Search for direct production of chargino pairs which decay via W-boson, with "mass splitting" near the electroweak scale, using data collected by the ATLAS experiment during Run 2 of LHC

Matteo Greco

Istituto Nazionale di Fisica Nucleare, Sezione di Lecce
Dipartimento di Matematica e Fisica "Ennio De Giorgi", Università del Salento

Supervisors: Prof. Edoardo Gorini, Dr. Margherita Primavera

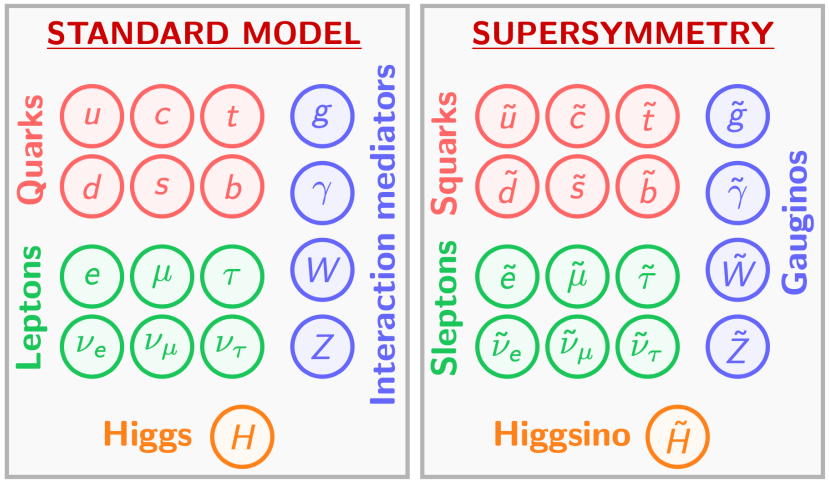
10th IDPASC School, 6-17 September, 2021

Introduction

Supersymmetry is an extension of the Elementary Particle Standard Model (SM)

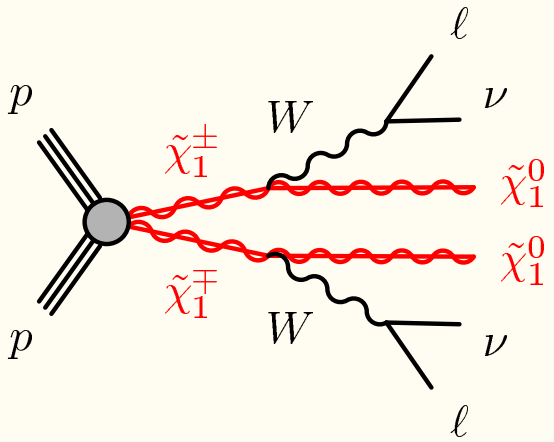
- For each SM particle it is associated a **superparter**, which differs of 1/2 in spin
- New quantic number, R-parity: $\boldsymbol{R = (-1)^{3(B-L)+2S}}$

$$R = \begin{cases} +1 & \text{for SM particles} \\ -1 & \text{for SUSY particles} \end{cases}$$



Minimal Supersymmetryc Standard Model (MSSM) is the simplest formulation:

- four **charginos** $\widetilde{\chi}_j^{\pm}$ $(j = 1, 2)$ and four **neutralinos** $\widetilde{\chi}_i^0$ $(i = 1, 2, 3, 4)$
- R-parity is conserved $\rightarrow$ $\boldsymbol{\widetilde{\chi}_1^0}$ **= Lightest Supersymmetric Particle (LSP)** $\rightarrow$ stable and a candidate of dark matter

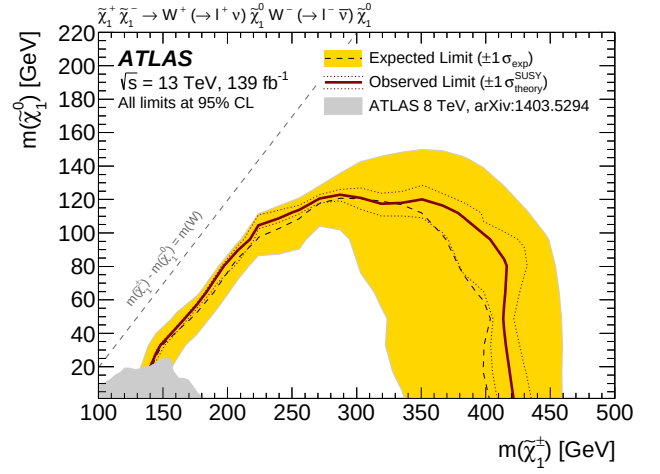


- Direct production of $\widetilde{\chi}_1^{\pm}$ pairs:

$$pp \rightarrow \widetilde{\chi}_1^+ \widetilde{\chi}_1^- \rightarrow W^+ \widetilde{\chi}_1^0 W^- \widetilde{\chi}_1^0$$

