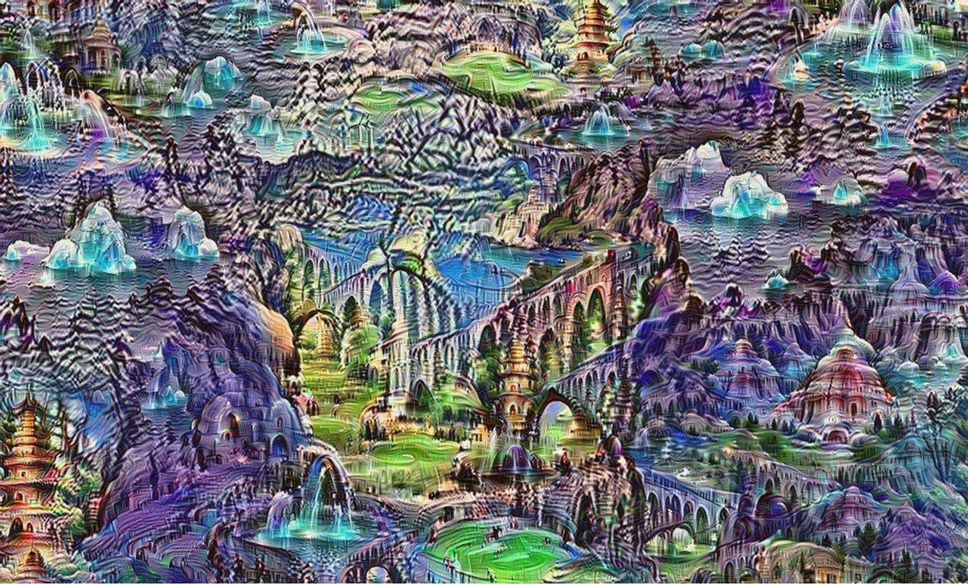
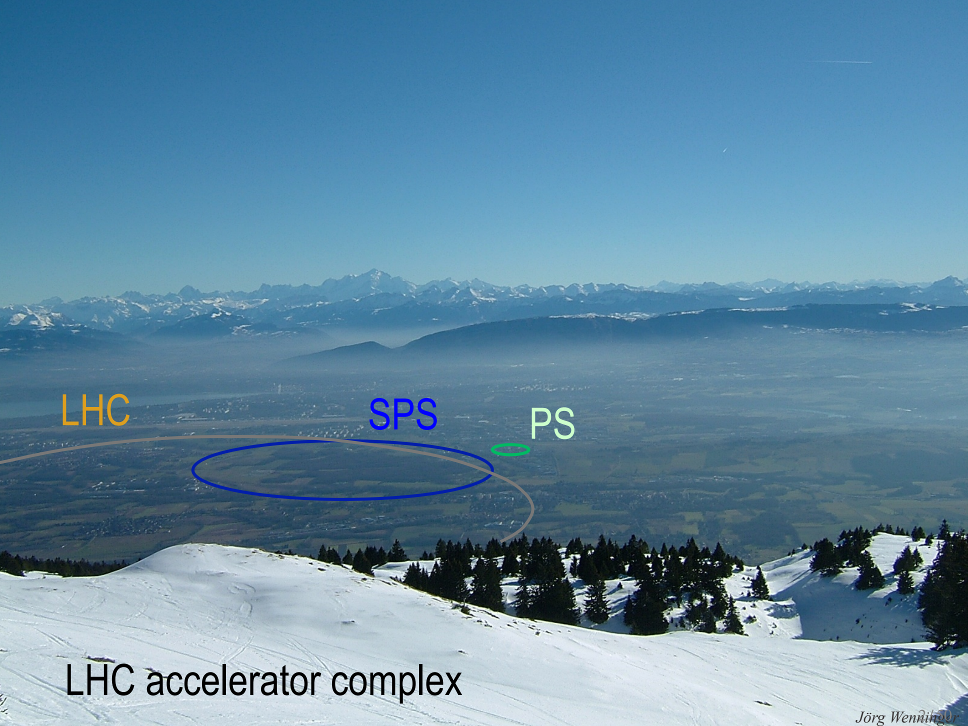
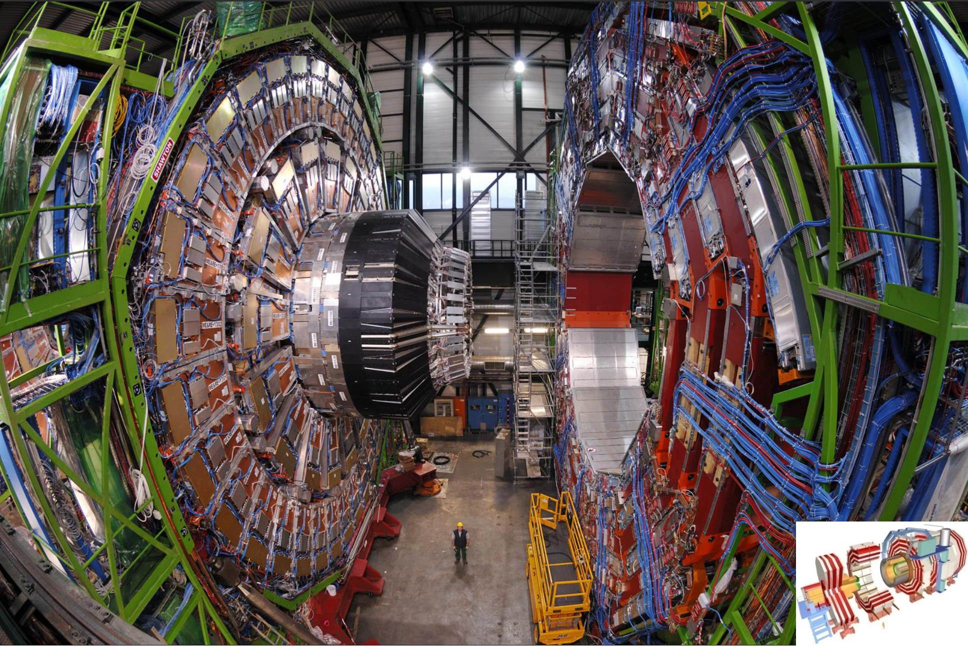


Application of MVA in new physics searches

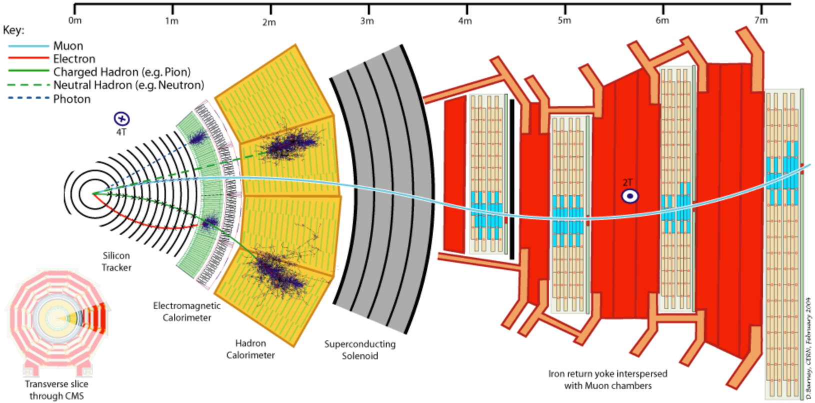
Lesya Shchutska (UF)
MLHEP School
St. Petersburg, Russia


LHC
SPS
PS
LHC accelerator complex
Jörg Wenninger



CMS in the cavern before closing

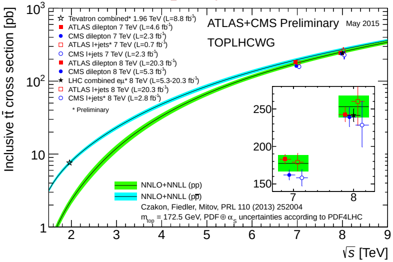
Schematic view



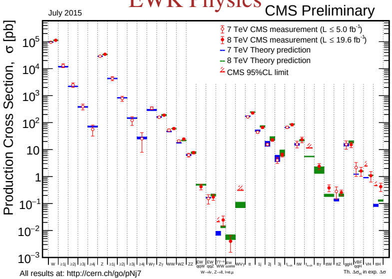
- 4π general purpose detector: registers all particles in transverse plane
- spectrometer with a large magnetic field (3.8 T): precise measurement of charged particles
- electromagnetic calorimeter for photon/electron measurements
- hadron calorimeter with moderate energy resolution: $\sim 10\%$ above 100 GeV
- muon system: identification (also improves momentum resolution for TeV muons)

