Searches for New Physics: Les Houches Recommendations for the Presentation of LHC Results

S. Kraml[1], B.C. Allanach[2], M. Mangano[3], H.B. Prosper[4], S. Sekmen[3,4] (editors),
C. Balazs[5], A. Barr[6], P. Bechtle[7], G. Belanger[8], A. Belyaev[9,10], K. Benslama[11],
M. Campanelli[12], K. Cranmer[13], A. De Roeck[3], M.J. Dolan[14], T. Eifert[15], J.R. Ellis[16,3],
M. Felcini[17], B. Fuks[18], D. Guadagnoli[8,19], J.F. Gunion[20], S. Heinemeyer[17],
J. Hewett[15], A. Ismail[15], M. Kadastik[21], M. Krämer[22], J. Lykken[23] F. Mahmoudi[3,24],
S.P. Martin[25,26,27], T. Rizzo[15], T. Robens[28], M. Tytgat[29], A. Weiler[30]

Abstract

We present a set of recommendations for the presentation of LHC results on searches for new physics, which are aimed at providing a more efficient flow of scientific information between the experimental collaborations and the rest of the high energy physics community, and at facilitating the interpretation of the results in a wide class of models. Implementing these recommendations would aid the full exploitation of the physics potential of the LHC.

[1] *Laboratoire de Physique Subatomique et de Cosmologie, UJF Grenoble 1, CNRS/IN2P3, INPG, 53 Avenue des Martyrs, F-38026 Grenoble, France*

[2] *DAMTP, CMS, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, United Kingdom*

[3] *Physics Department, CERN, CH 1211 Geneva 23, Switzerland*

[4] *Department of Physics, Florida State University, Tallahassee, Florida 32306, USA*

[5] *Monash University, School of Physics, Melbourne Victoria 3800 Australia*

[6] *Department of Physics, Keble Road, Oxford, OX1 3RH, United Kingdom*

[7] *Universit"at Bonn, Physikalisches Institut, Nussallee 12, 53115 Bonn, Germany*

[8] *LAPTH, Université de Savoie, CNRS, B.P.110, F-74941 Annecy-le-Vieux Cedex, France*

[9] *NExT Institute: School of Physics and Astronomy, Univ. of Southampton, UK*

[10] *Particle Physics Department, Rutherford Appleton Laboratory, UK*

[11] *Physics Department, University of Regina,3737 Wascana Parkway, Regina–SK, Canada*

[12] *University College London, Gower Street, WC1B 6BT London, UK*

[13] *Center for Cosmology and Particle Physics, New York University, New York, NY 10003, USA*

[14] *Institute for Particle Physics Phenomenology, University of Durham, Durham DH1 3LE, UK*

[15] *SLAC National Accelerator Laboratory, 2575 Sand Hill Rd, Menlo Park, CA 94025, USA*

[16] *Theoretical Particle Physics and Cosmology Group, Department of Physics, King's College London, London WC2R 2LS, UK*

[17] *Instituto de Física de Cantabria (IFCA), CSIC-Universidad de Cantabria, E–39005 Santander, Spain*

[18] *Institut Pluridisciplinaire Hubert Curien/Département Recherches Subatomiques, Université de Strasbourg/CNRS-IN2P3, 23 Rue du Loess, F-67037 Strasbourg, France*

[19] *Laboratoire de Physique Théorique, Université Paris-Sud, Centre d'Orsay, F-91405 Orsay-Cedex, France*

[20] *Department of Physics, University of California at Davis, Davis CA, USA*

[21] *National Institute of Chemical Physics and Biophysics (NICPB), Tallinn, Estonia*

[22] *Institute for Theoretical Particle Physics and Cosmology, RWTH Aachen University, D-52056 Aachen, Germany*

[23] *Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, IL 60510*

[24] *Clermont Université, Université Blaise Pascal, CNRS/IN2P3, LPC, BP 10448, 63000 Clermont-Ferrand, France*

[25] *Department of Physics, Northern Illinois University, DeKalb IL 60115, USA*

[26] *Fermi National Accelerator Laboratory, P.O. Box 500, Batavia IL 60510, USA*

[27] *Kavli Institute for Theoretical Physics, University of California, Santa Barbara CA 93106, USA*

[28] *IKTP, TU Dresden, Zellescher Weg 19, 01069 Dresden, Germany*

[29] *Ghent University, Dept. of Physics and Astronomy, Proeftuinstraat 86, B-9000 Gent, Belgium*

[30] *Deutsches Elektronen-Synchrotron DESY, Notkestraße 85, D-22607 Hamburg, Germany*

1 Introduction

The LHC has very successfully begun to explore the TeV energy scale, and will be the energy frontier machine for the foreseeable future. Everyone who has had a hand in bringing this scientific and technological marvel to fruition deserves considerable credit and our thanks: the physicists and engineers who conceived, designed, and built it; those who operate the machine and its experiments; those who produce experimental results; those who try to understand them, and the public and its representatives whose generous support has enabled all this to happen.

