RUN-2 Physics Activities at CMS/LHC

6th Annual INPP Meeting
15 November 2019

G. Daskalakis
The LHC plan

Run 2
$<\mu> \sim 37$
$L \sim 2 \times 10^{34} / \text{cm}^2/\text{sec}$

Run 3
$<\mu> \sim 55$
$L \sim 2 \times 10^{34} / \text{cm}^2/\text{sec}$

Run 4-5
$<\mu> \sim 140-200$
$L \sim 7.5 \times 10^{34} / \text{cm}^2/\text{sec}$
End of RUN-2

CMS Integrated Luminosity, pp, $\sqrt{s} = 13$ TeV

Data included from 2015-06-03 08:41 to 2018-10-26 08:23 UTC

- LHC Delivered: 163.02 fb$^{-1}$
- CMS Recorded: 150.53 fb$^{-1}$

2018
68.2 fb$^{-1}$

2017
49.8 fb$^{-1}$

2016
40.8 fb$^{-1}$

2015
4.2 fb$^{-1}$

Date

Total Integrated Luminosity (fb$^{-1}$)
931 papers from CMS

931 collider data papers submitted as of 2019-11-12
**EXOTICA:** \( Z' \to e^+e^- , \mu^+\mu^- \) Searches

**The team:**
Georgios Daskalakis  
IIHE-ULB, RAL,  
University of Notre Dame,  
Purdue University

---

**final paper**

data: 2016-2017-2018  
13 TeV, 140 fb\(^{-1}\)

Analysis is finished and approved.

Still few things to be checked before the final publication.

There is a long paper (legacy) to be written.  
We might include non-resonance searches.
EXOTICA: $Z' \rightarrow e^+e^-, \mu^+\mu^-$ Searches

Analysis Strategy

- Understand ID efficiency in data and MC from Z peak to high $E_T$
- Normalize cross sections to Z peak
  All ET independent effects are included
- Develop a simple, ET independent ID
- Maintain a robust high efficient trigger
  Keep low $E_T$ threshold
- Understand mass scale and resolution
- Measure and understand backgrounds
- Set limit on the cross section of various new physics models
- Parameterization of signal and background shapes
EXOTICA: $Z'\rightarrow e^+e^-, \mu^+\mu^-$ Searches

 CMS Preliminary $137 \text{ fb}^{-1} (13 \text{ TeV})$

Data
- $\gamma^*/Z \rightarrow e^+e^-$
- $t\bar{t}, tW, WW, WZ, ZZ, \tau\tau$
- Jets

 CMS Preliminary $140 \text{ fb}^{-1} (13 \text{ TeV})$

Data
- $\gamma^*/Z \rightarrow \mu^+\mu^-$
- $t\bar{t}, tW, WW, WZ, ZZ, \tau\tau$
- Jets
- Total MC (NR)
EXOTICA: $Z'\rightarrow e^+e^-, \mu^+\mu^-$ Searches

Limits on the masses of spin-1 $Z'_{SSM}$ and $Z'_\Psi$ bosons, assuming a signal width of 0.6% (3.0%) of the resonance mass for $Z'_\Psi$ ($Z'_{SSM}$).

<table>
<thead>
<tr>
<th>Channel</th>
<th>$Z'_{SSM}$</th>
<th>$Z'_\Psi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e,e$</td>
<td>4.72</td>
<td>4.11</td>
</tr>
<tr>
<td>$\mu^+\mu^-$</td>
<td>4.89</td>
<td>4.29</td>
</tr>
<tr>
<td>$e,e + \mu^+\mu^-$</td>
<td>5.15</td>
<td>4.56</td>
</tr>
</tbody>
</table>
Gauge-Mediated SUSY with $2\gamma$ & MET

**data**: 2016
13 TeV, 36 fb$^{-1}$

**The team**: A. Kyriakis, I. Topsis-Giotis, G. Paspalaki (student)
Florida State University
University of Notre Dame

**Signature**: Characteristic events with jets, two photons and large $E_T^{\text{miss}}$

- Final states in pp collisions with high-$E_T$ photons and significant $E_T^{\text{miss}}$ emerge naturally from a variety of new physics scenaria, particularly in models of supersymmetry (SUSY) broken via gauge mediation and including a stable, weakly interacting lightest supersymmetric particle (LSP)