Success of the Standard model

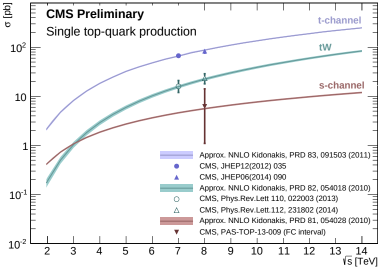
Top Physics



EWK Physics



Single top

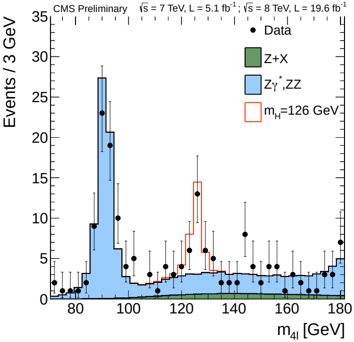


1. $\sigma_{t\bar{t}} = 241.5 \pm 1.4(\text{stat.}) \pm 5.7(\text{syst.}) \pm 6.2(\text{lumi.})$ pb
2. $m_t = 172.38 \pm 0.14(\text{stat.}) \pm 0.64(\text{syst.})$ GeV

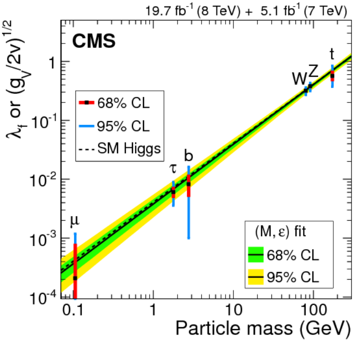
Stunning agreement with SM expectations!

Higgs at CMS

6.7σ in H$\rightarrow 4\ell$



Higgs couplings

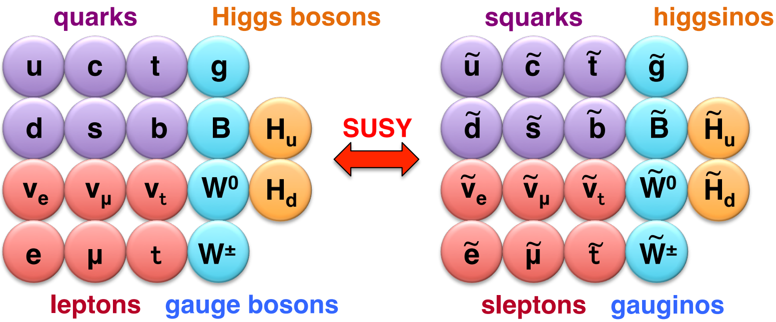


Moving from discovery to precision measurement!

Supersymmetry: searches for new physics

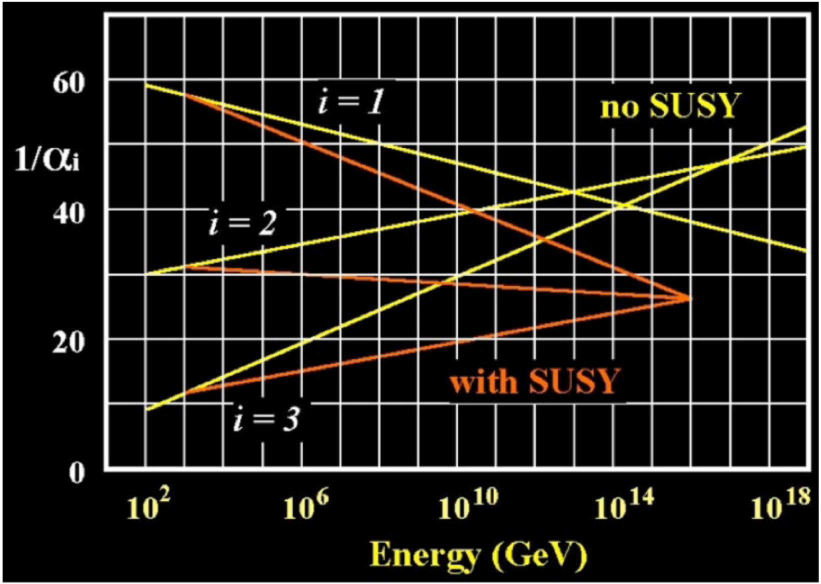
One important very basic symmetry (have not been seen (*yet?*))

- for each $\frac{1}{2}$-integer spin particle (fermion) there is an integer spin partner (boson) and vice versa
  - complete spectrum of partners to standard model particles
  - their spins are different by $\frac{1}{2}$ unit
  - they are heavier (or else we'd have seen them already).



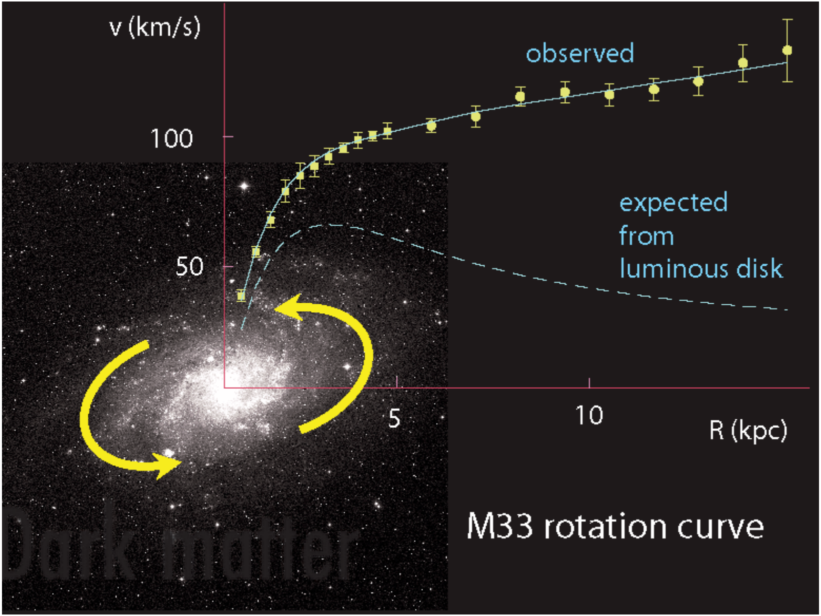
SUSY implications

SUSY unifies the strengths of all forces at $\sim 10^{16}$ GeV



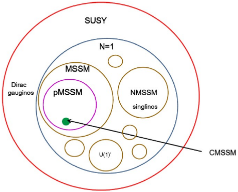
SUSY implications

Explains 25% of the energy in the universe: the **dark matter**



Beautiful but not minimal theory

- SUSY is a broken symmetry: masses of superpartners are not fixed by theory
- a parameter space which is impossible to fully exclude but to only constrain

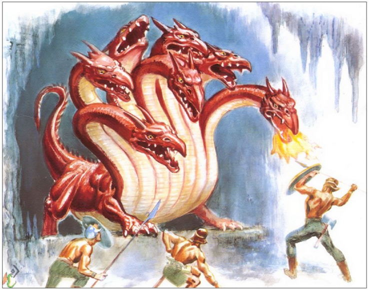


Within the MSSM only:

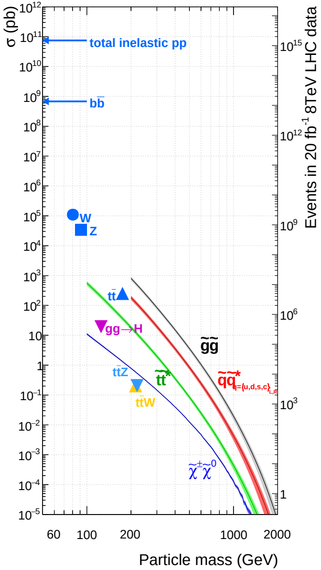
- MSSM: 109 parameters
- pMSSM: 19 parameters
- CMSSM: 5 parameters

Complementary strategies are required to maximally constrain the parameters:

- direct and indirect dark matter detection experiments
- study of the rates of the rare processes (e.g. heavy-flavor physics)
- precision SM production cross section measurements (e.g. $t\bar{t}$ production)