The LHC was designed as a machine of discovery. There are high hopes that ground-breaking discoveries will indeed occur and shed light on electroweak symmetry breaking (be it via the Higgs mechanism or some other new dynamics) and new physics beyond the Standard Model (SM) of electroweak and strong interactions. It is of highest priority to our community to exploit fully the physics potential of the LHC. One aspect of this exploitation is the interpretation of LHC results in the contexts of different models of new physics. This is crucial if we are to unravel the correct new physics model, determine its parameters, and move beyond the SM.

The ATLAS and CMS collaborations are providing detailed experimental results [1, 2] of searches in many different channels. They are also providing interpretations in terms of popular models, such as the CMSSM[1], or in terms of Simplified Models[2]. These results are being used to test as large a variety of beyond-the-SM (BSM) scenarios as possible. For example, the searches for supersymmetry (SUSY), including [5–10], were interpreted in a number of different SUSY-breaking schemes, see e.g. [11–16], as well as in the weak-scale "phenomenological" MSSM [17, 18]. The sensitivity to light stops was investigated in [19–21], while implications of compressed SUSY spectra were analyzed in [22]. Interpretations were also made for non-SUSY models, for instance for the minimal universal extra dimension (UED) model in [23]. Similar non-collaboration efforts to interpret Higgs search results [24,25] in a large variety of BSM scenarios are also underway. These examples illustrate the community's interest in the LHC experimental results—interest that will surely grow as results become more comprehensive and readily available.

A systematic way of presenting LHC results will also greatly facilitate the comparison and combination of analyses within and across the LHC collaborations, as well as the assessment of the physics potential of future facilities. Furthermore, agreement on a set of recommendations and their implementation would be a further step towards a more comprehensive approach to the storage, persistence and future use of LHC results.

In this report, we therefore propose a set of recommendations for the presentation of LHC results aimed at maximizing its scientific return. Many of the experimental publications already implement several of the basic recommendations we make. But, as we shall see, our recommendations go substantially beyond current practice. Our wish is to stimulate discussions among the whole community and work towards an agreement on a common standard for the presentation of results. The goal is to help the community make the most of an extraordinary scientific opportunity.

---

[1]Constrained Minimal Supersymmetric Standard Model, see e.g. [3].

[2]Simplified Models are designed as an effective-Lagrangian description of a small number of accessible new particles. This approach has a long heritage; for a recent paper advocating it see e.g. [4].

2   Guiding Principles

As mentioned, it is our purpose here to formulate a set of recommendations that could act as guidelines for the documentation and use of the LHC results, in a form that would be most useful to the community at large, and that would help to maximize the scientific output of the LHC. Our recommendations are intended to respect the intellectual property rights of the collaborations and be concrete, practical and clear, as well as not being burdensome for the scientists performing the experimental analyses. By and large, several of these recommendations are rather obvious, and are already implemented in the publications of the LHC experiments. On the other hand, others are more ambitious. We present them here to stimulate further discussion between the experiments and the community, in the hope that they could be eventually adopted as part of the common practice.

With this in mind, our recommendations are guided by the following principles:

- *What* has been observed should be clear to a non-collaboration colleague.

- *How* it has been observed should be clear to a non-collaboration colleague.

- An interested non-collaboration colleague should be able to use and (re-)interpret results without the need to take up the time of collaboration insiders.

The latter principle implies that all ingredients (e.g., data, experimental systematics, cuts, procedures and so forth) in the analysis should be completely and unambiguously specified. We are not, of course, arguing that scientific discourse either within a collaboration or between collaboration members and those outside should be curtailed. On the contrary, this is vital to maintain the intellectual vibrancy of the field, and our suggestions are intended to make this more efficient and to reduce the burden on collaboration members.

To this aim, we think it useful to distinguish between experimental results and their interpretation. We suggest that the term *experimental result* be used exclusively to mean the empirical outcome, such as an event count or the measurement of some physical quantity. The experimental results themselves should be independent of any hypothesized new physics model. The term *interpretation* is the act of comparing the experimental results to model predictions. If full details of the experimental results employed for a given interpretation are readily available, their interpretation in different new physics models becomes possible. While the design of an analysis may have been inspired or guided by a specific physics model, ideally, the experimental results themselves should be independent of any such model so that the results can, in principle, be subject to different interpretations.

We also emphasize that it is important for the legacy of the LHC that its experimental results can be used in the future, even after the LHC has shut down and the collaborations have been disbanded. A coherent strategy for data preservation and re-use is discussed in [26].

These considerations, along with the principles listed above, have guided the recommendations in this document. We note again that many of these recommendations are already implemented in experimental analyses and publications: we hope that this document will facilitate discussions of—and serve as a guide for—best practice.

3 Recommendations

In the following we discuss our recommendations, which we present in four broad categories: analysis description, detector modeling, analysis dissemination and analysis design. Moreover, we include some recommendations regarding the interpretation of the results. Where appropriate, we split our recommendations into options:

(a) "crucial" recommendations, defined as actions that we believe should be undertaken immediately, and

(b), (c) "desirable steps", i.e. actions that would help, but whose implementation is recognized as requiring major efforts and a longer timescale.