- Final state with 2 leptons ($\ell = e, \mu$) of *Opposite Sign* (OS) + missing transverse energy E_T^{miss} (due to ν and $\widetilde{\chi}_1^0$)
- Leptons can be *Same Flavour* (SF) or *Different Flavour* (DF)

State of art and current analysis



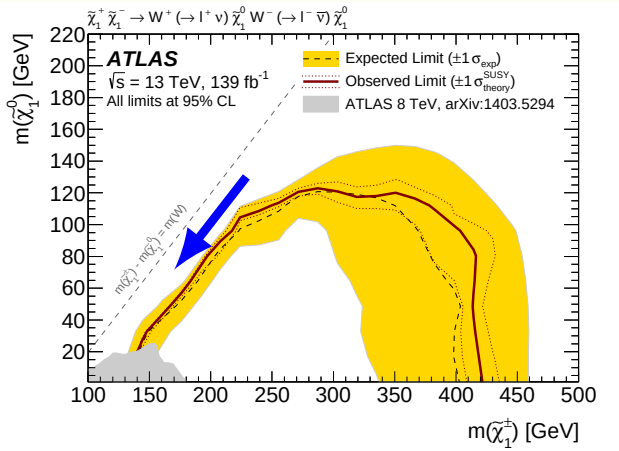
- Published results → Exclusion contour in the $(m(\widetilde{\chi}_1^{\pm}), m(\widetilde{\chi}_1^0))$ plane, by using data collected by ATLAS during Run2 of LHC, corresponding to 139 fb^{-1} of integrated luminosity and a center of mass energy of 13 TeV [1].
- Exclusion limits at 95% of CL have been set for chargino masses up to 420 GeV, in the hypothesis of massless neutralino

[1] ATLAS Collaboration, *Search for electroweak production of charginos and sleptons decaying into final states with two leptons and missing transverse momentum in $\sqrt{s}$=13 TeV pp collisions using the ATLAS detector*, Eur. Phys. J. C 80 123, 2020

- Current analysis's objective: study of the "less compressed" region of the mass spectrum $(m(\widetilde{\chi}_1^{\pm}), m(\widetilde{\chi}_1^0))$, with *mass splitting*:

$$\Delta m(\widetilde{\chi}_1^{\pm}, \widetilde{\chi}_1^0) = m(\widetilde{\chi}_1^{\pm}) - m(\widetilde{\chi}_1^0) \sim m_W$$

- In order to explore this region machine learning techniques (ML) have been developed → *Boosted Decision Tree* (BDT) with *Gradient Boosting*

