Statistical methods in CMS searches

Amnon Harel, on behalf of the CMS collaboration
University of Rochester

Abstract
A review of the statistical methods used in the first CMS searches for new physics at 7 TeV, from 2010 to January 2011.

1 Introduction

In 2010, the Large Hadron Collider (LHC) started producing pp collisions at a center of mass energy of 7 TeV. By the year's end, a data sample corresponding to over 40 pb^{-1} was recorded in the CMS detector, a multipurpose high-energy particle detector. The CMS collaboration analyzed this data for evidence of physics beyond the standard model (SM). The results of these first searches are consistent with the SM, and limits were placed on the corresponding new physics scenarios. We will review the statistical methods used to set these first limits and to rule out evidence of new physics.

2 The W' search

In Ref. [1] we report a search for the production and decay of a heavy copy of the W boson. Specifically, we search for a W' boson that has W-boson like couplings to fermions and does not couple to other gauge bosons. A previous search at the Fermilab Tevatron Collider ruled out $M_{\text{W'}} < 1.1$ TeV.

We simulate the signal using the PYTHIA V6.422 event generator [2], and scale the production cross section to match next-to-next-to-leading order (NNLO) calculations. We select events with an isolated electron, and a p_T imbalance ($\not\!\!E_T$). The p_T imbalance is reconstructed using the particle-flow technique [3]. The main background processes are W+jets and multijet production. We derive the distributions of key observables for both processes using data-driven techniques. We then fit a linear sum of these distribution to the distribution observed in collision data, with the additional smaller background contributions accounted for according to simulation, as shown in Fig. 1.

Next, we define for each event its visible mass,

$$(M_T)^2 = 2E_T^e \not\!\!E_T \left(1 - \cos\phi_{e,\not\,E_T}\right), \tag{1}$$

and look for an excess of high M_T, as shown in Fig. 2. For each W' mass $M_{\text{W'}}$, we defined a priori a search region consisting of M_T values above some minimal value and set limits on the effective cross section of a hypothetical W' boson. No data are observed in any of the search regions.

We use a Bayesian limit setting procedure following Ref. [4] which addresses this canonical scenario: Poisson statistics in each bin of M_T, no interference between the signal and background contributions, and the systematic uncertainties are easily factorized.

The statistical problem is then that for each M_T bin we have:

$$N_{\text{pred}} = b + \mathcal{L}\epsilon\sigma_{\text{eff}}, \tag{2}$$

where $\mathcal{L}$ is the integrated luminosity, ϵ is the selection efficiency for that bin, and σ_{eff} is the effective cross section, i.e. the production cross section (σ) times the branching fraction into the observed channel (B). Then $N_{\text{pred}}(M_T)$ is given for the null (SM) hypothesis ($\sigma_{\text{eff}} = 0$) and for the alternative (SM+signal) hypothesis. We use a constant prior for σ_{eff}, often described as "flat" in HEP papers:

$$f(\sigma_{\text{eff}}) = \begin{cases} \text{const} & \sigma_{\text{eff}} \in [0, \sigma_{\max}] \\ 0 & \text{otherwise} \end{cases}, \tag{3}$$

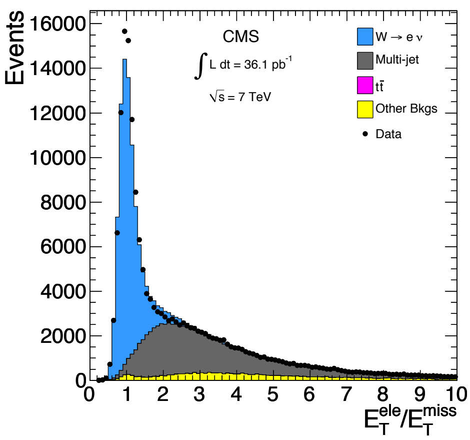

Fig. 1: Sample composition in the W' search and the distributions of the observable used in its fit.

