

# Machine Learning-Enhanced Feasibility Studies on the Production of New Particles with Preferential Couplings to Third Generation Fermions at the LHC

PhD(c). Cristian Fernando Rodríguez Cruz<sup>1</sup>  
c.rodriguez45@uniandes.edu.co

Research advisors:

Prof. Andrés Florez<sup>1</sup> ( ca.florez@uniandes.edu.co )

Prof. J. Jones-Pérez<sup>2</sup> ( jones.j@pucp.edu.pe )

<sup>1</sup>Universidad de los Andes

<sup>2</sup>Pontificia Universidad Católica del Perú

January 21, 2026

# Outline

## 1 Introduction

- Standard Model of Particle Physics
- Deficiencies of the SM
- LHC and Beyond the SM Physics

## 2 Phenomenological Framework

- Monte-Carlo Pipeline
- Hypothesis Testing & Statistical Significance
- Machine Learning in High Energy Physics

## 3 $U(1)_{T_R^3}$ Model

- Production Channel
- Search Channel
- Kinematic Variables
- Gradient Boosting

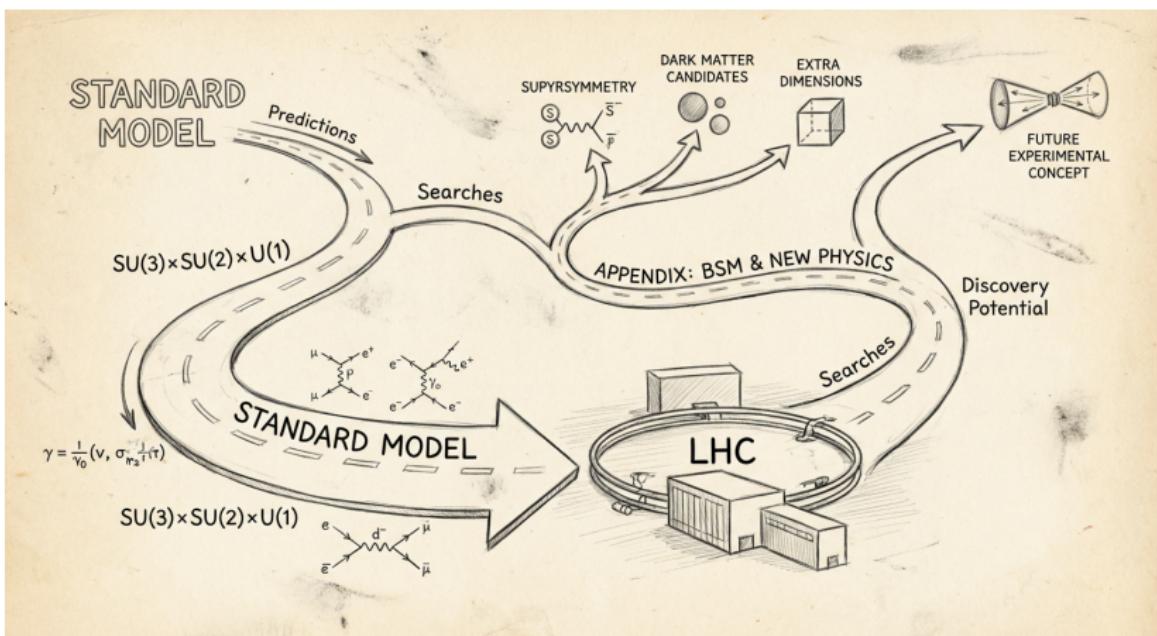
## 4 $U_1$ Leptoquark Model

- Kinematic Feature Eng.
- Single Channels Sensitivity Reach
- Combined Sensitivity Reach

## 5 $Z'$ Interferences

## 6 Summary and Conclusions

# Introduction



# Standard Model of Particle Physics: A Successful Framework

## Core Theoretical Structure

### • Fermion Sector:

- 3 generations
- Left-handed doublets:

$$Q_L = \begin{pmatrix} q_u \\ q_d \end{pmatrix}_L ; \quad L_L = \begin{pmatrix} \ell \\ \nu_\ell \end{pmatrix}_L$$

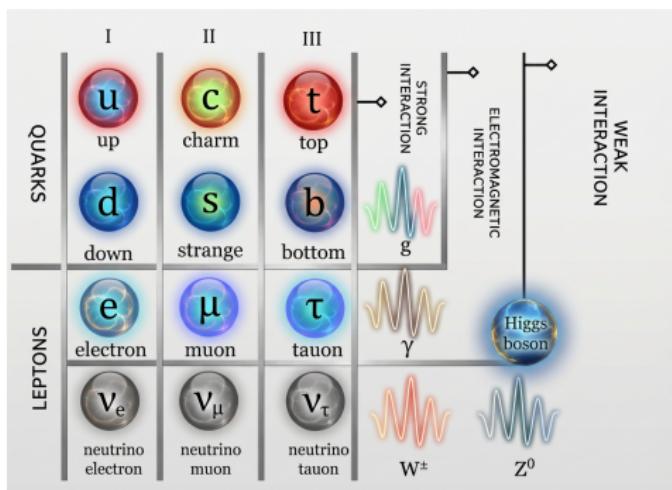
- Right-handed singlets:  $q_{uR}, q_{dR}; \ell_R$
- Flavor structure from CKM mixing
- Initially,  $\nu_R$  not required (no mass)

### • Gauge Group:

- $SU(3)_C \times SU(2)_L \times U(1)_Y$
- Strong force:  $SU(3)_C$  (QCD)
- Electroweak:  $SU(2)_L \times U(1)_Y$

### • Higgs Mechanism:

- $SU(2)_L \times U(1)_Y \rightarrow U(1)_{EM}$
- Masses to  $W^\pm, Z^0$  bosons
- Massless  $\gamma$
- Fermion masses via Yukawa couplings
- Higgs boson  $h$  discovered (LHC 2012)



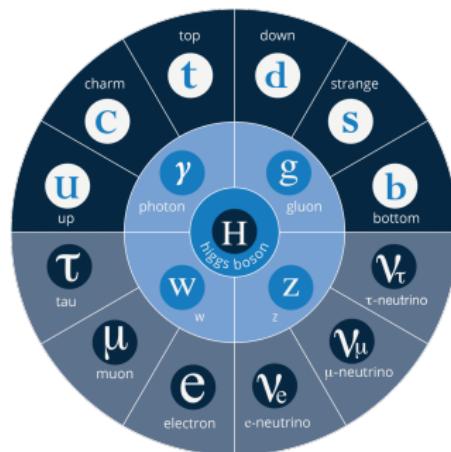
## Key Features

- **Particle spectrum:** 12 fermions + 5 bosons
- **QFT:** Renormalizable, gauge invariant, anomaly-free
- **Tested experimentally:** Exceptional agreement with data
- **Predictive power:** Successfully tested at LHC

# Deficiencies of the SM

The SM cannot be complete:

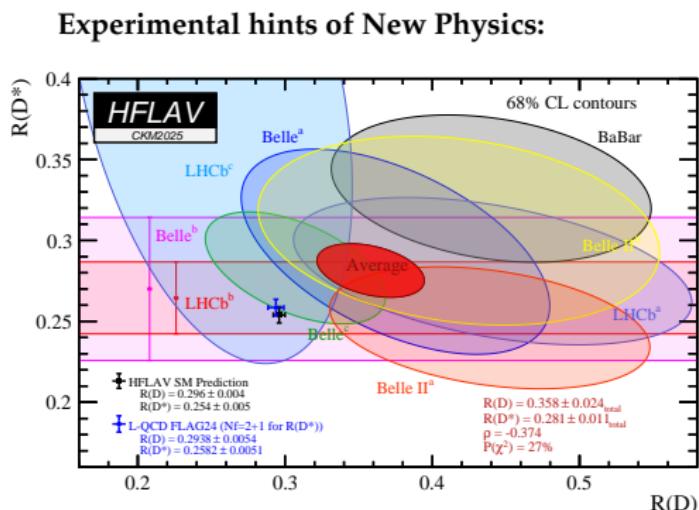
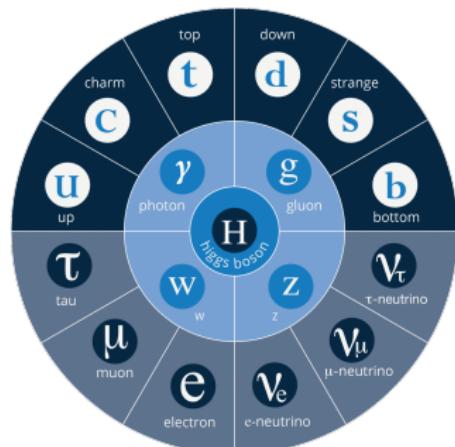
- Unexplained  $\nu$  masses
- No explanation for DM
- Matter-antimatter asymmetry unsolved
- Gravity not included



# Deficiencies of the SM

The SM cannot be complete:

- Unexplained  $\nu$  masses
- No explanation for DM
- Matter-antimatter asymmetry unsolved
- Gravity not included



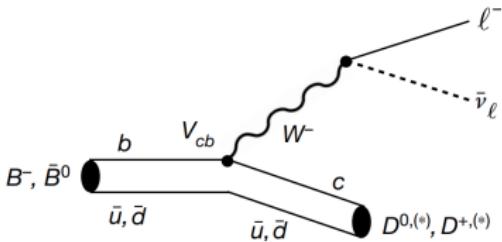
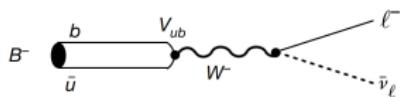
Recent measurements of the  $R(D)$  and  $R(D^*)$  ratios show a  $\sim 3\sigma$  deviation from SM predictions, suggesting possible lepton universality violation.

# $R(D)$ and $R(D^*)$ anomalies

## Flavor Observables

$$R(D) = \frac{\text{BR}(B \rightarrow D \tau \nu)}{\text{BR}(B \rightarrow D \ell_{(e,\mu)} \nu)}, \quad R(D^{(*)}) = \frac{\text{BR}(B \rightarrow D^{(*)} \tau \nu)}{\text{BR}(B \rightarrow D^{(*)} \ell_{(e,\mu)} \nu)}$$

## Standard Model (Tree-level)



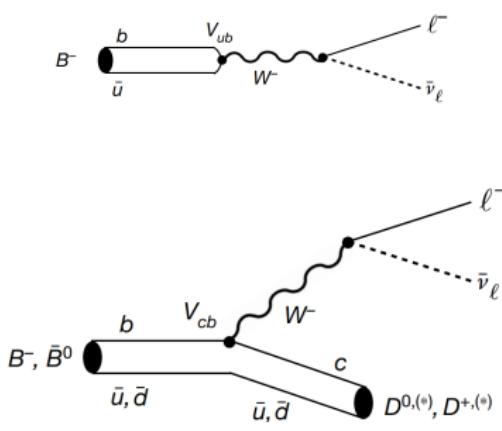
$B^- \rightarrow D^0 \ell^- \bar{\nu}_\ell$  via  $W^-$  exchange

# $R(D)$ and $R(D^*)$ anomalies

## Flavor Observables

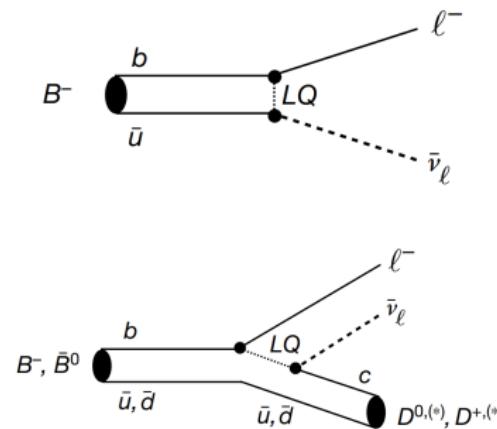
$$R(D) = \frac{\text{BR}(B \rightarrow D \tau \nu)}{\text{BR}(B \rightarrow D \ell_{(\epsilon, \mu)} \nu)}, \quad R(D^{(*)}) = \frac{\text{BR}(B \rightarrow D^{(*)} \tau \nu)}{\text{BR}(B \rightarrow D^{(*)} \ell_{(\epsilon, \mu)} \nu)}$$

## Standard Model (Tree-level)



$B^- \rightarrow D^0 \ell^- \bar{\nu}_\ell$  via  $W^-$  exchange

## Leptoquark Mediated



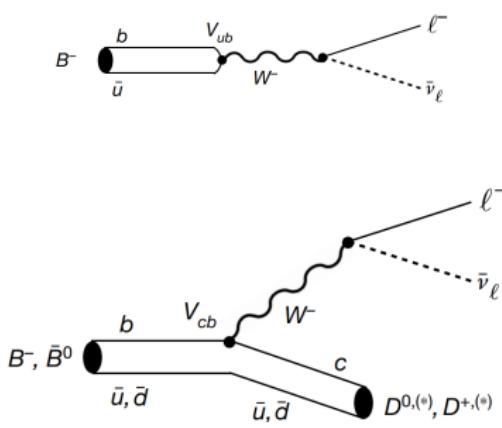
$B^- \rightarrow D^0 \ell^- \bar{\nu}_\ell$  via leptoquark exchange

# $R(D)$ and $R(D^*)$ anomalies

## Flavor Observables

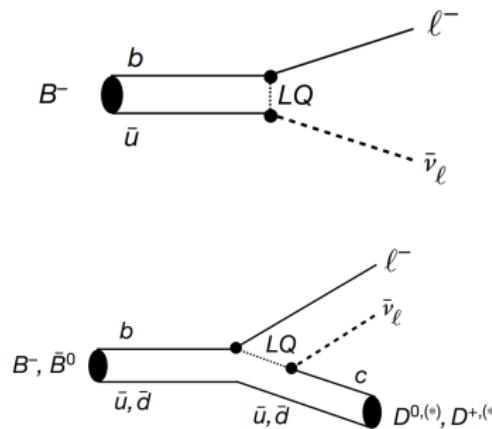
$$R(D) = \frac{\text{BR}(B \rightarrow D \tau \nu)}{\text{BR}(B \rightarrow D \ell_{(\epsilon, \mu)} \nu)}, \quad R(D^{(*)}) = \frac{\text{BR}(B \rightarrow D^{(*)} \tau \nu)}{\text{BR}(B \rightarrow D^{(*)} \ell_{(\epsilon, \mu)} \nu)}$$