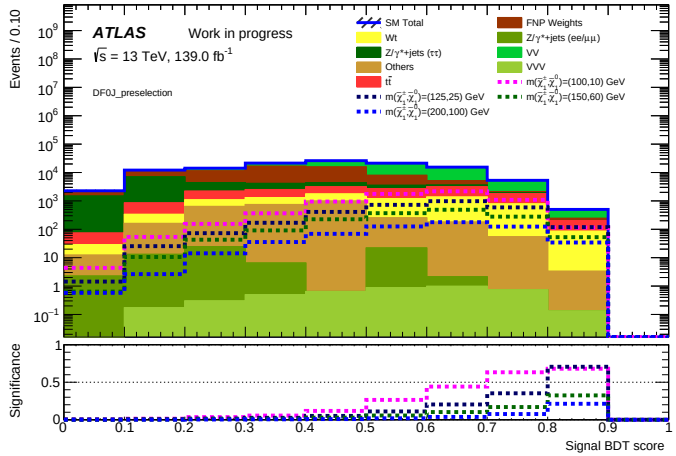


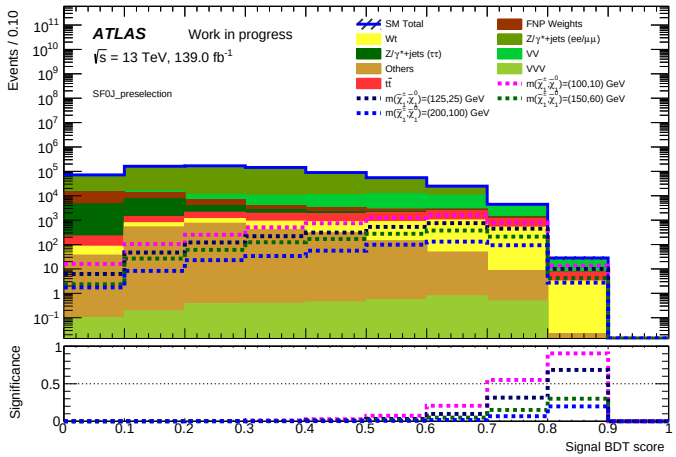
- 4 BDT variables to improve the signal-bkg discrimination: BDT Signal, BDT VV and BDT Top (dibosons, VV, and top, $t\bar{t} + Wt$, are the main bkgs), BDT Others (for the remaining bkgs)
- Two analyzed channels:
  - DF0J: DF leptons + no hadronic jets
  - SF0J: SF leptons + no hadronic jets

Signal Regions

- Definition of *Signal Regions* (SRs), by maximizing SUSY signal and by rejecting SM background
- Each SRs is a BDT Signal bin:
  - ◇ DF0J: 16 SRs starting from BDT Signal = 0.81
  - ◇ SF0J: 8 SRs starting from BDT Signal = 0.77

Regioni di Segnale (SR)	SR-DF-0J	SR-SF-0J
n_{jet}	=0	=0
$m_{\ell\ell}$ [GeV]	-	<76.2 or >106.2
E_T^{miss} significance	>8	>8
m_{T2} [GeV]	>50	>50
n_{bjet}	=0	=0
BDT Others	-	< 0.01
bin 1	∈[0.81 ,0.8125)	∈[0.77 ,0.775)
bin 2	∈[0.8125 ,0.815)	∈[0.775 ,0.78)
bin 3	∈[0.815 ,0.8175)	∈[0.78 ,0.785)
bin 4	∈[0.8175 ,0.82)	∈[0.785 ,0.79)
bin 5	∈[0.82 ,0.8225)	∈[0.79 ,0.795)
bin 6	∈[0.8225 ,0.825)	∈[0.795 ,0.80)
bin 7	∈[0.825 ,0.8275)	∈[0.80 ,0.81)
bin 8	∈[0.8275 ,0.83)	∈[0.81 ,1)
bin 9	∈[0.83 ,0.8325)	
bin 10	∈[0.8325 ,0.835)	
bin 11	∈[0.835 ,0.8375)	
bin 12	∈[0.8375 ,0.84)	
bin 13	∈[0.84 ,0.845)	
bin 14	∈[0.845 ,0.85)	
bin 15	∈[0.85 ,0.86)	
bin 16	∈[0.86,1)	





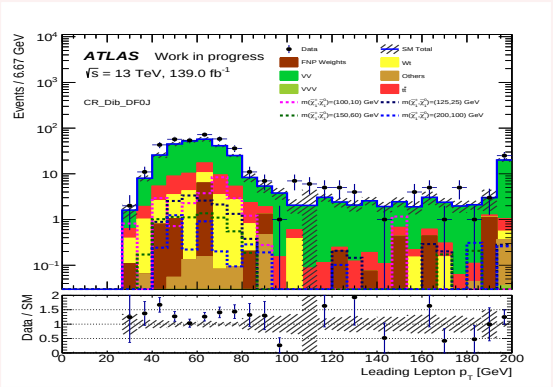
Signal BDT distribution for each of the analyzed channel, by applying only preselection cuts

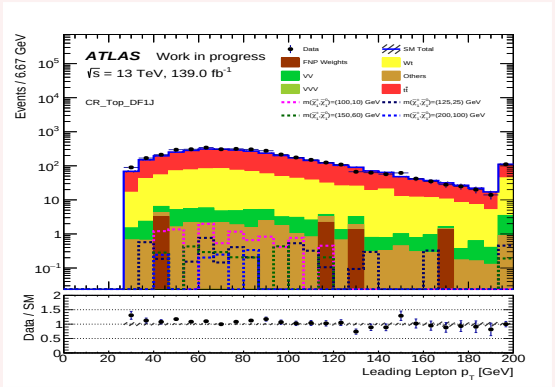
- SR "unblinding" → after having correctly estimated SM bkgs