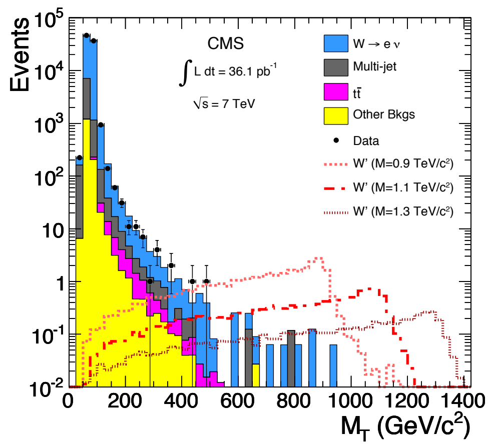

Fig. 2: Data, sample composition, and examples of signal models for the W' search.

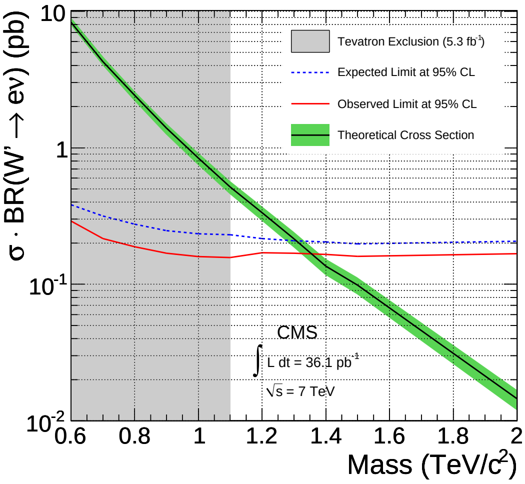

Fig. 3: Limits on W' bosons.

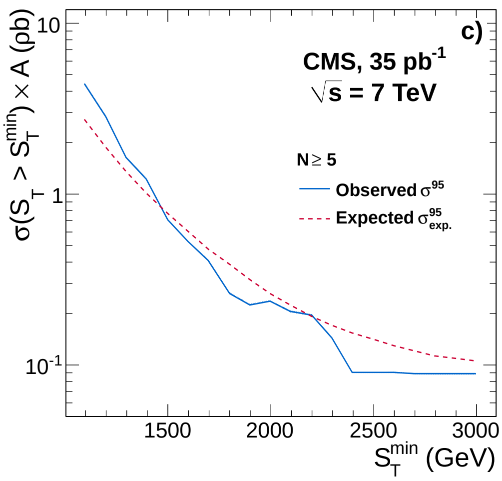

Fig. 4: An example of model-independent limits on microscopic black hole production.

with $\sigma_{\rm max}$ chosen to be large enough so that the results are not sensitive to its exact value. We use log-normal priors for the nuisance parameters, and integrate them out. In particular, the signal normalization uncertainties from the fit are summarized into one number, which is a typical approximation. The resulting limits are shown in Fig. 3 together with the $\sigma_{\rm eff}$ predicted for each $M_{\rm W'}$. From their intersection we rule out at 95% CL the existence of such W' bosons with $M_{\rm W'} < 1.36$ TeV.

3 Other Bayesian limits

The same statistical treatment was used in the early CMS searches for 1st [5] and 2nd [6] generation leptoquarks and for microscopic black holes [7]. The latter also contains model-independent limits on the effective cross section times acceptance for the different final-state particle multiplicities: ≥ 3, ≥ 4, and ≥ 5 (see example in Fig. 4). Several models are considered with rotating or non-rotating black holes, with or without a stable non-interacting remnant, and with differing values of the Planck scale in the bulk, the number of extra dimensions, and of the minimal black hole mass. In all cases, the model independent limits were only 10% worse than the full model-dependent limits, as in each model there is one particle multiplicity that dominates the limits.

