TÉCNICO LISBOA

LABORATÓRIO DE INSTRUMENTAÇÃO E FÍSICA EXPERIMENTAL DE PARTÍCULAS
*partículas e tecnologia*

# Search for top squarks in the four-body decay mode with single lepton final states in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$

Dario Vaccaro

13 July 2023

# Overview: What is Supersymmetry?

Symmetries are fundamental concepts in Quantum Field Theory.

- The Standard Model (SM) describes the matter as composed by fermions, and the interactions as mediated by bosons. Fermions have half-integer spin, bosons have integer spin.
- It is well known that the space-time respects certain fundamental symmetries, such as Lorentz invariance (more generally Poincaré invariance), CPT, and others ...
- It is reasonable to argue a new kind of symmetry between matter and interactions, which is a symmetry between bosons and fermions.
- SUSY predicts that every boson (fermion) has a correspondent fermion (boson) particle, called *sparticle*.

# Overview: Why Supersymmetry?

## Limitations of the Standard Model

Nonetheless the SM describes very accurately phenomena over several orders of magnitude, certain theoretical and experimental reasons suggest it is not the most fundamental model of the nature:

- Neutrino oscillations imply neutrinos have mass. This is not included in the SM.
- The matter-antimatter asymmetry observed in the universe needs a theoretical explanation.
- From astronomic motivations, it is thought that universe is composed by about 5% of matter, and the remaining is dark matter and dark energy. These two are not included in the SM.
- Currently there is no confirmed quantum gravity theory. Found it is probably the dream of every theoretical physicist.

# Overview: Why Supersymmetry?

## Theoretical and Phenomenological motivations

There are several reasons to think about SUSY as a possible remedy of the previous issues:

Profound motivations:

- "If something can happens, it will be happen".
- It's mathematically elegant and beautiful.

Other motivations:

- Hierarchy Problem: it explain in elegant way the magnitude scale of the Higgs mass, by solving the hierarchy problem.
- Dark Matter: it provides candidate particles to describe dark matter.
- Gauge coupling unification: Gauge couplings scale with energy. All the couplings meet at a certain magnitude value in supersymmetric models.

# SUSY: Lagrangian

The construction of the supersymmetric Lagrangian follows the same recipe as the SM Lagrangian:

- Start by choosing the gauge group G of the symmetries respected by the model:
$$U(1) \times SU(2) \times SU(3)\,.$$
- Group SM fields into superfields. Two SM fields cannot be grouped in a single superfield since they do not share quantum numbers.
- Write the most general Lagrangian invariant under G which couples all these fields:
$$\mathcal{L} = \mathcal{L}_{KE} + \mathcal{L}_{INT} + \mathcal{L}_{W}\,,$$
which are respectively the kinetic, the interaction, and the superpotential terms.

# SUSY: Broken Symmetry

SUSY has to face a big problematic issue.

- To solve the hierarchy problem, particles must have the same mass as the supersymmetric ones. That's is in contrast with experimental data, since we have never detected one single supersymmetric particle.

To address the problem, it is necessary to introduce a spontaneous symmetry breaking.

**SUSY is a broken symmetry!**

# SUSY: MSSM

The easiest way to address the problem is with the Minimal SuperSymmetric Model (MSSM).

- It is the simpler supersymmetric model. Other theories have more parameters and particles.
- Introduces a "soft" supersymmetry breaking term into the Lagrangian.
- Terms are soft because they break the supersymmetry but not so much as to reintroduce the quadratic divergence which motivated Supersymmetry to start.

Then, the MMSM Lagrangian is

$$\boxed{\mathcal{L} = \mathcal{L}_{KE} + \mathcal{L}_{INT} + \mathcal{L}_{W} + \mathcal{L}_{soft}\,.}$$

There are 124 free parameters... John Von Neumann once said "With four parameters I can fit an elephant, and with five I can make him wiggle his trunk". I can't imagine what he could do with 124!