Recommendations without such sub-division are understood as "crucial".

3.1 Analysis Description

As noted above, our guiding principle is that an interested non-collaboration colleague should be provided with all of the necessary information that is needed to use published results without having to consult collaboration insiders (although it may be wise to do so anyway). We thus recommend that the experimental publications contain a description of the analyses as clear and explicit as possible. Basic object definitions — for example, what constitutes an isolated electron — should be specified, so that the analysis may be reliably reproduced. Definitions of the important variables for the analysis should be precisely stated, because different definitions or conventions may exist for selection variables such as E_T^{miss}, m_{eff}, H_T, M_{T2}, α_T.

It is crucial that the analysis description provides sufficient information to validate an implementation of the analysis by users. In this regard, providing *cutflows*, i.e., the number of events obtained after each stage of the event selection for a given data or Monte Carlo (MC) event set, would provide valuable assistance. Since non-collaboration colleagues do not have access to the experimental data, nor the MC event set simulated with an official collaboration detector simulation, they do not have direct means to perform an exact, one-to-one synchronization and validation. It would help substantially to adopt the common practice of providing cutflows for a set of MC events, as well as experimental data, for a physics process that can be easily reproduced. Relevant information defining these MC events, e.g. the underlying physics model and processes, and the details of tools used in pre-detector event generation, including version information, should be specified. We note that guidelines for using event generators already exist in `MCNET` [27], see also [28], and we re-emphasize to adopt them.

Access to all this necessary information will be facilitated if it is tabulated, rather then described in the text. If limits on publication length do not allow the inclusion of all relevant information in the publication itself, the remaining details could be provided as *auxiliary information* alongside the publication. It would further greatly help to provide the relevant information in figures (coordinates of points in a graph, events in a histogram, etc.) in a digital form that is easily readable, e.g., as lists of numbers, as self-contained functions or as `ROOT` objects, etc. We thus summarize

> **Recommendation 1a:** *Provide a clear, explicit description of the analysis in publications. In particular, the most crucial information such as basic object definitions and event selection should be clearly displayed in the publications, preferably in tabular form, and kinematic variables utilised should be unambiguously defined. Further information necessary to reproduce the analysis should be provided, as soon as it becomes available for release, on a suitable common platform.*

We note that it is already common practice in the LHC experiments to provide useful auxiliary information for the longer papers[3], e.g., in `Rivet` [29], on `HEPdata` [30] and/or collaboration twiki pages [1, 2]. The `inSPIRE` [31] project may help to build a coherent information system, with detailed searchable and citeable entries. The ultimate goal should be to store all analysis information systematically in a common public archive based, e.g., at CERN. This brings us to

> **Recommendation 1b:** *The community should identify, develop and adopt a common platform to store analysis databases, collecting object definitions, cuts, and all other information, including well-encapsulated functions, necessary to reproduce or use the results of the analyses, and as required by other recommendations.*

The analysis database should also be capable of storing any analysis-related software that may be provided alongside the analysis. Although this is not listed as a recommendation, it would be extremely useful, and elegant, to design the infrastructure of analyses in a highly modular fashion so that cuts that define event selections, or perhaps even object definitions, or codes that perform kinematic reconstructions, or that compute the variables on which an analysis is based, are all encapsulated in well-defined functions that are independent of the provenance of the data they use. In fact, several such functions decoupled from the internal software infrastructure are already being made public by various analyses, and these functions can be added to the analysis database systematically in the future. In the case of complex analyses such as multi-variate analyses (MVAs), details about the MVA functions can help validate the tools developed by the user before applying them to models other than those used in the published analysis.

As mentioned, `Rivet` and `HEPdata` provide examples of such a platform, possibly supported by the `inSPIRE` indexing and searching infrastructure. Their functionality could be adapted to accommodate further needs, emerging from the discussions on the implementation of Recommendation 1b. The continued development of such tools should be encouraged, and could benefit from the support of initiatives such as the LHC Physics Centre at CERN (LPCC).

3.2 Detector Modeling

A) Efficiency maps: Analyses often use different definitions of analysis objects. For example, the definition of a candidate electron in one analysis may use a different definition of isolation than in another, or one analysis may use a cut on an MVA function to define an electron candidate while another applies cuts to several measured quantities. A well-understood way to shield a potential analyst from unnecessary complexity

---

[3]It is understood that, in order not to delay the publication while all the supplementary information is being prepared, not all of this information may be available immediately with the release of a paper, in particular when shorter articles such as Letters are concerned.

is to provide efficiency maps for each candidate object. Indeed experiments do provide efficiency maps along with some analyses, and we strongly encourage this practice. For a reliable use of efficiency maps,

- the definition of the object for which an efficiency is provided, e.g. an offline isolated electron, missing energy trigger, etc.,

- the definition of the efficiency, e.g. in which fiducial volume, or after which cuts an efficiency is defined,

- the final state topology for which an efficiency is defined

should be given precisely. Furthermore, it is very helpful if the efficiencies are presented in formats that can be implemented easily, such as lists of numbers or mathematical functions or standard digitized forms that are easy to interface with simulators or analysis codes. We thus arrive at

> **Recommendation 2a:** *Provide histograms or functional forms of efficiency maps wherever possible in the auxiliary information, along with precise definitions of the efficiencies, and preferably provide them in standard electronic forms that can easily be interfaced with simulation or analysis software.*

These standard electronic forms could rely on a platform similar to that discussed under Recommendation 1b, for example `Rivet/HEPdata` data and routines.