- Models with general gauge mediation (GGM) can have a wide range of features, but typically entail a gravitino LSP and a next-to-lightest supersymmetric particle (NLSP) commonly taken to be a neutralino or a stau.
Gauge-Mediated SUSY with $2\gamma$ & MET

Background estimation methods:
- based on control samples in data.
Gauge-Mediated SUSY with 2γ & MET

Limits on the gluino/squarks pair production cross section are set. Using NLO + NLL cross section calculations → constrain the masses of gluinos, squarks and neutralinos in the GGM framework.

The team:
Georgios Anagnostou, Georgios Daskalakis, (student?)

Searching simultaneously for both a heavy top partner $T'$ and a new gauge boson $W'$.

Method ingredients:
*Input*: MET & visible particles, the topology under investigation
*Output*: the $T'$ & $W'$ masses

The method is documented in a paper under preparation (G. Anagnostou) using a Delphes production of 200M events.
B2G : T’T’ Searches

2D mass reconstruction - simulation
B2G : T’T’ Searches

Data 2016 – control region

The analysis is almost ready for 2016 dataset.

The search will be extended to 2017/2018 datasets.
**HIGGS : ttH (H→bb)**

**The idea:**
Study of the di-leptonic ttH( H→bb) channel reconstructing the Higgs mass and using a **data-driven background** prediction method.

**Higgs mass reconstruction**: is performed by simultaneously solving analytically the ttH dileptonic decay system while scanning the $M_t$ vs $M_W$ mass plane searching for solutions. (details in G. Anagnostou talk)

**Data-driven background**: From events with exactly 2 b-tagged jets (ttbar enriched + Higgs contamination small) predict the **shape** and **normalization** of the $m_{bb}$ distribution of events with exactly 3 / 4 b-tagged jets by applying probability weights.

**The team**: Charis Kleio Koraka (student) Niki Saoulidou (UoA), Georgios Anagnostou, Georgios Daskalakis
**HIGGS : ttH (H→bb)**

**Higgs mass reconstruction** : is performed by simultaneously solving analytically the ttH dileptonic decay system while scanning the $M_t$ vs $M_W$ mass plane searching for solutions.

**For each event:**
- Two out of the four b-jets are assigned to the Higgs boson
- The remaining two b-jets, are assigned to the t/anti-t decays.
- The kinematic mass reconstruction of the tt-bar system is performed
- For each solution, one reconstructs the event and calculates $x_1, x_2$ (proton momentum fractions) of the ttH system:

\[
x_1 = \frac{E_{t\bar{t}H} + p_{t\bar{t}H_z}}{2} \quad x_2 = \frac{E_{t\bar{t}H} - p_{t\bar{t}H_z}}{2}
\]

The b-jet, lepton, $m_{t_{top}}$ and $m_W$ combination is chosen based on the **highest PDF weight**.

- The weights are evaluated using the LHAPDF-6.1.2 version & the pdf set CT10.
  - It returns the product $xf(x,Q)$, where $Q$ is the energy scale, $x$ is the parton momentum fraction and $f(x,Q)$ is the parton distribution function (here we consider gg interactions).
  - The weight of each solution corresponds to the ratio $w = xf(x,Q)/x$.
- $Q = 235 \ [2(m_{top} + m_{Higgs})/2]$. 
HIGGS : ttH (H→bb)

Data-driven background

Motivation
● Predict the main background in a data-driven way

Assumptions
● Probability of tagging a jet independent of the rest of the jets in the event
● The variables used to parameterize tagging efficiency ($p_T$, $|\eta|$, flavor) are sufficient to describe the b-tagging dependencies.

Implementation
● Jet b-tagging not performed by a direct cut on the discriminant.
● Rather, probability of each jet being b-tagged is estimated using parameterized efficiencies: $\epsilon(p_T, |\eta|, \text{flavor})$

Result
● Predict the shape and normalization of the distribution of interest for a certain b-tag multiplicity from events with lower btag multiplicities, applying probability weights.

