (code creation, parton-level event, hadron-level event and reconstructed object generation) are driven by input cards provided by the user. All cards have a comment section which describes in detail their content and syntax. Here we just mention their functionalities and how the generation proceeds, as outlined in Fig. 1. For sake of simplicity, we consider the procedure to be followed by a web user, who can control the full generation process on one of the MadGraph clusters, via a web interface. The steps undertaken by a user working locally on his/her own computer, are the same, but in this case the simulation is initiated by calling scripts from the command line. In this case, the exact syntax to be used can be found in the corresponding `README` files and in our on-line documentation.

The first phase is the creation of the code specific to the process that the user requests. This can be done directly by filling a web form or by editing the process card, `proc_card.dat`, to identify the process or processes to be included in the code, together



**Figure 1:** Flow chart of the MadGraph/MadEvent event generation. Each step is controlled by a specific card. Details are given in the text.

with the model and the maximum order of the couplings to appear in the tree-level diagrams. Labels can be specified to identify multi-particles (such as jets) which defines summations over physical particles. Examples of process cards are given in the Section 7. Once completed, the process card is read by MadGraph and a process-specific code (packaged in a gzipped tar file) is created and can either be downloaded in order to run the simulation locally, or be run directly on the cluster using a web interface.

Two cards are needed to perform the event generation. The first one is the parameter card, `param_card.dat`, which contain the numerical values of the necessary parameters for a specific model. The parameter card has a format compliant with the SUSY Les Houches Accord and it is dependent on the physics model. At variance with the previous version, secondary parameters and widths are no longer computed in MadEvent, but instead passed through the parameter card also for the Standard Model. Such a card can be generated with an independent code, which we refer to as a “Model Calculator”. There are calculators available on the web for the SM, MSSM and 2HDM models, which, starting from the parameters in the Lagrangian (primary parameters) calculate all secondary parameters (such as masses, widths, and auxiliary parameters) needed by MadGraph to perform the

cross section integration. We stress that the calculation of all such parameters should be done by a Model Calculator since they must be consistently performed at tree level. The second card to be completed is the run card, `run_card.dat`, where all the information regarding the event generation (number of events requested, random seed, ...), the collision (energy, beam type, pdf's, scales) and the acceptance of the detector ( $p_T^{min}$ ,  $\eta_{max}$ , ...) is passed. Both cards can be generated using a web form, or uploaded. Once the cards are completed, the event generation can start.

As a result of the run, the user will obtain a data file with the unweighed events in the Les Houches format, a set of plots of kinematical variables, and a ROOT file.

The parton-level events can then be passed to Pythia by specifying in the `pythia_card.dat` the desired options. Events at the hadron level are then passed to PGS, whose parameters are set in the `pgs_card.dat`. On the web the Pythia and PGS cards can be completed at the same time as the parameter and run cards or at a later time. In the first case the generation goes all the way through the reconstructed objects at the detector level, and the information (events and plots) of intermediate stages of the simulation is kept.

The MadGraph/MadEvent package can be used at different levels, depending on the user's needs/expertise. We have identified three main categories:

- A **web user** performs the full generation (from the process specification to the Pythia or PGS events) over the Web, downloading only the input cards and the results (events and plots). No code download or installation is needed as all phases are handled via a web interface and the generation is performed on one of our clusters. Each user manages a personal database where all the codes/runs/events are stored for future use. This is the simplest level of accessing MadGraph/MadEvent, and it is suitable to both theorists and experimentalists who might want a quick determination of a cross section or a fair amount of events to conduct preliminary studies using the available physics models. Code creation and all other tools on the web page are accessible to anybody upon web registration.

For a web user it is enough to understand the logic and the steps of the code, Fig. 1, and to correctly manage the information which is passed through the input cards. This is explained in detail in our on-line documentation.

- A **local user** creates the code specific to the process of interest over the web and then downloads and runs it locally on her/his personal computer or cluster. This allows somewhat more flexibility (for instance, special selection cuts can be implemented, as well as different PDF's, dynamic scale choices, and so on), as well as the possibility for the production of large samples of events. In this case, it is suggested to run in parallel over a cluster with the PBS [23] or CONDOR [24] batching systems.
- A **developer** downloads the full MG/ME package to exploit all functionalities locally. This is certainly the preferred option not only for the users making detailed and/or sophisticated simulation studies, but also for developers of new tools or models in the MG/ME framework. We underline in passing that experiments or larger collaborations that have at their disposal a computer farm running a batching system, such as

Dir name	Content
Cards	steering cards
Source	proc. independent source files
SubProcesses	proc. dependent source files and dirs for the subprocesses
Events	LHE event, plots and ROOT files
bin	executables and scripts (csh and Perl)
lib	libraries and PDF data
HTML	web pages

**Table 1:** Organization of the process directory (Template). When the code specific to a user’s process is generated over the web, it is packaged into `MadEvent.tar.gz` which has the structure above.

PBS or CONDOR, might also install the web server software, which manages all the generation steps (code creation, event generation, interfacing to Pythia and PGS, as well as ROOT-based analysis) and user databases automatically. The corresponding web server software and tech support can be obtained from the authors.

The installation of the MG/ME package is straightforward since no external libraries are needed. The interested reader should refer to the README file contained in the main directory, obtained after decompressing the `MG_ME_V4.X.Y.tar.gz` archive file. The main directory is organized in subdirectories whose contents we briefly describe below.

- **MadGraphII** : The source code (Fortran) and executable (`mg2`) for the matrix-element generator. `mg2` generates not only a code for the matrix element squared for a given process, but also all subprocesses contributing to it. When interfaced with MadEvent it also generates the mappings needed for the phase space integration, including the Single-Diagram-Enhanced expressions. `mg2_sa` is the “stand alone” version of MadGraph producing the code for the matrix element squared only.
- **Template** : The general structure of a process dependent code, as shown in Table 1. To create the code corresponding to a new process, a copy of the full directory should be made (at the main directory level) and the script `./bin/newprocess` invoked from it.
- **Models** : It contains the available models (each of them is a directory) including the user model template (`usrmod`). The user developing a new model makes a copy of the `usrmod` directory at this level, giving a name to it that identifies the model (which can later be set in `proc_card.dat`).
- **HELAS** : The HELAS library [25] source code.
- **DECAY** : The source code to decay unstable particles directly from parton-level LHE files.



### 3. The Standard Model implementation

#### 3.1 Standard Model interactions

The Standard Model of particles and interactions, based on the  $SU(3)_c \times SU(2)_L \times U(1)_Y$  gauge symmetry has been available since the first versions of both MadGraph and more recently of MadEvent. There is, however, one important differences w.r.t. the previous version of the package, regarding how the couplings of the models are handled. As was already mentioned in the previous section, the task of computing from the parameters in the Lagrangian (primary parameters) all the secondary parameters (masses, widths and dependent parameters) needed by MadGraph is left to an external program, the SM Calculator. The output of the SM Calculator is a parameter card, `param_card.dat`, which contains the numerical values of the main couplings (primary and secondary) of a specific model. The parameter card has a format compliant with the SUSY Les Houches Accord.

A simple example is given by the EW parameters that characterize the gauge  $SU(2)_L \times U(1)_Y$  interactions and its breaking: in the Standard Model there are five relevant parameters,  $\alpha_{em}, G_F, \sin \theta_W, m_Z, m_W$  of which only three are independent at tree level. Various schemes differing by the choice of the parameters considered independent are used in the literature. In the SM Calculator, the default is to take  $G_F, m_Z, m_W$  as inputs and derive  $\alpha_{em}, \sin \theta_W$ , but other choices are available. As a result a consistent and unique set of values of the couplings appearing in the Feynman rules is derived and used for the computation of the amplitudes.

Another sometimes important feature of our SM implementation, is the possibility of distinguishing between the kinematic mass (pole mass) for the quarks and that entering in the Yukawa coupling definition ( $\overline{MS}$  mass). For the latter, the user can choose to evolve the mass to the scale corresponding to the Higgs mass, which leads to an improvement of the perturbative expansion.

Finally, we mention that various versions of the Standard Model are actually available for specific studies. For example, in the “minimal SM” (`sm`) the CKM matrix is diagonal while in the `sm_ckm` model a mixing between the first and second generation is allowed (Cabibbo angle). Another example is the `sm_nohiggs` model where the Higgs has been eliminated and the EWSB sector behaves as a non-linear sigma-model.

#### 3.2 Higgs Effective Theory

The Higgs effective field theory (`heft`) model is an ‘extension’ of the Standard Model (SM), where the Higgs boson couples directly to gluons (and photons) [26, 27, 28]. In the SM these couplings are present through a heavy (top) quark loop. For a not too heavy Higgs ( $m_h < 2m_t$ ), it is a good approximation to take the mass of the heavy quark in the loop to infinity<sup>1</sup>. This results in effective couplings between gluons and Higgs bosons.

---

<sup>1</sup>For this approximation to hold, not only should the Higgs mass be smaller than twice the top mass, also all other kinematic variables, such as the transverse momentum of the Higgs boson, should be smaller than  $2m_t$  [29].

The effective vertices can be derived from the effective dimension five Lagrangian

$$\mathcal{L}_h = -\frac{1}{4}g_h G_{\mu\nu}^a G_{\mu\nu}^a \Phi, \quad (3.1)$$

where  $G_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf^{abc}A_\mu^b A_\nu^c$ . The coupling constant  $g_h$  is given by

$$g_h = \frac{\alpha_s}{3\pi v}. \quad (3.2)$$

Due to the non-abelian nature of the  $SU(3)_C$  color group the effective vertices do not only include two, but also three and four gluons coupling to the Higgs boson. Since MadGraph can work only with three- and four-point vertices, the four-gluon interactions in the **heft** model are obtained by rewriting the QCD four-gluon interaction in terms of three-point vertices with an extra *non-propagating* internal tensor particle,  $T$  [30, 31]. This trick can be easily understood by noting that the usual (text-book) form of the four-gluon interaction is the sum of three terms, whose color and Lorentz structure correspond to  $2 \rightarrow 2$  diagrams where a color octet tensor is exchanged in the  $s, t, u$  channels. With the introduction of this extra particle, the four-gluon-Higgs vertices can be reduced to diagrams with at most four-point vertices. To get the standard diagrammatic visualization of four-gluon and four-gluon-Higgs vertices it is sufficient to contract the  $T$  particle lines to a single point.

The gluon couplings to a pseudo-scalar Higgs are also implemented. The name of the pseudo-scalar Higgs in MadGraph is **h3** (*i.e.*, the same as in the 2HDM and MSSM models). The effective dimension five Lagrangian for the pseudo-scalar Higgs coupling to the gluons is

$$\mathcal{L}_A = \frac{1}{2}g_A G_{\mu\nu}^a \tilde{G}_{\mu\nu}^a \Phi_A, \quad (3.3)$$

where  $\tilde{G}_{\mu\nu}^a$  is the dual of  $G_{\mu\nu}^a$ ,  $\tilde{G}_{\mu\nu}^a = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}G_{\rho\sigma}^a$ . The effective coupling constant  $g_A$  is given by

$$g_A = \frac{\alpha_s}{2\pi v}. \quad (3.4)$$

The pseudo-scalar Higgs has only effective couplings to two or three gluons. The four-gluon-pseudo-scalar Higgs vertex is absent due to the anti-symmetry of the epsilon tensor  $\epsilon^{\mu\nu\rho\sigma}$ . If a mixed Higgs with no definite CP parity is needed, it is sufficient to change the couplings of the Higgs to the gluons. First generate the process with the SM Higgs, then, after downloading the code, change the coupling in the `./Source/Model/couplings.f` file. The coupling constant is defined as a two-dimensional object, where the first and second elements are the CP-even and CP-odd couplings of the Higgs to the gluons, respectively. The HELAS subroutines automatically use the correct kinematics for odd-, even- or mixed CP Higgs's coupling to the gluons. At present, the implementation allows production of only one Higgs-boson. The effective couplings of two Higgs bosons to gluons are available in HELAS, but not yet included in the HEFT model.

#### 4. Going beyond the Standard Model

In the current version a few new physics models have been added and fully tested: the minimal supersymmetric extension of the Standard Model (MSSM) and the general two

Higgs doublet model (2HDM). In addition, a framework for setting up new models is available. Here, we briefly describe the main features of these new implementations.

#### 4.1 The 2HDM implementation

The two-Higgs-doublet model (2HDM) has been extensively studied for more than twenty years, even though it has often been only considered as the scalar sector of larger models like the MSSM [32] or Little Higgs models [33]. The generic 2HDM considered here may display by itself an interesting phenomenology justifying its study. As a non exhaustive list, let us mention new sources of  $CP$  violation in scalar-scalars interactions [34], tree-level flavor changing neutral currents (FCNCs) due to non diagonal Yukawa interactions, dark matter candidates [35] or Higgs bosons lighter than the LEP bound [36].