B) Public fast detector simulation: A fast detector simulation provides an approximate mapping from the pre-detector data to the post-reconstruction data. Publicly available fast detector simulations exist, like `PGS` [32] or `DELPHES` [33], which perform quite well and are generally found to reproduce ATLAS and CMS results with reasonable precision. Continued development, support and validation of these tools is of high value.

> **Recommendation 2b:** *The community should take responsibility for providing, validating and maintaining a simplified simulation code for public use, reproducing the basic response of the LHC detectors. The validation and tuning of this tool should be based on comparisons with actual performance plots, and/or other inputs, made available by the experiments along the lines of Recommendation 2a. Limits of validity should be investigated and clearly documented.*

We propose an open Workshop, bringing together the developers of the existing tools, the experts from the experiments, and the potential users, to discuss possible means of implementing this recommendation.

For completeness we note that publishing unfolded results provides an approach alternative to the need for processing MC data through a detector simulation. Unfolding and correcting to the particle level is certainly a preferred approach in many cases, e.g. in the presentation of cross-sections or distributions, but in practice it is not always viable or desirable in the case of BSM searches.

3.3   Analysis dissemination

A) Basic requirements: It is extremely important that all the crucial numbers regarding experimental results be summarized in a clear, concise, yet complete manner, preferably in tables. Experimental publications routinely provide these numbers, nevertheless here we encourage the maintenance of this good practice and list the minimum required information.

> **Recommendation 3a:** *Provide all crucial numbers regarding the results of the analysis, preferably in tabulated form in the publication itself. Further relevant information, like fit functions or distributions, should be provided as auxiliary material.*

In the case of a single-bin counting experiment these numbers include:

- D — the number of observed events in the signal region,

- δS — the systematic error on the expected number of signal events across the parameter space of the new physics model considered,

- B — the background estimate,

- δB — the estimated background uncertainty,

- $\mathcal{L}$ — the integrated luminosity estimate, and

- $\delta\mathcal{L}$ — the integrated luminosity uncertainty.

If the background estimate is the result of extrapolating from a control region (e.g, from a side-band) to the signal region, the following should also be provided (perhaps in the auxiliary information):

- Q — the observed number of events in the control region

- δQ — the uncertainty in Q

- k — the ratio of expected background in the control region to the expected background in the signal region. If the uncertainty δk on k is not negligible, it should also be included.

In the case of a multi-bin analysis, the above numbers should be given for each bin.

An important complication to address is how to account for systematic uncertainties. For the single-bin analysis, numbers should be reported with and without the inclusion of systematic uncertainties. The same holds for theoretical uncertainties of various types: it would be useful if the experiments also provided results obtained without the inclusion of the theoretical uncertainties for well-specified theoretical inputs (such as parton distribution functions (PDFs), top mass, etc.). In particular, since theoretical uncertainties are not static, this has the advantage of facilitating their re-assessment at a later stage in a straightforward manner. Systematic uncertainties on the signal, δS, should be given separately for detector specific sources, and for SM theory uncertainties, such as PDFs. The systematic uncertainty stemming purely from the calculation of the signal model prediction should be left out.

Whilst this method suffices for a single-bin analysis, it is not adequate for multiple-bin analyses, because it does not account for statistical dependencies between bins. One

way to account for the lowest-order statistical dependencies, i.e., linear correlations, is to provide the correlation matrix. However, this approach breaks down if the uncertainties are large or if errors are highly asymmetric.

Since it is common practice to include systematic uncertainties by integrating over systematic parameters, we include the following recommendation, which provides a straightforward way to perform this integration (or to profile).

> **Addendum to 3a:** *For multi-bin results, provide an ensemble of sets of the numbers B, δB, $\mathcal{L}$, $\delta\mathcal{L}$, Q, k, etc in the auxiliary information. These would be created by sampling from the various experiment-specific systematic effects, such as the jet energy scale, jet energy resolution, etc. Results should be quoted without inclusion of systematic/theoretical uncertainties external to the experiment.*

B) The full likelihood: The statistical model of an analysis provides the complete mathematical description of that analysis. The statistical model, through the probability density $p(o|x)$, relates the observed quantities o to the parameters x, describing the prediction in a model-independent way. By definition, the likelihood for a given set of theoretical model parameters x is the probability density over the observables o evaluated at their observed values O. (For clarity, we denote observed quantities, e.g. O, by upper case symbols and parameters, e.g. o, x, by lower case symbols.)

