Searches for new phenomena at the LHC: backgrounds and optimization

Lukas Vanelderen
Universiteit Gent

Bigger picture
W/Z + jets studies
SUSY with b-jets

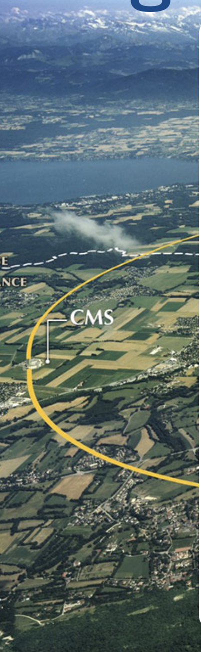
Bigger picture

W/Z + jets studies

SUSY with b-jets

Large Hadron Collider

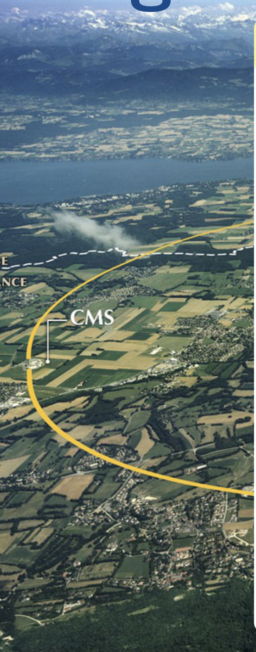
Large Hadron Collider



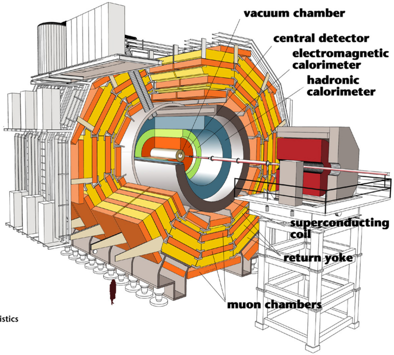
- proton-proton, 4 collision points
- $\sqrt{s} = 7$ TeV
- general purpose detectors:
  ATLAS, CMS
- dedicated to specific physics:
  LHCb, ALICE
- a rich physics program:
  - Higgs
  - new phenomena:
    SUSY, W' and Z',
    extra dimensions, 4th generation, etc.
  - precision measurements of SM*

*Standard Model

Large Hadron Collider



CMS



Detector characteristics

Width: 22m
Diameter: 15m
Weight: 14'500t

- 2010: 36 pb^{-1} collected
- 2011: 5.2 fb^{-1} collected

similar numbers for ATLAS

Searches for new phenomena at the LHC

2010, 2011:
Many searches, no discoveries

Desperation is not justified yet
Still much work before making the following statement

"we can say with a high level of confidence that the LHC data does not contain hints for new phenomena"

We are making great progress...

Tremendous progress

Precision measurements of the SM

- new constraint SM parameters
- indirect constraints on new phenomena
- improvements/better tunes of MC simulation

Better calculations

- NLO is taking over

Improved understanding of detector

=> better background estimates for searches

Tremendous progress

Precision measurements of the SM

- new constraint SM paramete
- indirect constraints on new p
- improvements/better tunes of MC simulation

more to come:
W/Z + jets studies

Better calculations

- NLO is taking over

Improved understanding of detector

=> **better background estimates for searches**

Tremendous progress

Higgs searches
Searches for new phenomena
- exclusion regions grow

a more general interpretation of searches
- slowly moving away from cMSSM
- use of more general models (e.g. pMSSM)
- use of simplified models

=> **better understanding of potential new phenomena**

Let's put the pieces together

Combine all available information about SM processes and new phenomena to interpret the results and design the next generation of searches*

*if it still doesn't work:
- wait for more data
- wait for a new machine
- try turning it off an on again

Let's put the pieces together

Combine all available information about SM processes and new phenomena to interpret the results and design the next generation of searches*

more to come:
optimization of
a SUSY search with *b*-jets
(novel Bayesian technique)

*if it still doesn't work:
- wait for more data
- wait for a new machine
- try turning it off an on again

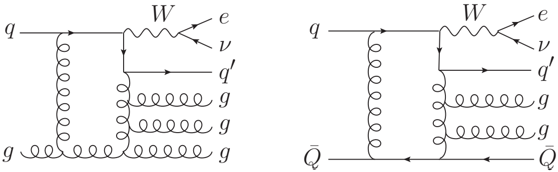
Bigger picture
W/Z + jets studies
SUSY with b-jets

An early W/Z + jets study with CMS:

"Jet Production Rates in Association with W and Z Bosons in pp Collisions at $\sqrt{s} = 7$ TeV"

CMS-EWK-10-012
arXiv:1110.3226 [hep-ex]
submitted to JHEP

Why?