In the “full” version of the model (`2hdm_full`), no particular restrictions are imposed on the interactions allowed by gauge invariance, except electric charge conservation. Many diagrams involving tree-level FCNCs and violating the  $CP$  symmetry are thus present. The user who is not interested in these phenomena should use the “simplified” version of the model (`2hdm`), where the number of generated diagrams is in general much smaller.

The following naming convention is used: `h+` and `h-` stand for the positively and negatively charged Higgs bosons and `h1`, `h2` and `h3` stand for the neutral ones. Since the  $CP$  invariance of the potential is not assumed, the neutral bosons are not necessarily  $CP$  eigenstates and the standard naming convention in this case (*i.e.*, `h1` being the lightest one and `h3` the heaviest one) is used.

TwoHiggsCalc is the calculator associated with the model. It has been written in C and is accessible from a web interface. It has been designed to compute input values for the 2HDM extension of MadGraph/MadEvent but it can also be used as an independent tool. Starting from various parameters of the Lagrangian, such as the vacuum expectation values (vevs) or the Yukawa couplings, the program computes useful secondary physical quantities at leading order such as the scalar mass spectrum, the mixing matrix, the total decay widths and the branching ratios.

TwoHiggsCalc reads input and writes out results in a specific format close to the “SUSY Les Houches Accord 1.0” convention for SUSY parameters [37]. This format can later be read by MadEvent to perform numerical calculations for 2HDM processes. A `README` file describing this modified version of the LHA format used as input convention is available. To ease the use of TwoHiggsCalc, a web form has been designed to automatize the parameter card writing process. Numerical values for the parameters (units being fixed when needed) can be entered on this form. Some simple algebraic expressions can also be used. The `+`, `-`, `*`, `/` operators and the reserved keyword `PI`, *e.g.*, `PI/2+3*PI/2`, are correctly interpreted.

In the general 2HDM, one has the freedom to choose a specific basis for entering parameters. All the possible choices are physically equivalent (see *e.g.* [38] for a discussion). TwoHiggsCalc and the 2HDM model both assume that the parameters are given in a particular basis, called the “Higgs basis” where only one Higgs doublet gets a vacuum expectation value. An independent program, Gen2HB, has been written to convert parameters given in an arbitrary basis (where both Higgs doublets get vevs), called “generic”, to parameters

in the Higgs basis. See [34] for more information on basis invariance and on the notation used.

The scalar potential in the Higgs basis reads

$$\begin{aligned}
V = & \mu_1 H_1^\dagger H_1 + \mu_2 H_2^\dagger H_2 - \left( \mu_3 H_1^\dagger H_2 + \text{h.c.} \right) \\
& \lambda_1 \left( H_1^\dagger H_1 \right)^2 + \lambda_2 \left( H_2^\dagger H_2 \right)^2 \\
& + \lambda_3 \left( H_1^\dagger H_1 \right) \left( H_2^\dagger H_2 \right) + \lambda_4 \left( H_1^\dagger H_2 \right) \left( H_2^\dagger H_1 \right) \\
& + \left[ \left( \lambda_5 H_1^\dagger H_2 + \lambda_6 H_1^\dagger H_1 + \lambda_7 H_2^\dagger H_2 \right) \left( H_1^\dagger H_2 \right) + \text{h.c.} \right].
\end{aligned}$$

All parameters in front of quartic terms and the charged Higgs mass are input parameters, while  $\mu_1$ ,  $\mu_2$  and  $\mu_3$  are fixed by minimization constraints and by the vev extracted from the observed SM parameters.  $\lambda_1$  to  $\lambda_4$  are real while  $\lambda_5$  in general is complex. However, since only the phase differences between  $\lambda_5$ ,  $\lambda_6$ ,  $\lambda_7$  and  $\mu_3$  matter, the phase of  $\lambda_5$  can always be rotated out. It is thus considered as a real parameter while  $\lambda_6$  and  $\lambda_7$  are a priori complex.

In the same basis, the Yukawa interactions read

$$\begin{aligned}
\mathcal{L}_Y = & \frac{\overline{Q}_L \sqrt{2}}{v} \left[ (M_d H_1 + Y_d H_2) d_R + (M_u \tilde{H}_1 + Y_u \tilde{H}_2) u_R \right] \\
& + \frac{\overline{E}_L \sqrt{2}}{v} [(M_e H_1 + Y_e H_2) e_R].
\end{aligned}$$

Yukawa couplings are expected to be given in the physical basis for fermions, *i.e.*, in the basis where the mass matrix is diagonal. Since in the Higgs basis only the first Higgs doublet gets a non zero vev, the  $M$  matrices are completely fixed by the physical fermion masses and CKM mixing matrix (restricted to Cabibbo angle) while the  $Y$  matrices (giving the couplings of the second Higgs doublet) are a priori free. For these matrices, the first index refers to doublet generation while the second refers to the singlet generation. For example, Y2B stands for the complex Yukawa couplings of the second Higgs doublet to the second generation quark left doublet and to the bottom singlet.

In the generic basis, similar expressions are assumed. For the scalar potential all parameters in front of quartic terms are inputs as well as  $\tan(\beta)$ , the norm of  $\mu_3$  and the phase of  $v_2$ . The overall vev is again extracted from SM parameters while mass terms parameters, like  $\mu_1$ ,  $\mu_2$  and the phase of  $\mu_3$ , are fixed by the minimization constraints.  $\lambda_1$  to  $\lambda_4$  are real parameters,  $\lambda_5$ ,  $\lambda_6$  and  $\lambda_7$  are a priori complex. Like in the Higgs basis, the Yukawa couplings must be given in the physical basis for fermions. Since the mass matrices are fixed, only the Yukawa coupling matrices of the second Higgs doublet ( $\Gamma$ ), is required. The other one is going to be automatically evaluated to match observed fermion masses and CKM mixing matrix (restricted to Cabibbo angle). For the  $\Gamma$  matrix, the first index refers to doublet generation while the second one refer to the singlet generation. For example, G2B stands for the complex Yukawa couplings of the second Higgs doublet to the second generation quark left doublet and to the bottom singlet.

Given the above parameters and some SM parameters, TwoHiggsCalc computes the following quantities

Block	Comment
SMINPUTS	From 1 to 4, SM parameters, see the SM section for more details
MGSMPARAM	Extra block with $\sin\theta_W$ and $M_W$ , see the SM section for more details
MGYUKAWA	“Yukawa” masses used in the Yukawa couplings evaluation
MGCKM	The full CKM matrix
BASIS	Basis choice, must be 1 (Higgs basis) for <b>MadEvent</b> !
MINPAR	Scalar potential parameters in the Higgs basis
YUKAWA2	Yukawa couplings of the second Higgs doublet
MASS	All SM particles masses, plus the five new Higgs boson masses
TMIX	The scalar mixing matrix
DECAY	For all the Higgs bosons, top, $W^\pm$ and $Z$

**Table 2:** LHA blocks used in the 2HDM implementation

- Scalar particles mass spectrum
- Normalized mixing matrix of neutral scalars (called  $T$  in [34])
- Decay widths for all scalars as well as for  $W$  and  $Z$  bosons and the top quark. All widths are evaluated at tree-level using the same couplings as in MadEvent. Below threshold formulas are included for the scalar decays into two vector bosons and the one loop driven scalar decay into two gluons is also computed.

The LHA blocks and parameters used by MadEvent are given in Table 2. All blocks in the table are provided by TwoHiggsCalc. Note that if parton density functions (PDFs) are used in the MadEvent run, the value for  $\alpha_s$  at  $M_Z$  and the order of its running is given by the PDF. Otherwise  $\alpha_s(M_Z)$  is given by block **SMINPUTS**, parameter 3, and the order of running is taken to be 2-loop. The scale where  $\alpha_s$  is evaluated can be fixed or evaluated on an event-by-event basis like in the SM.

## 4.2 The MSSM implementation

One of the most popular extensions of the Standard Model is TeV scale supersymmetry. Supersymmetry solves the problem of quadratically divergent corrections to the Higgs boson mass by the introduction of new bosonic particles having the same couplings as the Standard Model fermions, and new fermions having the same couplings as the Standard Model bosons, thus cancelling the loop contributions to the Higgs mass to all orders. The Minimal Supersymmetric Standard Model, MSSM, represents the minimal particle content for a supersymmetric extension of the Standard Model together with the maximum coupling space allowed by so-called “soft supersymmetry breaking terms” in the effective low-energy Lagrangean. These are constructed not to introduce new divergencies in any couplings, and therefore maintain the cancellations of quadratically divergent corrections to the Higgs mass. For an introduction to supersymmetry and the MSSM, see *e.g.*, Refs. [39, 40].

The implementation of the MSSM particles and vertices into MadGraph II was made in Ref. [41, 42], following the conventions of Refs. [43] and [44]. Specifically, it is restricted

to the minimal supersymmetric model conserving  $R$ -parity, without  $CP$ -violating phases and with diagonal CKM and MNS matrices. Higgs Yukawa couplings as well as mixing between right- and left-handed sfermions are implemented only for the third generation. However, no specific supersymmetry breaking scheme is assumed, so the spectrum and couplings of the supersymmetric particles can be produced with any spectrum generator regardless of the assumptions going into its calculations. The spectrum and couplings of the particles are read through SUSY Les Houches Accord files [37].

In order to consistently calculate decay widths and the dependent parameters, a model calculator (see Section 2) for the MSSM is available. MSSMCalc takes a SUSY Les Houches Accord (SLHA) file [37] from any Spectrum generator as input, and produces a MadEvent readable file, `param_card.dat`, with the missing Standard Model parameters, as well as decay widths for all supersymmetric particles (calculated at leading order by Sdecay [45]), the Higgs particles and the top,  $W^\pm$  and  $Z$  particles. Care has been taken to ensure that the parameters used in the calculation of decay widths are as similar as possible to the parameters used in MadEvent, since the correct total decay widths are vital to get the correct tree-level cross-sections for processes involving decaying particles.

In the default run mode, MSSMCalc uses the Standard Model parameters given in the SUSY Les Houches accord ( $\alpha_{em}$ ,  $G_F$  and  $M_Z$ ) to calculate the parameters  $\sin\theta_W$  and  $M_W$ , which are stored in a MadEvent specific block `MGSMPARAM` in the resulting `param_card.dat`. The  $b$  quark pole mass is calculated from the  $\overline{MS}$  mass at 2-loop order. Another option is to extract the Standard Model parameters (and the vacuum expectation value ratio  $\tan\beta$ ) from the chargino and neutralino mixing matrices, in order to ensure unitarity of ino-ino scattering at high energy. In this mode, also the Yukawa masses of the third generation fermions are extracted from the third generation sfermion mixing matrices. For a thorough discussion of this option, see section II C of [41].

The strong coupling  $\alpha_s$  is calculated in MSSMCalc using 2-loop renormalisation group running in the  $\overline{MS}$  scheme, at the scale specified in the `GAUGE` block statement. The value used for the strong coupling  $g$  in the decay width calculations is stored for comparison in the block `GAUGE`, parameter 3. Note however, that the value of  $\alpha_s$  used in MadEvent is given by the choice of parton distribution function and the scale chosen in the run.

If there are blocks missing in the SLHA file which are necessary for MadEvent, MSSMCalc will produce a `param_card.dat` file containing error messages.

The SUSY Les Houches blocks and parameters used by MadEvent are given in Table 3. All blocks in the table should be provided by the user (and are indeed provided by most MSSM spectrum generators), except for the `MGSMPARAM` and the `DECAY` blocks which are produced by the parameter calculator MSSMCalc. Note that if parton density functions (PDFs) are used in the MadEvent run, the value for  $\alpha_s$  at  $M_Z$  and the order of its running is given by the PDF. Otherwise  $\alpha_s(M_Z)$  is given by block `SMINPUTS`, parameter 3, and the order of running is taken to be 2-loop. The scale where  $\alpha_s$  is evaluated is however always given by the “scale” parameter in the `run_card.dat`.

### 4.3 The User Model

Beside the models already implemented in the MadGraph structure, it is also important to

Block	Comment
SMINPUTS	Except for 5, the $b$ quark $\overline{MS}$ mass Extra block with $\sin\theta_W$ and $M_W$ , written by MSSMCalc Including 5, the $b$ quark pole mass
MGSMPARAM	
MASS	
NMIX, UMI, VMIX	
STOPMIX, SBOTMIX, STAUMIX	
ALPHA	Only parameters 1 ( $\mu$ ) and 2 ( $\tan\beta$ ) Only the third generation parameter 3 3 Only the third generation parameter 3 3 For all SUSY particles, Higgs bosons, top, $W^\pm$ and $Z$
HMIX	
AU, AD, AE	
YU, YD, YE	
DECAY	

**Table 3:** SHLA blocks used by SUSY MadEvent. See the text for details.

let the possibility to the non-expert to be able to test his or her own models. For example, adding a  $Z'$  boson or a  $t'$  fermion to Standard Model and all their possible couplings to other particles should be now straightforward. For this purpose a new framework has been developed to provide users with an easy and safe, yet flexible, tool for implementing their own models. It is important to mention that the user model is limited by the MadGraph and HELAS assumptions for the Lorentz and color structure of the vertices.