## Standard Model (Tree-level)



$B^- \rightarrow D^0 \ell^- \bar{\nu}_\ell$  via  $W^-$  exchange

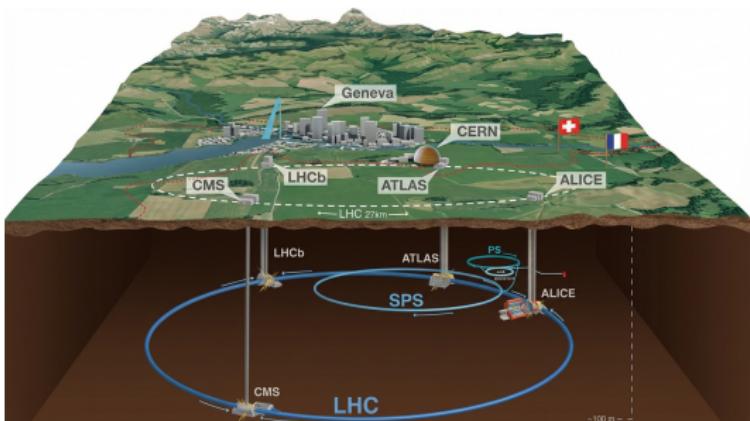
## Leptoquark Mediated



$B^- \rightarrow D^0 \ell^- \bar{\nu}_\ell$  via leptoquark exchange

How can we test this hypothesis?

# The Large Hadron Collider (LHC)

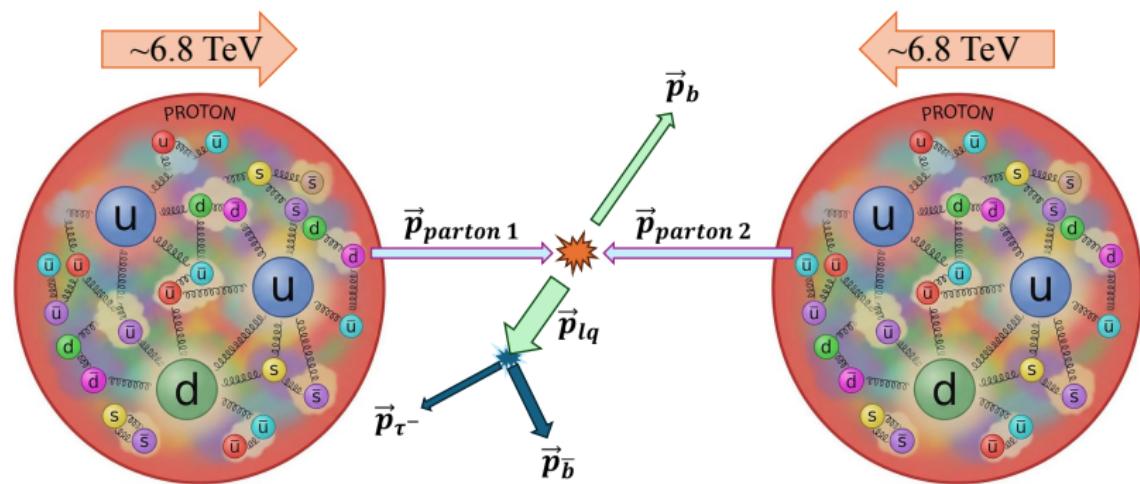


World's largest particle collider:

- Located at CERN, at the France-Switzerland border
- 27 km circumference ring
- Currently running proton-proton collisions at  $\sqrt{s} = 13.6$  TeV
- Designed to probe the TeV scale
- Four main experiments: ATLAS, CMS, LHCb, ALICE
- High luminosity upgrade (HL-LHC) planned for 2029
- Key tool for testing SM and exploring new physics scenarios
- Now searching for physics **beyond the Standard Model**

# The Quark-Gluon Sea

**Partons** are the fundamental constituents inside protons: valence quarks ( $uud$ ) and a **sea** of virtual quark-antiquark pairs and gluons.

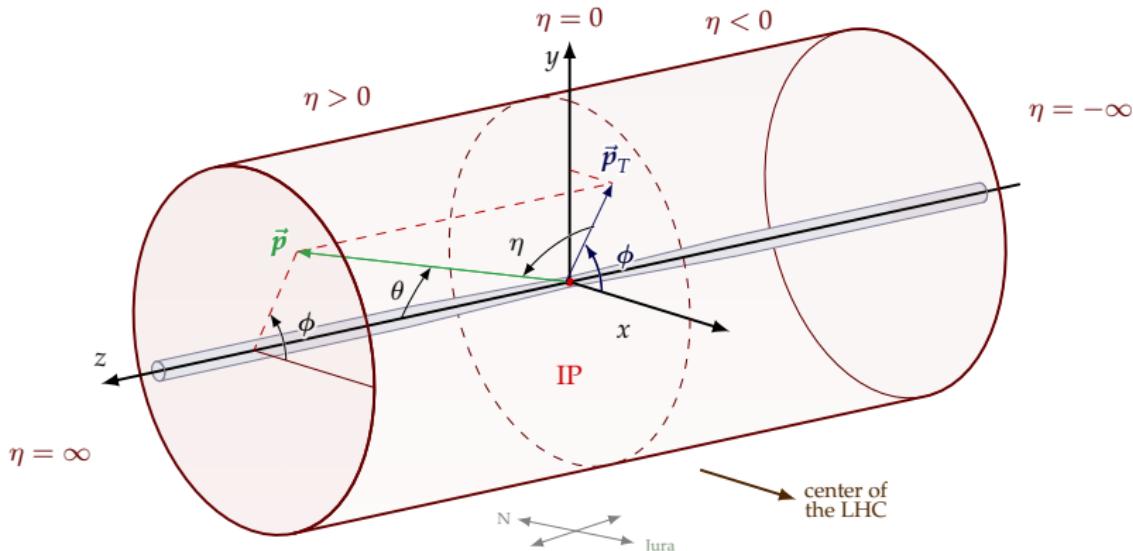


The interacting partons are typically **sea quarks or gluons**, which carry only a fraction of the proton's momentum but dominate the cross-section at high energies.