- important backgrounds to many
  - Higgs searches
  - searches for new phenomena
  - top quark studies
- validation/tuning of (N)LO matrix element + parton shower MC predictions
- probes proton structure

What?

- 2010 proton-proton data
- 36 pb^{-1}, $\sqrt{s} = 7$ TeV
- ratios of $\sigma(W/Z + \geq n \text{ jets})$, $n = 1, 2, 3, 4$
- corrected for detector effects
- compared to Leading Order calculations

What?

4 Channels, 4 independent measurements:

- $W(\to e\nu)$
- $W(\to \mu\nu)$
- $Z(\to ee)$
- $Z(\to \mu\mu)$

I will present all results,

What?

4 Channels, 4 independent measurements:

- $W(\to e\nu)$
- $W(\to \mu\nu)$
- $Z(\to ee)$
- $Z(\to \mu\mu)$

I will present all results,
but only the analysis for $W(\to \mu\nu)$

(Note: not all figures/tables are officially approved)

W/Z + jets studies

$W(\rightarrow \mu\nu)$ analysis

data, reco and selection

signal extraction

unfolding

systematics

general results

data, reco and selection

- **simple muon triggers**
  - $p_T > 15$ GeV

data, reco and selection

- simple muon triggers
- event reconstruction with Particle Flow (PF)

data, reco and selection

- **simple muon triggers**
- **event reconstruction with Particle Flow (PF)**
- **selection of W candidates**

  leading muon
  - $p_T > 20$ GeV, $|\eta| < 2.1$,
  - loose identification, isolation, vertex requirements
  - match with trigger object

data, reco and selection

- **simple muon triggers**
- **event reconstruction with Particle Flow (PF)**
- **selection of W candidates**

  leading muon

  - $p_T > 20$ GeV, $|\eta| < 2.1$
  - loose identification, isola ents
  - match with trigger objec

eff measured with Tag and Probe (TnP) in $Z(\rightarrow \mu\mu)$ data vs.
- muon p_T and η
- jet multiplicity

data, reco and selection

- **simple muon triggers**
- **event reconstruction with Particle Flow (PF)**
- **selection of W candidates**
- Z-**veto**

  no second muon

  - $p_T > 10$ GeV, $|\eta| < 2.4$
  - well-reconstructed ("global muon")
  - $60 < M_{\mu\mu} < 120$ GeV

data, reco and selection

- **simple muon triggers**
- **event reconstruction with Particle Flow (PF)**
- **selection of W candidates**
- Z**-veto**
- **neutrino energy reconstruction**

  missing transverse energy

$$\vec{E}_T^{miss} = - \sum_{\text{all part}} \vec{p}_{T,i}$$

data, reco and selection

- **simple muon triggers**
- **event reconstruction with Particle Flow (PF)**
- **selection of W candidates**
- Z**-veto**
- **neutrino energy reconstruction**
- W **boson mass reconstruction**

transverse mass

$$m_T = \sqrt{2 p_T^{\mu} E_T^{miss} \left[1 - cos(\Delta\Phi(\mu, E_T^{miss})\right]}$$

$m_T > 20$ GeV

eff measured in simulation, cross checked with TnP in $Z(\rightarrow \mu\mu)$ data

data, reco and selection

- **simple muon triggers**
- **event reconstruction with Particle Flow (PF)**
- **selection of W candidates**
- Z**-veto**
- **neutrino energy reconstruction**
- W **boson mass reconstruction**
- **jet definition:**
    - cluster PF-particles
    - leading muon not included
    - anti-k_T, R=0.5
    - calibration (L1 FastJet, L2, L3)
    - loose identification criteria