The aim here is not to give a detailed step-by-step description on how to implement a new model (which is available in the README file in the `usrmod` directory), but to present the philosophy adopted in the design of the framework. In practice, a model is defined by a list of particles with given quantum numbers and their interactions. The implementation of a model proceeds in three simple steps.

First, three files have to be provided: i) the list of particles; ii) the list of the interactions (the complete Standard Model, which is already present by default, plus the user's new inputs) iii) the list of possible new parameters present in non-SM couplings expressions. As a second step a Perl script, `ConversionScript.pl`, is invoked that takes the above information and creates all the files needed by MadGraph/MadEvent. As a final step, the user must modify by hand the `couplings.f` file to set the values of the couplings and provide the formulas to go from the parameters of the model to those appearing in the Feynman rules.

Let us consider the implementation of a  $Z'$  vector boson as an example. In the particle list file (`particle.dat`), an extra  $Z$  appears as

```
#Name anti_Name Spin Linetype Mass Width Color Label Model
#xxx  xxxx   SFV   WSDC   str str ST0   str PDG code
...
#MODEL EXTENSION
zp    zp     V     W     ZPMASS ZPWID S   ZP   32
```

where the needed information is provided in a simple syntax. In this example we have a  $Z'$  (zp) of vector type (V), with mass and width ZPMASS and ZPWID. We made it a color

singlet by setting color to S. The linetype W and the label ZP are used in the Feynman diagrams for cosmetics only.

As described above, in the second file (`interactions.dat`) the interactions are specified. In this file the names of the particles as defined above have to be used. In this example the vertex  $Z'd\bar{d}$ , *i.e.*, the vertex between the  $Z'$  and two down quarks, is shown.

```
# USRVertex
d d ZP GZPD QED
```

The name GZPD is chosen arbitrarily (but paying attention not to choose a name already in use); QED indicates the type of interaction.

If there are new parameters related to coupling expressions, the user can edit a third file `VariableName.dat`. For example, to include the couplings between  $Z'$  and down quarks, the expression of the coupling appearing in the Feynman rule can be written as

$$g_{Z'd\bar{d}} = -\sqrt{4\pi\alpha} \left[ C_1 \frac{1}{s_W c_W} \left( -\frac{1}{2} + \frac{s_W^2}{3} \right) \right] \frac{1 + \gamma_5}{2} + C_2 \frac{s_W}{3c_W} \frac{1 - \gamma_5}{2}, \quad (4.1)$$

where  $s_W$  and  $c_W$  are  $\sin \theta_W$  and  $\cos \theta_W$ , respectively.  $C_1$  and  $C_2$  are one for a SM like  $Z'$ . This can, of course, also be written as

$$g_{Z'd\bar{d}} = C'_1 \frac{1 + \gamma_5}{2} + C'_2 \frac{1 - \gamma_5}{2} \quad (4.2)$$

where no link with the parameters in the SM is explicit. The declaration of these parameters appears as a list in `VariableName.dat`

```
C1 #first variable name
C2 #other variable name
```

The second step is to run the `ConversionScript.pl`. The script uses the above three files as input and creates the parameter card, `param_card.dat` and the file `couplings.f`, as well as other files needed by MadEvent. The third and final step is to include in these files the numerical values related to the model. The user model parameter card is in LHA format and is similar as the SM one, but also includes the information about the new particles and coupling parameters. The masses and widths of the new particles, as well as the new coupling parameters, need to be set by the user to their correct numerical values. In the `couplings.f` file, expressions for the coupling strengths have to be provided. The Perl script already takes care of the formats for the couplings in such a way that they are in HELAS compliant. For our example, the expression in Eq. (4.1) can be implemented as

```
GZPD(1)=dcmplx(C1*( -ez*(-Half + sin2w/Three)),Zero)
GZPD(2)=dcmplx(C2*( -ey/Three),Zero)
```

The possibility of testing the couplings values and/or other parameters like the masses and the widths is also available. The program `testprog` can be compiled and run and it prints the values for all masses and couplings that will be used by MadGraph. A second program `couplingsvalues` writes out the names and the corresponding values of couplings in a format that can be read by external tools like the width and decay calculator BRIDGE [46].



## 5. Matching of jet production by parton showers and matrix elements

For many processes at the LHC, the major backgrounds include multijet production, either pure QCD production or jet production in association with weak vector bosons or top quarks. Also when generating signal processes, such as Higgs boson production, it is often important to understand the jet activity, *i.e.* the probability for extra jet production, to be able to discriminate from the backgrounds.

Parton showering, *e.g.* by Pythia, is well known to give a good description of parton emissions when the emitted partons are close in phase space. To describe hard and well separated partons, matrix element generators such as MadGraph/MadEvent should be used. However, even in that case, it is necessary to perform parton showering and hadronization in order to get a realistic description of the event, in particular when detector simulation is needed. In order to combine these two descriptions, it is essential to use some kind of matching between them to avoid double counting of emissions in overlapping phase space regions.

One such matching scheme is the CKKW algorithm [13, 14]. This scheme uses a  $k_\perp$  measure to separate emissions into two phase space regions, a low- $k_\perp$  region described by the parton showers and a high- $k_\perp$  region described by the matrix elements. In order to get a smooth transition between the two regions, the matrix element emissions are treated as similarly as possible to parton shower emissions. This is done by clustering the event using the  $k_\perp$  jet clustering algorithm [47, 48] to find its corresponding “parton shower history”, *i.e.*, the sequence of parton emissions which would have been necessary for the parton shower to generate the event. The event is then reweighted by using the values of  $k_\perp$  at the clustering nodes as scales in the running  $\alpha_s$  couplings (as is done in the parton shower), and a Sudakov suppression factor is applied to get the probability for this particular event to be generated without further emissions above the  $k_\perp$  cutoff scale. After that the event is showered, allowing only shower emissions below the cutoff scale.

An alternative way to accomplish an equivalent result has been proposed by Mangano [49]. In his approach, the multiparton events are reweighted by  $\alpha_s$  factors in the same way as in the CKKW prescription, but no Sudakov reweighting is performed. Instead, the event is showered, and then discarded if the showering generates emissions harder than the phase space cutoff.

In the current MadGraph/MadEvent release, both of these matching schemes are implemented. For the MLM matching option, the showering and vetoing of events with too large emissions are implemented in the Pythia interface included in the Pythia-PGS package (see Section 6.3), while the CKKW matching, including both  $\alpha_s$  and Sudakov reweighting, is so far implemented only at the matrix element creation level.

In the original version of the MLM matching procedure, which is implemented in Alpgen [7], the phase space separation is defined in terms of cone jets. The MadGraph/MadEvent implementation allows the user to choose whether to use cone jets or  $k_T$  clustered jets, enabling a more immediate comparison to the original CKKW formulation as implemented in Sherpa [5].

A comparison between event generators implementing matching of matrix elements

and parton showers is underway [19], including MadGraph/MadEvent, AlpGen, Sherpa, HELAC [50] and Ariadne [51].  $W^\pm$ +jets production at Tevatron and the LHC is there used as a case study.

It should also be noted that a similar matching method for matching of MadEvent matrix elements to Pythia showers has been implemented independently by Mrenna [15] and has been successfully used to describe Tevatron  $W$ +jets backgrounds.

## 6. Tools

The parton-level events generated with MadEvent are stored in the so-called “Les Houches accord event file format” [52]. Within this format enough information for each event is available so that some of the data analysis usually performed during the event generation phase, such as plotting, estimating PDF’s errors or scale variations can be deferred to a later stage. To this aim, simple routines have been developed to perform some tasks “off-line”, *i.e.*, directly on the event files produced. The main reason is to improve versatility and save time. Generating events is a CPU expensive activity, which, in some cases, can take many hours. Therefore, it is not desirable to have to rerun codes only for making new plots or switching from one scale choice to another. Another important advantage in working directly with the events is that the tools developed are “independent” of how the events were generated and can be used with any event set in the Les Houches format. In this respect, the applications presented below can be used with events produced by any matrix element generator. The expert user is invited to develop his/her own tools and make them available to the physics community.

### 6.1 Analysis platforms: MadAnalysis and ExRootAnalysis

Two platforms for performing studies on the event samples at all stages of the simulations (parton, hadron and detector level) are available. Here we describe them briefly.

MadAnalysis is a fortran-based analysis tool which allows you to select events (by setting further acceptance or selection cuts) and produce plots for events read in LHEF (parton level) and LHC Olympics (detector level) formats. Plots are simple ASCII files, output either in a minimal format (just title and comments and  $xy$  list) which can be read by GnuPlot [53] with a simple script (provided), or in TopDrawer [54] format. A perl script to overlay plots of the corresponding quantities for two event sets (or the same set at different stages of the simulation chain) is also provided.

Particles, objects and reconstructed objects (like jets) in the final state can be organized in classes, on which cuts and plots are based. Particles or objects in each class are ordered with respect to a one-particle quantity variable (the default is the transverse momentum). For instance, it is natural (at the parton level) to define any light quark and gluon as a jet. The user selects the cuts and the details of the plots by editing a `ma_card.dat`. Typical quantities that can be plotted are one-particle quantities, such as transverse momentum ( $p^T$ ) and the rapidity ( $y$ ) or pseudo-rapidity ( $\eta$ ), or multi-particle quantities, such as two-body or three-body invariant masses ( $m(i, j)$ ,  $m(i, j, k)$ ), or “distances” (such as  $R(i, j)$  or  $k_T(i, j)$ ) for each pair of final-state particles. A rather exhaustive library of kinematic

functions is available to the user, as well as the possibility of defining new functions of the momenta of any number of particles in a very user-friendly way.

The `ExRootAnalysis` package allows to store and analyze events in a ROOT tree format [55]. Normally, ROOT allows to store data in several different formats. However, there are some crucial differences, which makes the ROOT tree format more attractive than others:

- information is stored in arrays of objects, which enables efficient storage (compression using the ROOT gzip algorithm) and retrieval;
- the same C++ classes are used for creating the ROOT tree and for analysing the stored data.

ROOT tree objects are created from particles generated by MadGraph and Pythia or objects produced by PGS (in most of the cases physics objects like jets, electrons, *etc.*), in order to perform analysis in a ROOT environment.

The `ExRootAnalysis` package can be subdivided into several subsystems: basic framework of few classes providing event loop, event selection and basic operations with a ROOT tree file; modules selecting events and objects to be analysed at per event and per object level; and modules analyzing selected events.

For example, a selector module can select and group partons generated by MadGraph into several classes (such as leptons, jets, top quarks, *etc.*) according to their status and particle identification number. Any number of classes can be specified in the configuration file. After the partons have been classified, an analysis module can be used to produce series of standard plots for each class of partons.

Documentation on the content of the ROOT tree is available on the web [56].

## 6.2 Decaying unstable particles in the SM and beyond

MadEvent includes a tool, Decay, that performs the decays of unstable particles in the Standard Model directly on the parton-level events (weighted or unweighted). At present, a total of 68 decay modes are included for  $\tau$ ,  $W$ ,  $Z$ ,  $t$  and  $h$  decays. The advantage of using Decay is obvious: from the point of view of MG/ME the generation of events is faster for a simple final state. When a detailed knowledge of the spin-correlations is not needed, Decay, which is only keeping the diagonal terms in the spin-correlation matrix in the helicity basis, is an accurate and very efficient tool to get a multi-particle final state. A generalization and considerable improvement of Decay has been recently released by Meade and Reece [46], under the name of BRIDGE. BRIDGE can take an arbitrary model as input and compute the widths (and hence branching ratios) for all the kinematically allowed two-body and three-body tree level decays. It can also decay generated events, choosing decay modes randomly according to the branching ratios.

## 6.3 Hadronization and detector simulation using Pythia and PGS

One objective in the latest development of MadGraph/MadEvent has been to facilitate hadronization and detector simulation. We have therefore made a package available which

includes the hadronization package Pythia 6.4 [11] and the fast detector simulation package PGS 4 [57], as well as the parton density package LHAPDF [58], a stripped version of CERNLIB, and the utility STDHEP [59] which is used for communication between Pythia and PGS. The package also includes main programs for running Pythia and PGS, and an interface for reading MadEvent files into Pythia. This package is available for download, or can be used directly in the on-line event generation to simulate fully hadronized and detector reconstructed events.

The MadEvent-Pythia interface reads the Les Houches Event file output of MadEvent, and communicates model parameters such as particle masses to Pythia. The behaviour of Pythia is determined by the input file `pythia_card.dat`. In the case where it finds particles in the events which are considered massless in MadEvent but massive in Pythia, such as electrons or muons, the particles are given their Pythia mass, and momenta are redistributed to account for the changes. The interface also allows for matching of jet production by Pythia parton showers with multiparton samples generated by MadEvent, see Section 5.