# Kinematic Variables

From Spherical coordinates,

$$\begin{cases} \text{Pseudorapidity: } \eta = -\ln \tan(\theta/2) \\ \text{Transverse momentum: } p_T = p \sin(\theta) \\ \text{Azimuthal angle: } \phi \\ \text{Deposited energy: } E \end{cases}$$

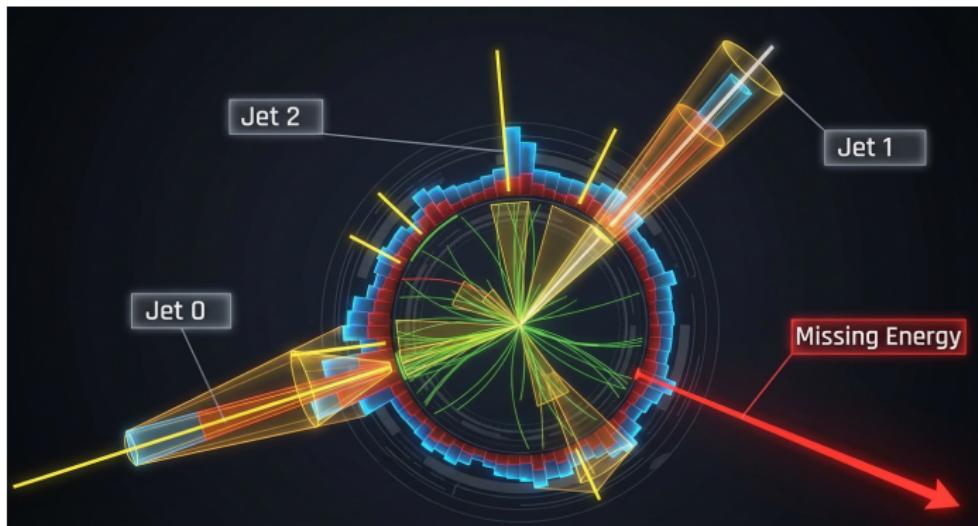


# Jet Clustering & Missing Transverse Momentum

**Jet clustering** groups collimated particle showers into jets to reconstruct the original quarks/gluons from the hard scatter.

**Missing  $p_T$**  appears as transverse momentum imbalance, signaling undetected particles like neutrinos:

$$\vec{p}_T^{\text{miss}} = - \sum_{\text{visible prods}} \vec{p}_T$$



# What Do We Look For at the LHC?

## Three Complementary Approaches

### • SM Parameter Determination

- Precisely measure fundamental SM parameters
- Test consistency of the SM framework
- Reduce theoretical uncertainties

### • Indirect Searches (Precision Tests)

- Measure SM processes with high precision
- Look for deviations in predicted distributions
- Constrain new physics through virtual effects

### • Direct Searches

- Look for new particles in final states (resonances, excesses)
- Examples: SUSY particles,  $Z'$ , leptoquarks, dark matter mediators

## Key Types of Measurements

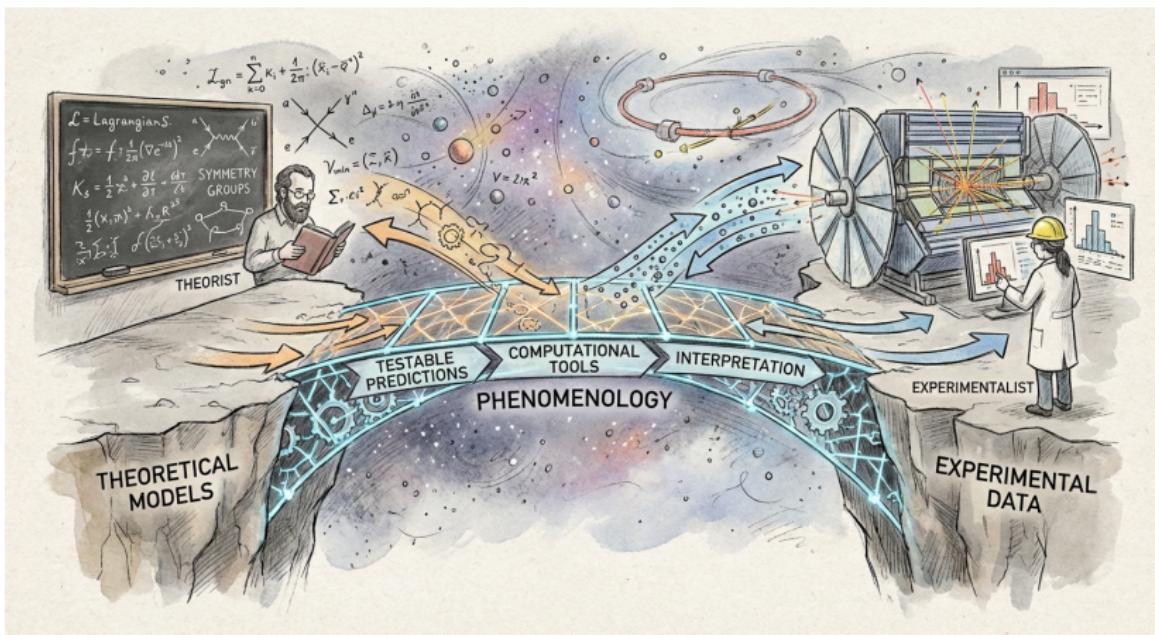
### Direct Signatures:

- Resonant mass peaks
- Excess events over SM background
- Missing transverse energy (MET)
- Unusual kinematic features

### Precision Observables:

- Differential cross sections
- Angular correlations
- Rare decay rates
- Lepton flavor universality ratios
- Charge-parity (CP) asymmetries

# Phenomenological Framework



# From Theory to Simulation: FeynRules



## Input: .fr Model File

- Lagrangian  $\mathcal{L}$  terms
- Particle definitions: F, V, S fields
- Gauge symmetries:  $SU(N)$ ,  $U(1)$
- Parameters: masses, couplings
- Mixing matrices, constraints

## Output: UFO Format

- Complete Feynman rules (vertices)
- Lorentz and color structures
- Parameter definitions
- Ready to simulate

## Features

- ↗ Standardized UFO format
- ↗ Flexible for BSM models
- ↗ Community-driven
- ↘ NLO complexity
- ↘ Poor debugging
- ↘ Performance issues

## Why Simulate New Physics Models?

- **Predict signals** for experimental searches
- **Test theoretical consistency** (unitarity, constraints)
- **Optimize analyses** before data collection
- **Interpret potential discoveries** from LHC data
- **Compare predictions** across different models

# Madgraph-Pythia-Delphes

Ecosystem for Event Simulation

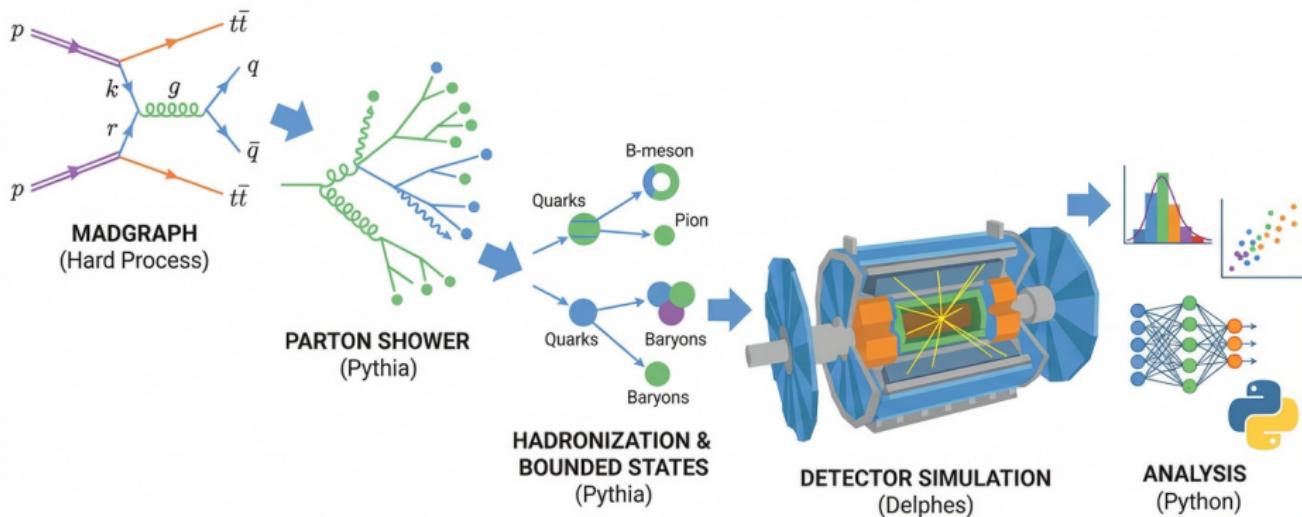
**Goal:** Produce realistic simulated data comparable to experimental observations

