University of California
Santa Barbara

Search for new physics with final states containing a same-sign dilepton pair at a center-of-mass energy of 13 TeV with the CMS detector

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Physics

by

David McAlister Barry

Committee in charge:

Professor Charles McThornbody, Chair
Professor Russell Hammond
Professor Alfred Alfredo
Professor Jackmerius Tacktheritrix

Month 2019

The Dissertation of David McAlister Barry is approved.

Professor Russell Hammond

Professor Alfred Alfredo

Professor Jackmerius Tacktheritrix

Professor Charles McThornbody, Committee Chair

March 2019

Search for new physics with final states containing a same-sign dilepton pair at a center-of-mass energy of 13 TeV with the CMS detector

Acknowledgements

I would like to acknowledge and thank my advisor, Claudio Campagnari, as well as friends, family, and colleagues. I would like to thank the Vim editor.

Curriculum Vitæ

David McAlister Barry

Education

2019	Ph.D. in Physics (Expected), University of California, Santa Barbara.
2017	M.A. in Physics, University of California, Santa Barbara.
2014	B.Sc. in Physics, Texas A&M University, College Station, TX.

Publications

- B. Hashemi, N. Amin, K. Datta, D. Olivito, and M. Pierini, *LHC analysis-specific datasets with Generative Adversarial Networks* [arXiv:1901.0528] **In progress**
- CMS Collaboration, *Search for standard model production of four top quarks in final states with same-sign and multiple leptons in proton-proton collisions at* $\sqrt{\mathrm{s}} =$ 13 TeV [PAS TOP-18-003] **In progress**
- CMS Collaboration, *Search for physics beyond the standard model in events with two same-sign leptons or at least three leptons and jets in proton-proton collisions at* $\sqrt{\mathrm{s}} = 13$ TeV [PAS SUS-19-008] **In progress**
- CMS Collaboration, *Search for standard model production of four top quarks with same-sign and multilepton final states in proton-proton col- lisions at* $\sqrt{\mathrm{s}} = 13$ TeV, *Eur. Phys. J.* **C78** (2018) [arXiv:1710.1061]
- CMS Collaboration, *Search for physics beyond the standard model in events with two leptons of same sign, missing transverse momentum, and jets in proton-proton collisions at* $\sqrt{\mathrm{s}} = 13$ TeV *Eur. Phys. J.* **C77** (2017) [arXiv:1704.0732]

Abstract

Search for new physics with final states containing a same-sign dilepton pair at a center-of-mass energy of 13 TeV with the CMS detector

by

David McAlister Barry

Two related searches for Standard Model and beyond the Standard Model physics with a final state containing a pair of same-charged leptons and jets are performed using a sample of $\sqrt{s} = 13$ TeV data corresponding to an integrated luminosity of 137 fb^{-1}, collected by the CMS detector between 2016 and 2018. The first inclusive search observes no excess above the Standard Model and thus places constraints on R-parity violating and R-parity conserving supersymmetric models with pair production of gluinos and squarks. Gluino masses are excluded up to 2.1 TeV, while top and bottom squarks are excluded up to 0.9 TeV. The second search measures the cross-section of the production of four top quarks within the Standard Model using both cut-based and multivariate approaches. The observed (expected) significance of the multivariate approach is 2.6 (2.7) standard deviations, with a measured cross-section of $12.6^{+5.8}_{-5.2}$ fb, consistent with the Standard Model prediction of $12.0^{+2.2}_{-2.5}$ fb. These results are translated into constraints on the Yukawa coupling of the top quark, as well as constraints on heavy scalar or pseudoscalar production in a type II 2HDM scenario.

Contents

Chapter 1

Introduction

1.1 Info

Some facts

- Standard model missing some stuff
- LHC smashes protons together very fast
- CMS detector is very big
- Analyze CMS data to find stuff beyond the standard model (e.g., SUZY)
- Did not find anything

The results presented in this thesis correspond to the published results in Refs. [1, 2, 3, 4]

Chapter 2

Theory

Chapter 3

The CMS Experiment

Chapter 4

Analysis

Chapter 5

Summary

Appendix A

Appendix Title

A.1 Toy profile likelihood

This section presents an introduction to the profile likelihood method (used for the statistical results of this thesis) with a concrete toy example. The toy example consists of one signal, one background, and one shape-based nuisance parameter. Complete details of the method are in Reference [5].