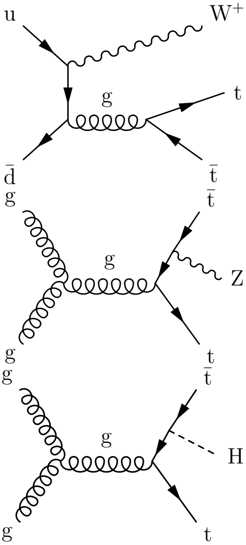
- **direct SUSY particle production in the pp collisions at the LHC and their detection in ATLAS or CMS experiments**



Challenges

- SUSY particle production cross sections are 10 and more orders of magnitude lower than full pp collision rate at the LHC
- comparable to other rare SM processes, e.g.:
  - $\sigma(t\bar{t}W) = 203$ fb
  - $\sigma(t\bar{t}Z) = 206$ fb
  - $\sigma(t\bar{t}H) = 129$ fb
  - expect around 4000 events produced in full 8 TeV dataset
- compared to $\sigma(t\bar{t}) = 252$ pb: 3 orders of magnitude to go
- $t\bar{t}$+X processes are great proxies and test-cases for possible SUSY discovery
  - $t\bar{t}W$ and $t\bar{t}Z$ story is told in the next slides

Discovering new processes in standard model



The same full 8 TeV dataset ($\sim$20 fb^{-1}), and the search for $t\bar{t}Z$ and $t\bar{t}W$.

In **CMS**:

1. **June 30, 2014**: "Measurement of top quark-antiquark pair production in association with a W or Z boson in pp collisions at $\sqrt{s} = 8$ TeV"
2. **May 08, 2015**: "Measurement of top quark pair production in association with a W or Z boson using event reconstruction techniques"

And in **ATLAS**:

1. **July 4, 2014**: "Evidence for the associated production of a vector boson (W, Z) and top quark pair in the dilepton and trilepton channels in pp collision data at $\sqrt{s} = 8$ TeV collected by the ATLAS detector at the LHC"
2. **July 22, 2015**: "Measurement of the $t\bar{t}W$ and $t\bar{t}Z$ production cross sections in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector"

What is the difference between the two?

First results

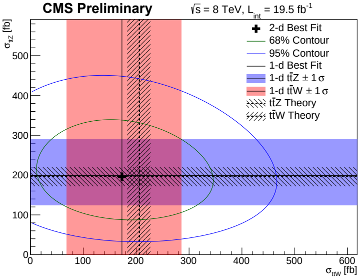
First analyses by CMS and ATLAS at 7 and 8 TeV used a cut-based approach in the most sensitive channels, observed $t\bar{t}Z$ at $\sim 3\sigma$ significance and achieved $\sim 2\sigma$ sensitivity to $t\bar{t}W$

CMS

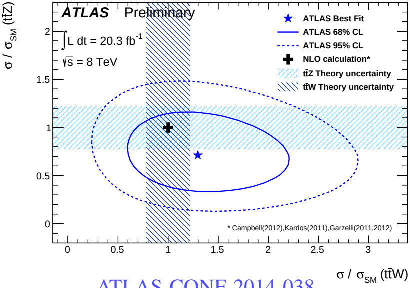
Channels used	Process	Cross section	Significance
2ℓ	$t\bar{t}W$	170^{+90}_{-80}(stat.)$^{+70}_{-70}$(syst.) fb	1.6 σ
3ℓ+4ℓ	$t\bar{t}Z$	200^{+80}_{-70}(stat.)$^{+40}_{-30}$(syst.) fb	3.1 σ

ATLAS

Process	Measured cross-sections	Observed σ	Expected σ
$t\bar{t}Z$	150^{+58}_{-54}(total) = 150^{+55}_{-50}(stat.) ± 21(syst.) fb	3.1	3.7
$t\bar{t}W$	300^{+140}_{-110}(total) = 300^{+120}_{-100}(stat.)$^{+70}_{-40}$(syst.) fb	3.1	2.3



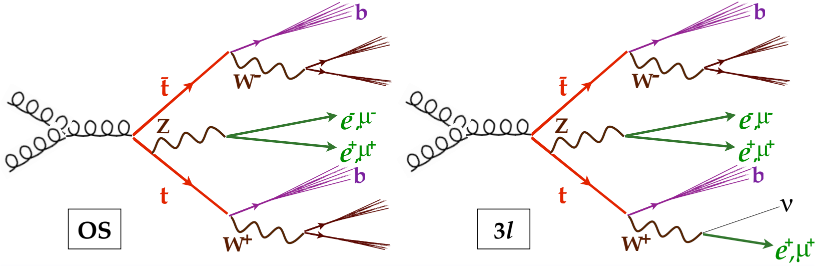
EPJ C74 (2014) 3060



ATLAS-CONF-2014-038

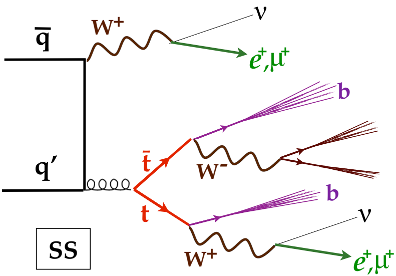
$t\bar{t}Z$ signatures

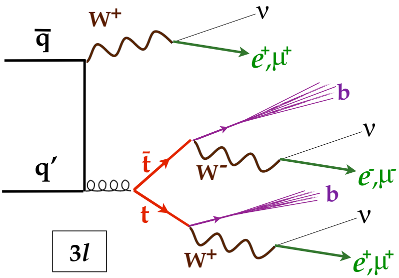
- explore a presence of a Z boson:
  - use all final states with $Z \to \ell\ell$ ($\ell = $ e or μ)
- and a fact of having b-jets, missing transverse energy (E_{T}^{miss}), and/or light flavor jets
- can use 2ℓOS, 3ℓ, 4ℓ signatures
  - first CMS C&C analysis used only 3ℓ, 4ℓ
  - first ATLAS paper covered 2ℓOS, 2ℓSS and 3ℓ



$t\bar{t}W$ signatures

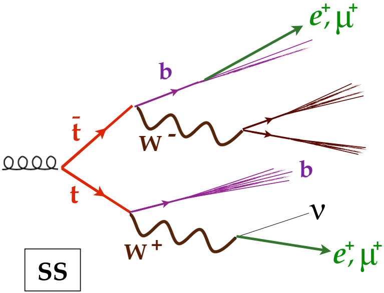
- explore leptonic W decays
- CMS looked at 2ℓSS
- ATLAS used 2ℓSS and 3ℓ final states
- in addition: b jets, light flavor jets, E_T^{miss}

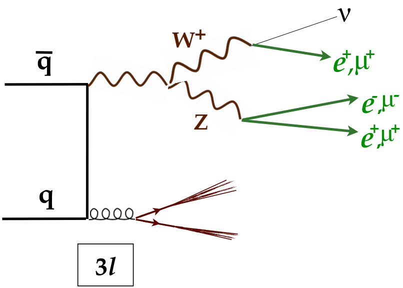




Backgrounds

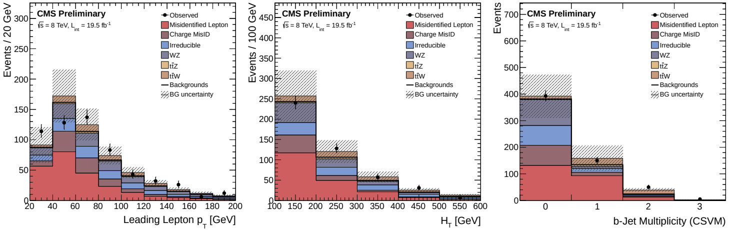
- SM processes with fewer particles (and orders of magnitude larger cross sections):
  - $t\bar{t}$, Z, WZ, ZZ
- to fake $t\bar{t}V$ need to have extra objects:
  - either leptons - from b decays or misidentified jets
  - or jets - from initial state radiation (ISR) or pile up
- these backgrounds have different importance in various search channels:
  - first categorize in lepton multiplicity