data, reco and selection

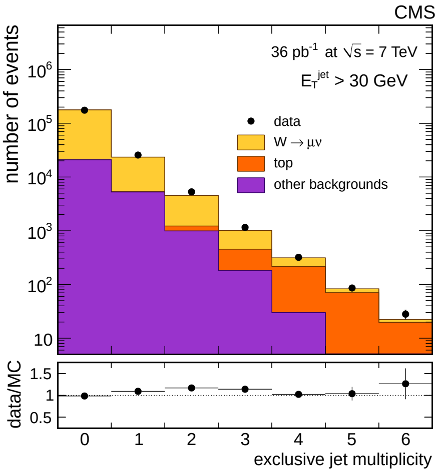
- **simple muon triggers**
- **event reconstruction with Particle Flow (PF)**
- **selection of W candidates**
- Z**-veto**
- **neutrino energy reconstruction**
- W **boson mass reconstruction**
- **jet definition:**
- b**-tagging**
  - "Track Counting High Efficiency" algorithm
  - "Medium working point"
    (mistag eff. $\sim$1%, b-tag eff. $\sim$50%)

data, reco and selection

Jet multiplicity

Steep !

A lot of top !



W/Z + jets studies

$W(\to \mu\nu)$ analysis

- data, reco and selection
- signal extraction
- unfolding
- systematics

general results

Signal extraction

assign data to “jet bins”

“exclusive jet bins” for 0 ... 3 jets

“inclusive jet bin” for 4 jets

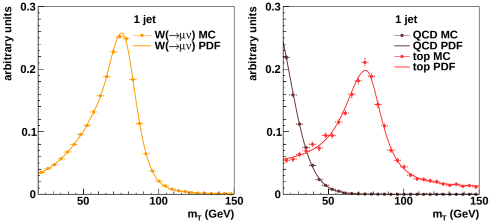
Determine $W(\rightarrow \mu\nu)$ per jet bin with Maximum Likelihood fit

- unbinned
- functional forms
- calibrated mainly on data
- control samples to test assumptions
- weight events $1/\epsilon$ (TnP)

Signal extraction

$W(\to \mu\nu)$ vs. QCD:

transverse mass, m_T



Shapes for fit:

Cruijff functions

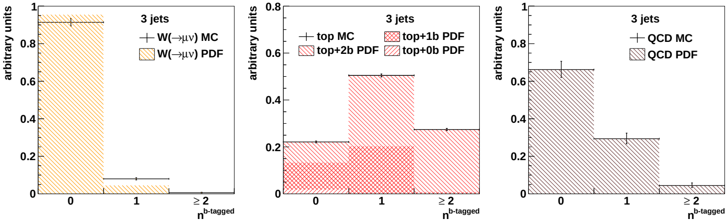
most important shape parameters floated

Similar for other jet bins

Signal extraction

$W(\to \mu\nu)$ vs. top:

b-tagged jet multiplicity

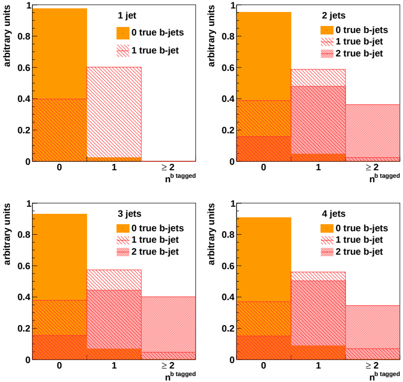


Signal extraction

$W(\rightarrow \mu\nu)$ vs. top:

distribution of
b-tagged jet multiplicity
determined by

- jet multiplicity
- true *b*-jet multiplicity
- average mistag eff
- average *b*-tag eff



$\rightarrow$ determine true *b*-jet multiplicity per jet bin
$\rightarrow$ $W(\rightarrow \mu\nu) \sim$ 0 true *b*-jets
$\rightarrow$ top $\sim$ 1,2 true *b*-jets

Signal extraction

$W(\rightarrow \mu\nu)$ **vs. top:**

Calibration through:

average mistag eff

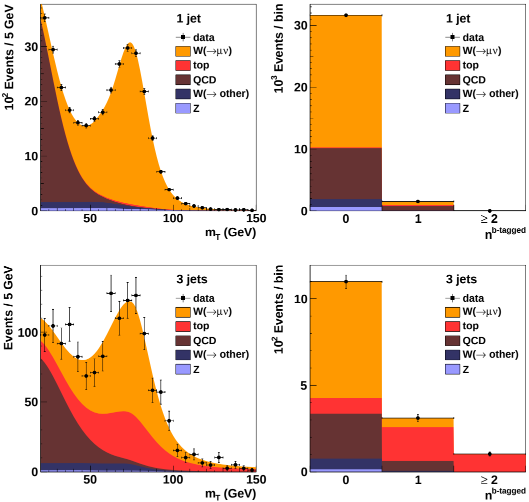
from top and $W(\rightarrow \mu\nu)$ MC simulation,
corrected for differences between data and MC (SF)

average b-tag eff

pure control sample of top quark pairs in data
(double leptonic decays)