The outputs of the Pythia main program are a binary file with the full hadronic event information in the STDHEP output format and a text file in the Les Houches Event file format with information on resonances, jets, leptons, and missing energy. Jet clustering for the latter file is performed either using a  $k_T$ -clustering routine or a cone jet routine.

The PGS main program reads the STDHEP file from the Pythia generation, and performs detector simulation using the detector setup given in the input file `pgs_card.dat`. The output is a text file in the PGS4 LHC Olympics format with information on the trigger status of the event, jets, leptons and missing energy.

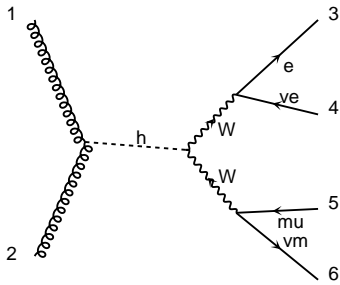
All the output files can be read by the ExRootAnalysis package to get ROOT files and plots for the different stages. They can also be read by the MadAnalysis package for simple analysis purposes (see Section 6.1). The events are consistently numbered in the output files throughout the chain.

## 7. Applications to hadron collider physics

In this section we present several examples of the types of studies that can be performed with the MadGraph/MadEvent package. The aim is to show how easily the various new features can be used to perform signal and background analyses with both theoretical and experimental aims. Even though the discussion is kept concise for space reasons, we stress that all the results obtained in the following could stand in a dedicated publication and that some of them are original and presented here for the first time.

### 7.1 Higgs search in $pp \rightarrow h \rightarrow W^+W^-$ : signal and backgrounds

For a Higgs boson of moderate mass,  $140 \text{ GeV} < m_h < 170 \text{ GeV}$ , the golden discovery channel is via gluon-gluon production and successive decay into a pair of leptonically decaying  $W$  bosons, Fig. 2.



**Figure 2:** The diagram corresponding to the process  $pp \rightarrow h \rightarrow W^+W^- \rightarrow e^-\mu^+\bar{\nu}_e\nu_\mu$

This process has a large cross section and also special kinematics characteristics, notably the angular distance between the leptons coming from the  $W$ 's decays, which can be exploited to improve the signal over background ratio. Reducible backgrounds are dominated by processes which involve single top and  $t\bar{t}$  production, which typically yield a larger number of jets in the final state and can be therefore suppressed by imposing a jet veto. Many studies have been performed on this channel and its backgrounds both at the experimental and the Monte Carlo level, sometimes with different generators, selection cuts and

approximations [60].

In the following we discuss how such a study can be done using MG/ME. Our aim is to show how easy it is to generate all the necessary event samples in one go, from parton-level to detector objects, while keeping all the relevant non-trivial features of the signal (such as spin correlations) and the background (resonant and non-resonant contributions). In particular, the example makes use of the following features:

- Inclusion of more than one process in the MadGraph generation (see Fig. 3). The different processes included in this way are then produced in proportion to their cross section, so the event file produced automatically gets the correct mix of unweighted events.
- The possibility to generate the backgrounds corresponding to a given final signature (in this case  $e^\pm\mu^\mp b\bar{b}$ ), by including all the classes of diagrams (including resonant and non-resonant diagrams) consistently at once. The information on the relative contributions from the various resonances is available on event-by-event basis.
- The interfaces to Pythia for decay and hadronization, PGS for detector simulation and MadAnalysis for plotting, analyzing and comparing observables at the various stages of the simulation.

We consider that both  $W$ 's decay leptonically, and for simplicity we restrict to the different flavor case,  $e^\pm\mu^\mp$ . As our purpose is illustrative we use (a simplified version of) the selection cuts used in Ref. [61]. we choose the angular distance between the two leptons in the transverse plane,  $\Delta\phi$ , as the discriminating variable. As it is well-known, this variable encodes the fact that in the signal the two leptons tend to be produced close in phase space, due to the constraints coming from angular momentum conservation and the purely left-handed couplings of the  $W$  to the leptons.

The selection cuts used in the  $\Delta\phi$  distributions at the parton and PGS level shown in Fig. 4 are:

- leptons:  $25 \text{ GeV} < p_T^\ell < 50 \text{ GeV}$ ,  $|\eta^\ell| < 2.4$ ,  $\Delta R(\ell, \ell') > 0.4$ , and  $m(\ell, \ell') < 80 \text{ GeV}$ ;

```

# Begin PROCESS # This is TAG. Do not modify this line

pp>h> e- mu+ ve~ vm @1 # Signal Process
QCD=0                # max qcd order
QED=4                # max qed order
HIG=1                # max hig coupling order
end_coup              # the coupling list is over

pp > e- mu+ ve~ vm/h @2 # Irriducible background pp > W+W-> e- mu+ ve~ vm
QCD=0                # max qcd order
QED=4                # max qed order
end_coup              # the coupling list is over

pp > bb~e-mu+ve~vm/h @3 # Reducible background non-, single-, double-resonant
QCD=2                # max qcd order
QED=4                # max qed order
end_coup              # the coupling list is over

done                  # the process list is over

# End PROCESS # This is TAG. Do not modify this line
# Begin MODEL # This is TAG. Do not modify this line
heft
# End MODEL # This is TAG. Do not modify this line
# Begin MULTIPARTICLES # This is TAG. Do not modify this line
p uu~dd~ss~cc~g
# End MULTIPARTICLES # This is TAG. Do not modify this line

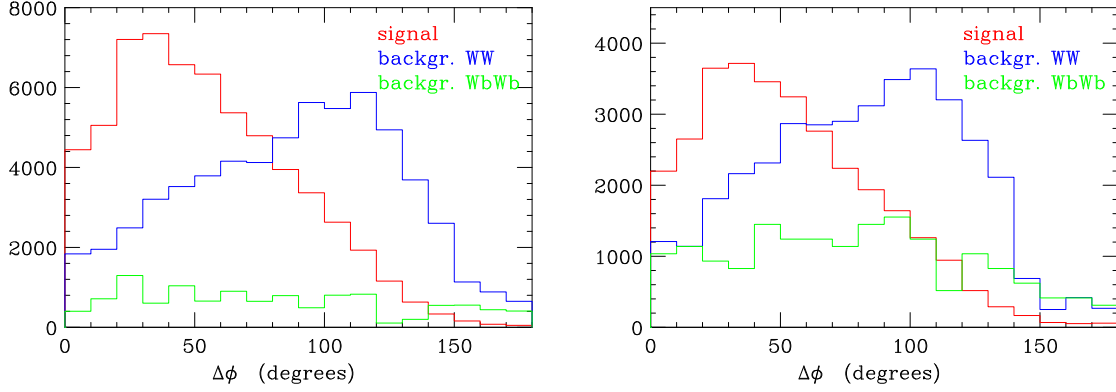
```

**Figure 3:** The `proc_card.dat` used for signal  $pp \rightarrow h \rightarrow W^+W^- \rightarrow e^-\mu^+\bar{\nu}_e\nu_\mu$  and the irreducible and reducible backgrounds. By a careful choice of the couplings and by vetoing the Higgs as an intermediate particle, the backgrounds are correctly included. The third process include non-resonant, single-top resonant and  $t\bar{t}$  double-resonant production.

- missing  $E_T$ :  $E_T^{\text{miss}} > 30$  GeV;
- jet veto: no jets with  $p_T^j > 30$  GeV and  $|\eta^j| < 3$ .

We note that there is not much difference in the  $\Delta\phi$  distributions in going from the parton-level events (where the Higgs has no transverse momentum and the two  $W$ 's might have a very small one) to the fully simulated case. However, by plotting the  $p_T$  of the leptons we checked that the initial state radiation has a non-negligible impact on the distributions, moving the average  $p_T$  of the leptons considerably.

Another very interesting feature to study is how the relative contributions coming from  $WbWb$  final states with top-quark non-resonant, single-resonant and double-resonant diagrams change as the jet veto is applied. This is a very good examples of a MC generation that has to be done carefully [62]. In fact single-resonant contributions are completely negligible and  $WbWb$  production is completely dominated by  $t\bar{t}$  in generic areas of the phase space. On the other hand, the request of a jet veto dramatically enhances the non-



**Figure 4:**  $\Delta\phi$  distributions of the two leptons in  $W^+W^-$  events at the parton (left) and hadron (right) level.

resonant and single-resonant contributions which account for more than 50% of the final  $WbWb$  background.

## 7.2 Higgs CP properties: $pp \rightarrow hjj$ in the HEFT

In this example we show how the Higgs effective theory implementation can be used to study the CP properties of a Higgs boson by looking at the angular distributions of the jets in the process  $pp \rightarrow hjj$  [63, 64]. Our study is done at the parton level, but could be promoted to the hadron level without effort.

We assume that the Higgs boson couples mainly to heavy quarks. In this case the Higgs boson will then be mainly produced by gluon fusion through a top quark loop. As discussed in Sec. 3.2, for Higgs boson masses smaller than two times the top quark mass,  $m_h \lesssim 2m_t$ , we can send the mass of the top quark in the loop to infinity  $m_t \rightarrow \infty$ , to a very good approximation. Effectively, this means that we contract the top quark loops in the Feynman diagrams and get effective  $ggh$ ,  $gggh$  and  $ggggh$  vertices. In Fig. 5 three diagrams with these effective vertices contributing to  $hjj$  production are presented. These vertices are implemented into the HEFT model of MadGraph, both for scalar and pseudo-scalar Higgs bosons.

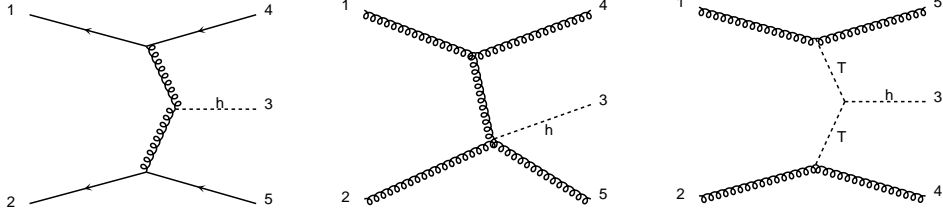
To investigate the CP properties of the Higgs boson we look at the angle  $\Delta\phi_{jj}$  [65], *i.e.*, the angle between the transverse momenta of the two jets. We ask for one very forward and one very backward jet by applying the following cuts on the jets:

$$p_T(j) > 20 \text{ GeV}, \quad \Delta R_{jj} > 0.4, \quad |\eta_{j1} - \eta_{j2}| > 4, \quad \eta_{j1} \cdot \eta_{j2} < 0. \quad (7.1)$$

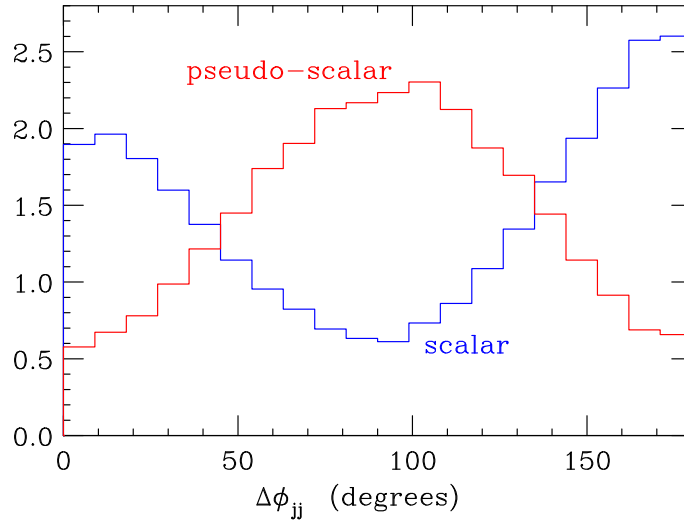
These cuts lead to a signal comparable with Higgs boson production through W boson fusion. In Fig. 6,  $\Delta\phi_{jj}$  is plotted with these cuts for a Higgs boson mass of  $M_h = 120 \text{ GeV}$ , for the pure scalar and pure pseudo-scalar cases.

## 7.3 Spin of a new resonance from lepton angular distributions

Here we present the angular distributions for different intermediate particles in the process



**Figure 5:** Selection of the diagrams contributing to  $hjj$  production. The dashed line corresponds to the auxiliary tensor particle  $T$ . For a pseudo-scalar Higgs the diagram on the *rhs* does not contribute.



**Figure 6:**  $\Delta\phi_{jj}$  distribution for  $pp \rightarrow hjj$  production for a scalar and a pseudo-scalar Higgs boson through gluon fusion. The plots are normalized. Cuts are as in (7.1),  $M_h = 120$  GeV.