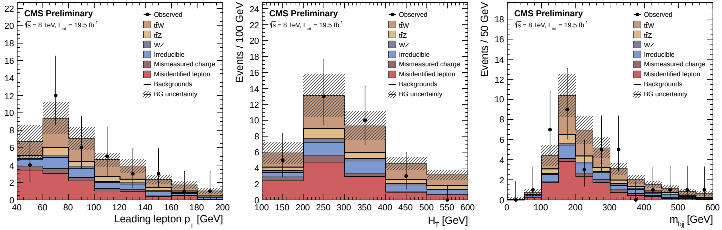
How it looks like: same-sign dileptons

Preselection



- lepton kinematics: leading lepton p_{T} is harder in signal
- hadronic activity: more jets in signal $\Longrightarrow$ larger $H_{\mathrm{T}} = \sum_{\mathrm{jets}} p_{\mathrm{T}}^{\mathrm{jet}}$
- t quark mass reconstruction of 3 jets: poor mass resolution

Full selection



arXiv:1406.7830

Same-sign dileptons: result

End up with a selection like:

- 2 same-sign leptons with $p_\mathrm{T} > 40$ GeV
- at least 3 jets with $p_\mathrm{T} > 30$ GeV, at least one of them a b jet
- hadronic activity $H_\mathrm{T} > 155$ GeV

$\Longrightarrow$ box-like cuts

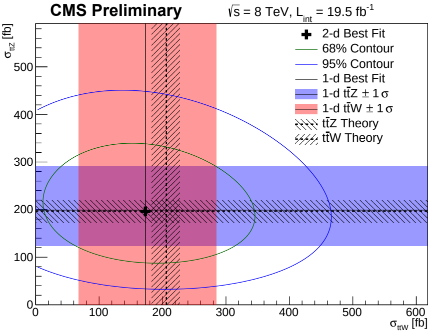
- expect around 15 signal events in all categories with about 25±7 background
- while initially have 4000×0.25×0.25 = 250 2ℓSS signal $t\bar{t}W$ events
- need to cut very hard to arrive at reasonable S/B ratio

	$\mu^+\mu^+$	$e^+\mu^+$	e^+e^+	$\mu^-\mu^-$	$e^-\mu^-$	e^-e^-
$t\bar{t}W$(expected)	2.8 ± 0.4	5.1 ± 0.5	2.2 ± 0.3	1.1 ± 0.2	2.3 ± 0.3	1.0 ± 0.2
Misidentified lepton	1.0 ± 0.6	4.1 ± 2.1	1.6 ± 0.9	0.7 ± 0.4	3.0 ± 1.5	1.7 ± 0.9
Mismeasured charge	—-	0.4 ± 0.1	0.7 ± 0.2	—	0.4 ± 0.1	0.7 ± 0.2
Irreducible	0.7 ± 0.4	1.6 ± 0.9	0.9 ± 0.5	0.5 ± 0.3	1.4 ± 0.7	0.7 ± 0.4
WZ	0.1 ± 0.1	0.4 ± 0.1	0.1 ± 0.1	0.1 ± 0.1	0.4 ± 0.1	0.2 ± 0.1
$t\bar{t}Z$	0.6 ± 0.3	0.9 ± 0.5	0.5 ± 0.3	0.4 ± 0.2	1.0 ± 0.5	0.5 ± 0.3
Total background	2.4 ± 0.7	7.4 ± 2.3	3.9 ± 1.1	1.7 ± 0.5	6.1 ± 1.8	3.7 ± 1.1
Total expected	5.2 ± 0.8	12.5 ± 2.4	6.1 ± 1.1	2.8 ± 0.5	8.4 ± 1.8	4.7 ± 1.1
Observed	6	12	5	1	6	6

- similar approach in other lepton multiplicities

arXiv:1406.7830

Final result



- very much compatible with SM
- evidence for $t\bar{t}Z$ production
- quite low sensitivity to $t\bar{t}W$

Channels used	Process	Cross section	Significance
2ℓSS	$t\bar{t}W$	$170^{+90}_{-80} \pm 70$fb	1.6
3ℓ+4ℓ	$t\bar{t}Z$	$200^{+80}_{-70}{}^{+40}_{-30}$fb	3.1
2ℓSS+3ℓ+4ℓ	$t\bar{t}W+t\bar{t}Z$	$380^{+100}_{-90}{}^{+80}_{-70}$fb	3.7

arXiv:1406.7830

First ATLAS analysis

- use 2ℓSS, 2ℓOS, 3ℓ channels
  - **3ℓ and 2ℓSS**: comparable signal and background $\Longrightarrow$ C&C
  - **opposite-sign dilepton**: small signal in huge background $\Longrightarrow$ neural network
    - a channel was not present in previous C&C search!
- main sensitivity of the search is from cut&count analyses
  - combined result is comparable with CMS cut&count paper

	Trilepton and same-sign dilepton channels		
Process	Signal Strength	Observed σ	Expected σ
$t\bar{t}V$	$0.91^{+0.27}_{0.24}$	4.6	4.6
$t\bar{t}W$	$1.31^{+0.62}_{-0.50}$	3.0	2.3
$t\bar{t}Z$	$0.72^{+0.33}_{-0.28}$	2.8	3.4

	Opposite-sign dilepton channel		
Process	Signal Strength	Observed σ	Expected σ
$t\bar{t}V$	$0.77^{+0.63}_{-0.56}$	1.4	1.7
$t\bar{t}W$	$0.57^{+2.48}_{-2.30}$	0.3	0.4
$t\bar{t}Z$	$0.77^{+0.69}_{-0.59}$	1.4	1.5

	Simultaneous fit of two signal strengths in all channels			
Channel	$\mu_{t\bar{t}Z}$	$\mu_{t\bar{t}W}$	Observed σ	Expected σ
trilepton and same-sign dilepton	$0.70^{+0.30}_{-0.28}$	$1.37^{+0.62}_{-0.51}$	4.1	4.1
opposite-sign dilepton	0.77 ± 0.65	0.71 ± 2.41	0.4	0.6
combination	$0.71^{+0.28}_{-0.26}$	$1.30^{+0.59}_{-0.48}$	4.4	4.4

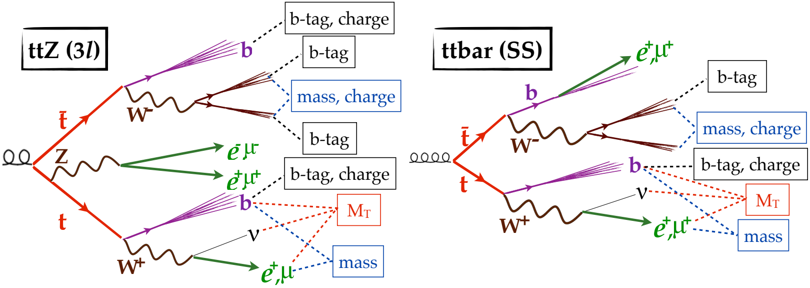
ATLAS-CONF-2014-038

$t\bar{t}$V reapproached at CMS

Instead of relying on simple kinematic properties of objects in the event:

1. **employ physics knowledge**: be aware of the presence of top quarks
   - reconstruct a $t\bar{t}$ pair in $t\bar{t}$Z, $t\bar{t}$W and $t\bar{t}$ events
2. **acknowledge complicated kinematics**: use multivariate techniques
   - construct a linear discriminant (**MatchLD**) with reconstructed masses and object properties
   - perform a training on the $t\bar{t}$ (background) MC

Note: cannot easily separate the gain between the two



Event reconstruction: full input list to MatchLD

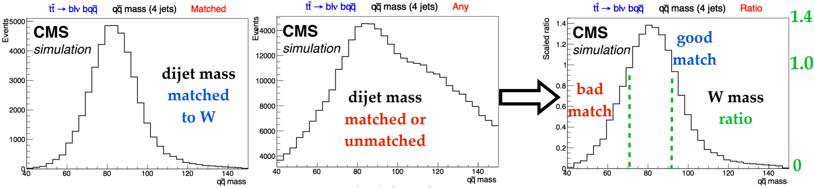
- extensively use the knowledge of event topology
- prefer a lot of non-linear and **physics-motivated** variables and a **linear discriminant**
  - some inputs are another MVA results, e.g. b-tagging score
- rather than fewer simple 4-vectors and a **deep learning neural network**