The likelihood is the appropriate starting point for any detailed interpretation of experimental results. However, many published analyses use likelihoods implicitly rather than explicitly. One problem with using likelihoods implicitly (for example, when results are expressed as $O \pm \delta O$) is that possible non-Gaussian tails are ignored. If the uncertainties are small this is not an issue. However, if the uncertainties are large the likelihood should be modelled explicitly. Given the likelihood, all the standard statistical approaches are available for extracting information from it. We therefore suggest the following.

> **Recommendation 3b:** *When feasible, provide a mathematical description of the <u>final</u> likelihood function in which experimental data and parameters are clearly distinguished, either in the publication or the auxiliary information. Limits of validity should always be clearly specified.*

Often, likelihood functions are constructed in several steps involving several hierarchies of fitted functions. Here, we define the <u>final</u> likelihood function as the last step in this process: it may be expressed in terms of an integral or maximisation over the product of several (possibly fitted) functions, which may be Gaussian or Poisson distributed, for example.

Since providing the full likelihood requires a very good understanding of its construction as well as of how to make correct use of it, a first approach towards Recommendation 3b may be to provide likelihoods in an approximate (simplified) form, again with the range of validity—typically the region where the signal hypothesis is defined—clearly specified.

It is, moreover, worth making the distinction between likelihoods parametrized only in terms of parameters that modify the normalization of distributions (e.g. cross-sections and branching ratios) versus those that parametrize acceptance (e.g. masses, spins, couplings of intermediate particles, etc.). The former is conceptually simple, as the distributions are trivially related to the parameters. The latter is very difficult, since the

distributions depend non-trivially on the parameters. Parametrizing acceptance over model parameters often requires interpolation or even extrapolation, which are both a legitimate cause for concern. Clearly, it is up to the experiments to make the judgement as to which parameters they feel the likelihood can be properly parametrized.

It would also be very useful and practical if the likelihood was provided in addition in a digital form. There already exists a generic, unified framework, `RooStats` [34], used by many LHC analyses, which allows one to model the probability density functions and likelihoods required as an input for any statistical inference technique, and also provides a set of major statistical techniques as `c++` classes with coherent interfaces to the statistical model. Publication of likelihoods in a systematic fashion under a standard digital format would also make combination of results much more feasible.

> **Recommendation 3c:** *Additionally provide a digitized implementation of the likelihood that is consistent with the mathematical description.*

We also note at this point that the `RECAST` [35] project would allow one to obtain the signal contribution to the likelihood for an arbitrary theoretical model, thus allowing one to build a higher-level framework for analysis re-interpretation.

3.4 Interpretation of experimental results

So far our recommendations concern generally the presentation of experimental results, irrespective of whether they report a signal or are used to set exclusion limits. Let us now turn to the interpretation of these results, the presentation of confidence intervals, parameter inference and limit setting.

Many different forms of experimental limit exist. Commonly, one-sided limits are derived in the absence of a signal observation, as is currently the case, but this will switch to two-sided limits (constraints) in case of a discovery. Limits may be quoted in various different schemes (such as Feldman-Cousins, CL_s, etc). It is crucial that the limit setting procedure be explicitly defined in order to permit an informed comparison of the quoted confidence level.

The *shape* (steepness) of the confidence level is essential information, e.g., for analyses that combine different experimental searches. It is therefore important that constraints are shown at several, rather than just one, confidence levels. Moreover, for the correct statistical interpretation, the expected constraints should be given in addition to the observed ones. Of course a more informative option would be to here, too, implement Recommendation 3b and publish the final likelihoods.

Regarding uncertainties, as mentioned earlier, it would be useful if confidence intervals were (also) presented for fixed input PDF's and other theoretical input, all explicitly tabulated. Moreover, when the interpretation of the experimental results is done in a "model-independent" way in terms of $\sigma \times \mathrm{BR} \times$ acceptance, the modeling of the acceptance should be precisely described. We sum this up as

> **Recommendation 4:** *In the interpretation of experimental results, preferably provide the final likelihood function (following Recommendations 3b/3c). When this is not possible or desirable, provide a grid of confidence levels over the parameter space. The expected constraints should be given in addition to the observed ones, and whatever sensitivity measure is applied must be precisely defined. Modeling of the acceptance needs to be precisely described.*

Note that Recommendation 4 in principle applies to any (re-)interpretation study, irrespective of whether it is done by an LHC collaboration or by non-collaboration groups. Needless to say, the model under investigation must be precisely defined.

As an aside we note that when conducting searches for supersymmetry or other new physics, experimental collaborations often use grids of models for which signal cross-sections, acceptances, efficiencies, and exclusions are evaluated, and then used to set limits by interpolation. It would be useful if these grid models were documented and fully specified in terms of model inputs, spectrum information (e.g., SLHA files), predicted signal cross-sections, acceptance×efficiency after selections and cuts, etc.. Also, since the tools provided by theorists constantly evolve, it is useful to document which tools and versions thereof have been used.