$p\bar{p} \rightarrow X \rightarrow \mu^+\mu^-$ , where  $X$  is a  $s$ -channel resonance. We consider the three cases where  $X$  is a spin-0, spin-1 or a spin-2 particle.

This study exploits the possibilities of the user model framework to introduce new particles and interactions and also makes use of the full simulation chain, from parton level events to detector reconstruction.

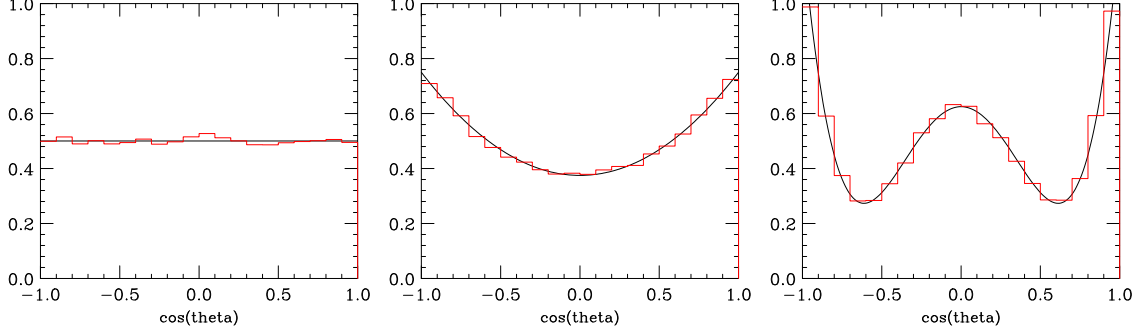
To minimize the effects of a non-zero transverse momentum of the intermediate state  $X$ , we use the angle  $\theta$  introduced by Collins and Soper to study spin correlations [66]. The angle  $\theta$  is defined as follows. Let  $p_A$  and  $p_B$  the momenta of the incoming hadrons in the rest frame of the muon pair. If the transverse momentum of the muon pair is non-zero, then  $p_A$  and  $p_B$  are not collinear. Let us define an axis in such a way that it bisects the angle between  $p_A$  and  $p_B$ . The angle  $\theta$  is defined to be the angle between this axis and the  $\mu^+$  momentum in the muon pair rest frame.

For (leading-order) parton level results, where we do not include extra radiation, the transverse momentum of the muon pair is zero. Then the angle  $\theta$  is the same as the more



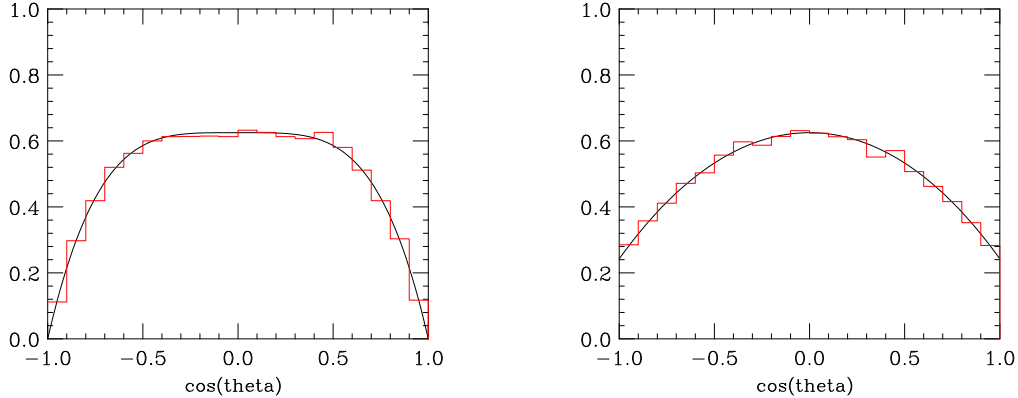
commonly used angle  $\theta^*$ , *i.e.*, the angle between the  $\mu^+$  momentum in the muon pair rest frame and the beam direction in the lab frame.

In Fig. 7 the cosine of this angle is plotted for the three different cases of the spin of the particle  $X$ . This figure includes only  $q\bar{q}$  initial states.



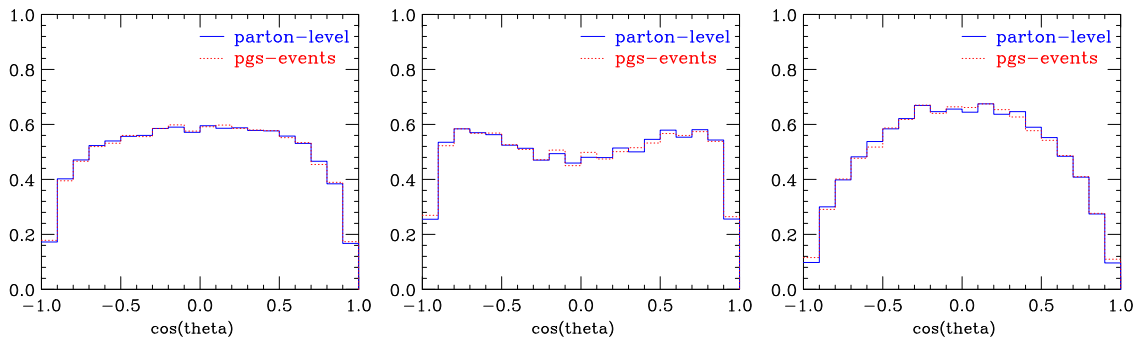
**Figure 7:** The normalized cross section as a function of  $\cos\theta$  in  $q\bar{q} \rightarrow X \rightarrow \mu^+\mu^-$ . *Left* for a spin-0 particle, *center* for a spin-1 particle, *right* for a spin-2 particle. No cuts applied.

The spin-2 particle can also be created by gluon fusion, which dramatically impacts the distribution. The angular distributions of the muons depend, in general, on the way the  $X$  resonance was produced. In Fig. 8 the muons angular distribution is plotted for the gluon initial state and for the sum of the gluon and quark initial states, for a spin-2 resonance with a mass of 1 TeV.

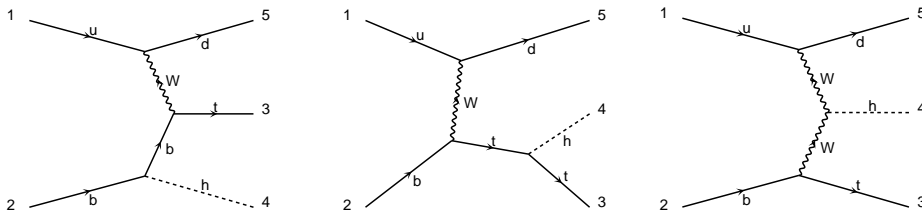


**Figure 8:** The normalized cross section as a function of  $\cos\theta$  in  $gg \rightarrow \text{spin-2} \rightarrow \mu^+\mu^-$  (*left*) and  $pp \rightarrow \text{spin-2} \rightarrow \mu^+\mu^-$  (*right*). The mass of spin-2 particle is 1 TeV. No cuts applied.

Including initial state radiation, showering and hadronization, does not modify the lepton distributions significantly. If we use PGS to simulate detector response we can get relatively reliable results for the CMS experiment at the LHC. In Fig. 9, reconstructed events generated with PGS are plotted. In the same figure the parton level results are plotted with an acceptance cut for the rapidity of the muons  $|\eta_{\mu^+, \mu^-}| < 2.4$ . Due to this cut the events with  $\theta \approx 0$  (*i.e.*, both muons in the same forward direction) and  $\theta \approx \pi$  (*i.e.*, muons in opposite forward-backward directions) are not detected. We see that for a 1 TeV resonance the lepton distributions hardly change at all.



**Figure 9:** The normalized cross section as a function of  $\cos\theta$  in  $pp \rightarrow X \rightarrow \mu^+ \mu^-$  for the reconstructed events after simulation of extra radiation using Pythia and detector response using PGS (dotted, red) and parton level events with rapidity cut  $|\eta_{\mu^+, \mu^-}| < 2.4$  (solid, blue). Left for a spin-0 particle, center for a spin-1 particle, right for a spin-2 particle. Mass of the  $X$  particle is 1 TeV.



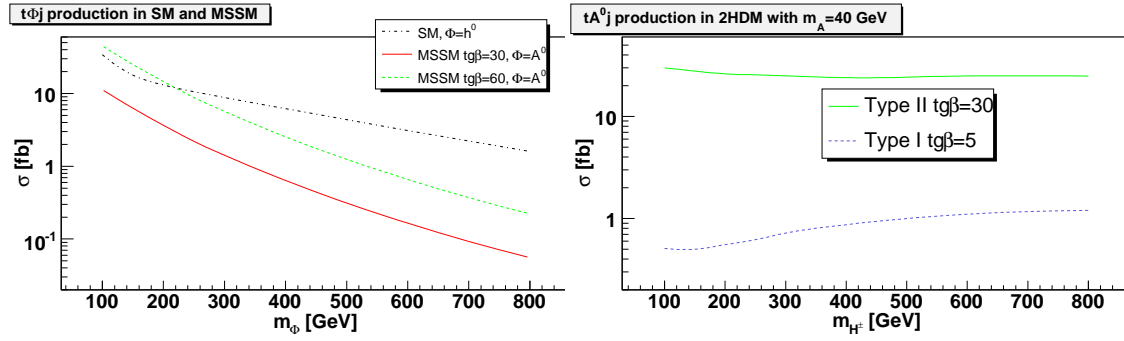
**Figure 10:** The diagrams contributing to single-top and Higgs associated production. The first diagram is negligible in the SM and in type I 2HDMs while the second one is negligible in the MSSM and in type II 2HDMs.

## 7.4 Single-top associated Higgs production in a generic 2HDM

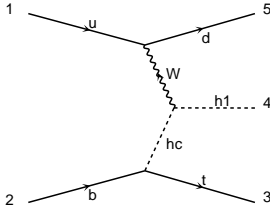
In this example we make use of the 2HDM implementation in MadGraph/MadEvent and of the associated calculator.

The  $t\bar{t}$  associated Higgs boson production is well known to have a non-negligible cross section at LHC ( $\simeq 1$  pb) for moderate values of the Higgs mass ( $m_h \simeq 100$  GeV). This can be estimated starting from the large  $t\bar{t}$  pair production cross section and then accounting for the large top-Higgs Yukawa coupling and phase space suppression. Since the single top production, in particular in the  $t$ -channel, also has a sizeable rate at LHC (more or less a third of the  $t\bar{t}$  production) it could naively be expected that the single-top associated production of the Higgs boson would be promising. It has been shown, however, that the cross section for single top in association with the Higgs is sizeably smaller than expected [67, 68].

The single-top associated production of the SM Higgs boson appears to be of the order of 100 fb for  $m_h \simeq 100$  GeV instead of the naive estimation of  $\simeq 300$  fb obtained by scaling the single-top cross section by  $\sigma(t\bar{t}h)/\sigma(t\bar{t})$ . The main reason for this is a particularly strong destructive interference between the two dominant amplitudes, associated with the diagrams shown in Fig. 10. It can be shown that each diagram contains a term proportional to  $m_t$  which violates unitarity at high energies and cancels in their sum.



**Figure 12:** Cross sections of single top associated production of different Higgs bosons. On the left, the SM Higgs boson production cross section is shown together with the MSSM pseudoscalar  $A^0$  production cross section as a function of their masses for two different  $\tan\beta$  values. On the right, the 2HDM pseudoscalar  $A^0$  production cross section is shown as a function of  $m_{H^\pm}$  for type I and type II Yukawa couplings. A minimal  $p_T$  of 20 GeV and a maximal rapidity of 2.5 is assumed for the jets. The factorisation and renormalisation scales are both set equal to  $m_\phi$ . The  $b$  quark mass involved in the Yukawa coupling as well as its pole mass (in order to deal consistently with unitarity cancellations) are equal to the running mass at  $m_\phi$ . The PDF used is CTEQ6L1.



**Figure 11:** Extra diagram contributing to single-top associated Higgs production in the MSSM and 2HDM.

So, even though the sign of this interference can be guessed from unitarity requirements, its absolute magnitude at moderate energies, compared to the masses involved (*e.g.*, the top mass), is surprising. The integrated squared amplitude for each of these diagram separately is indeed up to three times larger than the total cross section. The total SM cross section at LHC as a function of  $m_h$  is shown in Fig. 12.