Signal extraction

Results



W/Z + jets studies

$W(\rightarrow \mu\nu)$ analysis

- data, reco and selection
- signal extraction
- **unfolding**
- systematics

general results

Unfolding

Jet resolution effects
$\rightarrow$ events migrate between jet bins
unfolding corrects for such detector effects

(SVD unfolding algorithm)

W/Z + jets studies

$W(\rightarrow \mu\nu)$ analysis

- data, reco and selection
- signal extraction
- unfolding
- systematics

general results

Systematics

Jet counting

- jet energy callibration (L2, L3)
- pile-up subtraction

Unfolding

- algorithm
- MC simulation for unfolding matrix
- jet resolution

signal extraction

- b-tag and mistag eff
- m_T shape parameters
- selection efficiency

Systematics

Jet counting

- jet energy callibration (L2,
- pile-up subtraction

effects studied in MC

Unfolding

- algorithm
- MC simulation for unfolding matrix
- jet resolution

repetition of measurement

signal extraction

- b-tag and mistag eff
- m_T shape parameters
- selection efficiency

toy MC

W/Z + jets studies

- $W(\to \mu\nu)$ analysis
  - data, reco and selection
  - signal extraction
  - unfolding
  - systematics
- general results

Results

Results quoted on “particle level”
(= corrected for detector effects)

Signal extraction and unfolding
$\rightarrow \sigma(V + n \text{ jets})$
$\rightarrow \sigma(V + \geq n \text{ jets})$
$\rightarrow$ ratios, the final results

Results

Compared against two sets of predictions:

PYTHIA

- exact leading order for W + 0,1 hard parton
- more hard partons described approximately with parton shower

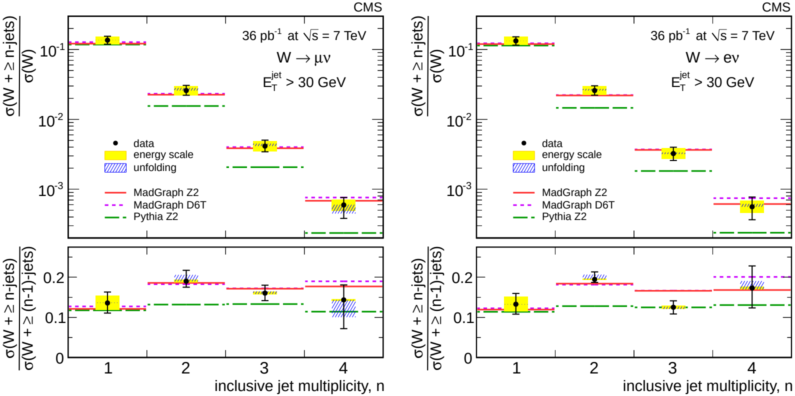
MADGRAPH + PYTHIA

- exact leading order for W + 0,...,5 hard partons
- matched to parton shower with MLM

includes parton shower, hadronization, Underlying Event (UE) (D6T, Z2)

Results

"single ratios" $W(\rightarrow \mu\nu)$ and $W(\rightarrow e\nu)$

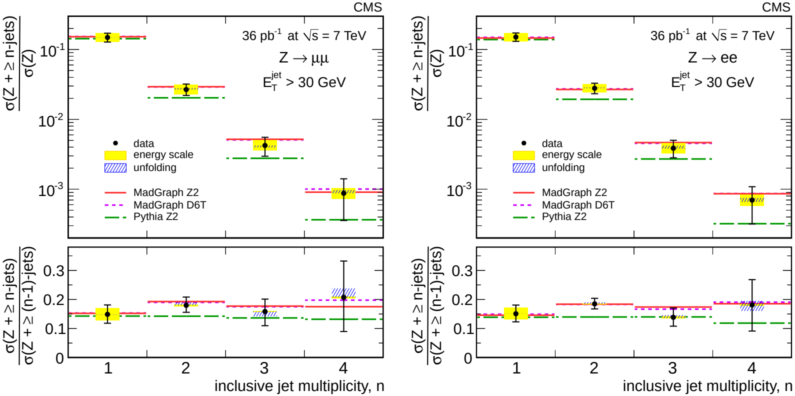


MadGraph + Pythia performs very well

no sensitivity to UE tunes (Z2 and D6T)