**Input:** BSM model in UFO format, collider setup, process definitions, parameters

**Pipeline:** UFO model → LHE (parton) → HEPMC (hadron) → ROOT (detector)

**Output:** Simulated events in ROOT format, ready for physics analysis

**Automated:** Cross-sections, widths, branching ratios calculated automatically



# Statistical Significance

The statistical significance for discovery is a parametric test defined as:

$$\kappa = \frac{\langle t \rangle_B - \langle t \rangle_{S+B}}{\sigma_{S+B}} \quad (1)$$

where  $t = -2 \ln[\mathcal{L}(n|S + B)/\mathcal{L}(n|B)] \sim \chi^2$  is the optimal test statistic.

We denote  $\mathcal{L}$  as the likelihood for Poisson-distributed event counts  $n$  in each bin and gaussian systematic uncertainties.

For binned data analysis, this simplifies to:

$$\kappa = \frac{\sum_i s_i w_i}{\sqrt{\sum_i (s_i + b_i + \delta_{\text{sys}}^2) w_i^2}} \quad (2)$$

where  $s_i, b_i$  are signal/background events in bin  $i$ ,  $w_i \sim \ln(1 + s_i/b_i)$  are optimal weights, and  $\delta_{\text{sys}}$  is the systematic uncertainty.

**Exclusion criterion:** A hypothesis with  $\kappa \leq 1.69\sigma$  can be excluded at 95% CL.

# Optimizing Sensitivity

The event yield in each bin is  $N_i = \sigma_i \cdot \mathcal{L} \cdot \epsilon_i$ , where:

- $\sigma_i$ : cross section
- $\mathcal{L}$ : integrated luminosity
- $\epsilon_i$ : selection efficiency

The significance becomes:

$$\kappa = \frac{\sum_i \sigma_{s_i} \epsilon_{s_i} w_i}{\sqrt{\sum_i (\sigma_{s_i} \epsilon_{s_i} + \sigma_{b_i} \epsilon_{b_i} + \delta_{\text{sys}}^2) w_i^2}} \sqrt{\mathcal{L}} \quad (3)$$

## Strategies to improve $\kappa$ :

- Increase luminosity  $\mathcal{L}$
- Optimize cuts to enhance  $\epsilon_s/\epsilon_b$
- Increase  $\sigma_s$  via higher  $\sqrt{s}$  for heavier states
- Use multivariate discriminants (BDT, DNN)
- Reduce systematic uncertainties  $\delta_{\text{sys}}$
- Exploit correlations between bins

# Machine Learning Revolution in HEP

## Why ML in Particle Physics?

- High-dimensional data:**  
Complex detector outputs
- Complex patterns:**  
Jet and event classification
- Rare signals:**  
Need optimal background rejection
- Automation:**  
Reduce manual feature engineering
- Computational efficiency:**  
Speed up workflows

## Traditional vs ML Approaches

### Traditional:

- Manual cut-based analysis
- Simple kinematic variables
- Univariate methods
- Limited feature engineering

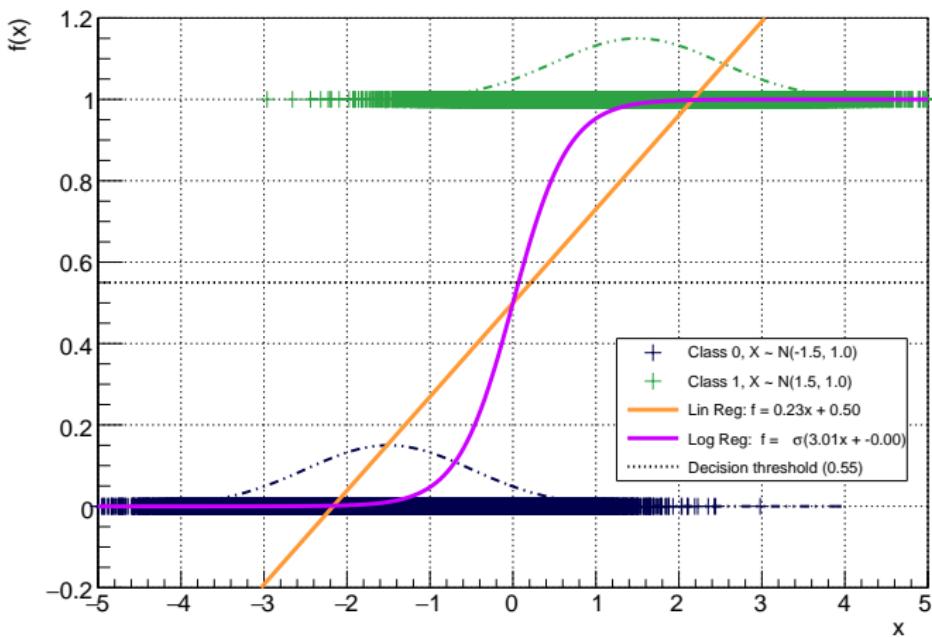
### Machine Learning:

- Multivariate classifiers (BDT, NN)
- Deep learning on raw data
- Anomaly detection for BSM searches