A.1.1 Terminology

Bayes' theorem can be stated in a more applicable way as

$$p(k|D) = \frac{p(D|k)p(k)}{p(D)} \tag{A.1}$$

where $p(A|B)$ is the (conditional) probability of A given B. Here, k represents a model, which consists of a set of parameters (parameters of interest, as well as annoying/nuisance parameters). Typically, the single parameter of interest is a signal "strength" μ (how much signal is there actually, compared to what is nominally in the simulation – $\mu =$

$\mu_{\rm obs}/\mu_{\rm SM}$), and a set of (many) nuisance parameters technically specified as a vector $\vec{\theta}$, but often just θ. The posterior probability distribution function (pdf), $p(k|D)$ is the distribution of parameters given the observed data D. The **likelihood** $p(D|k)$ gives the likelihood of getting some data D given a particular model encoded in k. $p(k)$ is a prior distribution of models, often taken to be "flat" in μ. Lastly, the overall constant $p(D)$ is ignorable when dealing with differences in likelihoods.

Analysis observables are typically binned into many regions, and compared with an observed count (data). Integral data event counts (N) obey Poisson statistics, where λ governs the underlying rate of a process: $p(N|\lambda) = e^{-\lambda}\frac{\lambda^N}{N!}$. (Although, when statistics are large, Gaussian approximations can be made in order to simplify computations.) For independent bins, the likelihood is a product over each of the bins.

This sets the stage for the toy example, which clarifies the meaning of the word **profile**.

A.1.2 Toy example

Figure A.1 shows a toy distribution of jet multiplicity for background, signal, and data. The background component also contains a systematic uncertainty band corresponding to a shape nuisance parameter. This shape nuisance parameter prefers to increase yields at higher number of jets.

As an intermediate goal to most statistical results/interpretations of the data, we want to compute the likelihood which will be a 2D function of the signal strength μ and the value of the systematic variation θ. $\theta = 0$ will give the normal background yields, while $\theta = -1$ and $\theta = 1$ will give the 1σ down and up variations, respectively. From above, $\theta = 1$ will absorb the signal, as it increases the background yield at high number of jets.

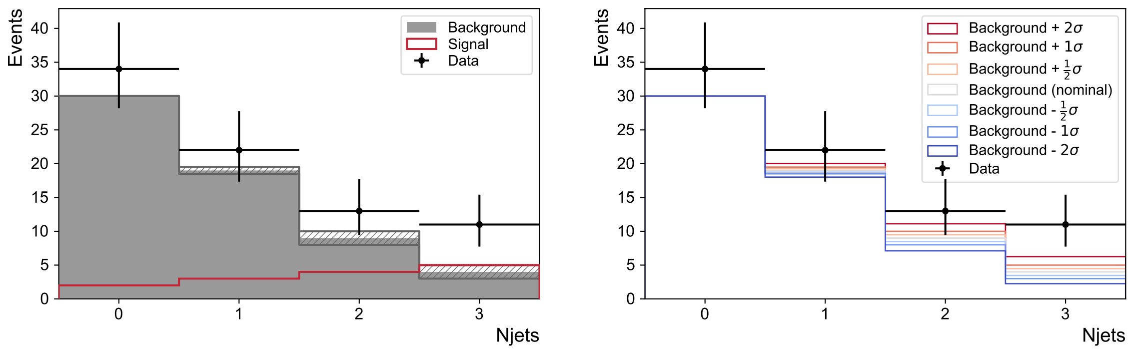

Figure A.1: Toy distribution (left) and single shape nuisance variation on background component (right)

The likelihood function is the binned probability to get the data from a b (background-only) or $s+b$ (background and signal) poisson distribution, accounting for the probability of nuisances θ with the pdf $p(\theta)$:

$$\mathcal{L}(\text{data}|\mu,\theta) = \prod_{j\in\text{bins}} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)} p(\theta) \tag{A.2}$$

here n_j represents the data count in bin j, b_j the background count, and s_j the signal count. Technically, b_j is a function of θ; that is, $b_j(\theta)$ depends on the value of the systematic variation where $b_j(0)$ is the nominal background yield.