# Searching for SUSY: Channels

There are different modes of searching for SUSY, such as

- Supersymmetric top decay modes: Search for stop is an interesting avenue because the stop could be the lightest squark. We focus on the situation where stop is next-to-lightest supersymmetric particle (NLSP) and neutralino is the LSP.
- Supersymmetric tau decay modes: Stau is probably the lightest supersymmetric lepton. It does not depend on QCD, so limits set on stau impose limits on MSSM independent of colour sector.
- The structure of MSSM requires 2 Higgs superfields. Then, there are 5 different Higgs fields, and we can search their decay modes.

From now on, we will focus on a particular stop four-body decay mode.

# Searching for SUSY: Top squarks in the four-body decay

We are going to discuss about the following paper:

J. High Energ. Phys. 2023, 60 (2023).
https://doi.org/10.1007/JHEP06(2023)060

Here, a search for the pair production of the lightest supersymmetric partner of the top quark, the top squark ($\tilde{t}_1$), in the four-body decay mode, is presented.

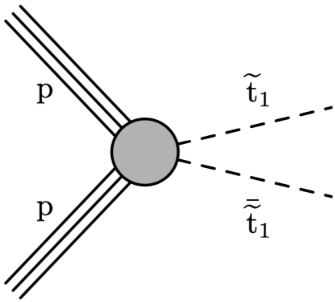
- In the paper the group combine the previous 2016 result of CMS at $\sqrt{s} = 13\,\text{TeV}$ with data recorded at 2017 and 2018. The total integrated luminosity for the combined 2016–2018 analysis is $138\,\text{fb}^{-1}$.

# Searching for SUSY: Why the stop four-body decay?

1. If $R$-parity is conserved, supersymmetric particles would be produced in pairs, and their decay chains would end with the lightest supersymmetric particle (LSP), often considered to be the lightest neutralino $\tilde{\chi}_1^0$.
2. When the symmetry is broken, the scalar partners of an SM fermion acquire a mass different from the mass of the SM partner, with the mass splitting between scalar mass eigenstates being proportional to the mass of the SM fermion. Since the top quark is the heaviest fermion of the SM, the lighter supersymmetric scalar partner of the top quark, the top squark, could therefore be the lightest squark.
3. If SUSY is realized in nature, cosmological observations imply that the lightest top squark is almost degenerate with the LSP neutrino.
4. In this scenario, because the mass difference between the top squark, $\tilde{t}_1$, and the neutralino, $\tilde{\chi}_1^0$, is smaller than the mass of the W boson, other decays modes are either forbidden or supressed. This motivates the search for the four-body decay mode!

# Stop Four-body decay: stop-antistop production

We want to justify the last assumption, and show that other decay modes are forbidden.




We consider a collision between two protons producing a final state composed by stop-antistop pair.

# Stop Four-body decay: First stop decay

- In our model, the neutralino is LSP, and the quark stop is NLSP. Stop decays to a top and a neutralino (R-parity conservation and baryon number conservation):

$$\tilde{t}_1 \to t\,\tilde{\chi}^0_1\,.$$

- Because the LSP and NLSP mass difference is smaller than the top quark mass, the decay of the stop to a real top and a neutralino is not kinematically allowed. The top quark could only be virtual, and it decays in $W$ boson and $b$ quark: we have a three-body decay:

$$\tilde{t}_1 \to \tilde{\chi}^0_1\,W^+\,b\,.$$

- If mass difference is also smaller than the W boson mass, the three-body decay is not kinematically allowed as well. Then, the $W$ boson will decay, and the final state is

$$\boxed{\tilde{t}_1 \to \tilde{\chi}^0_1\,f\,\nu_f\,b}$$

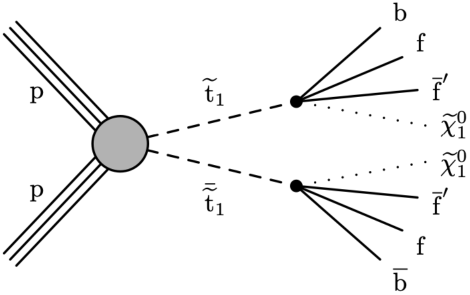
# Stop Four-body decay: Second stop decay

- We consider the situation where the second stop undergoes a four body decay as well, but with two quarks in the final state:

$$\overline{\tilde{t}_1} \to \tilde{\chi}_1^0\, q\, \overline{q}\, b$$

- The reason of this final states, where one top squark decays into a lepton, is motivated by the decrease of the contributions from the multijet background in this mode, while increasing the selection efficiency with the other top squark decaying hadronically.
- The selected jet, attributed to initial-state radiation (ISR) of a parton, is required to have high transverse momentum ($p_T$). Both neutralinos and the neutrino (in the previous decay) escape undetected, leaving high missing transverse momentum.