3.5   Higgs Searches

Given the special role of Higgs searches, we make a specific and separate recommendation for them. Higgs bosons are searched for in many different possible topologies, each of which are predicted to be present at some level, dictated by Higgs branching ratios in the Standard Model or new physics model. Many Higgs searches may be interpreted within the Standard Model itself, but both the branching ratios and the production cross-sections and distributions (and indeed the number of Higgs particles) may differ in new physics models. For this reason, it is important to include the channel-by-channel information in the primary Higgs search papers. Of course this does not preclude an additional combination of channels assuming some model, e.g., the Standard Model.

> **Recommendation 5:** *For Higgs searches, provide all relevant information on a channel-by-channel basis for both production and decay processes.*

Many publications on Higgs searches are already consistent with this recommendation in that constraints for individual production/decay modes are presented. The procedure of doing so as a function of the Higgs mass is crucial, especially in the context of multiple (or composite) Higgs boson models. Indeed, different Higgs models weight various possible production mechanism and decay distributions differently. It is moreover very instructive to give the best-fit signal strengths, $\sigma/\sigma_{\rm SM}$, as function of the SM Higgs boson mass for all available channels, along with error bands, as this facilitates testing deviations from SM couplings (see e.g. the discussion in [36]).

In order to make the channel-by-channel data available for an analysis of the Higgs sector of (nearly any) arbitrary model, it is also important that the expected sensitivity be reported for each channel in addition to the observed data. This permits the selection of the potentially strongest search channel, and allows to retain the correct statistical interpretation of, e.g., 95% C.L. exclusion bounds [37]. Finally, in setting limits or analyzing positive signals it is crucial to give details of acceptance systematics related thereto. These aspects are covered by Recommendation 4. Needless to say, all of the other Recommendations also apply.

3.6   Exclusive Analysis Design

It is a common approach to confront a new physics model with multiple analyses. In such cases, a likelihood is assigned to each analysis separately, and then all likelihoods are combined. This combination is typically easy only if the statistical data from each analysis are independent: then, one may combine the analyses by simply taking the

product of all likelihoods. The correlations that may arise in the systematic errors can be dealt with by following the addendum to Recommendation 3a. We realize that it is not possible to build an experimental search program that consists fully of disjoint analyses, and understand that avoiding overlaps between various inclusive analyses, e.g., those based on complicated kinematic variables, is far from trivial. Our recommendation applies to simpler cases, e.g. when building analyses based on simple variables like object multiplicity, or when defining different search regions in a single analysis. It is our intend to emphasize that more information is typically available in the combination of multiple disjoint analyses than via a single analysis.

Recommendation 6: *When relevant, design analyses and signal regions that are based on disjoint sets of events.*

4 Executive Summary of Recommendations

We here summarize our recommendations. We remind the reader that whenever we split into several steps, options **(a)** should be understood as "crucial" recommendations, while **(b), (c)** are "desirable steps". For completeness we also note that the ordering of Recommendations 1–6 does not imply prioritizing.

1. (a) Provide a clear, explicit description of the analysis in publications. In particular, the most crucial information such as basic object definitions and event selection should be clearly displayed in the publications, preferably in tabular form, and kinematic variables utilised should be unambiguously defined. Further information necessary to reproduce the analysis should be provided, as soon as it becomes available for release, on a suitable common platform.

   (b) The community should identify, develop and adopt a common platform to store analysis databases, collecting object definitions, cuts, and all other information, including well-encapsulated functions, necessary to reproduce or use the results of the analyses, and as required by other recommendations.

2. (a) Provide histograms or functional forms of efficiency maps wherever possible in the auxiliary information, along with precise definitions of the efficiencies, and preferably provide them in standard electronic forms that can easily be interfaced with simulation or analysis software.

   (b) The community should take responsibility for providing, validating and maintaing a simplified simulation code for public use, reproducing the basic response of the LHC detectors. The validation and tuning of this tool should be based on comparisons with actual performance plots, and/or other inputs, made available by the experiments along the lines of Recommendation 2a. Limits of validity should be investigated and clearly documented.

3. (a) Provide all crucial numbers regarding the results of the analysis, preferably in tabulated form in the publication itself. Further relevant information, like fit functions or distributions, should be provided as auxiliary material.
   Addendum:
   For multi-bin results, provide an ensemble of sets of the numbers B, δB, $\mathcal{L}$, $\delta \mathcal{L}$, Q, k, etc in the auxiliary information. These would be created by sampling from the various experiment-specific systematic effects, such as the jet energy scale, jet energy resolution, etc. Results should be quoted without inclusion of systematic/theoretical uncertainties external to the experiment.

   (b) When feasible, provide a mathematical description of the <u>final</u> likelihood function in which experimental data and parameters are clearly distinguished, either in the publication or the auxiliary information. Limits of validity should always be clearly specified.