Control and Validation Regions

- Definition of CRs to well estimate the main SM backgrounds which affect the analysis:
  - ◇ Dibosons (VV): regions characterized by 0 jet and 0 bjet, high VV purity and low signal contamination
  - ◇ Top ($t\bar{t} + Wt$): regions characterized by 1 jet and 1 bjet, high top purity and low signal contamination
- From CRs scale factors μ for the main backgrounds considered (VV and top) are extracted and applied to SRs: $\mu_{top} = 1.06 \pm 0.03$ and $\mu_{VV} = 1.39 \pm 0.08$

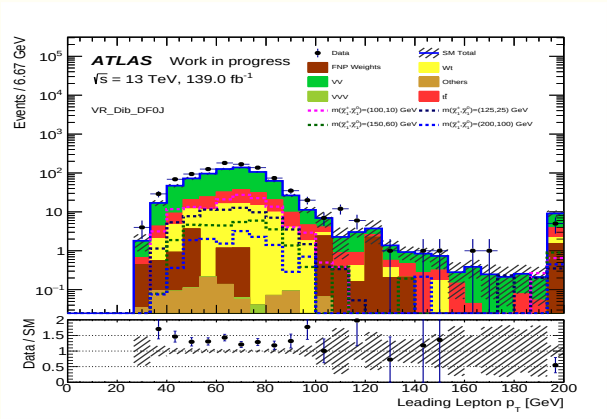
Region	CR_Dib_DF0J	CR_Dib_SF0J	CR_Top_DF1J	CR_Top_SF1J
isSF	=0	=1	=0	=1
n_{jet}	=0	=0	=1	=1
n_{bjet}	=0	=0	=1	=1
$m_{\ell\ell}$	>11	>11	>11	>11
$\lvert m_{\ell\ell}\text{-}m_Z \rvert$	-	>15	-	>15
E_T^{miss} significance	>8	>8	>8	>8
m_{T2}	>50	>50	>50	>50
BDT Signal	0.2-0.65	0.2-0.65	0.5-0.7	0.7-0.75
BDT Top	<0.1	<0.1	-	-
BDT VV	>0.2	>0.2	-	-
BDT Others	-	<0.01	-	<0.01





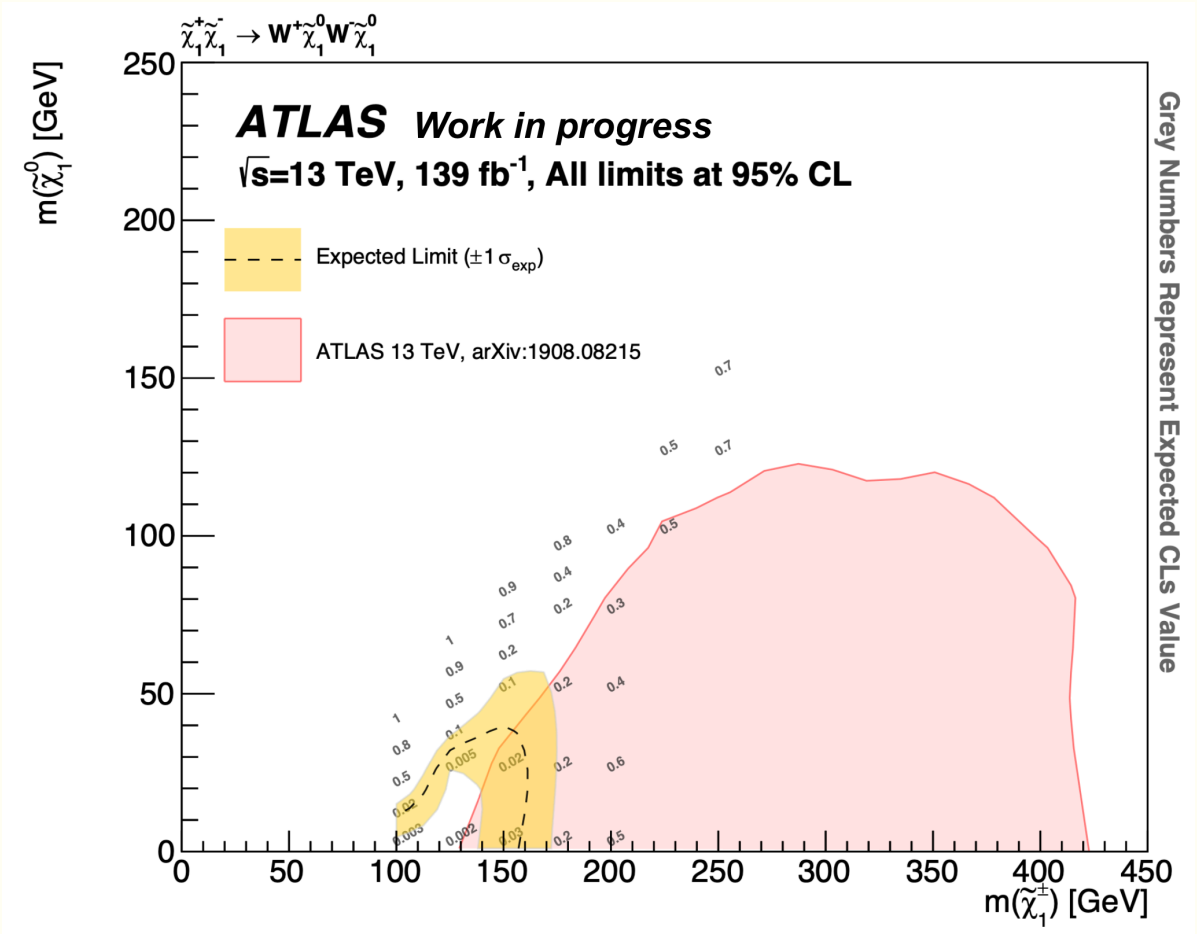
- We have defined VRs to validate scale factors from CRs