# Stop Four-body decay: full diagram




$$\overline{\tilde{t}_1} \to \tilde{\chi}_1^0 \, q \, \overline{q} \, b$$

$$\tilde{t}_1 \to \tilde{\chi}_1^0 \, f \, \nu_f \, b$$

# Stop Four-body decay: signature

In order to exclude SM processes and identify clearly the four-body stop decay, the experimental signature for the process is chosen to be:

- One single charged lepton in the final state.
- Presence of large transverse momentum missing, $p_T^{\text{miss}}$, due to the neutralinos and neutrino.
- Presence of jets in final state, due to the hadronization of the final quarks.

# Stop Four-body decay: MC simulations

Background and signal are generated using MC generators. The processes contributing to the background are:

- W + jets, Z $\rightarrow \nu\overline{\nu}$ + jets, and multijet processes, generated at LO.
- The $t\bar{t}$ process, generated at NLO.
- Single top and associated $tW$ production, and diboson events, simulated at NLO.
- Additional backgrounds such as $t\bar{t}$ produced in association with a Z boson, W boson, or photon, referred to as $t\bar{t}X$, generated at NLO.

All SM MC events are passed through a full simulation of the CMS apparatus, where the response of the detector is modeled using the Geant4 software. Generated events are processed using the same version of the CMS event reconstruction software used as for data.

# Analysis: Event reconstruction

Data reconstruction

- Data and simulated events are reconstructed using the CMS particle-flow (PF) algorithm. The primary vertex (PV) is taken to be the vertex corresponding to the hardest scattering in the event, evaluated using tracking information alone.

Jets identification

- Jets are reconstructed by applying the anti-$k_T$ clustering algorithm to PF candidates with a distance parameter of 0.4. Jets are required to satisfy $p_T > 30\,\mathrm{GeV}$, and $|\eta| < 2.4$. The tagging of b jets is performed by taking into account information from the secondary vertex and using a deep neural network algorithm.

The missing transverse momentum vector is computed as the negative vector sum of all PF candidates in the event.

# Analysis: Event reconstruction

Electrons identification

- Electron candidates are reconstructed from energy deposits in the ECAL and matched charged particle tracks in the inner tracker. To reduce the number of misidentified electrons, additional constraints on the shape of the electromagnetic shower in the ECAL, the quality of the match between the trajectory of the track, and the ECAL energy deposit around the electron, and the relative HCAL deposit in the electron direction are applied.
- Electrons are required to have $p_T$ above 5 GeV and $|\eta| < 2.5$, with a veto on electron candidates in the ECAL gap region $(1.4442 < |\eta| < 1.5660)$.

# Analysis: Event reconstruction

Muons identification

- Muon candidates are reconstructed by combining the information from the silicon tracking systems and the muon spectrometer in a global fit that assigns a quality to the matching between the tracker and muon systems and imposes minimal requirements on the track to reduce the misidentification of muons.
- Muons are required to pass the selection requirements of $p_T > 3.5\,\mathrm{GeV}$ and $|\eta| < 2.4$.

# Analysis: Event selection

The event selection is performed firstly with the online selection, due to the trigger, and then with the offline selection, by applying preselecting cuts to reduce the contribution of the main background processes and then by using a boosted decision trees (BDTs) algorithm in order to define the signal selection.

## Trigger

- The data events collected by the trigger system are required to have both $p_T^{\text{miss}}$ and $H_T^{\text{miss}}$ above 120 GeV, where $H_T^{\text{miss}}$ is the magnitude of the missing transverse momentum calculated only from jets.
- In order to maintain the performance of the online selection with increased luminosity from the late runs of 2017 onward, the condition $H_T > 60$ GeV is also required, where $H_T$ is defined as the scalar $p_T$ sum of all jets in the event.