Such a strong cancellation also takes place in the MSSM [68], but in this case more aspects have to be taken into account. To illustrate this, let us consider the decoupling regime where the mixing between  $h^0$  and  $H^0$  is small. For the light Higgs boson, the situation is similar to the SM. The  $WW h^0$  coupling is close to the SM value and in the limit of large  $\tan\beta$  where  $h^0$  couples mainly to  $b$ , the  $b\bar{b}h^0$  coupling is close to the  $t\bar{t}h$  of the SM. The amplitude of the diagram where the Higgs boson is emitted from the initial state  $b$  quark appears to be slightly suppressed compare to the one associated to the diagram where it comes from the final state  $t$  quark (up to a factor 2 for identical Yukawa couplings), but the overall impact on the physical cross section stays small, or is even positive, since the negative interference previously described in the SM is also decreased by the same factor. For the heavier scalar  $H^0$  the coupling  $WW H^0$  almost vanishes and one could then expect an enhancement of the total cross section. However there is an additional diagram involving a charged Higgs boson (see Fig. 11) which has to be taken into account due to the large  $W^\pm H^\mp H^0$  coupling. This diagram leads to a amplitude of roughly the same order and the same sign as the one involving only the  $W$  boson and

thus no particular enhancement is observed. The situation is similar for the pseudoscalar  $A^0$  (see Fig. 12) even though in this case the  $WWA^0$  coupling is strictly zero due to  $CP$  invariance.

Another possibility to get an increase of the cross section for the latter process would be to consider a 2HDM where the pseudoscalar  $A^0$  is relatively light and where the charged Higgs pair is much heavier (*e.g.*, see [36]) so that the negative interference cannot occur (*i.e.*, the amplitudes associated with the SM like diagrams in Fig. 10 are dominant). This, of course, cannot occur in the MSSM where the masses of  $H^\pm$  and  $A^0$  are linked at tree level through the mass relation  $m_{H^\pm}^2 = m_{A^0}^2 + m_{W^\pm}^2$ . The resulting cross section for  $m_{A^0} = 40$  GeV is plotted in Fig. 12 both in case of type I and of type II 2HDM as a function of the charged Higgs mass. An enhancement at high  $m_{H^\pm}$  is observed in the 2HDM type I case but the overall cross section stays much smaller than the SM one due to the reduced top quark Yukawa coupling. In type II models, the effect of varying  $m_{H^\pm}$  is quite small. At low  $m_{H^\pm}$ , the diagram involving the charged Higgs boson should contribute but its squared amplitude is more or less of the same order of magnitude as the negative interference it creates, so that its total contribution is negligible. Like in the MSSM, for all values of  $m_{H^\pm}$ , the overall cross section appears to be slightly suppressed, more or less by a factor 2, compared to the SM when the pseudoscalar is emitted from the initial state  $b$  quark.

To conclude, the cancellations in single top associated Higgs production expected from the unitarity of the models considered together with the suppressions due to different emission configurations are very effective in reducing the cross section to a value close the SM one, which is probably too small to be successfully measured at LHC.

## 7.5 Comparison of strong SUSY pair production at the SPS points

As a simple example showing some of the power of the SUSY implementation in MadGraph/MadEvent 4, we have chosen to compare results from inclusive strong SUSY pair production for the ten Snowmass (or SPS) benchmark parameter points [69]. In particular, the example makes use of the following new features:

- Implementation of the SUSY particle and interaction content [41] and the possibility to read SLHA files with mass, coupling and mixing information.
- Inclusion of more than one process in the MadGraph generation. The different processes included in this way are then produced in proportion to their cross section, so the event file produced automatically gets the correct mix of unweighted events.
- The possibility to define multi-particle labels also for scalar particles.
- The interfaces to Pythia for decay and hadronization, PGS for detector simulation and ROOT for event analysis.

### 7.5.1 Setup and generation

To easily be able to see the relative importance of different groups of subprocesses, we generate the process using the `proc_card.dat` shown in Fig. 13. In this card we differentiate

between gluino pair production, squark of the first and second families pair production, third family squark pair production, associate production of gluinos and squarks, and associate production of third family squarks with first and second family squarks. In the generation all these subprocesses (in total 497 different subprocesses, since we make a distinction among the different squark flavors) are automatically generated in the correct proportions according to their relative cross sections.

We next generate events at the LHC for each of the ten SPS points, and run the events through Pythia 6.409 [11], which decays the SUSY particles and performs parton showering and hadronization, and finally through PGS 4 [57], a general-purpose detector simulation. We used the default LHC setup for PGS 4. The parameter cards for the points were generated using SoftSusy 2.0.5 [70] and Sdecay 1.1a [45] and are available from the MG/ME 4 web sites. The results of the generation are:

1. The total cross section for strong SUSY pair production.
2. Cross sections for all the individual subprocesses, and summed cross sections for the different groups of subprocesses.
3. Files with unweighted events at parton level, decayed and hadronized level and detector reconstruction level.
4. ROOT files for event analysis at all three levels.

### 7.5.2 Comparison of the results for the SPS points

The cross sections and relative importance of the contributing subprocesses for the different SPS points are presented in Table 4. The relative cross sections for the different subprocesses are calculated by MadEvent at leading order, but we should be able to take the results as a good indication of which processes are most important to take into account at each point.

A striking feature of Table 4 is that between 68% (SPS1a) and 92% (SPS2) of the total cross section is due to only the production of valence quark partners, gluinos, and in some cases top partners, constituting  $\sim 25$  processes out of almost 500. This is of course due to the dominance of the valence  $u$  and  $d$  quarks over other quark species in the parton distributions of the proton, which is larger the heavier the squarks are.

We also look at kinematical event distributions for the different points, after decay, hadronization and detector simulation. One of the most suggestive distributions is the total transverse event energy,  $H_T$ , defined as  $H_T = \sum p_{\perp}^{\text{jets}} + \sum p_{\perp}^{\text{leptons}} + \cancel{E}_{\perp}$ , since it gives an indication of the mass scale of the produced particles. In Fig. 14 the  $H_T$  distributions are shown for a reduced set of SPS points. We can clearly see how the position of the peak (or peaks) is correlated with the masses of the particles produced, see Table 5. For the ten SPS points, the peak positions are at about 65-85% of the sum of the masses of the produced particles, with a slightly increasing ratio as the masses increase. The exception is the very low mass stop pair production peak at SPS5, where the peak position is below 40% of the sum of the produced stop masses. This is due to the large proportion of the

```

# Begin PROCESS # This is TAG. Do not modify this line
pp > gogo    @1 # Process
QCD=2        # max qcd order
QED=0        # max qed order
end_coup     # the coupling list is over

pp > S1S1    @2 # Process
QCD=2        # max qcd order
QED=0        # max qed order
end_coup     # the coupling list is over

pp > S2S2    @3 # Process
QCD=2        # max qcd order
QED=0        # max qed order
end_coup     # the coupling list is over

pp > S1go    @4 # Process
QCD=2        # max qcd order
QED=0        # max qed order
end_coup     # the coupling list is over

pp > S2go    @5 # Process
QCD=2        # max qcd order
QED=0        # max qed order
end_coup     # the coupling list is over

pp > S1S2    @6 # Process
QCD=2        # max qcd order
QED=0        # max qed order
end_coup     # the coupling list is over

done         # the process list is over
# End PROCESS # This is TAG. Do not modify this line
# Begin MODEL # This is TAG. Do not modify this line
mssm
# End MODEL # This is TAG. Do not modify this line
# Begin MULTIPARTICLES # This is TAG. Do not modify this line
p uu~dd~ss~cc~bb~g
S1 ulul~urur~dldl~drdr~slsl~srsr~clcl~crcr~
S2 b1b1~b2b2~t1t1~t2t2~
# End MULTIPARTICLES # This is TAG. Do not modify this line

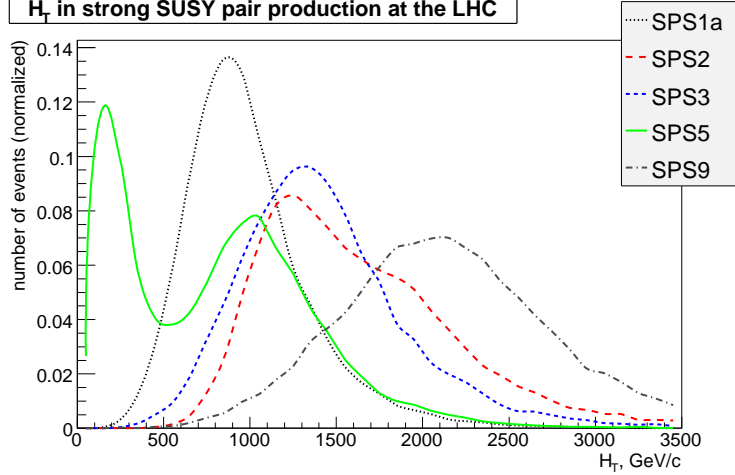
```

**Figure 13:** The `proc_card.dat` used for the generation of all strong  $2 \rightarrow 2$  SUSY processes. The processes are grouped such that gluino pair production is in the first group, first and second generation squark pair production is in the second group, third generation in the third and associated production in the last three groups. Please note that further subdivisions have been made in Table 4.

event energy carried away by the two LSPs, which largely balance in the detector so that the total missing energy detected is relatively small. If we add  $2m_{\chi_1^0}$  to the peak positions, they reach between 85% and 100% of the mass of the produced particles, again with a

SPS pt.	$\sigma$ (pb)	$\sigma_{\text{NLO}}$ (pb)	$\tilde{g}\tilde{q}$	$\tilde{q}\tilde{q}$	$\tilde{g}\tilde{g}$	$\tilde{t}\tilde{t}^*$	$\tilde{b}\tilde{b}^*$	$\tilde{g}\text{ sq}^{(*)}$	$\text{sq}^{(*)}\text{sq}^{(*)}$
1a	42.6	52.7	45	11	12	3.4	1.4	9.7	17
1b	3.39	4.51	44	21	6.6	2.4	1.3	6.4	18.5
2	1.48	2.70	29	3.1	56+7.0	0.55	0.073	3.1	0.8
3	3.81	4.87	44	21	6.9	2.5	1.1	6.1	18
4	10.0	13.9	47	14	13	2.6	1.5	8.3	17
5	35.5	39.2	29	9.2	6.1	33+4.7	1.1	5.4	11
6	14.5	20.5	45	15	9.7	2.3	1.2	8.4	17
7	3.08	4.00	45	22	7.5	1.0	0.7	6.2	17
8	2.34	3.52	50	13	23	0.4	0.3	6.4	6.7
9	0.425	0.506	42	31	5.4	2.5	1.0	3.9	14

**Table 4:** Cross sections for the ten SPS points and the contributions, in %, from different subprocesses for each point. In the process definitions,  $\tilde{q} = \tilde{u}_{L,R}, \tilde{d}_{L,R}$ . For  $\tilde{g}\tilde{g}$  and  $\tilde{t}\tilde{t}^*$  production, the dominant channel is gluon fusion, with  $q\bar{q}$  annihilation contributing to less than 10% of the numbers. The exceptions are point 2 and 5, respectively, where we have separated the  $gg$  and  $q\bar{q}$  channels by the “+” sign. The last two columns show the contributions from associated gluino-(anti)squark production where the processes in column 4 are excluded, and (anti)squark-(anti)squark production where the processes in column 5, 7 and 8 are excluded. The NLO cross sections are calculated using Prospino 2 [71, 72, 73].



**Figure 14:**  $H_T$  distribution for SPS points 1a, 2, 3, 5 and 9. The peak or peaks of the distribution is correlated with the masses of the produced SUSY particles.

slightly rising trend going to higher masses.

In the present study, we did not make use of the most important feature of MadGraph/MadEvent: its efficiency in calculating multiparticle final states. In particular, we used Pythia to perform the decay of the SUSY particles produced. For the scalar quarks this should be a reasonable approximation, since no spin correlations are expected (al-

SPS point	1a	2	3	5	9
Peak position(s)(GeV)	800	1200 / 1900	1300	200 / 1100	2200
Particle masses (GeV)	$\tilde{g}$ : 600 $\tilde{q}$ : 550 $\chi_1^0$ : 100	$\tilde{g}$ : 780 $\tilde{q}$ : 1550 $\chi_1^0$ : 120	$\tilde{g}$ : 930 $\tilde{q}$ : 830 $\chi_1^0$ : 160	$\tilde{t}_1$ : 260 $\tilde{g}$ : 720 $\tilde{q}$ : 650 $\chi_1^0$ : 120	$\tilde{g}$ : 1290 $\tilde{q}$ : 1260 $\chi_1^0$ : 200

**Table 5:** The positions of the peaks in the  $H_T$  distributions and the masses of the particles mainly produced, as well as the LSP, for the SPS points 1a, 2, 3, 5 and 9.

though there might still be effects from interference or Breit-Wigner curve distortions in some parameter regions). However, if we want to study angular distributions where spin correlations might be important, the gluino decay should be done in MadEvent. For this kind of refinements, efficiency will start to become an important factor. It is therefore better to use different treatments for the different parameter points, in order to make an optimized choice for which processes to include at each point. So this should be seen rather as a preparatory study for such a more elaborate analysis.