Region	VR_Dib_DF0J	VR_Dib_SF0J	VR_Top_DF1J	VR_Top_SF1J	VR_Top_DF0J	VR_Top_SF0J
isSF	=0	=1	=0	=1	=0	=1
n_{jet}	=0	=0	=1	=1	=0	=0
n_{bjet}	=0	=0	=1	=1	=0	=0
$m_{\ell\ell}$	>11	>11	>11	>11	>11	>11
$\lvert m_{\ell\ell}\text{-}m_Z \rvert$	-	>15	-	>15	-	>15
E_T^{miss} significance	>8	>8	>8	>8	>8	>8
m_{T2}	>50	>50	>50	>50	>50	>50
BDT Signal	0.65-0.81	0.65-0.77	0.7-1	0.75-1	0.5-0.81	0.5-0.77
BDT Top	<0.1	<0.1	-	-	-	-
BDT VV	>0.2	>0.2	-	-	<0.15	<0.15
BDT Others	-	<0.01	-	<0.01	-	<0.01



Expected exclusion limit

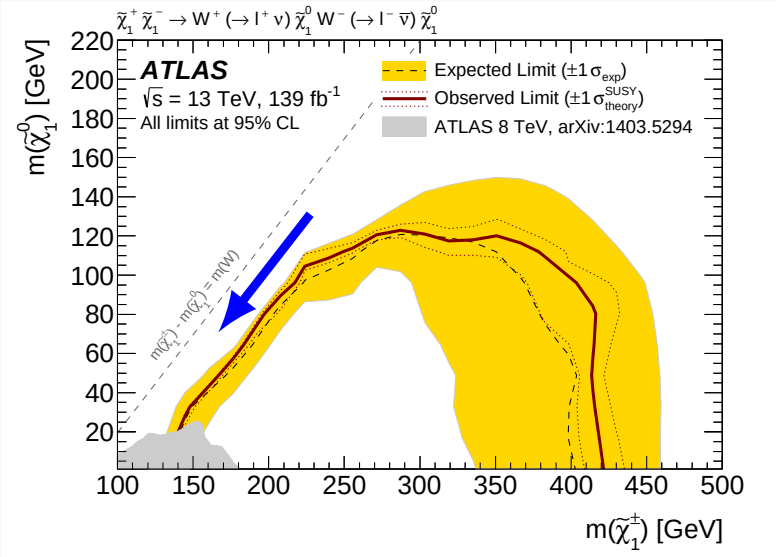
- Expected exclusion limit for chargino pair production decaying via W-boson
- SR "unblinding" is now in progress → if any significant excess of events wouldn't be observed → improvement of the current exclusion contour



THANKS FOR YOUR ATTENTION

BACKUP

The "less compressed" region



- Current analysis objective: study of the *less compressed* region of the mass spectrum ($m(\tilde{\chi}_1^\pm)$, $m(\tilde{\chi}_1^0)$), with *mass splitting*:

$$\Delta m(\tilde{\chi}_1^\pm, \tilde{\chi}_1^0) = m(\tilde{\chi}_1^\pm) - m(\tilde{\chi}_1^0) \sim m_W$$

- Data collected by the ATLAS experiment during Run 2 of LHC, corresponding to 139 fb^{-1} of integrated luminosity and to a center of mass energy of $\sqrt{s} = 13$ TeV
- The most relevant backgrounds in the analysis are di-bosons (VV) and top ($t\bar{t} + Wt$)