Results

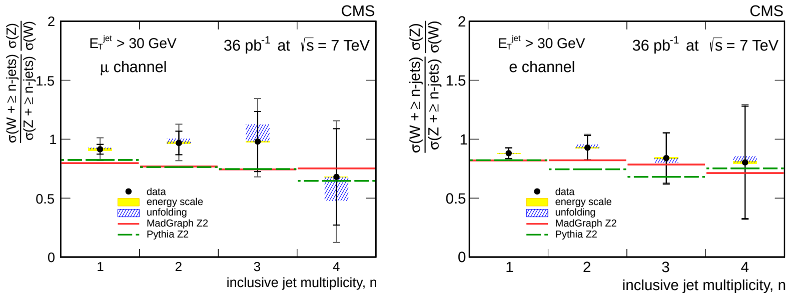
"single ratios" $Z(\to \mu\mu)$ and $Z(\to ee)$



MadGraph + Pythia performs very well
no sensitiviy to UE tunes (Z2 and D6T)

Results

“double ratios”



Results

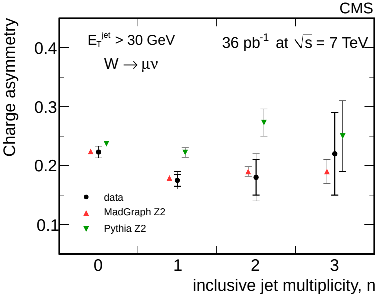
W Charge asymmetry

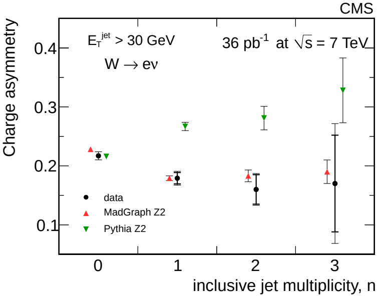
$$A_W^n = \frac{\sigma\left(W^+\right) - \sigma\left(W^-\right)}{\sigma\left(W^+\right) + \sigma\left(W^-\right)}$$

alternative signal extraction: bins of

- jet multiplicity
- W charge

excellent performance of MADGRAPH + PYTHIA





Results

Berends-Giele scaling

Describe cross sections as function of n jets:

$$\alpha + \beta n = \frac{\sigma\left(V+ \geq n \text{ jets}\right)}{\sigma\left(V+ \geq (n+1) \text{ jets}\right)}, \quad n = 1, 2, 3, 4$$

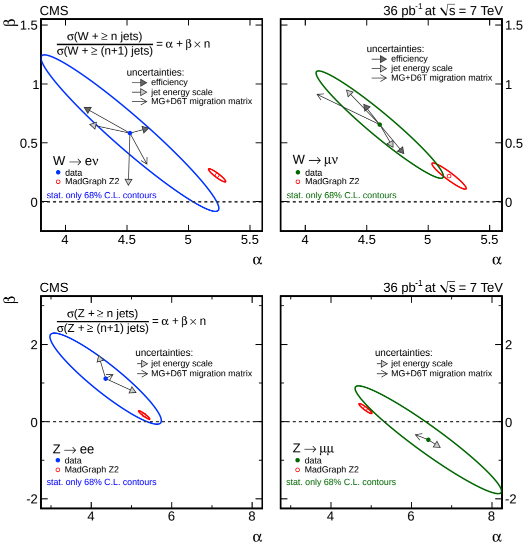
Berends Giele hypothesis: $\beta = 0$

Measure α and β with all-in-one fit:

- all jet bins simultaneously
- eff correction (weighted events)
- unfolding (smearing matrix)
- impose relation above

Results

Berends-Giele scaling



Bigger picture
W/Z + jets studies
SUSY with b-jets

SUSY with *b*-jets

Inclusive searches

What?
searching for the unknown

How?
search for an excess of events w.r.t. SM predictions
in one or more pre-defined "Search Regions"

Challenges?
reasonable definition of the Search Region(s)
SM background predictions

Search region examples

Case 1:
- Search for a specific model
- single search region

- 1 search is the best choice,
- defining search region is "easy",
  e.g. maximize $\frac{S}{\sqrt{S+B}}$