From events with exactly 2 b-tagged jets (ttbar enriched + Higgs contamination small) predict the shape and normalization of the $m_{bb}$ distribution of events with exactly 3 / 4 b-tagged jets by applying probability weights.
HIGGS: \( \text{ttH (H} \rightarrow \text{bb)} \)

From the \( N_{b\text{-tags}} = 2 \) distribution can we predict the \textbf{shape} and \textbf{normalization} of the \( N_{b\text{-tags}} = 4 \) distribution?

\textbf{TRF} (Tagging Rate Function) method, used already by D0 & ATLAS.

First we need to estimate the probability of each jet to be \( b \)-tagged. Parameterize the probability as:

\[
\varepsilon(p_T, |\eta|, \text{flavor})
\]

Probability of \( b \)-tagging a true \( b \)jet, \( c \)-jet or light jet parametrized as a function of \( p_T \) and \( \eta \).

Jets are considered \( b \)tagged if \( b\text{tag} > 0.8 \).
HIGGS: \( ttH \ (H\rightarrow bb) \)

For a given requirement of b-tagged jets (\( M \)) in the event, all the possible combinations of labeling \( M \) jets as “tagged” out of the \( N \) jets of the event are considered.

For each combination, a weight is assigned to each jet corresponding to the probability of it being b-tagged using the efficiency \( \varepsilon(p_T, \eta, \text{flavor}) \).

Multiplying all jet weights and adding for all combinations yields the probability for that event to contain the selected number \( M \) of b-tagged jets.

![Diagram showing CMS Simulation results for b-tagged jet multiplicities and comparison between Direct Cut and TRF prediction]

We can observe a perfect closure between the direct cut and the TRF prediction!

Everything works well at simulation level. What about a more realistic scenario?
**HIGGS : ttH (H→bb)**

**Using truth information**
- Use all jets and all events to predict the number of N btags & the distribution of interest.
- True b-jets identified using hadron flavor.
- Produce & use tagging probability maps for each hadron flavor.

**Using reconstructed information (Both MC & Data)**
- Use the 4 leading jets of the events in the control region (exactly 2 btagged jets) to predict the distribution of interest.
- As “true” b-jets are considered the 2 btagged ones in the control region.
- After removing the “true” b-jets per event, produce the approximate c+light jets b-tagging probability map from the remaining jets.

**A very rough idea on significances ...**

<table>
<thead>
<tr>
<th>(%) systematic error of Bkg prediction</th>
<th>Lumi = 140 fb⁻¹</th>
<th>Lumi = 300 fb⁻¹</th>
<th>Lumi = 3000 fb⁻¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIG-17-026 Post Fit</td>
<td>1.7</td>
<td>2.1</td>
<td>2.9</td>
</tr>
<tr>
<td>Our approach (2% systematic)</td>
<td>1.2</td>
<td>1.7</td>
<td>4.8</td>
</tr>
</tbody>
</table>

**Table:**

<table>
<thead>
<tr>
<th>Event yield</th>
<th>TRF prediction</th>
<th>Direct Cut (b-tags = 3)</th>
<th>Ratio (TRF/DC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3522</td>
<td>3681</td>
<td>0.97 ± 0.02</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Event yield</th>
<th>TRF prediction</th>
<th>Direct Cut (b-tags = 4)</th>
<th>Ratio (TRF/DC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>120</td>
<td>133</td>
<td>0.9 ± 0.1</td>
<td></td>
</tr>
</tbody>
</table>

**Work is going on !**
An Analysis Note with all the details of our approach is under preparation.
TOP : W helicities from ttbar events

Motivation:

i) The measurement is sensitive to the Wtb vertex structure; new physics from anomalous Wtb couplings

\[ F_0 = 0.681 \pm 0.012 \text{ (stat)} \pm 0.023 \text{ (syst)}, \]
\[ F_L = 0.323 \pm 0.008 \text{ (stat)} \pm 0.014 \text{ (syst)}, \]
\[ F_R = -0.004 \pm 0.005 \text{ (stat)} \pm 0.014 \text{ (syst)} \]

ii) New methodology to improve systematic uncertainties w.r.t. 7 & 8 TeV analyses

The team:
M. Soares, J. Brochero (CIEMAT)
A. Stakia (PhD student),
G. Anagnostou,
G. Daskalakis

Analysis quite advanced. Shown at TOP meetings.