   (c) Additionally provide a digitized implementation of the likelihood that is consistent with the mathematical description.

4. In the interpretation of experimental results, preferably provide the final likelihood function (following Recommendations 3b/3c). When this is not possible or desirable, provide a grid of confidence levels over the parameter space. The expected constraints should be given in addition to the observed ones, and whatever sensitivity measure is applied must be precisely defined. Modeling of the acceptance needs to be precisely described.

5. For Higgs searches, provide all relevant information on a channel-by-channel basis for both production and decay processes.

6. When relevant, design analyses and signal regions that are based on disjoint sets of events.

5 Conclusions

This document presents a set of recommendations for the presentation of LHC results on searches for new physics, which are aimed at providing a more efficient flow of scientific information and at facilitating the interpretation of the results in wide classes of models. It originated from discussions at the Les Houches "Physics at TeV Colliders 2011" workshop [38] and was thoroughly discussed and refined, with valuable input from representatives of the ATLAS and CMS collaborations, in a dedicated miniworkshop organized by the LHC Physics Centre at CERN [39]. The target of these recommendations are physicists both within and outside the LHC experiments, interested in the best exploitation of the BSM search analyses.

The added value for the experiments, and the whole HEP community, in extending the scope of the information made available about the experimental results, is a faster and more precise feedback on the implications of these results for a broad range of theoretical scenarios. Correlations and consistency checks among the findings of different experiments, at the LHC and elsewhere, will be facilitated, and will provide crucial input in the choice of the best research directions in both the near and far future, at the LHC and elsewhere. Improving the way the results of the LHC searches are documented and stored furthermore provides a forum to explore alternative approaches to long-term data archiving.

The tools needed to provide extended experimental information will require some dedicated efforts in terms of resources and manpower, to be supported by both the experimental and the theory communities. Practical solutions towards the development of these tools and the implementation of the proposed recommendations will be addressed in dedicated workshops and working groups.

References

[1] https://twiki.cern.ch/twiki/bin/view/AtlasPublic

[2] https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResults

[3] S. S. AbdusSalam, B. C. Allanach, H. K. Dreiner, *et al.*, Benchmark Models, Planes, Lines and Points for Future SUSY Searches at the LHC, Eur. Phys. J. C **71** (2011) 1835, arXiv:1109.3859 [hep-ph].

[4] D. Alves *et al.* [LHC New Physics Working Group Collaboration], Simplified Models for LHC New Physics Searches, arXiv:1105.2838 [hep-ph].

[5] S. Chatrchyan *et al.* [CMS Collaboration], Search for Supersymmetry at the LHC in Events with Jets and Missing Transverse Energy, Phys. Rev. Lett. **107** (2011) 221804 [arXiv:1109.2352 [hep-ex]].

[6] G. Aad *et al.* [ATLAS Collaboration], Searches for supersymmetry with the ATLAS detector using final states with two leptons and missing transverse momentum in $\sqrt{s} = 7$ TeV proton-proton collisions, arXiv:1110.61899 [hep-ex], to appear in Phys. Lett. **B**.

[7] G. Aad *et al.* [ATLAS Collaboration], Search for squarks and gluinos using final states with jets and missing transverse momentum with the ATLAS detector in $\sqrt{s} = 7$ TeV proton-proton collisions, arXiv:1109.6572 [hep-ex], to appear in Phys. Lett. **B**.

[8] G. Aad *et al.* [ATLAS Collaboration], Search for supersymmetry in final states with jets, missing transverse momentum and one isolated lepton in $\sqrt{s} = 7$ TeV pp collisions using 1 fb^{-1} of ATLAS data, Phys. Rev. D **85** (2012) 012006, arXiv: 1109.6606 [hep-ex]].

[9] S. Chatrchyan *et al.* [CMS Collaboration], Search for new physics with same-sign isolated dilepton events with jets and missing energy, CMS-PAS-SUS-11-010.

[10] S. Chatrchyan *et al.* [CMS Collaboration], Search for new physics in events with opposite-sign dileptons and missing transverse energy, CMS-PAS-SUS-11-011.

[11] M. Farina, M. Kadastik, D. Pappadopulo, *et al.*, Implications of XENON100 and LHC results for Dark Matter models, Nucl. Phys. B **853** (2011) 607, arXiv:1104.3572 [hep-ph].

[12] O. Buchmueller, R. Cavanaugh, A. De Roeck, *et al.*, Supersymmetry in Light of 1/fb of LHC Data, arXiv:1110.3568 [hep-ph].

[13] A. Fowlie, A. Kalinowski, M. Kazana, L. Roszkowski and Y. L. S. Tsai, Bayesian Implications of Current LHC and XENON100 Search Limits for the Constrained MSSM, arXiv:1111.6098 [hep-ph].

[14] B. C. Allanach, T. J. Khoo and K. Sakurai, Interpreting a 1 fb^{-1} ATLAS Search in the Minimal Anomaly Mediated Supersymmetry Breaking Model, JHEP **1111** (2011) 132, `arXiv:1110.1119` [hep-ph].