Search region examples

Case 2:
- Search for a model with $\sim$100 free parameters (e.g. MSSM)
- single search region

- for different points of parameter space, different optimal search regions
- what is an overall reasonable search region?
- probably better to have more than 1 search region

Search region examples

Case 3:
- Search for a model with $\sim$100 free parameters (e.g. MSSM)
- multiple search regions

- same problems as in **Case 2**
- and, more regions =
  - more excess from signal
  - look-elsewhere-effect
  - problematic bkg predictions

Search region examples

new physics models
with multiple free parameters
$\times$
complex final states
(many discriminating variables)
$=>$
intuition falls short

SUSY with *b*-jets

- about searches
- optimization of searches
- SUSY with *b*-jets
- outlook

optimization of searches

A novel Bayesian approach

Under the hypothesis of a new physics model with free parameters, maximize the expected evidence, taking into account all available prior information

Documented in CMS-AN-2011-351

Many updates being documented

optimization of searches

$$\begin{aligned}
e &= \text{evidence} \\
H_{NP} &= \text{new physics hypothesis} \\
\theta &= \text{parameters of } H_{NP}
\end{aligned}$$

$$\begin{aligned}
E(e|H_{NP}) &= \int eP(e|H_{NP})\mathrm{d}e \\
&= \int\int eP(e|\theta, H_{NP})\pi(\theta)\mathrm{d}e\mathrm{d}\theta \\
&\approx \sum_i \int eP(e|\theta_i, H_{NP})\pi(\theta_i)\mathrm{d}e
\end{aligned}$$

optimization of searches

$$E(e|H_{NP}) = \sum_i \int e P(e|\theta_i, H_{NP}) \pi(\theta_i) \mathrm{d}e \mathrm{d}\theta_i$$

$\pi(\theta)$	prior probability that True State of Nature (TSN) is θ built from previous measurements
$P(e\|\theta_i, H_{NP})$	probability to observe evidence e if TSN is θ_i obtained from simulation

A multi-parameter model: pMSSM

pMSSM

- phenomenological MSSM*
- 19 free parameters at the SUSY scale
- well-motivated assumptions, reducing MSSM
- no correlations between sparticle masses

$\rightarrow$ **rich phenomenology**

*Minimal Supersymmetric extension of the Standard Model

$\pi(\theta)$ for pMSSM

Pre-LHC Measurements (PLMs)

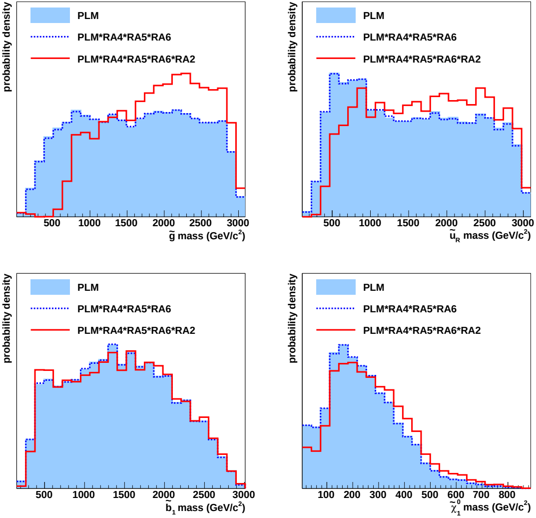
i	Observable	Limit	Likelihood function
1	Δa_μ	$(28.7 \pm 8.0) \times 10^{-10}$ $[e^+e^-]$ $(19.5 \pm 8.3) \times 10^{-10}$ [taus]	Weighted Gaussian average
2	$BR(b \to s\gamma)$	$(3.55 \pm 0.34) \times 10^{-4}$	Gaussian
3	$BR(B_s \to \mu\mu)$	$\leq 4.7 \times 10^{-8}$	Upper limit
4	$R(B_u \to \tau\nu)$	1.66 ± 0.54	Gaussian
5	m_t	173 ± 1.1	Gaussian
6	$m_b(m_b)$	$4.19^{+0.18}_{-0.06}$	Two-sided Gaussian
7	$\alpha_s(M_Z)$	0.117 ± 0.002	Gaussian
8	m_h	LEP & Tevatron (`HiggsBounds`)	$L_8 = 1$ if allowed. $L_8 = 10^{-9}$ if $m_h^{'}$ sampled from $Gauss(m_h, 1.5)$ is excluded.
9	SUSY mass	LEP & Tevatron (`micrOMEGAs`)	$L_9 = 1$ if allowed $L_9 = 10^{-9}$ if excluded