Variable	OS $t\bar{t}Z$	SS $t\bar{t}W$ and 3ℓ $t\bar{t}Z$	3ℓ $t\bar{t}W$
b-jet CSV	Y	Y	Y
Higher jet CSV from W	Y	Y	N
Lower jet CSV from W	Y	Y	N
Leptonic top b-jet charge	N	Y	N
Hadronic top b-jet charge	N	Y	N
Top b-jet charge	Y	N	Y
Anti-top b-jet charge	Y	N	Y
Sum of charges of jets from W	Y	Y	N
Mass of lepton and b-jet from leptonic top	N	Y	N
Mass of lepton and b-jet from top	N	N	Y
Mass of lepton and b-jet from anti-top	N	N	Y
M_T of $E_\mathrm{T}^\mathrm{miss}$ and lepton and b-jet from leptonic top	N	Y	N
Mass of two jets from W	Y	Y	N
Mass of b-jet and one jet from W from hadronic top	Y	Y	N
Mass of b-jet and two jets from W from hadronic top	Y	Y	N
Ratio of M_T to mass for jets from top or W	Y	Y	Y

Example in $t\bar{t}Z$

Example of constructing a MatchLD score:

- left: $W \rightarrow qq$ dijet mass distribution for correctly picked jets in $t\bar{t}$
- center: dijet mass distribution for any pair of jets in the $t\bar{t}$ event
- right: scaled ratio of the two - used as a score

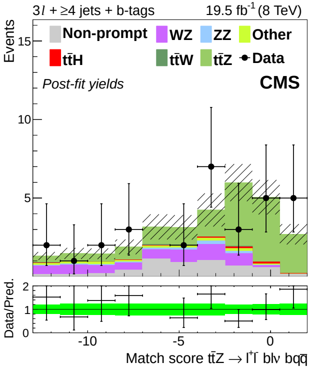


Application in data:

- test every permutation of jets and leptons with different “mother” hypotheses
- for each permutation take the product of values from ratio histograms: $\log(\prod_i \text{ratio}_i)$
  - this score has median at 1 for correct match, and < 1 for wrong reconstruction
- the highest MatchLD score out of all permutation is used as another discriminating variable

MatchLD score in 3ℓ for $t\bar{t}Z$

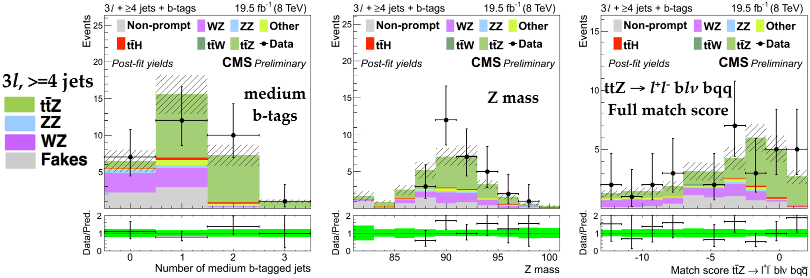
- for events where both b's from the top and both q's from the W are reconstructed as jets, 75% of 4 jet and 40% of $\geq$ 5 jet events have every jet **correctly matched** to its parent particle
- for correct matches, the average ratio is 1 (so the match score centers near 0)
- partial matches (all but one jet matched) allow to identify signal events where one of the quarks forms a jet outside our acceptance



3ℓ $t\bar{t}Z$ BDT

MatchLD itself is passed over to the BDT:

- train boosted decision tree with $t\bar{t}Z$ vs. WZ and $t\bar{t}$ MC
- also include M_T of system, and partial matches to $t\bar{t}Z$ system
  - $t\bar{t}Z \to \ell^+\ell^-(b\ell\nu)(bq)$
  - $t\bar{t}Z \to \ell^+\ell^-(\ell\nu)(bqq)$
  - $t\bar{t}Z \to \ell^+\ell^-(b\ell\nu)(qq)$



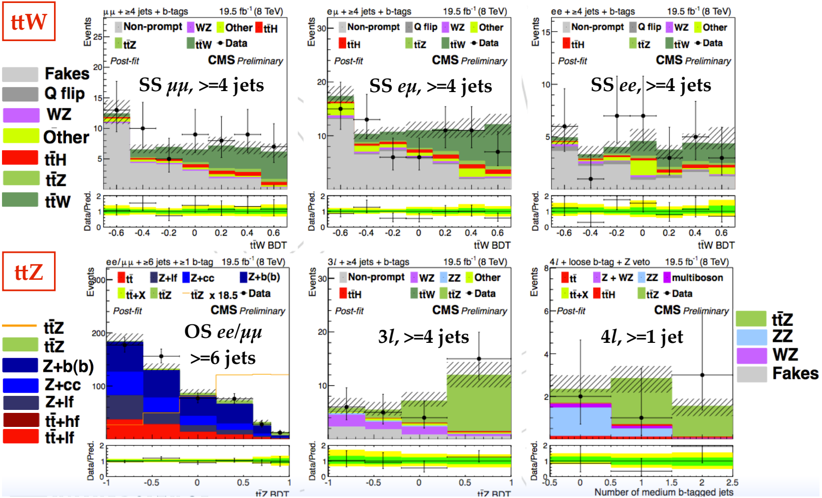
$t\bar{t}$W BDT inputs

- train boosted decision trees with $t\bar{t}$W vs. $t\bar{t}$ MC
- use a mix of kinematic and matching variables
  - kinematic variables from p_T of objects and jet b-tag (CSV)
  - event matching variables for $t\bar{t}$W and $t\bar{t}$ systems

BDT inputs: same-sign $t\bar{t}$W vs. $t\bar{t}$	3 jet	$\geq$4 jets
M_T of MET, leptons, and jets	**1**	**1**
MET	4	2
2[nd] highest lepton p_T	6	3
MatchLD score for $t\bar{t}$$(Bqq)$	**2**	**4**
Highest lepton p_T	5	5
2[nd] highest CSV value of a jet	8	6
$t\bar{t}$ matched top M_T from $b\ell\nu$	7	7
MatchLD score for $t\bar{t}$W(Bbq)	9	8
MatchLD score for $t\bar{t}$W$(Bbqq)$	-	9
$t\bar{t}$ matched top mass from $\ell_b qq$	3	-

- reconstructing events with a linear discriminant first allows for more input variables and better separation than a BDT alone, since BDTs are more limited by the statistics of training events

BDT output


ttW
Fakes
Q flip
WZ
Other
tt̄H
tt̄Z
tt̄W
μμ + ≥4 jets + b-tags
19.5 fb⁻¹ (8 TeV)
Non-prompt
Data
Post-fit
CMS Preliminary
SS μμ, >=4 jets
eμ + ≥4 jets + b-tags
SS eμ, >=4 jets
ee + ≥4 jets + b-tags
SS ee, >=4 jets
Events
Data/Pred.
tt̄W BDT
ttZ
tt̄+hf
tt̄+lf
Z+b(b)
Z+cc
Z+lf
tt̄+X
tt̄Z x 18.5
ee/μμ + ≥6 jets + ≥1 b-tag
OS ee/μμ
>=6 jets
3l + ≥4 jets + b-tags
3l, >=4 jets
ZZ
4l + loose b-tag + Z veto
Z + WZ
multiboson
4l, >=1 jet
tt̄Z BDT
Number of medium b-tagged jets

Results

$t\bar{t}Z$:

	Cross section (fb)		Signal strength (μ)		Significance	
Channels	Expected	Observed	Expected	Observed	Expected	Observed
OS	206^{+142}_{-118}	257^{+158}_{-129}	$1.0^{+0.72}_{-0.57}$	$1.25^{+0.76(+1.76)}_{-0.62(-1.16)}$	1.84	2.12
3ℓ	206^{+79}_{-63}	257^{+85}_{-67}	$1.0^{+0.42}_{-0.32}$	$1.25^{+0.45(+1.02)}_{-0.36(-0.62)}$	4.55	5.11
4ℓ	206^{+153}_{-109}	228^{+150}_{-107}	$1.0^{+0.77}_{-0.53}$	$1.11^{+0.76(+1.79)}_{-0.52(-0.86)}$	2.65	3.39
All	206^{+62}_{-52}	242^{+65}_{-55}	$1.0^{+0.34}_{-0.27}$	$1.18^{+0.35(+0.79)}_{-0.29(-0.51)}$	**5.73**	**6.44**