4 Dijet resonance search

CMS published a search for resonant dijet production [8]. Dijet resonances are common in models of new physics beyond the SM. Eight specific models are studied in the paper, which also describes a model-independent study based on three generic signal models: narrow resonances with quark-quark, quark-gluon and gluon-gluon final states.

We consider the two leading (largest p_T) jets as the dijet system. Events are collected using single-jet triggers. Only events where both jets have a pseudo-rapidity $|\eta| < 2.5$ and their unsigned pseudo-rapidity difference is $|\Delta\eta| < 1.3$ are used. For each event we reconstruct the invariant mass of the dijet system, m_{jj}. The observables used in the statistical analysis are the event counts in each of the predefined m_{jj} bins. The width of the m_{jj} bins corresponds to the experimental resolution on m_{jj}. A narrow resonance is one whose width is similar to or smaller than the experimental resolution on m_{jj}.

Non-resonant dijet production is described by a fit to the data of a smooth functional form that does not contain a peak. Three functional forms, used in similar searches in previous colliders, were considered. The best fit to the data (χ^2/N.D.O.F. $= 32/31$) was with the form

$$\frac{d\sigma}{dm_{jj}} = p_0 \frac{\left(1 - \frac{m_{jj}}{\sqrt{s}}\right)^{p_1}}{\left(\frac{m_{jj}}{\sqrt{s}}\right)^{p_2 + p_3 \ln\left(\frac{m_{jj}}{\sqrt{s}}\right)}}, \tag{4}$$

where p_i are the fitted parameters, and $\sqrt{s}$ is the collision energy (7 TeV), and is shown in Fig. 5.

To verify the fit's agreement with data and rule out evidence for dijet resonances, we find the biggest excess in the range 0.5 – 2.0TeV, which is for a resonance mass ≈ 0.9 TeV, and quantify its statistical significance. Its local significance, from the log likelihood ratio (LLR), is 1.7σ. We account for the "look elsewhere effect" (LEE) using ensemble tests, and find a similar or locally-more-significant fluctuation in almost half the pseudodatasets (PDSs), so that the overall significance is reduced to 0.02σ.

We set limits on resonance dijet production using an approximate Bayesian procedure. The statistics-only case is treated exactly, using the same method used in the W' search (see Section 2). In this analysis we define $\sigma_{\text{eff}} = \sigma BA$, where A is the acceptance, i.e., the probability that the resonance produces two jets that pass the selection criteria.

The systematic uncertainty is incorporated at each resonance mass by smearing the posterior probability density of σ_{eff} with a Gaussian whose width is set to the systematic uncertainty on the measured σ_{eff}. This is approximate, but here, it is also conservative. In particular, we verified frequentist coverage at 1 TeV for σ_{eff} equal to the limiting value, finding a coverage of $\approx 95\%$ without systematic uncertainties, and $> 98\%$ with systematic uncertainties.

The JES uncertainty is the dominant systematic uncertainty, yielding fractional uncertainties of roughly 20 to 40%, depending on the resonance mass. Other systematic uncertainties, due to the choice of background parametrization, jet energy resolution, and the integrated luminosity, yield fractional uncertainties of $\approx 10\%$ each. The systematic uncertainties increase the cross section limits by 15 to 50%, depending on the resonance mass and its parton content. They decrease the mass limits by $\approx 10\%$.

The limits on resonant dijet production are shown in Fig. 6. We rule out, at the 95% CL, string resonances of mass 0.5–2.5 TeV, excited quarks of mass 0.50–1.58 TeV, axigluons and colorons of mass 0.50–1.17 TeV and 1.47–1.52 TeV, and E_6 diquarks of mass 0.50–0.58 TeV, 0.97–1.08 TeV, and 1.45–1.60 TeV. References to the exact models used are available in the paper [8].