The goal is find maxima in the likelihood scan. However, numbers can get very large given the factorial and the exponential. Since the ln function is monotonically increasing, to make things numerically tractable and simpler, we take the (negative) log since minimization problems are easier than maximization.

$$-\ln\mathcal{L} = -\sum_{j=0}^{3} \left[n_j \ln(\mu s_j + b_j) - \ln(n_j!) - (\mu s_j + b_j) + \ln(p(\theta))\right] \tag{A.3}$$

Even though it won't matter in the end (since we ultimately care about relative differences in the likelihood) note that $\ln(n_j!)$ is just $\sum_{i=0}^{3} \ln(i)$ and it is independent of

μ and θ, so it can usually be pre-computed.

We can then calculate likelihood values over the μ-θ plane. Figure A.2 shows this two dimensional likelihood scan as a function of θ and μ, with contours overlaid, as well as the global maximum log likelihood (the minimum *negative* log likelihood). The plot also shows the maximum likelihood as a function of μ with a red line. Equivalently, this gives the nuisance parameter value that maximizes the likelihood for a fixed μ.

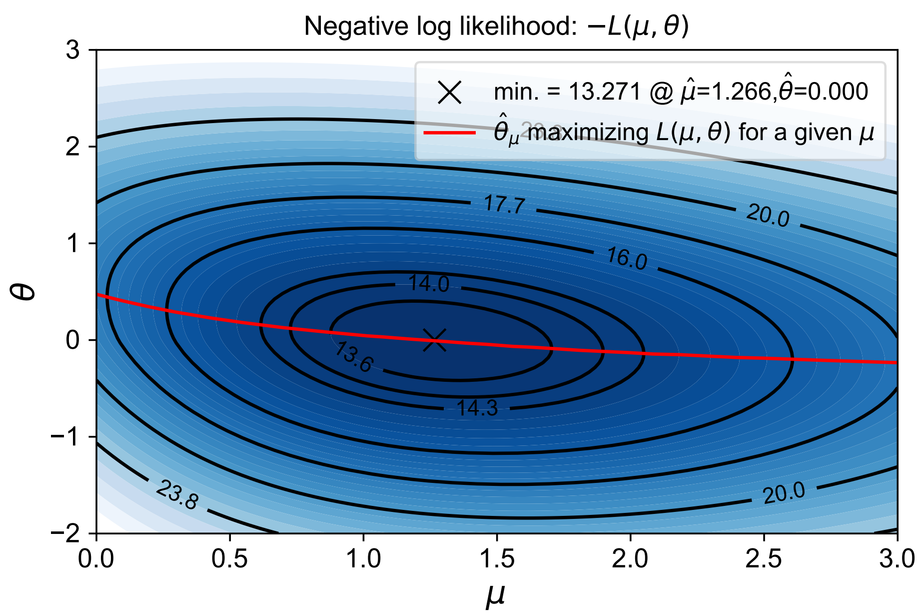

Figure A.2: Two dimensional likelihood scan as a function of θ and μ. The red line shows the maximum likelihood for a given μ.

From the two dimensional scan, we see the maximum global likelihood occurs when the signal strength parameter μ is 1.27. This roughly makes sense given the histogram templates which had 5 signal, 4 background, and 11 observed events in the last bin. This bin dominates the result due to the strong signal presence. With the fitted signal strength, $5 \times 1.27 + 4 \approx 11$.

For lower values of μ, the best likelihood values occur for increasing θ. This can be understood as a compensatory effect: when signal yields decrease, in order to have background+signal match data, we need to "borrow" some yields from the background nuisance, which pulls up yields at higher number of jets. Of course, there is a likelihood penalty to this due to the $p(\theta)$ term.

Figure A.3 shows two curves of $-\ln \mathcal{L}(\mu, \theta)$ for two slices of the scan ($\mu = \hat{\mu}$, and $\mu = 0$).