# Analysis: Event selection

Preselection

- We want a scenario in which two $\tilde{\chi}_1^0$'s escape the detention, and where the missing momentum is larger than for SM processes, then $p_T^{\text{miss}} > 280\,\text{GeV}$.
- To suppress the $W +$ jets background, we require $H_T > 200\,\text{GeV}$.
- To select the single-lepton topology, it is demanded exactly one identified electron or muon in the event, along with at least one jet. Then, dilepton topology $t\bar{t}$ events contribution is suppressed.
- To further improve the selection of signal over SM background events, at least one jet must have $p_T > 110$ GeV.
- In events with at least two jets, the azimuthal angle between the directions of the leading and second-highest-$p_T$ jets must be smaller than 2.5 radians, suppressing the SM multijet background.

# Analysis: Preselection Dominant Background

After the preselection, the situation is the following:

- The W + jets and $t\bar{t}$ processes are the main SM backgrounds, making up about 70 and 20%, respectively, of the total expected background.
- Another SM background process is the $Z \rightarrow \nu\bar{\nu}$+jets, since having jets, genuine missing momentum, and a jet misidentified as a lepton.
- The remaining background processes are diboson, single top quark, Drell-Yan (DY), multijet, and $t\bar{t}X$ production where $X$ is a vector boson. These processes are more suppressed because having a smaller cross section, a lower acceptance, or both.

# Analysis: Preselection Plots

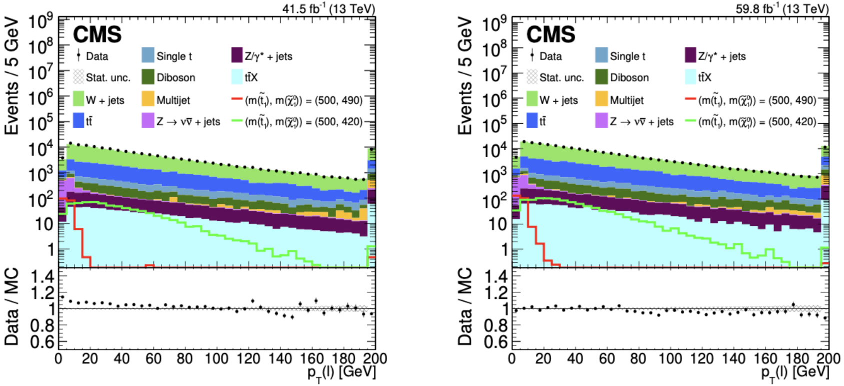



**Figure:** Distributions of $p_T(l)$ after the preselection from 2017 (left) and 2018 (right) data (points) and simulation (colored histograms).

# Analysis: Signal selection

After the preselection, the event selection continues using a BDT machine learning algorithm, in order to define the signal:

- Based on a BDT, to take advantage of the different correlations among the discriminating variables for the signal and background processes.
- For each event passing the preselection, the BDT discriminator value is evaluated. If it exceeds the determined threshold, the event is retained.
- Various BDTs are trained with different sets of discriminating variables, and a variable is included in the final set only if it significantly increases the figure of merit (FOM) obtained for any selection using the BDT output.

# Analysis: Background evaluation

The primary sources of background are estimated using data, and validated with MC simulations:

- Processes with a prompt lepton: W + jets and $t\bar{t}$. A lepton is defined non-prompt either when it does not originate from the PV, or when a jet is misidentified as a lepton.
- Processes contributing to the non-prompt background: Events where the lepton arises from the decay of heavy-flavor quarks or from misidentified hadrons, Z $\rightarrow \nu\bar{\nu}$ + jets, W + jets and $t\bar{t}$, where a jet is misidentified as a lepton, as well as the multijet background. There can also be events in which a genuine lepton (mainly from W + jets or $t\bar{t}$) escapes detection, while a non-prompt lepton is selected.

The background from other SM processes are estimated from simulations, such as single top quark, diboson, DY, and $t\bar{t}X$ production.