5 The CLs method

The CLs method is to exclude regions of phase space where

$$CL_s = \frac{CL_{s+b}}{CL_b} < 1 - \alpha, \tag{5}$$

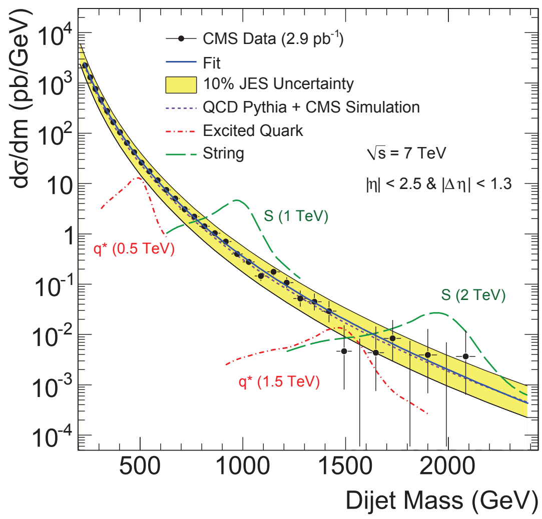

Fig. 5: Data, background fit, and examples of signal models for the dijet resonance search.

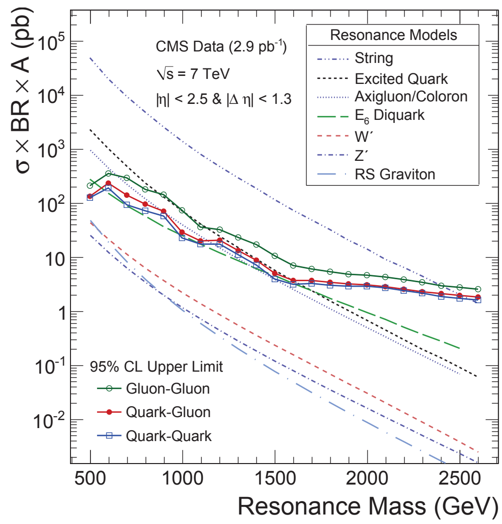

Fig. 6: Limits on dijet resonances.

where α is the desired confidence level, and CL_b and CL_{s+b} are the standard tail probabilities under the null and signal hypotheses (CL_b is the p value). It is recommended [9] to use the LLR as the observable. But often the systematic uncertainties are ignored when calculating the LLR observable, in keeping with the more general prescription [10].

The method's name is very descriptive, but also misleading, as the CLs exclusion region is not a confidence interval. The method is neither purely frequentist nor Bayesian, instead its motivation is practical — it seeks to modify the frequentist CL_{s+b} to avoid false exclusions when the experiment is insensitive to the signal, that it, it is a method of power-constraining frequentist limits. The CLs limit corresponds to the frequentist limits when the experiment is fully sensitive, and the method smoothly degrades the limits as the experiment's power decreases. Despite its shaky foundations in statistical theory, it has been producing sensible results for over a decade.

6 Search for quark compositeness

CMS searched for quark compositeness [11], which is expected to appear at low energies as a contact interaction. Quark contact interactions will enhance low-$|\eta|$ dijet production, in contrast to the SM production, where quantum chromodynamics predicts mostly high-$|\eta|$ jets from t-channel production.

We define the dijet centrality ratio R_η as the number of events where both leading jets are central, with $|\eta| < 0.7$, divided by the number of events where both are less central, with $0.7 < |\eta| < 1.3$. Except for these angular cuts, the event selection follows that of the dijet resonance search above.

The R_η observable is binned in m_{jj} using the same binning as in the dijet resonance search. The background is estimated from next-to-leading-order (NLO) calculations with non-perturbative corrections and with an offset in R_η to match the data in the low m_{jj} region, where no new physics is expected. The fitted offset is $-0.050 \pm 0.021(\text{stat.}) \pm 0.039(\text{syst.})$. Using an ensemble of PDS generated according to the background model, we find the two-sided p value of this offset to be 0.29. The data and background are shown in Fig. 7. At high m_{jj} the data is significantly less signal-like than the SM predictions. But overall, the data and background model are consistent. For example, fitting an offset over the entire m_{jj} range yields $-0.037 \pm 0.007(\text{stat.}) \pm 0.039(\text{syst.})$ with a two-sided p value of 0.34.