- In the previous scenario the analysis sensitivity has been optimized for a phase space with large chargino-neutralino mass splitting, by selecting events with two leptons in the final state, characterized by large invariant mass and *stransverse mass* (m_{T2}). In order to explore the *less compressed* region, machine learning techniques (ML) are being used → *Boosted Decision Tree* (BDT) with *Gradient Boosting*
- BDT variables are built by a training which uses a set of kinematic variables for signal-background discrimination (e.g. the transverse momentum of the two leptons, E_T^{miss}, m_{T2} and the angles in the transverse plane between E_T^{miss} and the two leptons)
- Use of four BDT variables to improve the signal-background discrimination: BDT Signal, BDT VV, BDT Top, BDT Others (for the remaining background)
- Two channel analyzed:
  - ◇ DF0J: DF leptons + no hadronic jets
  - ◇ SF0J: SF leptons + no hadronic jets

Research strategy

- **SUSY models**: the analysis is based on *simplified models*: few new particles are included, with masses as free parameters. The relative decay BRs are supposed to be 100%
- **Event selection**: selective criteria are applied on the relevant variables in order to define *Signal Regions* (SRs), by maximizing the SUSY signal and by rejecting the SM backgrounds
- **Background estimation**:
  - ◇ Monte Carlo (MC) simulations of the relevant SM backgrounds are normalized to data in *Control Regions* (CRs) $\rightarrow$ *scale factors* μ extrapolation to the SRs
  - ◇ the normalization is verified into *Validation Regions* (VRs)
  - ◇ CRs and VRs are ortogonal to SRs
- **Results**: SR "unblinding" and data-expected bkg comparison, after having optimized the background estimation strategy
- **Interpretation**:
  - ◇ if a significant excess of data over the expected background is observed $\rightarrow$ signal **evidence/discovery**
  - ◇ otherwise an **upper limit** on the SUSY production cross section is set and limits on some other considered model parameters are derived

SM backgrounds, which enter inside the SRs, are classified in two categories:

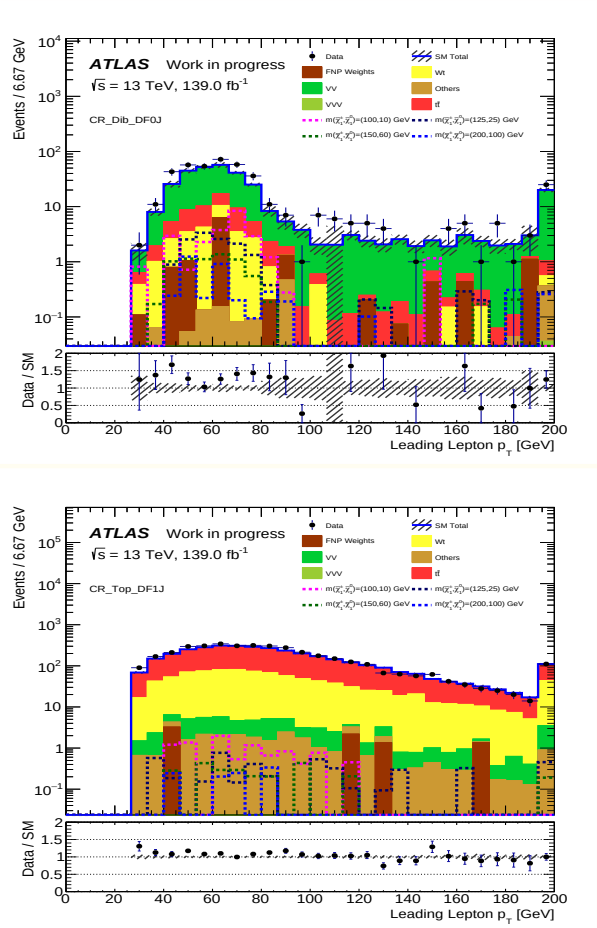
- **Riducible**: *fake-non-prompt lepton* background, estimated by using *fully data driven* techniques
- **Unreducible**: estimated via *semi data driven* techniques, based on CRs and VRs definition. In yellow in table are highlighted the most relevant unreducible backgrounds

SM backgrounds
$t\bar{t}$
$t\bar{t} + V$ $(V = W, Z)$
$t\bar{t} + WW$
tZ, $t\bar{t}t\bar{t}$, $t\bar{t}t$
Single top Wt
$Z/\gamma^* +$ jets
Dibosons VV, $(V = W, Z)$
Tribosons VVV $(V = W, Z)$
Higgs