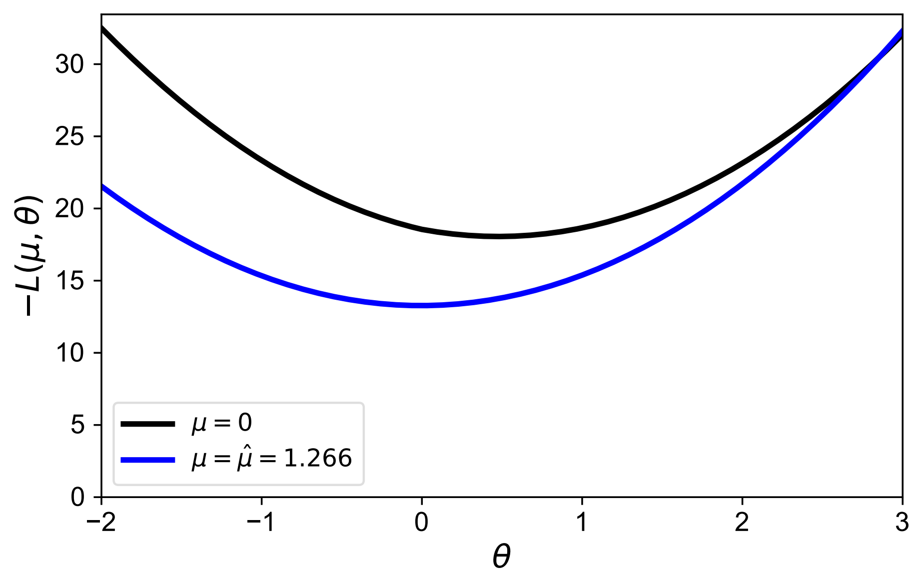

Figure A.3: Negative log likelihood as a function of θ for μ fixed to 0 and $\hat{\mu}$

Let's define the LHC profiled test statistic q_μ as

$$q_\mu = q(\mu) = -2 \ln \frac{\mathcal{L}(\mu, \hat{\theta}_\mu)}{\mathcal{L}(\hat{\mu}, \hat{\theta})} \tag{A.4}$$

where $\hat{\theta}_\mu$ is the θ that maximizes $\mathcal{L}$ for a particular μ. The pair $\hat{\mu}$ and $\hat{\theta}$ are the ones that globally maximize the likelihood. Thus, the denominator is a single number (the global extremum of the 2-dimensional likelihood scan). The numerator is a 1-dimensional function which gives the maximum likelihood as a function of μ. We have "profiled out" the nuisance parameter θ. Bear in mind the additional minus sign gymnastics we must do when switching between log likelihood and negative log likelihood:

$$q_\mu = 2\left[\mathrm{NLL}(\mu, \hat{\theta}_\mu) - \mathrm{NLL}(\hat{\mu}, \hat{\theta})\right] \tag{A.5}$$

Or, q_μ is calculated as twice the difference of the negative log likelihood values along the red curve from the two dimensional scan, and the global minimum negative log likelihood from the same scan. Figure A.4 shows a plot of $q_\mu = q(\mu)$