In each m_{jj} bin, R_η is distributed as a "Ratio of Poisson means", and we use the standard and extremely useful practice of conditioning this distribution on the total (inner + outer) number of events observed in that bin, simplifying it to a Binomial distribution [12]. We combine data from all m_{jj} bins

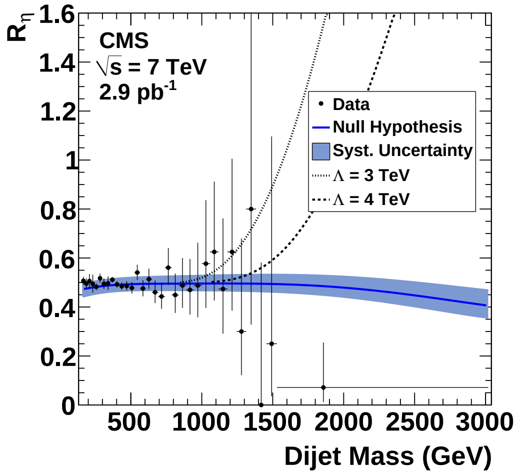

Fig. 7: Data, background model, and examples of signal models for the quark compositeness search.

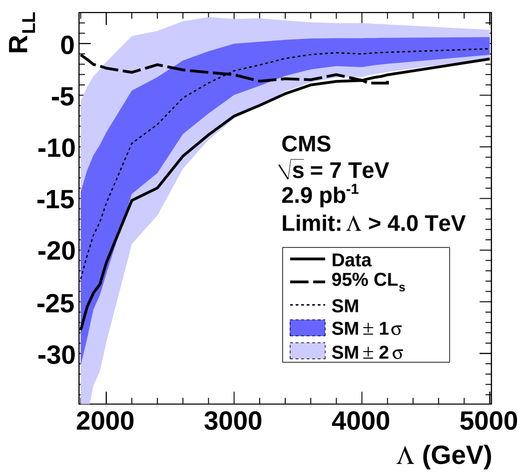

Fig. 8: Test statistics and limits on the scale of quark compositeness.

into one test statistic — the statistics-only LLR for the SM and SM-with-contact-interaction hypotheses. We use the CLs method to set limits on the scale of the contact interactions, Λ. Each Λ value is evaluated separately. The CL_b and CL_{s+b} tails are calculated by ensemble testing, with the nuisance parameters integrated out by varying them for each PDS. The low R_η values at high m_{jj} lead to low CL_b values which require evaluation of the extreme tails of CL_{s+b}, which proved difficult using this integration technique. Large ensembles were needed, some with $> 200\,000$ PDSs. To avoid bias from the choice of ensemble size we formalized stopping rules for the production of additional PDSs: either the Λ value is included/excluded by CLs at the 2σ level, or the statistical error on the CL_s value is $< 0.5\%$.

Fig. 8 illustrates the limit setting procedure. The intersection of the CLs and SM curves indicates the expected limit of $\Lambda > 2.9$ TeV, while the intersection of the CLs and data curves indicates the much higher observed limit of $\Lambda > 4$ TeV.

7 Stopped gluino search

CMS published a search for heavy, quasi-stable particles [13], in particular, for gluinos predicted by split-SUSY which give rise to charged R-hadrons. Such R-hadrons would stop within the CMS detector, and decay at a later time, in contrast to SM decays, whose timing is strongly correlated with LHC collisions.