Control Regions

- CRs are ortogonal to SRs and to VRs
- CRs are defined for the relevant backgrounds which affect the analysis:
  - ◇ Dibosons (VV): regions characterized by 0 jet and 0 bjet, high VV purity and low signal contamination
  - ◇ Top ($t\bar{t} + Wt$): regions characterized by 1 jet and 1 bjet, high top purity and low signal contamination

Region	CR_Dib_DF0J	CR_Dib_SF0J	CR_Top_DF1J	CR_Top_SF1J
isSF	=0	=1	=0	=1
n_{jet}	=0	=0	=1	=1
n_{bjet}	=0	=0	=1	=1
$m_{\ell\ell}$	>11	>11	>11	>11
$\lvert m_{\ell\ell}\text{-}m_Z \rvert$	-	>15	-	>15
E_T^{miss} significance	>8	>8	>8	>8
m_{T2}	>50	>50	>50	>50
BDT Signal	0.2-0.65	0.2-0.65	0.5-0.7	0.7-0.75
BDT Top	<0.1	<0.1	-	-
BDT VV	>0.2	>0.2	-	-
BDT Others	-	<0.01	-	<0.01
FNP	0.00±9.15	0.00±5.66	0.00±24.07	0.00±9.03
Wt	21.31±1.72	15.85±1.45	906.27±10.86	153.94±4.52
Zjets	0.00±0.00	1.43±3.71	0.00±0.00	0.00±1.32
Zttjets	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
VV	257.38±5.72	118.76±4.22	41.43±2.26	8.80±1.02
other	1.44±0.57	0.23±0.09	25.01±1.05	4.21±0.41
VVV	0.09±0.01	0.05±0.01	0.04±0.01	0.01±0.01
ttbar	35.06±1.19	29.01±1.09	2689.75±10.09	380.70±3.84
Total	315.28±11.01	165.33±8.18	3662.50±28.37	547.66±10.94
Data	424	208	3920	548
Purity	(81.64±3.38)%	(71.84±4.38)%	(98.18±0.72)%	(97.62±1.85)%
C1C1 100 10	24.49±3.49	8.94±2.28	9.38±2.50	2.88±1.22
contamination	(7.21±1.06)%	(5.13±1.33)%	(0.26±0.07)%	(0.52±0.22)%
C1C1 125 25	14.35±1.68	9.79±1.41	4.57±1.05	1.40±0.57
contamination	(4.35±0.53)%	(5.59±0.85)%	(0.12±0.03)%	(0.25±0.10)%
C1C1 150 60	7.03±1.13	3.85±0.84	2.54±0.76	0.19±0.19
contamination	(2.18±0.36)%	(2.28±0.51)%	(0.07±0.02)%	(0.03±0.03)%
C1C1 200 100	4.06±0.89	5.84±1.00	1.03±0.51	0.19±0.14
contamination	(1.27±0.28)%	(3.41±0.60)%	(0.03±0.01)%	(0.03±0.02)%

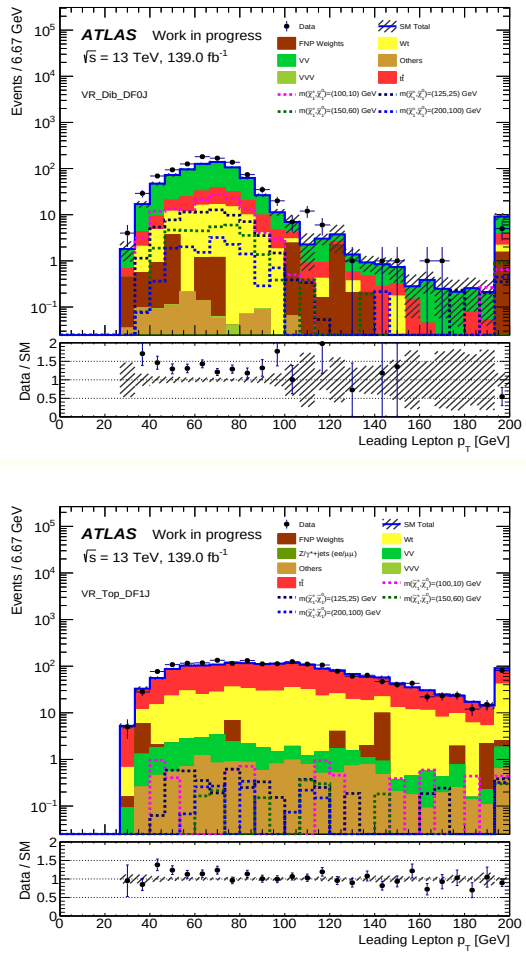


Data/Monte Carlo distributions of the leading lepton transverse momentum in CR_Dib_DF0J and in CR_Top_DF1J