## Typical Current Applications

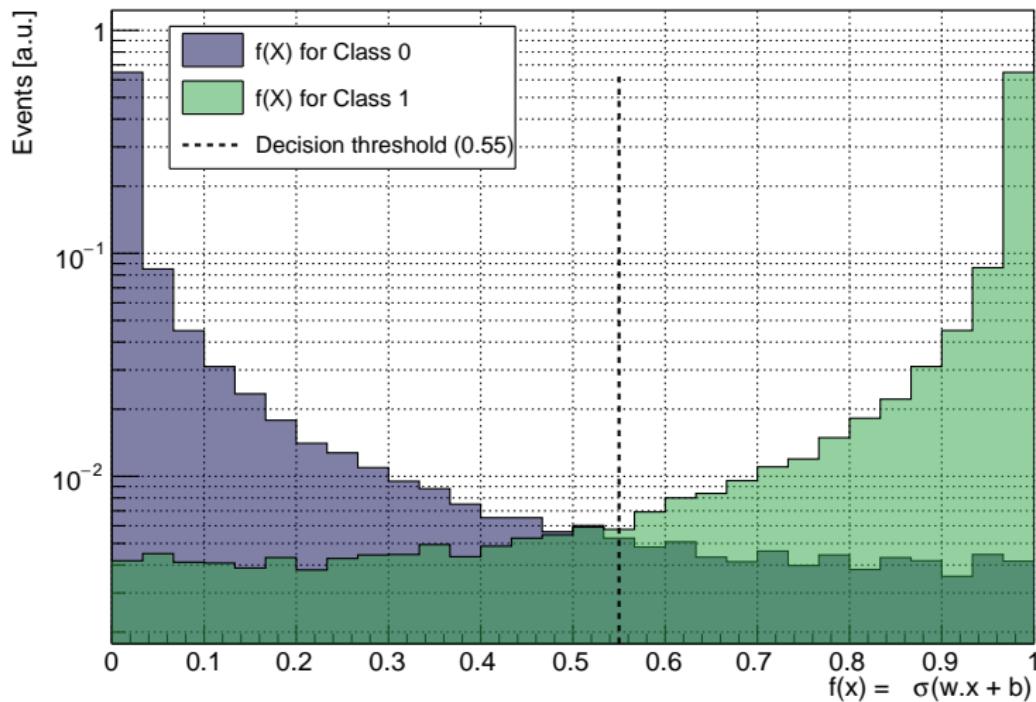
- Event classification:** Signal/background separation
- Jet tagging:**  $b/c/\tau$ -jet identification
- Object reconstruction:** Particle flow, tracking
- Anomaly detection:** Model-agnostic BSM searches
- Simulation acceleration:** Fast calorimeter simulation

# How does teh binary classifier work?

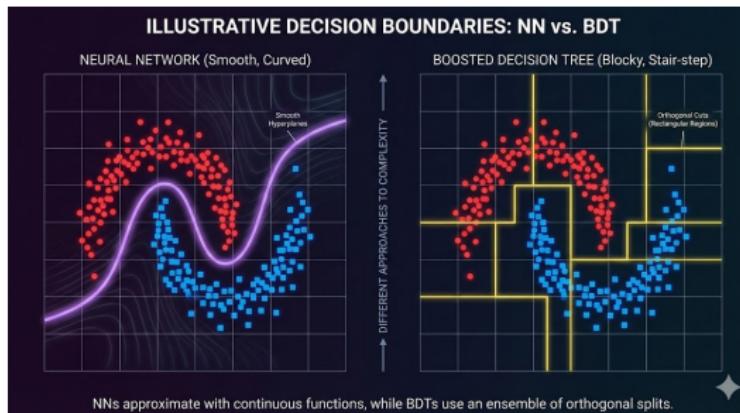


**Be aware:** The real performance is the same for both algorithms in terms of classification accuracy. The difference lies in interpretability and probabilistic output.

## A New Distribution Emerges

Distribution of  $f(x) = \sigma(w \cdot x + b)$ 

# Bossting Decision Trees vs Neural Networks



## BDTs

### Strengths:

- Interpretable:** Feature importance rankings
- Robust:** Handle missing data well
- Fast training:** Efficient on tabular data
- Low hyperparameter tuning:** Reasonable defaults work

### Limitations:

- Extrapolation:** Poor outside training range
- High-dimensional:** Can struggle with many features
- Discontinuous:** Piecewise constant predictions

## Deep NNs

### Strengths:

- Flexible:** Can learn complex non-linearities
- High-dimensional:** Excel with many features
- Continuous:** Smooth function approximations
- Transfer learning:** Pretrained models possible

### Limitations:

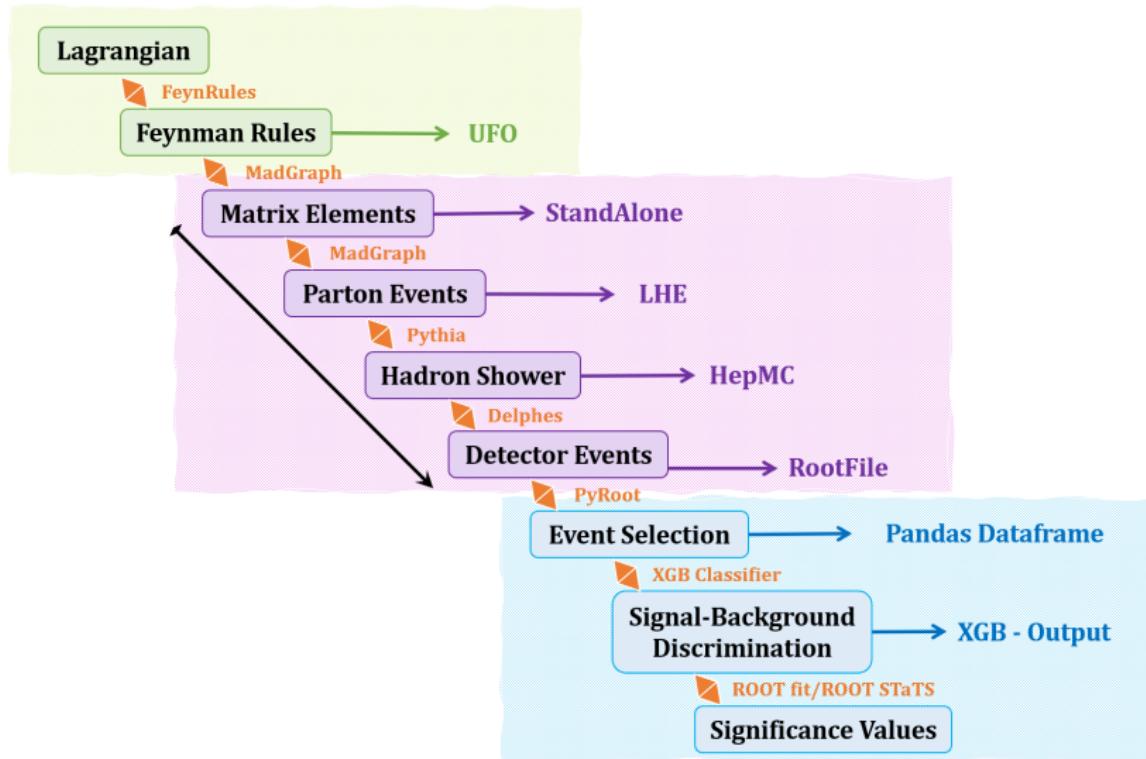
- Black box:** Difficult to interpret
- Data hungry:** Need large training sets
- Computational:** Require GPUs for large networks
- Sensitive:** Require careful hyperparameter tuning