The search used two observables. The first is the number of events within a time window. The window starts at 50 ns and ends at $1.256\tau_{\text{gluino}}$ (where τ_{gluino} is the gluino lifetime) after each LHC bunch crossing, and excludes 100 ns windows around subsequent bunch crossings. No signal excess was observed and the CLs method was used to derive limits on σ_{eff} for each τ_{gluino} hypothesis (see "Counting" in Fig. 9) and for different stopping scenarios.

The second observable was the time of the selected events, within those same time windows. The signal time-dependence is driven by the timing of the bunch crossing and by τ_{gluino}. The background is mostly from instrumental noises, and is time-independent. We calculate a likelihood as a function of the background amount (per LHC filling scheme) and σ_{eff}, and derive Bayesian limits from the posterior probability using uniform priors in both variables (see "Timing" in Fig. 9).

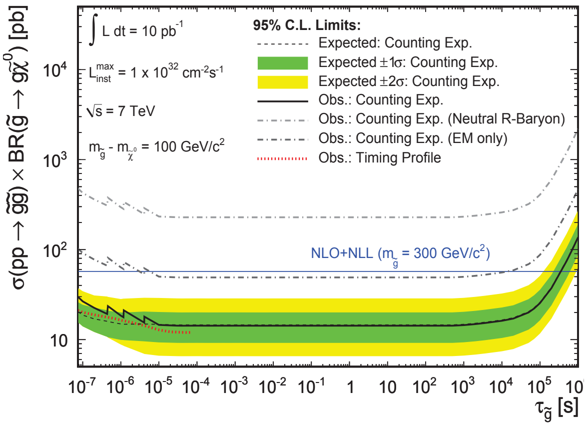

Fig. 9: Limits on stopped gluino production.

8 Summary

The statistical techniques used in the first CMS searches for new physics in pp collisions at $\sqrt{s} = 7$ TeV were reviewed. Consistency with the SM was typically evaluated using p values from ensemble tests. Limits were set using either Bayesian methods or the CLs method.

References

[1] CMS Collaboration, *Search for a heavy gauge boson W' in the final state with an electron and large missing transverse energy in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. B **698** (2011) 21.

[2] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, `http://dx.doi.org/doi:10.1088/1126-6708/2006/05/026` JHEP **2006** (2006) 026.

[3] CMS Collaboration, *Particle-flow event reconstruction in CMS and performance for jets, taus, and missing ET*, CMS PAS PFT-09-001 (2009).

[4] I. Bertram *et al.*, *A recipe for the construction of confidence limits*, technical report, 2000. FERMILAB-TM-2104.

[5] CMS Collaboration, *Search for pair production of first-generation scalar leptoquarks in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Rev. Lett. **106** (2011) 201802.

[6] CMS Collaboration, *Search for pair production of second-generation scalar leptoquarks in pp Collisions at $\sqrt{s} = 7$ TeV*, Phys. Rev. Lett. **106** (2011) 201803.

[7] CMS Collaboration, *Search for microscopic black hole signatures at the Large Hadron Collider*, Phys. Lett. B **697** 5 (2011) 434.

[8] CMS Collaboration, *Search for dijet resonances in 7 TeV pp collisions at CMS*, Phys. Rev. Lett. **105** (2010) 211801.

[9] A. L. Read, *Presentation of search results: the CLs technique*, J. Phys. G: Nucl. Part. Phys. **28** (2002) 2693.

[10] A. L. Read, *Modified frequentist analysis of search results (the CLs method)*, in Proceedings of the First Workshop on Confidence Limits, CERN, Geneva, Switzerland (2000) 81.

[11] CMS Collaboration, *Search for quark compositeness with the dijet centrality ratio in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Rev. Lett. **105** (2010) 262001.

[12] J. Przyborowski and H. Wilenski, *Homogeneity of results in testing samples from Poisson series with an application to testing clover seed for dodder*, Biometrika **31** (1940) 313.

[13] CMS Collaboration, *Search for stopped gluinos in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Rev. Lett. **106** (2011) 011801.