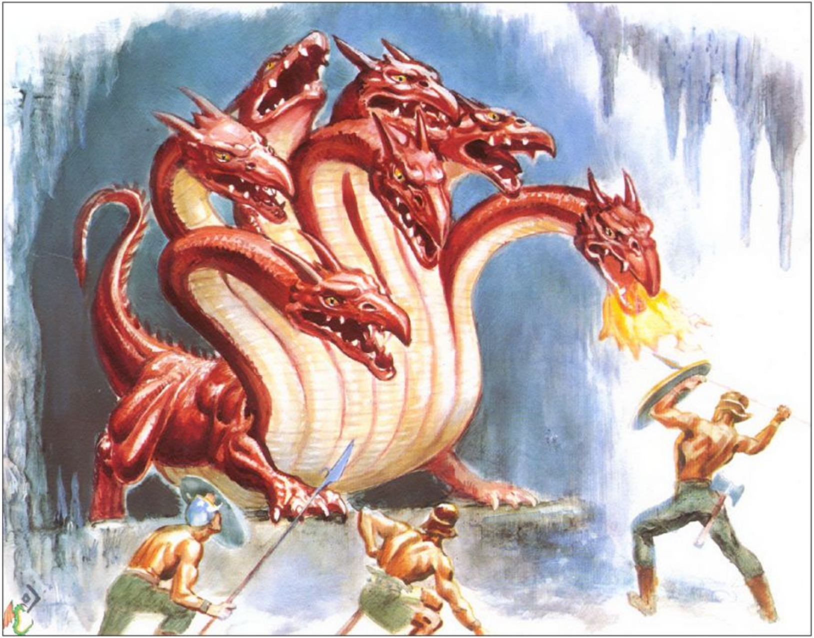
$t\bar{t}W$:

	Cross section (fb)		Signal strength (μ)		Significance	
Channels	Expected	Observed	Expected	Observed	Expected	Observed
SS	203^{+88}_{-73}	414^{+135}_{-112}	$1.0^{+0.45}_{-0.36}$	$2.04^{+0.74(+1.52)}_{-0.61(-1.05)}$	3.44	4.89
3ℓ	203^{+215}_{-94}	210^{+225}_{-203}	$1.0^{+1.09}_{-0.96}$	$1.03^{+1.07(+2.39)}_{-0.99(-1.92)}$	1.03	1.03
All	203^{+84}_{-71}	382^{+117}_{-102}	$1.0^{+0.43}_{-0.35}$	$1.88^{+0.66(+1.35)}_{-0.56(-0.95)}$	**3.54**	**4.81**

More specific and targeted approach makes a difference between the *evidence* and *discovery*!

1. $t\bar{t}Z$: from 3.1 (3.1)σ to 5.7(6.4)σ
2. $t\bar{t}W$: from 2.0 (1.6)σ to 3.5(4.8)σ

What about SUSY?

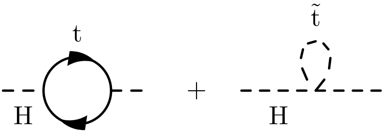


Again top-quark: superpartner

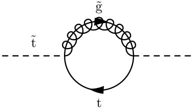
$\tilde{t}$ cancels out the largest divergency in the Higgs boson mass - from **t quark**:

1. 1-loop order: top contribution corrected by $\tilde{t} \rightarrow m_{\tilde{t}} \approx \mathcal{O}(100 \text{ GeV})$

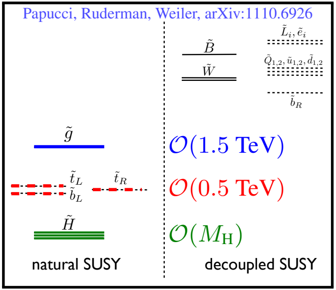
$\Delta m^2_{\text{H}} \propto$



2. 2-loop order: gluino enters $\tilde{t}$ mass $\rightarrow m_{\tilde{g}} \approx \mathcal{O}(1 \text{ TeV})$



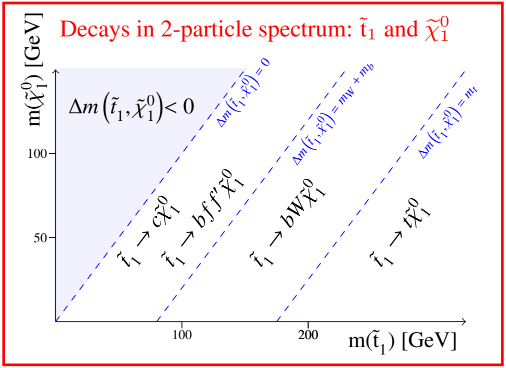
3. not too heavy $\widetilde{\text{b}}_{\text{L}}$: in the doublet with $\tilde{t}_{\text{L}}$



Plethora of ATLAS and CMS analyses looking for **gluinos**, **3$^{\text{rd}}$ generation squarks** and **gauginos**!

Topologies for Top squark search

Landscape is thoroughly combed in search of an elusive top-quark partner
- needed to stabilize a Higgs boson mass:



Also look at 3-particle spectra:

$\tilde{t}_1$, $\tilde{\chi}^\pm_1$, $\tilde{\chi}^0_1$:

$\tilde{t}_1 \to b\tilde{\chi}^\pm_1(\to W\tilde{\chi}^0_1)$

$\tilde{t}_1$, $\tilde{\chi}^0_2$, $\tilde{\chi}^0_1$:

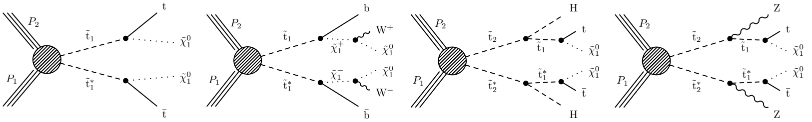
$\tilde{t}_1 \to t\tilde{\chi}^0_2(\to Z/H\tilde{\chi}^0_1)$

$\tilde{t}_2$, $\tilde{t}_1$, $\tilde{\chi}^0_1$:

$\tilde{t}_2 \to Z/H\tilde{t}_1(\to t\tilde{\chi}^0_1)$

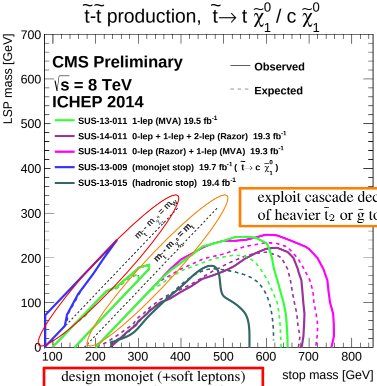
Each final state is targeted
by one or more dedicated
heavily optimized searches

Individual analyses
are combined to estimate
sensitivity in mixed scenarios



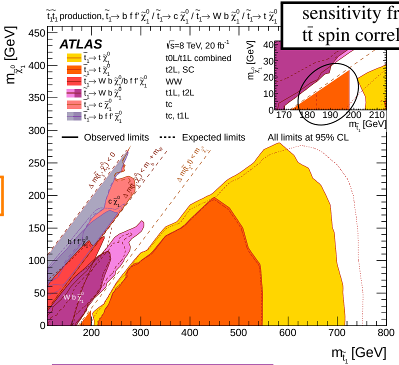
Top squark searches in two plots

- probed phase space extends up to 750 GeV in $\tilde{t}$ mass and 260 GeV in $\tilde{\chi}^0_1$ mass
- invent new tools to access difficult regions: $m_{\tilde{t}} - m_{\tilde{\chi}^0_1} \approx m_t$ or m_W



exploit cascade decays of heavier $\tilde{t}_2$ or $\tilde{g}$ to $\tilde{t}_1$

design monojet (+soft leptons) to catch $\tilde{t} \rightarrow c\tilde{\chi}^0_1$ ($\tilde{t} \rightarrow bff\tilde{\chi}^0_1$)



sensitivity from $t\bar{t}$ spin correlation

comprehensive combination of all results is produced

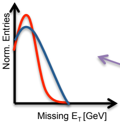
CMS-PAS-SUS-14-021 arXiv:1405.3886 arXiv:1506.08616 arXiv:1412.4742