## HEP Use Cases

**BDTs:** Preferred for  $\sim 10 - 100$  features

**NNs:** Preferred for low-level data (images, graphs)

# Feasibility Studies Workflow

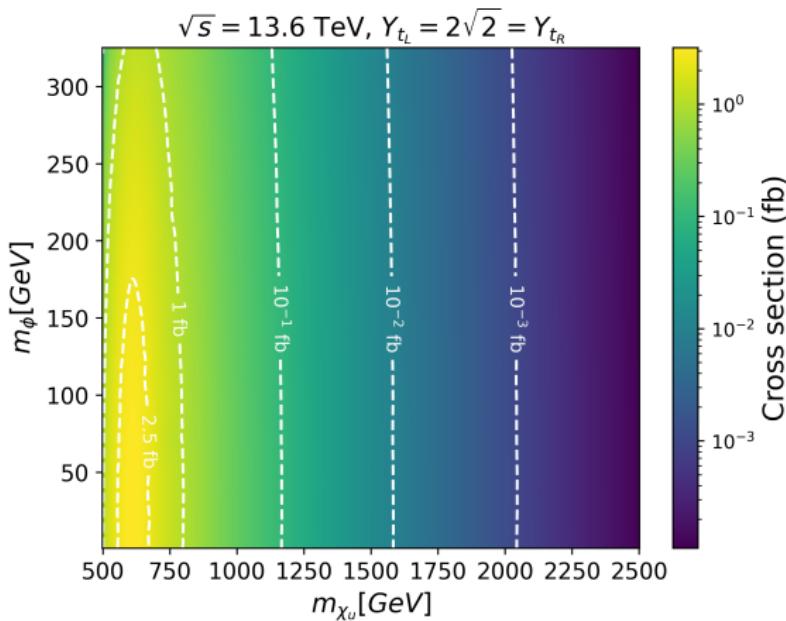


# $U(1)_{T_R^3}$ Model

To do ADD production channel diagrams

# Feasible Experimental Signatures

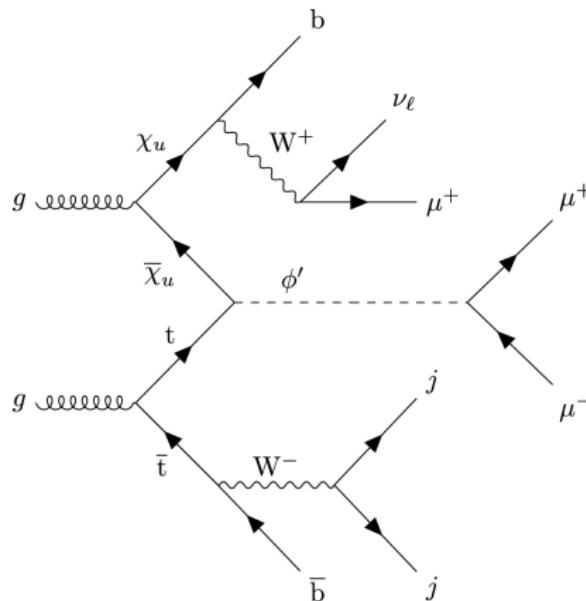
Cross Section



**Fig.:** Projected cross section (fb) plot for  $pp \rightarrow t\chi_u \phi'$  and subsequent decay as a function of  $m(\chi_u)$  and  $m(\phi')$ .

# Feasible Experimental Signatures

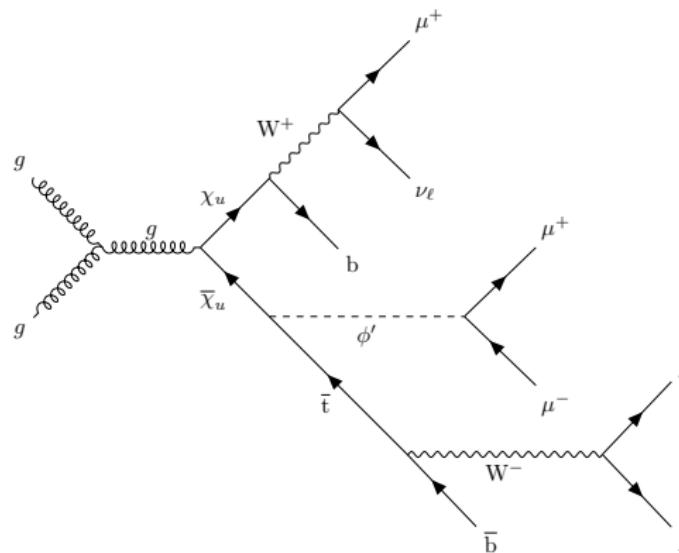
Representative Feynman diagram for the production of a  $\phi'$  boson in association with a  $\chi_u$  vector-like quark through the fusion of a top quark and  $\chi_u$  vector-like quark.



The  $\phi'$  decays to a pair of muons, the top quark decays fully hadronically, and the  $\chi_u$  decays semi-leptonically to muons, neutrinos and  $b$ -jets.

# Feasible Experimental Signatures

Representative Feynman diagram for the production of a  $\phi'$  boson in association with a  $\chi_u$  vector-like quark through the fusion of a gluon pair from incoming protons.

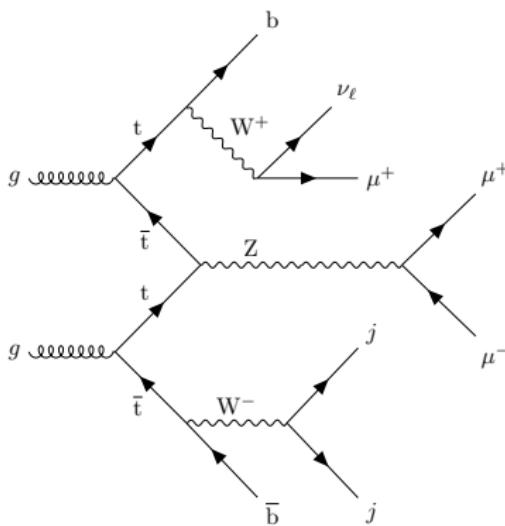


The  $\phi'$  decays to a pair of muons, the top quark that decays fully hadronically, and the  $\chi_u$  decay semi-leptonically to muons, neutrinos and jets.

# Feasible Experimental Signatures

## Background

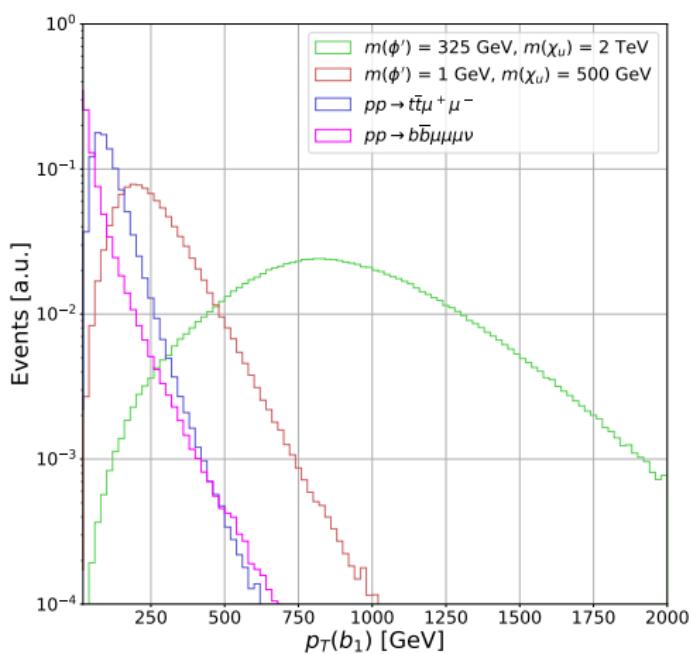
Representative Feynman diagram for a background event. A Z boson is produced in association with a top quark through the fusion of a top, anti top pair from incoming protons.