The CMS and ATLAS Collaborations

Data: full RUN-2
13 TeV, 140 fb\(^{-1}\)
TOP : W helicities from ttbar events

\[ \frac{1}{\Gamma} \frac{d\Gamma}{d\cos\theta} = \frac{3}{8} (1 - \cos\theta)^2 F_L + \frac{3}{8} (1 + \cos\theta)^2 F_R + \frac{3}{4} \sin^2 \theta F_0, \quad \theta \equiv \Theta^* \]

\[ F_0 = 0.687 \pm 0.005, \quad F_L = 0.311 \pm 0.005, \quad F_R = 0.0017 \pm 0.0001 \]

(Phys. Rev. D 81 (2010) 111503), \( m_t = 172.8 \pm 1.3 \) GeV

\[ \cos (\Theta^*) : \]

in the t-quark rest frame, the angle between the down-type fermion momentum in the W rest frame and the W momentum in the top-quark rest frame.

Can we do better by changing the ‘sensitive variable’?

- Based on \( \cos(\Theta^*) \) → Strong discriminant power

- \( \cos(\Theta^*) \) needs the reconstruction of the top process (tt̅ or single top)
- tt̅ kinematic fit introduces a dependency of top mass.

We propose a different approach to extract the W-helicity

1. \( \Delta \phi (\ell, \text{jet}) \)
2. \( M_{\ell b} \)
\[ \cos \theta^* = 2 \left( \frac{M_{lb}^2}{m_t^2 - m_W^2} \right) - 1 \]

Non uniform resolution versus \( \cos(\theta^*) \)

35.9 fb\(^{-1}\) (13 TeV)

Data
Top pairs
Single top + ttV
QCD Multijet
W/\(Z^+\)/\(Z^-\)
Diboson
Uncertainty

Lepton & Bjets matched to Gen Level DR<0.01
TOP : W helicities from ttbar events

1. Find optimal variable to extract the W-helicity
   - Discriminant power between $F_0$, $F_L$ and $F_R$
   - Good resolution
   - Different shape for backgrounds
2. $\Delta \Phi(\ell, \text{jet})$

1. Selection of the correct lepton-jet couple
   - Avoid any SM prior
   - Should work in $tW/t$-channel/$t\bar{t}$ system.
2. The solution: BDT

\[ |\Delta\phi(\text{lep, jet})| \text{ templates} \]

- $F_0 = 1$
- $F_L = 1$
- $F_R = 1$
TOP : W helicities from ttbar events

- We found a very good variable/configuration to extract the W-helicity without kinematic fits: $\Delta \Phi(\ell, \text{jet})$
- Three different regions have been defined in order to optimize the fit (mainly to constrain syst. unc.)
- Several test of the fit performance done. Results show very good precision so far.
- AN already created for this analysis. We are collecting all the details there.

### Ongoing

1. Further BDT optimizations: gain in 5% signal efficiency for the same number of correct lepton-b matching
2. Measurement of top mass dependency on $W$ helicity fractions
3. Theory uncertainties

### Starting

1. Data-driven method for QCD background
2. First look to real data
3. Re-processing of our ntuples to the latest 2016 recommendations

... more details in Anna’s presentation ..
Contributions to almost all Physics Groups of CMS

Run-2

13 TeV
140 fb$^{-1}$

Run-3

14 TeV
300 fb$^{-1}$
What next?

Pursuing parallel activities in three areas:
1. harvest of run 2 results
2. preparation for data taking & analysis in Run 3
3. preparation for HL-LHC


- Early analyses w/ full Run 2 data
- Physics planning Run 3
- Full Run 2 analyses for summer / fall ‘19
- Mainly measurements using best possible reco & calib (ultra-legacy)
- Preparation Run 3