[15] Y. Kats, P. Meade, M. Reece and D. Shih, The Status of GMSB After 1/fb at the LHC, `arXiv:1110.6444` [hep-ph].

[16] D. Grellscheid, J. Jaeckel, V. V. Khoze, P. Richardson and C. Wymant, Direct SUSY Searches at the LHC in the light of LEP Higgs Bounds, `arXiv:1111.3365` [hep-ph].

[17] S. Sekmen, S. Kraml, J. Lykken, *et al.*, Interpreting LHC SUSY searches in the phenomenological MSSM, JHEP **02** (2012) 075, `arXiv:1109.5119` [hep-ph].

[18] A. Arbey, M. Battaglia and F. Mahmoudi, Implications of LHC Searches on SUSY Particle Spectra: The pMSSM Parameter Space with Neutralino Dark Matter, Eur. Phys. J. C **72** (2012) 1847, `arXiv:1110.3726` [hep-ph].

[19] M. Papucci, J. T. Ruderman and A. Weiler, Natural SUSY Endures, `arXiv:1110.6926` [hep-ph].

[20] X.-J. Bi, Q.-S. Yan and P.-F. Yin, Probing Light Stop Pairs at the LHC, Phys. Rev. D **85** (2012) 035005, `arXiv:1111.2250` [hep-ph].

[21] N. Desai and B. Mukhopadhyaya, Constraints on supersymmetry with light third family from LHC data, `arXiv:1111.2830` [hep-ph].

[22] T. J. LeCompte and S. P. Martin, Compressed supersymmetry after 1/fb at the Large Hadron Collider, `arXiv:1111.6897` [hep-ph].

[23] S. Chang, K. Y. Lee and J. Song, Implications on the mUED model from the early LHC data on ℓ + MET signal, `arXiv:1112.1483` [hep-ph].

[24] G. Aad *et al.* [ATLAS Collaboration], Combined search for the Standard Model Higgs boson using up to 4.9 fb^{-1} of pp collision data at $\sqrt{s}$ = 7 TeV with the ATLAS detector at the LHC, `arXiv:1202.1408` [hep-ex].

[25] S. Chatrchyan *et al.* [CMS Collaboration], Combined results of searches for the standard model Higgs boson in pp collisions at $\sqrt{s}$ = 7 TeV, `arXiv:1202.1488` [hep-ex].

[26] R. Kogler, D. M. South and M. Steder, Data Preservation in High Energy Physics, `arXiv:1111.2788` [hep-ex].

[27] N. Lavesson and D. Grellscheid, MCNET Guidelines for Event Generator Authors and Users, `http://www.montecarlonet.org/GUIDELINES`

[28] J. M. Butterworth, F. Maltoni, F. Moortgat, *et al.*, The Tools and Monte Carlo working group Summary Report, `arXiv:1003.16438` [hep-ph].

[29] A. Buckley, J. Butterworth, L. Lonnblad, *et al.*, Rivet user manual, `arXiv:1003.0694` [hep-ph].

[30] The Durham HepData Project, `http://hepdata.cedar.ac.uk/`

[31] http://inspirehep.net/, see also http://www.projecthepinspire.net/

[32] J. Conway, R. Culbertson, B. Kilminster, et al., Pretty Good Simulator,
http://www.physics.ucdavis.edu/$\sim$conway/research/software/pgs/
pgs4-general.htm

[33] S. Ovyn, X. Rouby and V. Lemaitre, DELPHES, a framework for fast simulation
of a generic collider experiment, arXiv:0903.2225 [hep-ph].

[34] L. Moneta, K. Belasco, K. S. Cranmer, et al., The RooStats Project, PoS ACAT
2010 (2010) 057, arXiv:1009.1003 [physics.data-an].

[35] K. Cranmer and I. Yavin, RECAST: Extending the Impact of Existing Analyses,
JHEP **1104** (2011) 038, arXiv:1010.2506 [hep-ex],
http://recast.perimeterinstitute.ca/

[36] J. R. Espinosa, C. Grojean, M. Muhlleitner and M. Trott, Fingerprinting Higgs
Suspects at the LHC, arXiv:1202.3697 [hep-ph].

[37] P. Bechtle, O. Brein, S. Heinemeyer, G. Weiglein and K. E. Williams, HiggsBounds
2.0.0: Confronting Neutral and Charged Higgs Sector Predictions with Exclusion
Bounds from LEP and the Tevatron, Comput. Phys. Commun. **182** (2011) 2605,
arXiv:1102.1898 [hep-ph].

[38] G. Brooijmans, B. Gripaios, F. Moortgat, J. Santiago, P. Skands, et al., Les Houches
2011: Physics at TeV Colliders New Physics Working Group Report, arXiv:1203.
1488 [hep-ph].

[39] Workshop on "Searches for New Physics: Les Houches Recommendations for the
Presentation of LHC Results", CERN, 13 Feb 2012 [INDICO].