# Final Results: Results

The analysis is specifically trained for different $\Delta m$ regions, the difference between stop and neutralino's masses, thus adapting the signal selection to the evolution of the kinematical variables as a function

- There is good agreement between the observed and predicted numbers of events for all SRs. The largest difference is for $\Delta m = 10$ GeV, where there are 1.1 and 2.9 standard deviations excesses of data events over the predicted background for the 2017 and 2018 data respectively. The 2016 analysis had a similar excess for the same $\Delta m$ value, corresponding to 0.7 standard deviations. None of these excesses is statistically significant, so it is concluded that there is no evidence for direct top squark production.
- Limits of the previous analysis are improved. At low $\Delta m$ the top squark mass limit is 60 GeV higher, while at high $\Delta m$ the top squark mass limit is extended by 140 GeV.

# Final Results: Table of data

| Year | $\Delta m$ (GeV) | BDT > | $Y_{\rm p}^{\rm SR}$(W+jets) | $Y_{\rm p}^{\rm SR}$($\rm t\bar{t}$) | $Y_{\rm np}^{\rm SR}$ | $N^{\rm SR}$(Other) | $N^{\rm SR}$(Total) | $N^{\rm SR}$(Observed) |
|---|---|---|---|---|---|---|---|---|
| 2017 | 10 | 0.31 | 11.0 ± 2.4 | 2.2 ± 2.9 | 20.1 ± 3.5 | 5.4 ± 3.7 | 38.8 ± 6.3 | 49 |
| | 20 | 0.32 | 37.4 ± 4.6 | 3.3 ± 5.2 | 49.6 ± 7.0 | 18.4 ± 9.3 | 109 ± 14 | 116 |
| | 30 | 0.38 | 23.8 ± 3.8 | 0.0 ± 7.2 | 41.7 ± 6.1 | 19.4 ± 9.9 | 85 ± 14 | 86 |
| | 40 | 0.40 | 15.9 ± 2.6 | 0.0 ± 8.1 | 32.6 ± 5.5 | 20 ± 10 | 69 ± 15 | 66 |
| | 50 | 0.43 | 10.9 ± 2.0 | 0.0 ± 6.7 | 22.3 ± 4.0 | 17.9 ± 9.2 | 51 ± 12 | 48 |
| | 60 | 0.47 | 3.9 ± 0.8 | 0.0 ± 6.2 | 7.6 ± 2.2 | 10.3 ± 5.4 | 21.8 ± 8.5 | 23 |
| | 70 | 0.39 | 11.1 ± 2.0 | 8.9 ± 7.6 | 12.9 ± 2.9 | 19.7 ± 9.8 | 53 ± 13 | 50 |
| | 80 | 0.41 | 15.6 ± 4.3 | 10.3 ± 9.7 | 8.3 ± 2.2 | 17.1 ± 8.2 | 51 ± 14 | 51 |
| 2018 | 10 | 0.32 | 17.3 ± 4.3 | 0.0 ± 2.4 | 16.7 ± 3.6 | 7.1 ± 4.5 | 41.1 ± 7.6 | 77 |
| | 20 | 0.39 | 18.4 ± 2.8 | 0.3 ± 3.1 | 14.5 ± 3.4 | 6.3 ± 3.5 | 39.4 ± 6.4 | 57 |
| | 30 | 0.35 | 48.5 ± 8.1 | 9.1 ± 9.4 | 22.5 ± 4.8 | 33 ± 14 | 114 ± 19 | 127 |
| | 40 | 0.43 | 10.7 ± 3.1 | 3.4 ± 4.5 | 11.7 ± 2.9 | 12.3 ± 6.7 | 38.1 ± 9.1 | 49 |
| | 50 | 0.46 | 8.7 ± 3.0 | 3.4 ± 4.5 | 10.5 ± 2.8 | 10.3 ± 5.2 | 32.9 ± 8.0 | 36 |
| | 60 | 0.41 | 16.5 ± 4.7 | 16.2 ± 8.8 | 17.3 ± 3.8 | 22 ± 10 | 72 ± 15 | 61 |
| | 70 | 0.40 | 35.6 ± 8.7 | 15.2 ± 8.6 | 16.9 ± 5.2 | 30 ± 12 | 97 ± 18 | 96 |
| | 80 | 0.42 | 16.3 ± 3.7 | 10.9 ± 7.8 | 10.7 ± 4.3 | 21.5 ± 9.8 | 59 ± 14 | 41 |

**Figure:** The predicted number of $W^+$ jets, $\rm t\bar{t}$, non-prompt, and other $N^{\rm SR}$(Other) background events and their sum $N^{\rm SR}$(Total), in the eight SRs for the 2017 and 2018 data analysis. The first three predicted yields are derived from data, while the yields of the other background processes come from simulation.