Validation Regions

- VRs ortogonal to CRs and to SRs and kinematically closed to SRs

Region	VR_Dib_DF0J	VR_Dib_SF0J	VR_Top_DF1J	VR_Top_SF1J	VR_Top_DF0J	VR_Top_SF0J
isSF	=0	=1	=0	=1	=0	=1
n_{jet}	=0	=0	=1	=1	=0	=0
n_{bjet}	=0	=0	=1	=1	=0	=0
$m_{\ell\ell}$	>11	>11	>11	>11	>11	>11
$\lvert m_{\ell\ell}\text{-}m_Z \rvert$	-	>15	-	>15	-	>15
E_T^{miss} significance	>8	>8	>8	>8	>8	>8
m_{T2}	>50	>50	>50	>50	>50	>50
BDT Signal	0.65-0.81	0.65-0.77	0.7-1	0.75-1	0.5-0.81	0.5-0.77
BDT Top	<0.1	<0.1	-	-	-	-
BDT VV	>0.2	>0.2	-	-	<0.15	<0.15
BDT Others	-	<0.01	-	<0.01	-	<0.01
FNP	0.00±14.19	7.79±10.90	0.00±17.14	4.23±5.75	20.47±12.21	0.05±1.40
Wt	86.65±3.43	69.34±3.06	473.85±7.85	25.54±1.84	154.95±4.55	3.21±0.64
Zjets	0.00±0.00	67.53±34.13	0.02±0.02	0.04±0.03	0.00±0.00	0.00±0.00
Zttjets	0.00±0.00	0.24±0.24	0.00±0.00	0.00±0.00	0.50±0.50	0.00±0.00
VV	528.29±7.55	290.64±5.81	23.34±1.70	1.61±0.42	310.88±6.28	5.90±0.81
other	0.99±0.20	1.16±0.23	13.56±0.76	0.75±0.19	3.53±0.72	0.05±0.04
VVV	0.06±0.01	0.09±0.02	0.03±0.01	0.00±0.00	0.16±0.02	0.01±0.00
ttbar	106.91±2.05	102.95±2.04	1279.36±6.97	63.63±1.57	241.21±3.08	5.39±0.45
Total	722.90±16.56	539.74±36.48	1790.15±20.19	95.80±6.25	731.71±14.81	14.59±1.79
Data	972	593	1910	95	810	17
Purity	(73.08±1.97)%	(53.85±3.80)%	(97.94±1.04)%	(93.09±5.31)%	(54.14±1.09)%	(58.90±7.51)%
C1C1 100 10	133.82±7.53	94.63±6.48	5.94±1.66	0.60±0.43	134.31±7.59	1.69±0.85
contamination	(15.62±0.94)%	(14.92±1.34)%	(0.33±0.09)%	(0.62±0.44)%	(15.51±0.93)%	(10.38±5.35)%
C1C1 125 25	73.21±3.79	60.35±3.48	4.74±1.01	0.38±0.28	101.16±4.48	2.56±0.72
contamination	(9.20±0.51)%	(10.06±0.84)%	(0.26±0.06)%	(0.40±0.29)%	(12.15±0.58)%	(14.93±4.54)%
C1C1 150 60	33.89±2.46	24.73±2.10	1.67±0.60	0.43±0.30	46.77±2.90	1.55±0.52
contamination	(4.48±0.34)%	(4.38±0.47)%	(0.09±0.03)%	(0.44±0.32)%	(6.01±0.39)%	(9.60±3.44)%
C1C1 200 100	14.43±1.71	14.58±1.67	1.70±0.59	0.26±0.26	26.12±2.29	0.84±0.46
contamination	(1.96±0.24)%	(2.63±0.35)%	(0.09±0.03)%	(0.27±0.27)%	(3.45±0.31)%	(5.44±3.06)%



Data/Monte Carlo distributions of the leading lepton transverse momuntum in VR_Dib_DF0J and in VR_Top_DF1J

- From CRs scale factors μ for the main backgrounds considered (VV and to) are extracted to VRs and to SRs: $\mu_{top} = 1.06 \pm 0.03$ and $\mu_{VV} = 1.39 \pm 0.08$
- Once backgrounds have been correctly estimated, the SR "unblinding" follows