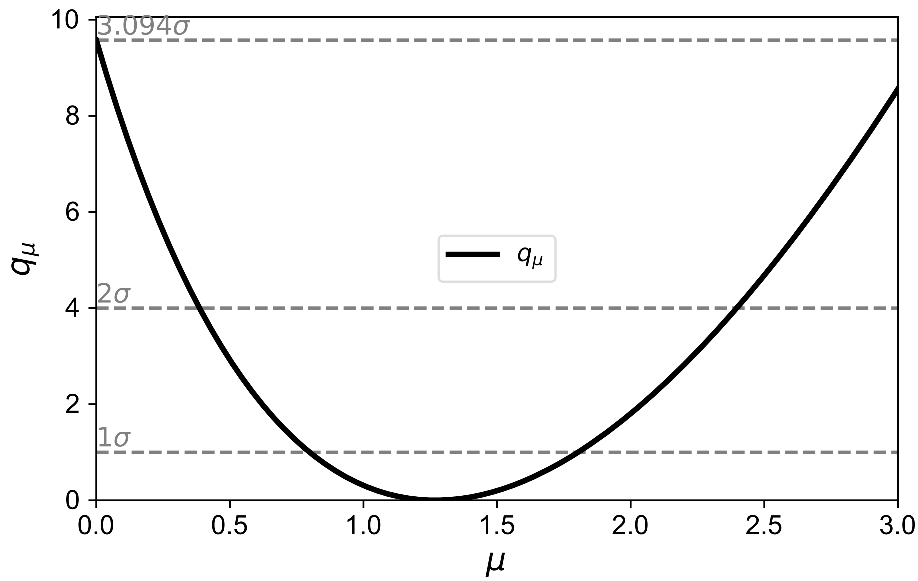

Figure A.4: LHC test statistic as a function of μ

Before proceeding, we now work with the asymptotic approximation in the case of large background. That is, the test statistic q_μ with data containing signal strength μ' follows a Gaussian distribution

$$q_\mu = \frac{(\mu - \hat{\mu})^2}{\sigma^2} + \mathcal{O}(1/\sqrt{N}) \tag{A.6}$$

where $\hat{\mu} \sim \mathcal{N}(\mu', \sigma)$, N is the sample size, and σ is the standard deviation of $\hat{\mu}$ which is extracted from the covariance matrix of nuisances θ. Ultimately, one finds that with this approximation, the test statistic follows a noncentral chi-square distribution for one degree of freedom. See the continuing discussion in Ref. [5]. This is exploited for huge computational gains, especially for SUSY scans which consist of hundreds to a few thousand mass points (number of hypotheses to test with the data and background predictions).

Returning to the curve in Figure A.4, we can identify the significance of the the data from one point. That is, how statistically significant would it be if just the background fluctuated to look like background+signal:

$$Z_{\text{obs}} = \sqrt{q_0} \tag{A.7}$$

And more generally, we can compute 1σ, 2σ, 3σ, etc. confidence bands on the fitted value of $\hat{\mu}$ by drawing lines at $q_\mu = 1, 4, 9, \ldots$. Why these values in particular? We turn to a normal distribution with a pdf of

$$f(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{A.8}$$

Identifying this as a likelihood pdf and ignoring all constant/offset terms independent of σ and μ, then

$$q_\mu = -2\ln(f) = 2\frac{(x-\mu)^2}{2\sigma^2} \tag{A.9}$$

We want $x \to \mu + k \cdot \sigma$ and k is 1,2,3,..., so

$$q_\mu \to \frac{(k\sigma)^2}{\sigma^2} = k^2 \tag{A.10}$$

In this toy example, based on Figure A.4, the observed significance of the result is 3.09σ and the fitted signal strength, μ, is approximately $1.3^{+0.4}_{-0.5}$. For "expected" quantities, one can consider the **asimov** dataset, which is constructed with background (or signal+background) expectation, incorporating fluctuations due to statistics as well as systematics encoded by the nuisance parameters. This is done to understand analysis results, for example, without looking at the real data.

Bibliography

[1] CMS Collaboration, *Search for standard model production of four top quarks in final states with same-sign and multiple leptons in proton-proton collisions at* $\sqrt{s} = 13$ TeV, .

[2] CMS Collaboration, *Search for physics beyond the standard model in events with two same-sign leptons or at least three leptons and jets in proton-proton collisions at* $\sqrt{s} = 13$ TeV., .

[3] CMS Collaboration, *Search for standard model production of four top quarks with same-sign and multilepton final states in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, *Eur. Phys. J.* **C78** (2018), no. 2 140, [arXiv:1710.1061].

[4] CMS Collaboration, *Search for physics beyond the standard model in events with two leptons of same sign, missing transverse momentum, and jets in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, *Eur. Phys. J.* **C77** (2017), no. 9 578, [arXiv:1704.0732].

[5] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J.* **C71** (2011) 1554, [arXiv:1007.1727]. [Erratum: Eur. Phys. J.C73,2501(2013)].