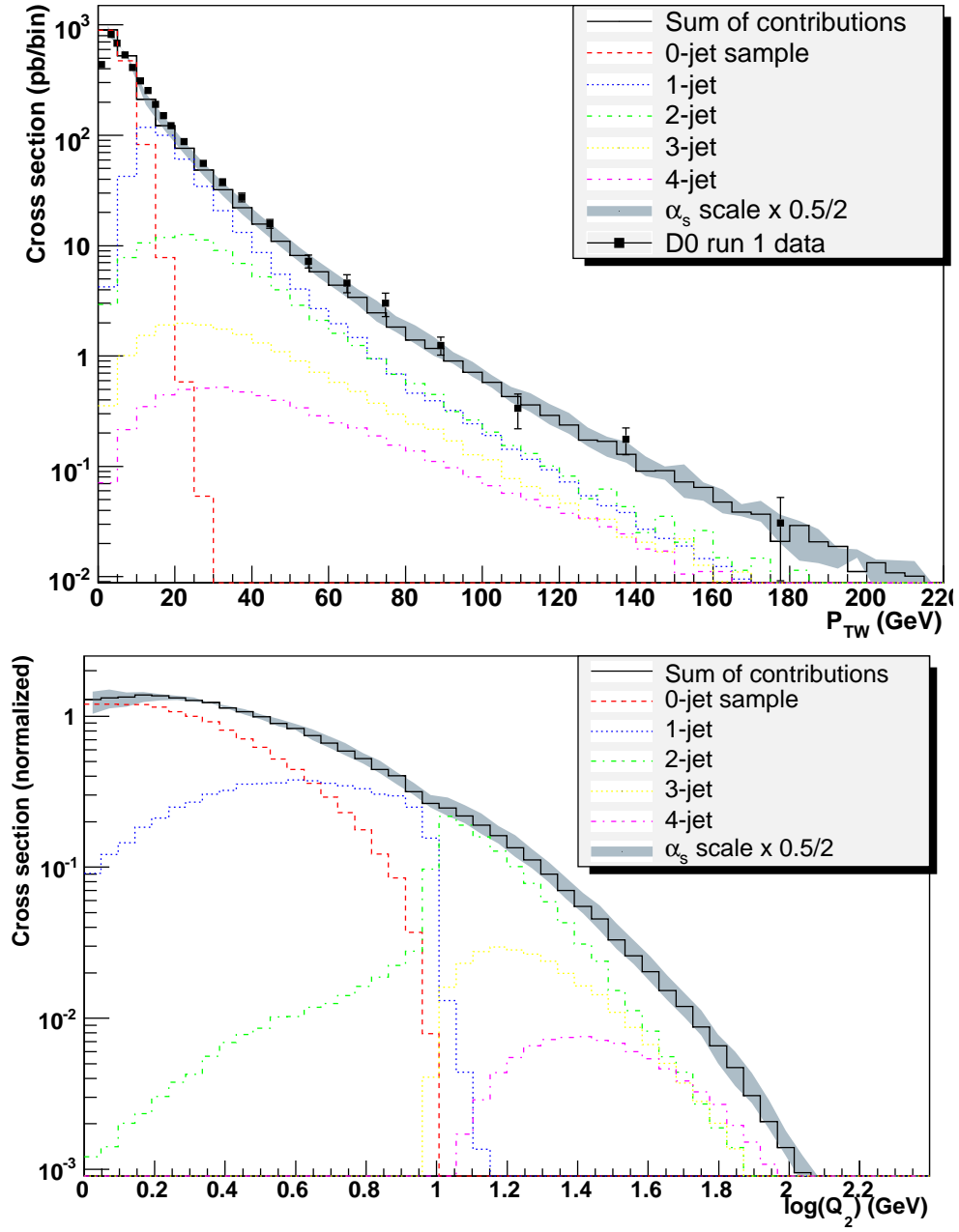
## 7.6 Inclusive $W$ +jets matched samples: comparison with the Tevatron data

At the Tevatron, as well as the LHC, many interesting signals include an isolated lepton, missing energy and hard jets. This means that weak boson production plus QCD jets is an important background. Until recently, this background was estimated using parton shower Monte Carlo's such as Pythia, where the largest jet multiplicity is  $W^\pm + 1$  jet, and additional jets are produced by parton showering. As discussed in Section 5, this gives a well-founded description of jets with small transverse momentum, and jets close in phase space, but for well separated hard jets it misses important non-logarithmic effects such as interference between diagrams. For that case, matrix element calculations are necessary. In order to simultaneously describe well-separated jets of different multiplicities, some kind of matching is needed between these descriptions.

$W^\pm$  and  $Z$  boson production with multiple jets is for several reasons an excellent testing ground for matching procedures; there is a simple “central”  $2 \rightarrow 1$  process, there is a hard scale set by the mass of the vector boson, and there is data from the Tevatron to compare to. A comparison between several implementations of matrix element–parton shower matching of  $W$  + jets production at hadron colliders is in progress [19].

In order to assess the robustness and flexibility of the method, there are several possibilities for variations. Keeping to the principal matching method used in MadEvent and Pythia, a modified MLM method using the  $k_\perp$ -clustered jet definition, the most prominent variable is the so-called matching scale, used as cutoff between the matrix element and parton shower descriptions for jet production. Other natural variables are the scales for the running of  $\alpha_s$ , which can be varied in the parton shower and/or in the matrix element generation, and the scale used in the parton densities, which might also be varied independently. In Fig. 15, we have chosen a fixed matching scale of 10 GeV and varied the scale





**Figure 15:** The  $p_{\perp}$  of the  $W$  boson and the parton-level differential  $1 \rightarrow 2$  jet rate for  $W$ +jets production at the Tevatron, using matching of matrix elements and parton showers with MadEvent and Pythia. Please refer to the text for further discussion of the plots.

for  $\alpha_s$  in the matrix element generation and the partons showers by a factor 2 up and down from the default scale, set by the  $k_{\perp}$  in the clustering corresponding to that power of  $\alpha_s$ .

Fig. 15a shows the  $p_{\perp}$  of the  $W$  boson in  $W$  production at the Tevatron, including hadronization of quarks and gluons but without underlying event simulation. The full black line shows the matched sum of the different jet multiplicity contributions, which are shown in different colors and line styles. For very low  $p_{\perp W}$  the pure parton shower dominates,

but from the matching scale and up, the different jet multiplicities become more and more important, until around 150 GeV, where all the four included multiplicities are of similar order of magnitude. The band in the plot corresponds to adding a factor  $\frac{1}{2} - 2$  in the arguments of  $\alpha_s$  in the matrix element generation only. It can be seen that the simulation with default parameters reproduces the shape of the Tevatron data [74] (the normalization has been fitted by 20%).

Fig. 15b shows the differential jet rate going from  $1 \rightarrow 2$  jets, after parton showers but before hadronization. It is included as an example, showing very clearly the transition between parton showers (in this case from the 0-jet and 1-jet samples), below the matching scale  $\log Q_2 = \log k_{\perp}^{\text{cut}} = 1$ , and matrix elements (from the 2-jet sample and upward) above the matching scale, as well as the relative smoothness of the transition. The jet rate is the variable used to perform the matching, although here only particles with  $|\eta| < 2.5$  are included in the jet definitions, explaining the low-end tails of the higher multiplicity samples.

## 8. Conclusions and Outlook

In this paper we have presented the new release of the MadGraph/MadEvent package which aims at providing a tool to perform the simulation of both signal (Standard Model and beyond) and backgrounds within one framework. It is worth summarizing the key points of our approach.

First, MG/ME is a user-driven Monte Carlo. Our package *does not contain a library of processes* but creates the code specific to the user's requests. It *contains an easy-to-extend library of physics models* which the user can choose from, including the Standard Model, the Higgs effective-theory, the MSSM, the most general 2HDM.

Second, event generation with MadGraph is available from the web, in addition to the downloadable source code. All functionalities of our package can be accessed directly via a web interface without the need of any installation/compilation. User's code creation, event generation and detector simulation can be controlled by filling (or uploading) simple input cards and the results in the form of event files and plots can be downloaded from the user's personal database. MG/ME event generation is based on an algorithm which parcels the overall work in many small independent jobs. At present, our resources include three medium-sized PC clusters that are available and open to the public for code creation and the generation of (limited-size) event samples. The web server software is portable and can be easily installed on any cluster running PBS or Condor. Extension to work with the Grid software are under development. The source code for this, and support for experimental as well as theoretical groups is available upon request to the authors.

MG/ME is designed to be a tool to ease the communication between theorists and experimentalists. Any theorist, model builder or phenomenologist, can implement and carefully test his/her own field theoretic model in MG/ME. The corresponding phenomenology at colliders, such as the LHC or an ILC, can be easily studied up to the simulated detector level by themselves and/or with the help of experimental colleagues. Both signal and background can be simulated within the same framework. Eventually, the experimental groups

can pick it up and promote it to a “exp-grade” analysis by using their detector specific tools.

The advantages of having “one framework for all” is manifest for both theorists (model builders) and experimentalists. For the former, it is important that the tool is very flexible, the learning curve is mild and no specific programming skills are needed. In particular, all the tedious and error-prone (and sometimes even difficult) tasks like matrix element creation, cross section integration and event generation, are automatically taken care of. On the experimental side, the insertion of a code in the simulation chain is normally a time-expensive process which involves testing and validation. For MG/ME this can be done, once for all, through a small set of simple Standard Model processes where comparison with available MC can be quickly done.

The further developments that are planned for the package are also towards building a framework suitable for theorists/experimentalists interactions.

In the current version, we have added a semi-automatic framework to implement new physics models. Even though this framework is quite general and very simple to use, it has two main drawbacks. First, the Feynman rules of a given model have to be available. This is trivial for very simple models but it can become rather tedious and cumbersome for richer or more complete models (such as SUSY). Second, the range of models that can be easily implemented is still limited by the kinematic form of the interactions in HELAS, which has to be similar to those present in the Standard Model. We are currently working on a package that starting from a generic Lagrangian generates the Feynman diagrams in a MadGraph compatible format and on a more flexible version of HELAS.

One feature that is under testing and will be soon included is the possibility of selecting the so-called decay chains in the generation of very rich multi-particle final states arising from the multi-staged decays of heavy particles in arbitrary models. By using a matrix element-based approach to decays, all information about possible correlations, such as those coming from the spin of the intermediate resonances, is available and can be exploited in the analysis.

Another line of development is towards building general analysis tools that can maximally exploit the information encoded in the matrix element to improve the accuracy or the sensitivity of measurements of key parameters (such as masses or couplings). These techniques have been successfully employed in several LEP and the Tevatron analyses, but always on a process specific basis. A general and automatic approach would be certainly welcome.

## Acknowledgements

We are thankful to the many people that during these years have directly or indirectly helped and/or supported the development MadGraph/MadEvent. In particular, Michelangelo Mangano and the Alpgen boys, Frank Krauss and the SHERPA kids, the Pythia and Herwig teams and in particular Torbjörn Sjöstrand, Steve Mrenna. We are in debt with our golden collaborator Kaoru Hagiwara for his continuous contributions to the project, from SMadGraph to the HELAS development. On the experimental side, we thank Henry

Frisch for being always among our most enthusiastic (and patient) users, Bruce Knuteson for being our most skilled hacker, Tony Liss, Sasha Nikitenko, and many many others among D0, CDF, ATLAS, and CMS physicists. We would also like to thank all the members of CP3 for the great atmosphere and environment that foster our efforts. Finally, we are grateful to David Lesny, Larry Nelson (UIUC), Fabrice Charlier, Thomas Keugten (UCL) and Roberto Ammendola, Filippo Palombi, Nazario Tantalò (Centro Fermi) for their restless and reliable technical support. Our human and computing resources are supported by the US National Science Foundation (Contract number NSF PHY 04-26272 ), Centro Fermi - Rome, Université Catholique de Louvain, Institut Interuniversitaire des Sciences Nucléaires, Belgian Interuniversity Attraction Pole P6/11 and the Swedish Research Council.