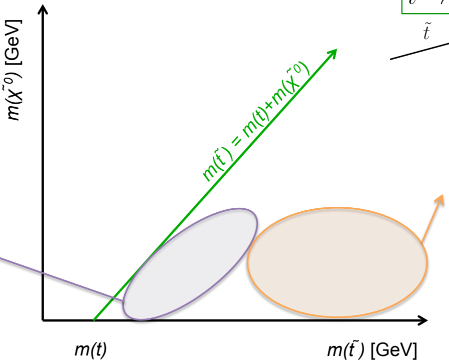
Search strategy

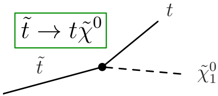
Multiple signal regions target different signal kinematics based on $m(\tilde{t})$ and $m(\tilde{\chi}_1^0)$

'low Δm'

Best sensitivity with looser requirements
→ small systematic uncertainty on the background is key

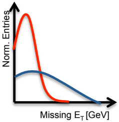






'high Δm'

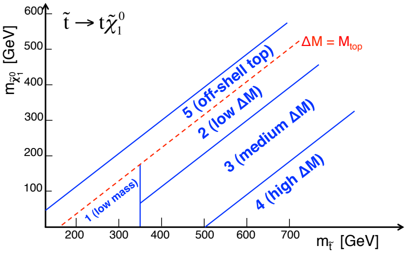
Best sensitivity with tighter requirements
background suppressed to O(10) events
→ low statistics regime



Search strategy

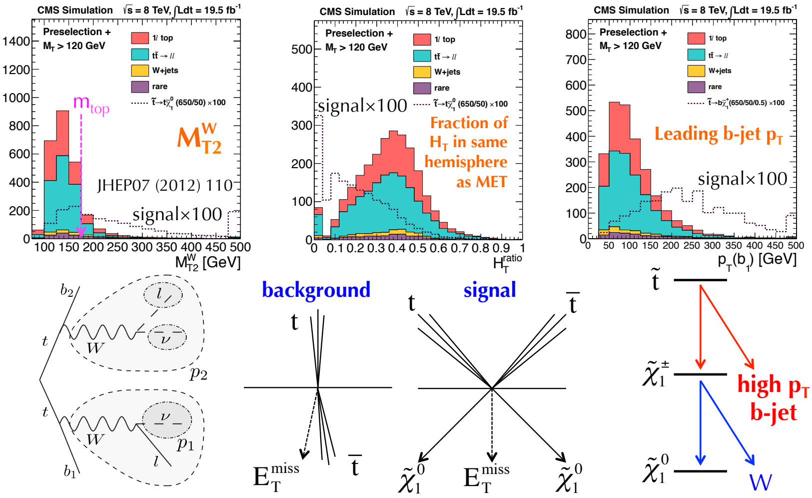
- Performed analysis in two different ways in terms of how the signal regions are defined
  - **Cut-based**: "Square" cuts on several variables
  - **BDT**: Combine variables into a BDT and define signal region by cutting on the output

Multiple signal regions to target different signal kinematics

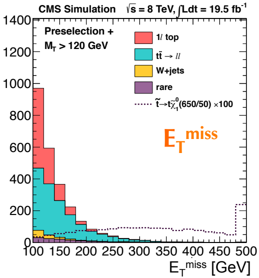


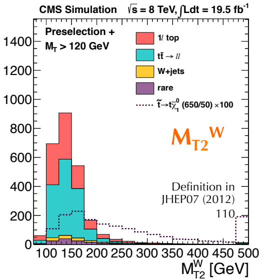
arXiv:1308.1586

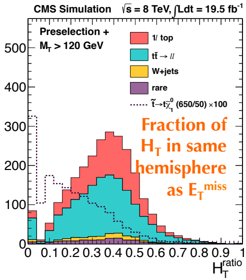
Example variables

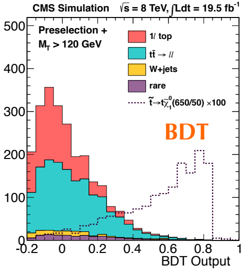


BDT output discriminant









Example BDT output discriminant used to define the signal regions

Signal regions definitions

Signal variables and regions used in two analyses

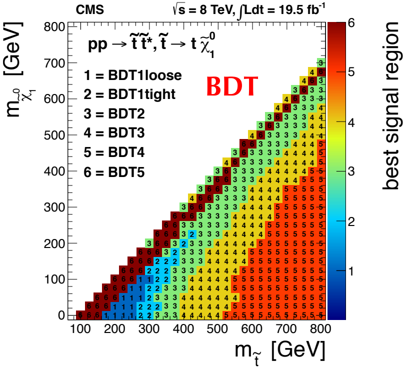
Selection	$\tilde{t} \to t\tilde{\chi}_1^0$		
	BDT	cut-based Low ΔM	cut-based High ΔM
E_T^{miss}(GeV)	yes	$> 150, 200, 250, 300$	$> 150, 200, 250, 300$
M_{T2}^{W} (GeV)	yes		> 200
min $\Delta\phi$	yes	> 0.8	> 0.8
H_T^{ratio}	yes		
hadronic top χ^2	(on-shell top)	< 5	< 5
leading b-jet p_T (GeV)	(off-shell top)		
$\Delta R(\ell$,leading b-jet)			
lepton p_T			

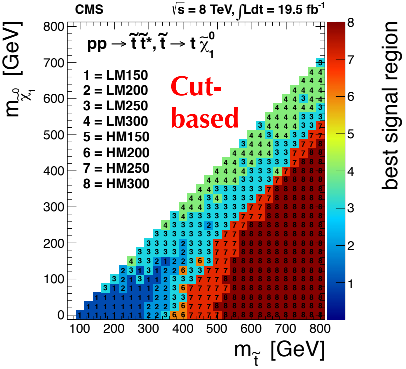
Note: for cut-based there is a low Δm and high Δm selection

Signal Region Choice for Interpretation

Map of best signal region on m($\tilde{t}$) vs m($\tilde{\chi}_1^0$) plane → signal regions selected using best expected limit

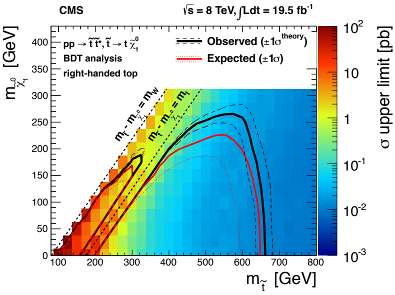
(These correspond roughly to the regions defined to target different signals)



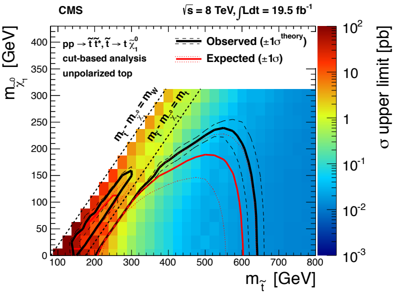


Interpretation of results for two analyses

BDT



Cut-based

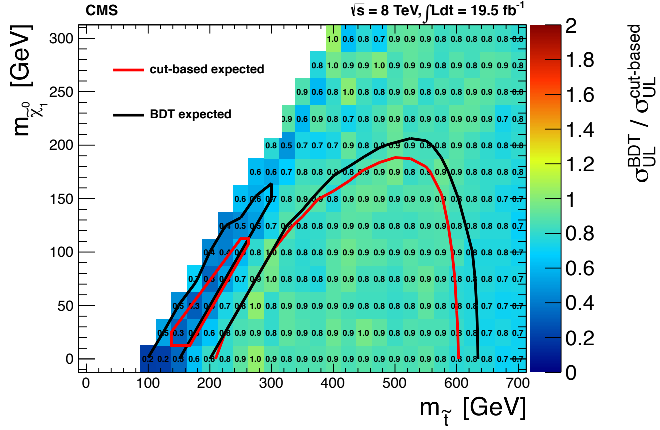


See a result of a clean comparison: MVA-based search goes to the corners not accessible to C&C:

low signal efficiency/high level of background ones!