# Final Results: Plots

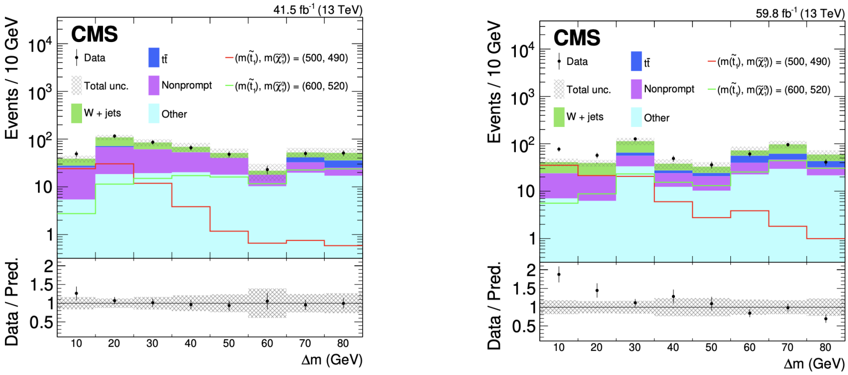



**Figure:** The observed yields in data (points) and the predicted background components (colored histograms) in the eight SRs for the 2017 (left) and 2018 (right) data.

# Final Results: Final Plot

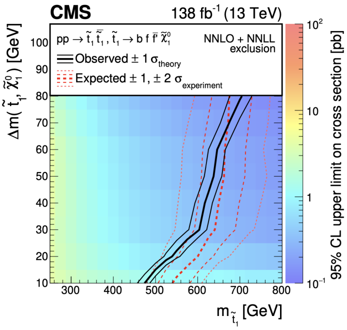



**Figure:** The 95% CL upper limits in the $(m(\tilde{t}_1), \Delta m)$ plane on the cross section for the production and four-body decay of the top squark using the combined 2016, 2017, and 2018 data. The color shading represents the observed upper limit for a given point in the plane, using the color scale to the right of the figure. The solid black and dashed red lines show the observed and expected 95% CL lower limits, respectively, on $m(\tilde{t}_1)$ as a function of $\Delta m$.

# Summary: Process and Setup

- Search for supersymmetry based on 2016, 2017, and 2018 CMS data collected from proton-proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$, integrated luminosity of $138\,\mathrm{fb}^{-1}$.
- Search for the direct pair production of top squarks in single-lepton final states: prompt four-body decay to $b\, f\, \overline{f}\, \tilde{\chi}^0_1$, charged lepton and its neutrino in the final state for the decay products of one $\tilde{t}_1$, two quarks for the other top squark.
- $R$ parity is conserved, mass difference between the lightest top squark and the lightest supersymmetric particle, taken to be the lightest neutralino, does not exceed the $W$ boson mass.
- Events are selected containing a single lepton (electron or muon), at least one high-momentum jet, and significant missing transverse momentum. The signal is selected based on a multivariate approach that is optimized for the difference between $m(\tilde{t}_1)$ and $m(\chi^0_1)$.
- Dominant background processes are W + jets, $t\overline{t}$, and events with nonprompt leptons, estimated using control regions in the data.

# Summary: Results

- The observed number of events is consistent with the predicted standard model backgrounds in all signal regions.
- Upper limits are set at the 95% confidence level on the $\tilde{t}_1\overline{\tilde{t}_1}$ production cross section as a function of the top squark and neutrino masses, within the context of a simplified model.
- Assuming a 100% branching fraction in the four-body decay mode, the search excludes top squark masses up to 480 and 700 GeV at $\Delta m = 10$ and 80 GeV respectively.
- The results summarized in this paper are among the best limits to date on the top squark pair production cross section, where the top squark decays via the four-body mode.

# Conclusion

Thank you, Einen schönen Tag noch, bom dia e buona giornata!