## References

- [1] T. Stelzer and W. F. Long, “Automatic generation of tree level helicity amplitudes,” *Comput. Phys. Commun.*, vol. 81, pp. 357–371, 1994, hep-ph/9401258.
- [2] F. Maltoni and T. Stelzer, “MadEvent: Automatic event generation with MadGraph,” *JHEP*, vol. 02, p. 027, 2003, hep-ph/0208156.
- [3] E. Boos *et al.*, “CompHEP 4.4: Automatic computations from Lagrangians to events,” *Nucl. Instrum. Meth.*, vol. A534, pp. 250–259, 2004, hep-ph/0403113.
- [4] A. Pukhov, “CalcHEP 3.2: MSSM, structure functions, event generation, batches, and generation of matrix elements for other packages,” 2004, hep-ph/0412191.
- [5] T. Gleisberg *et al.*, “SHERPA 1.alpha, a proof-of-concept version,” *JHEP*, vol. 02, p. 056, 2004, hep-ph/0311263.
- [6] W. Kilian, “WHIZARD 1.0: A generic Monte-Carlo integration and event generation package for multi-particle processes. Manual,” LC-TOOL-2001-039.
- [7] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, “ALPGEN, a generator for hard multiparton processes in hadronic collisions,” *JHEP*, vol. 07, p. 001, 2003, hep-ph/0206293.
- [8] C. G. Papadopoulos and M. Worek, “HELAC: A Monte Carlo generator for multi-jet processes,” 2006, hep-ph/0606320.
- [9] J. M. Campbell and R. K. Ellis, “An update on vector boson pair production at hadron colliders,” *Phys. Rev.*, vol. D60, p. 113006, 1999, hep-ph/9905386.
- [10] S. Frixione and B. R. Webber, “Matching NLO QCD computations and parton shower simulations,” *JHEP*, vol. 06, p. 029, 2002, hep-ph/0204244.
- [11] T. Sjostrand, S. Mrenna, and P. Skands, “Pythia 6.4 physics and manual,” *JHEP*, vol. 05, p. 026, 2006, hep-ph/0603175.
- [12] G. Corcella *et al.*, “HERWIG 6: An event generator for hadron emission reactions with interfering gluons (including supersymmetric processes),” *JHEP*, vol. 01, p. 010, 2001, hep-ph/0011363.
- [13] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, “QCD matrix elements + parton showers,” *JHEP*, vol. 11, p. 063, 2001, hep-ph/0109231.
- [14] F. Krauss, “Matrix elements and parton showers in hadronic interactions,” *JHEP*, vol. 08, p. 015, 2002, hep-ph/0205283.
- [15] S. Mrenna and P. Richardson, “Matching matrix elements and parton showers with HERWIG and PYTHIA,” *JHEP*, vol. 05, p. 040, 2004, hep-ph/0312274.
- [16] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, “Matching matrix elements and shower evolution for top-quark production in hadronic collisions,” 2006, hep-ph/0611129.
- [17] N. Lavesson and L. Lönnblad, “W + jets matrix elements and the dipole cascade,” *JHEP*, vol. 07, p. 054, 2005, hep-ph/0503293.
- [18] S. Hoche *et al.*, “Matching parton showers and matrix elements,” 2006, hep-ph/0602031.
- [19] e. a. J. Alwall, “Comparative study of various algorithms for the matching of parton showers and matrix elements in hadronic collisions.” *In preparation*.

- [20] S. Frixione, P. Nason, and B. R. Webber, “Matching NLO QCD and parton showers in heavy flavour production,” *JHEP*, vol. 08, p. 007, 2003, hep-ph/0305252.
- [21] P. Nason, “A new method for combining NLO QCD with shower Monte Carlo algorithms,” *JHEP*, vol. 11, p. 040, 2004, hep-ph/0409146.
- [22] P. Nason and G. Ridolfi, “A positive-weight next-to-leading-order Monte Carlo for Z pair hadroproduction,” *JHEP*, vol. 08, p. 077, 2006, hep-ph/0606275.
- [23] “PBS.” Web page at [http://www.altair.com/software/pbs\\_ser.htm](http://www.altair.com/software/pbs_ser.htm).
- [24] “Condor, high throughput computing.” Web page at <http://http://www.cs.wisc.edu/condor>.
- [25] H. Murayama, I. Watanabe, and K. Hagiwara, “HELAS: HELicity amplitude subroutines for Feynman diagram evaluations,” KEK-91-11.
- [26] M. A. Shifman, A. I. Vainshtein, M. B. Voloshin, and V. I. Zakharov, “Low-energy theorems for higgs boson couplings to photons,” *Sov. J. Nucl. Phys.*, vol. 30, pp. 711–716, 1979.
- [27] B. A. Kniehl and M. Spira, “Low-energy theorems in Higgs physics,” *Z. Phys.*, vol. C69, pp. 77–88, 1995, hep-ph/9505225.
- [28] S. Dawson and R. Kauffman, “QCD corrections to Higgs boson production: nonleading terms in the heavy quark limit,” *Phys. Rev.*, vol. D49, pp. 2298–2309, 1994, hep-ph/9310281.
- [29] V. Del Duca, W. Kilgore, C. Oleari, C. Schmidt, and D. Zeppenfeld, “H + 2 jets via gluon fusion,” *Phys. Rev. Lett.*, vol. 87, p. 122001, 2001, hep-ph/0105129.
- [30] F. Caravaglios, M. L. Mangano, M. Moretti, and R. Pittau, “A new approach to multi-jet calculations in hadron collisions,” *Nucl. Phys.*, vol. B539, pp. 215–232, 1999, hep-ph/9807570.
- [31] A. Pukhov *et al.*, “CompHEP: A package for evaluation of Feynman diagrams and integration over multi-particle phase space. User’s manual for version 3.3,” 1999, hep-ph/9908288.
- [32] J. F. Gunion, H. E. Haber, G. L. Kane, and S. Dawson, “The higgs hunter’s guide,” SCIPP-89/13.
- [33] N. Arkani-Hamed *et al.*, “The minimal moose for a little higgs,” *JHEP*, vol. 08, p. 021, 2002, hep-ph/0206020.
- [34] G. C. Branco, L. Lavoura, and J. P. Silva, *CP violation*. Oxford University Press, Oxford, England, 1999.
- [35] R. Barbieri, L. J. Hall, and V. S. Rychkov, “Improved naturalness with a heavy Higgs: An alternative road to LHC physics,” *Phys. Rev.*, vol. D74, p. 015007, 2006, hep-ph/0603188.
- [36] J. M. Gerard and M. Herquet, “A twisted custodial symmetry in the two-higgs-doublet model,” 2007, hep-ph/0703051. To be published in *Phys. Rev. Lett.*
- [37] P. Skands *et al.*, “SUSY Les Houches accord: Interfacing SUSY spectrum calculators, decay packages, and event generators,” *JHEP*, vol. 07, p. 036, 2004, hep-ph/0311123.
- [38] S. Davidson and H. E. Haber, “Basis-independent methods for the two-higgs-doublet model,” *Phys. Rev.*, vol. D72, p. 035004, 2005, hep-ph/0504050.
- [39] I. J. R. Aitchison, “Supersymmetry and the MSSM: An elementary introduction,” 2005, hep-ph/0505105.
- [40] S. P. Martin, “A supersymmetry primer,” 1997, hep-ph/9709356.

- [41] G. C. Cho *et al.*, “Weak boson fusion production of supersymmetric particles at the LHC,” *Phys. Rev.*, vol. D73, p. 054002, 2006, hep-ph/0601063.
- [42] K. Hagiwara *et al.*, “Supersymmetry simulations with off-shell effects for LHC and ILC,” *Phys. Rev.*, vol. D73, p. 055005, 2006, hep-ph/0512260.
- [43] J. F. Gunion and H. E. Haber, “Higgs bosons in supersymmetric models. 1,” *Nucl. Phys.*, vol. B272, p. 1, 1986.
- [44] T. Plehn, “Production of supersymmetric particles at high-energy colliders,” 1998, hep-ph/9809319.
- [45] M. Muhlleitner, A. Djouadi, and Y. Mambrini, “SDECAY: A Fortran code for the decays of the supersymmetric particles in the MSSM,” *Comput. Phys. Commun.*, vol. 168, pp. 46–70, 2005, hep-ph/0311167.
- [46] P. Meade and M. Reece, “BRIDGE: Branching ratio inquiry / decay generated events,” 2007, hep-ph/0703031.
- [47] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock, and B. R. Webber, “New clustering algorithm for multi - jet cross-sections in e+ e- annihilation,” *Phys. Lett.*, vol. B269, pp. 432–438, 1991.
- [48] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, “Longitudinally invariant k(t) clustering algorithms for hadron hadron collisions,” *Nucl. Phys.*, vol. B406, pp. 187–224, 1993.
- [49] M. L. Mangano, “Merging multijet matrix elements and shower evolution in hadronic collisions.” Available at <http://cern.ch/~mlm/talks/lund-alpgen.pdf>, 2004.
- [50] A. Kanaki and C. G. Papadopoulos, “Helac: A package to compute electroweak helicity amplitudes,” *Comput. Phys. Commun.*, vol. 132, pp. 306–315, 2000, hep-ph/0002082.
- [51] L. Lonnblad, “Ariadne version 4: A program for simulation of qcd cascades implementing the color dipole model,” *Comput. Phys. Commun.*, vol. 71, pp. 15–31, 1992.
- [52] J. Alwall *et al.*, “A standard format for Les Houches event files,” *Comput. Phys. Commun.*, vol. 176, pp. 300–304, 2007, hep-ph/0609017.
- [53] GnuPlot, web page at <http://www.gnuplot.info/>.
- [54] TopDrawer, available with manual at <http://www-group.slac.stanford.edu/beamphysics/codes/topdrawer/>.
- [55] ROOT, web page at <http://root.cern.ch/>.
- [56] ROOT tree structure, web page at <http://madgraph.phys.ucl.ac.be/Downloads/ExRootAnalysis/RootTreeDescription.html>.
- [57] J. Conway *et al.*, “PGS 4: Pretty Good Simulation of high energy collisions,” 2006, <http://www.physics.ucdavis.edu/~conway/research/software/pgs/pgs4-general.htm>.
- [58] M. R. Whalley, D. Bourilkov, and R. C. Group, “The Les Houches accord PDFs (LHAPDF) and LHAGLUE,” 2005, hep-ph/0508110. Web page at <http://hepforge.cedar.ac.uk/lhapdf/>.
- [59] L. Garren, “Stdhep, a common output format for monte carlo events.” Web page at <http://cepa.fnal.gov/psm/stdhep/>.

- [60] C. Buttar *et al.*, “Les Houches physics at TeV colliders 2005, standard model, QCD, EW, and Higgs working group: Summary report,” 2006, hep-ph/0604120.
- [61] G. Davatz, G. Dissertori, M. Dittmar, M. Grazzini, and F. Pauss, “Effective K-factors for  $gg \rightarrow H \rightarrow WW \rightarrow l\nu l\nu$  at the LHC,” *JHEP*, vol. 05, p. 009, 2004, hep-ph/0402218.
- [62] N. Kauer, “Top background extrapolation for  $H \rightarrow WW$  searches at the LHC,” *Phys. Rev.*, vol. D70, p. 014020, 2004, hep-ph/0404045.
- [63] V. Hankele, G. Klamke, D. Zeppenfeld, and T. Figy, “Anomalous Higgs boson couplings in vector boson fusion at the CERN LHC,” *Phys. Rev.*, vol. D74, p. 095001, 2006, hep-ph/0609075.
- [64] G. Klamke and D. Zeppenfeld, “Higgs plus two jet production via gluon fusion as a signal at the CERN LHC,” *JHEP*, vol. 04, p. 052, 2007, hep-ph/0703202.
- [65] T. Plehn, D. L. Rainwater, and D. Zeppenfeld, “Determining the structure of Higgs couplings at the LHC,” *Phys. Rev. Lett.*, vol. 88, p. 051801, 2002, hep-ph/0105325.
- [66] J. C. Collins and D. E. Soper, “Angular distribution of dileptons in high-energy hadron collisions,” *Phys. Rev.*, vol. D16, p. 2219, 1977.
- [67] W. J. Stirling and D. J. Summers, “Production of an intermediate mass Higgs boson in association with a single top quark at LHC and SSC,” *Phys. Lett.*, vol. B283, pp. 411–415, 1992.
- [68] F. Maltoni, K. Paul, T. Stelzer, and S. Willenbrock, “Associated production of higgs and single top at hadron colliders,” *Phys. Rev.*, vol. D64, p. 094023, 2001, hep-ph/0106293.
- [69] B. C. Allanach *et al.*, “The Snowmass points and slopes: Benchmarks for SUSY searches,” 2002, hep-ph/0202233.
- [70] B. C. Allanach, “SOFTSUSY: A C++ program for calculating supersymmetric spectra,” *Comput. Phys. Commun.*, vol. 143, pp. 305–331, 2002, hep-ph/0104145.
- [71] W. Beenakker, R. Hopker, M. Spira, and P. M. Zerwas, “Squark and gluino production at hadron colliders,” *Nucl. Phys.*, vol. B492, pp. 51–103, 1997, hep-ph/9610490.
- [72] W. Beenakker, M. Kramer, T. Plehn, M. Spira, and P. M. Zerwas, “Stop production at hadron colliders,” *Nucl. Phys.*, vol. B515, pp. 3–14, 1998, hep-ph/9710451.
- [73] T. Plehn., “Prospino 2.” Web page at <http://www.ph.ed.ac.uk/~tplehn/prospino/>.
- [74] B. Abbott *et al.*, “Differential cross section for  $w$  boson production as a function of transverse momentum in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.8$  tev,” *Phys. Lett.*, vol. B513, pp. 292–300, 2001, hep-ex/0010026.