Cut-based vs. BDT Result Comparison

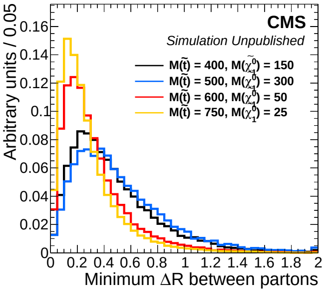
Ratio of expected cross section upper limit: BDT / cut-based



BDT is ∼10-30% better than cut-and-count in most of parameter space

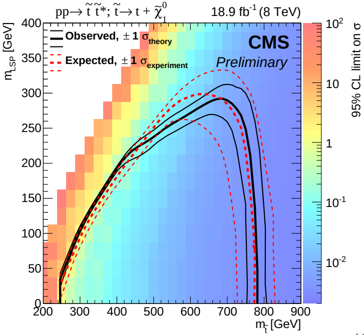
Another $\tilde{t}$ example

Targeted $\tilde{t}$ search pushes the boundaries:



- aims at all-hadronic final state for $\tilde{t} \to t\tilde{\chi}_1^0$ and $\tilde{t} \to bW\tilde{\chi}_1^0$
- operates in a $\tilde{t}$-search region with a range of boosts:
  - from unboosted top quarks at low ΔM
  - to merged top quarks at very high ΔM
- for this developed a new top quark reconstruction algorithm with varying cone size

- a set of MVA is trained for different kinematic regimes
- achieved sensitivity surpasses the one from previous 0ℓ+1ℓ combined: in both $\tilde{t}$ and $\tilde{\chi}_1^0$ masses
- up to 775 GeV for $\tilde{t}$ mass with $m_{\tilde{\chi}_1^0} < 200$ GeV
- up to 275 GeV for $\tilde{\chi}_1^0$ mass



CMS-PAS-SUS-13-023

More general new physics search?

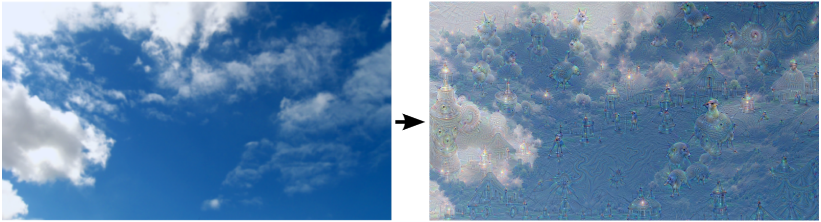
How do we define what we are looking for?

- in general, the tails of distributions:
  - high object multiplicity, high $E_{\mathrm{T}}^{\mathrm{miss}}$, H_{T} etc
- stealth SUSY:
  - rely on uncovering tags: hard initial state radiation, vector boson fusion topologies
- non-standard detector signatures:
  - multiply charged particles
  - disappearing tracks
  - displaced tracks
- anything else unthought of?

General question:

- how do we quantify what we see when applying complicated techniques

Non-physics example

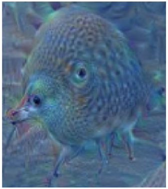


We ask the network: **“Whatever you see there, I want more of it!”**

This creates a feedback loop: if a cloud looks a little bit like a bird, the network will make it look more like a bird. This in turn will make the network recognize the bird even more strongly on the next pass and so forth, until a highly detailed bird appears, seemingly out of nowhere.

"Admiral Dog!"

"The Pig-Snail"



"The Camel-Bird"

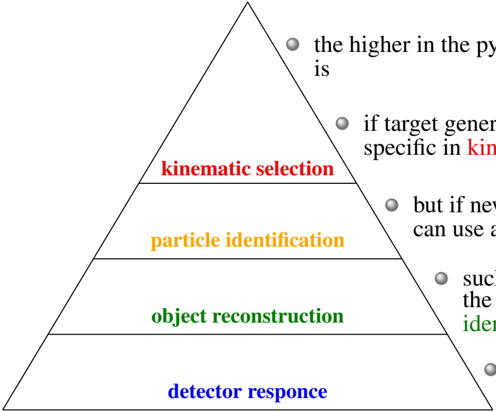
"The Dog-Fish"

inceptionism more of it some in Russian

Plenty of “new discoveries” in line!

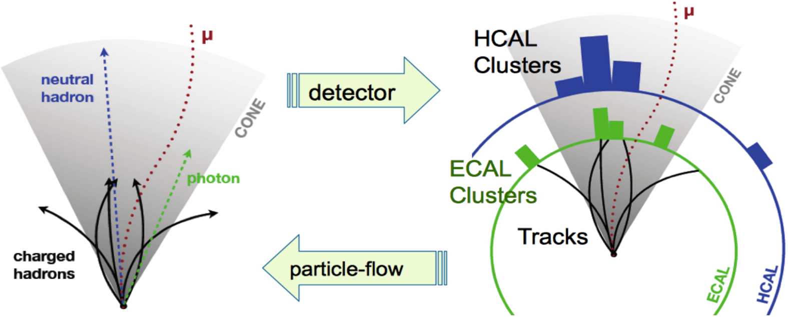
if apply advanced tools mindlessly

Resorting to lower levels



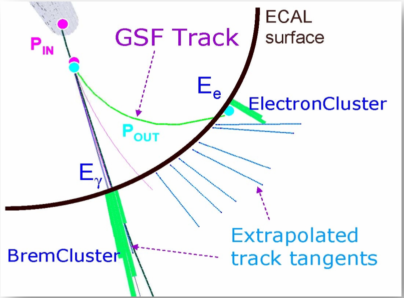
- the higher in the pyramid the more process-dependent a task is
- if target general new physics cannot be too specific in kinematics
- but if new physics decays to SM particles can use all advanced techniques there
- such hidden MVA are all over the place: electrons, muons, b-jet identification
- ensures low level of irreducible background

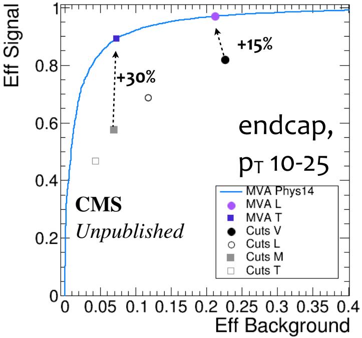
Particle Flow concept



- in practice we perform a sort of pattern recognition:
  - combine information from different subdetectors
  - and form a picture about which particles flew through a detector
- a separate optimization is done for each particle type based on physics knowledge

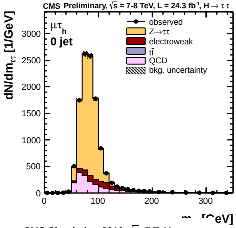
Example: electrons and photons





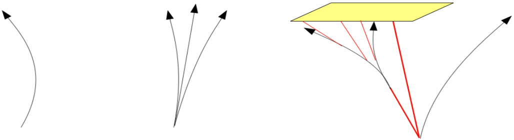
- cluster reconstruction in ECAL
  - common for both electrons and photons (electrons also reconstructed as photons)
  - designed to collect bremsstrahlung and conversions in extended phi region
- dedicated track reconstruction for electrons
  - "Gaussian sum" filter allows for tracks with large bremsstrahlung
- a range of detector-based variables to separate real electrons from misID jets
  - put together either in cut-based ID or MVA (BDT) ID: MVA can have 10-40% better efficiency with the same background rejection
  - huge gain in efficiency especially in multilepton events

Taus - even more complicated objects

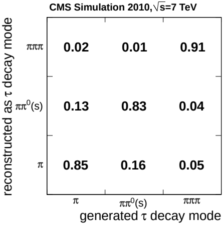


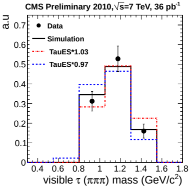
- not a stable particles: decays within the detector acceptance
- reconstructed in individual decay modes: higher level object
- requires both charged hadrons and electromagnetic objects
- series of MVA are trained to distinguish τ from jets or leptons:
  - automatically propagated to all analyses using τ

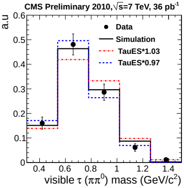
$\tau \to \pi\nu$ $\tau \to a_1\nu$ $\tau \to \rho\nu$



1-prong **3-prong** **1-prong+strip**







Conclusions

- To achieve a discovery big reduction rate of the backgrounds are needed
- Important to use all the features of our data to discriminate signal from background
- MVA are cautiously used in the new physics searches:
  - and usually only in well-motivated and understood scenarios
- But at the same time employed at many levels within the HEP framework:
  - Event level: Higgs searches, top events
  - Cone level: Tau vs jet reconstruction
  - Track level: particle identification
  - Lifetime and flavour tagging: b-tagging
- For a discovery, new tools are a two-way street:
  - can be used to enhance a cut-based hint
  - if signal seen first in MVA search: a cut-based analysis later to confirm the result

Plenty of discoveries are ahead: hopefully, mostly real ones!