$\pi(\theta)$ for pMSSM

CMS measurements

Analysis	Final state and signal region	N	B	δB	$\int \mathcal{L}$
RA2	Inclusive jets+H_T^{miss} HT $>$ 500, MHT $>$ 350	12	10	2.5	882 pb^{-1}
RA4	Single lepton+jets+E_T^{miss} MET $>$ 250, HT $>$ 500	73	66.66	20.2	1100 pb^{-1}
RA5	Same sign (SS) dileptons+jets+E_T^{miss} Inclusive leptons, HT $>$ 400, $E_T^{miss} > 120$	1	2.3	1.2	980 pb^{-1}
RA6	Opposite sign (OS) dileptons+jets+E_T^{miss} HT $>$ 300, $E_T^{miss} > 275$	8	4.2	1.3	980 pb^{-1}

$\pi(\theta)$ for pMSSM



$=$ a combined interpretation

SUSY with *b*-jets

SUSY with *b*-jets

documented in SUS-11-006

Search definition (previous version)

- PF object reconstruction
- at least 3 jets
  (anti-kt, $p_T > 50$ GeV, $|\eta| < 2.4$)
- no isolated muons or electrons
  ($p_T > 10$ GeV)
- b-jets: TCHEM
- search region defined by lower bounds on
  - $H_T = \sum |p_T^{jet}|$
  - E_T^{miss}
  - $\Delta\Phi(\vec{E}_T^{miss}, jets)$

SUSY with *b*-jets

Optimization

"evidence"

$$e = \begin{cases} 0 & \text{if } S/\sqrt{S+B} < 3 \\ 1 & \text{if } S/\sqrt{S+B} > 3 \end{cases}$$

$\rightarrow E(e|H_{NP}) \approx$ probability for discovery under H_{NP}

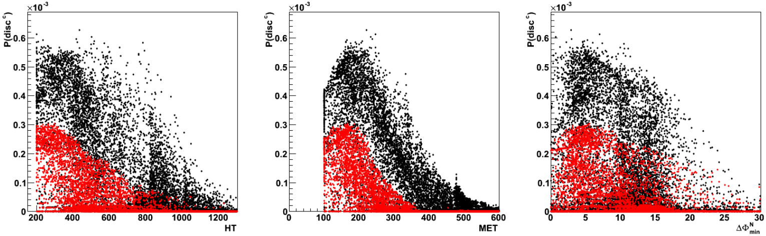
SUSY with *b*-jets

Optimization

Discovery probability vs.

- lower limit on H_T
- lower limit on E_T^{miss}
- lower limit on $\Delta\Phi^N_{min}$

red: 1 fb^{-1}, black: 5 fb^{-1}



SUSY with b-jets

Optimization

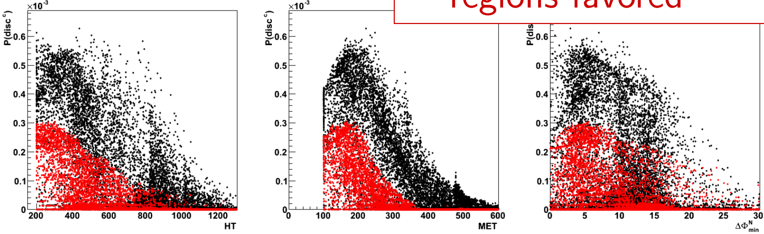
Discovery probability vs.

- lower limit on H_T
- lower limit on E_T^{miss}
- lower limit on $\Delta\Phi^N_{min}$

red: 1 fb^{-1}, black: 5 fb^{-1}

Conclusions

- broad optima
- "low" E_T^{miss} and H_T regions favored



SUSY with *b*-jets

Outlook

Proposal for the 2012 Analysis

- identical object definitions and variables
- multiple independent search regions
- shape analysis
- alternative Bayesian interpretation

... work in progress

Bigger picture
W/Z + jets studies
SUSY with b-jets

Summary

Summary

A lot of progress on many fields

e.g.:
- SM measurements (W/Z + jets)
- new physics interpretation of measurements
- design of searches (SUSY with *b*-jets)

Put the pieces together,
Design the next generation of searches!