The Z boson subsequently decays to a pair of muons and the two spectator top quarks decay semi-leptonically and purely hadronically to muons, neutrinos and jets, resulting in the same final states as the signal event.

# Feasible Experimental Signatures

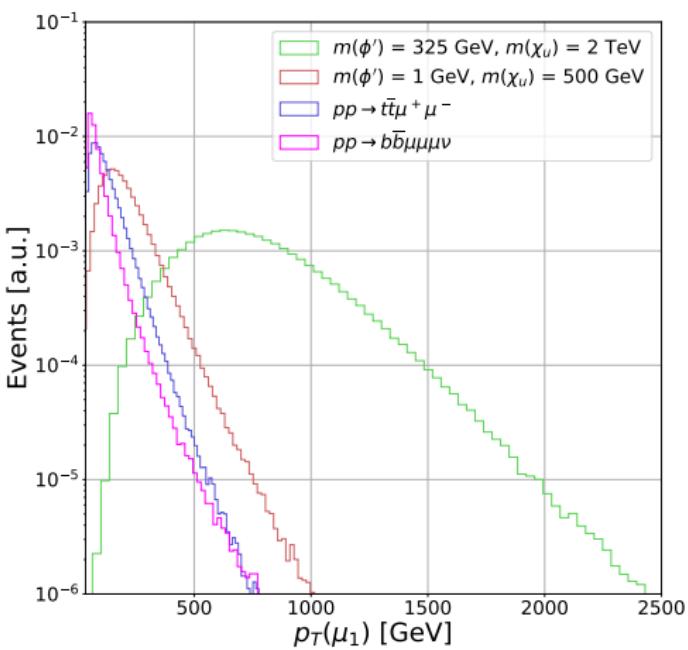
## Kinematic Variables



**Fig.:** Transverse momentum distribution of the leading b-quark jet candidate. The distributions are shown for the two main SM background processes and two signal benchmark points.

# Feasible Experimental Signatures

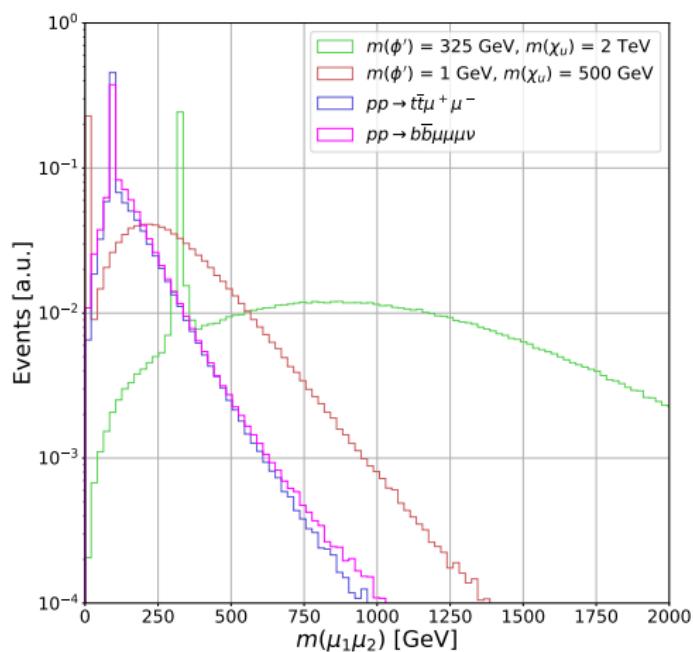
## Kinematic Variables



**Fig.:** Transverse momentum distribution of the leading muon candidate. The distributions are shown for the two main SM background processes and two signal benchmark points.

# Feasible Experimental Signatures

## Kinematic Variables

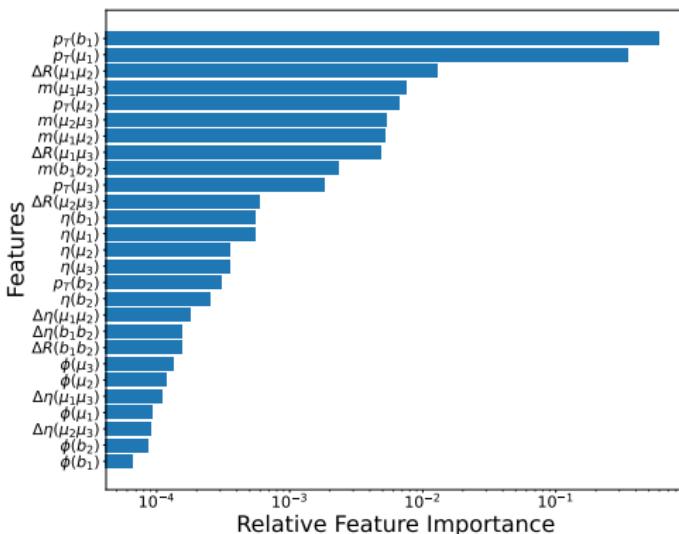


**Fig.:** Invariant mass distribution of the muon pair with the highest and second highest transverse momentum. The distributions are shown for the two main SM background processes and two signal benchmark points.

## Gradient Boosting

## Feasible Experimental Signatures

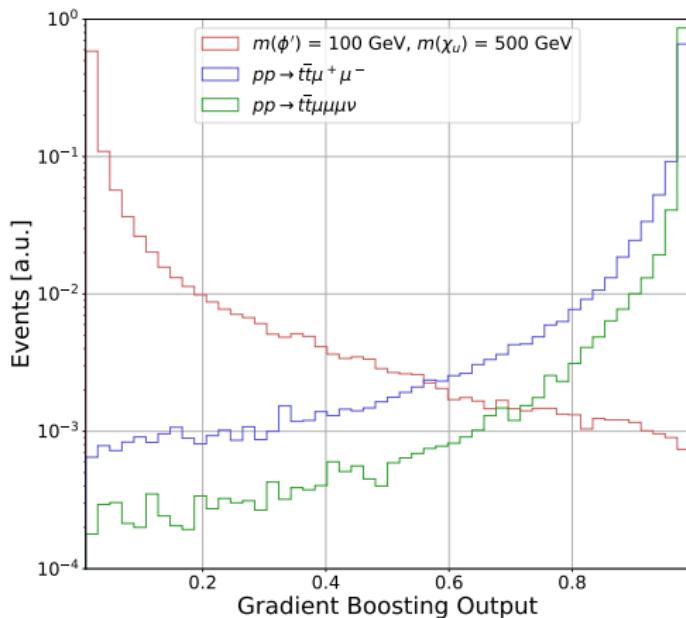
## Gradient Boosting



**Fig.:** Relative importance of features in training for a benchmark signal scenario with  $m(\phi') = 325 \text{ GeV}$  and  $m(\chi_u) = 2000 \text{ GeV}$ .

# Feasible Experimental Signatures

## Gradient Boosting



**Fig.:** Output of the gradient boosting algorithm for a benchmark  $m(\phi') = 100$  GeV and  $m(\chi_u) = 500$  GeV signal, and dominant backgrounds. The distributions are normalized to unity.

# Signal Significance

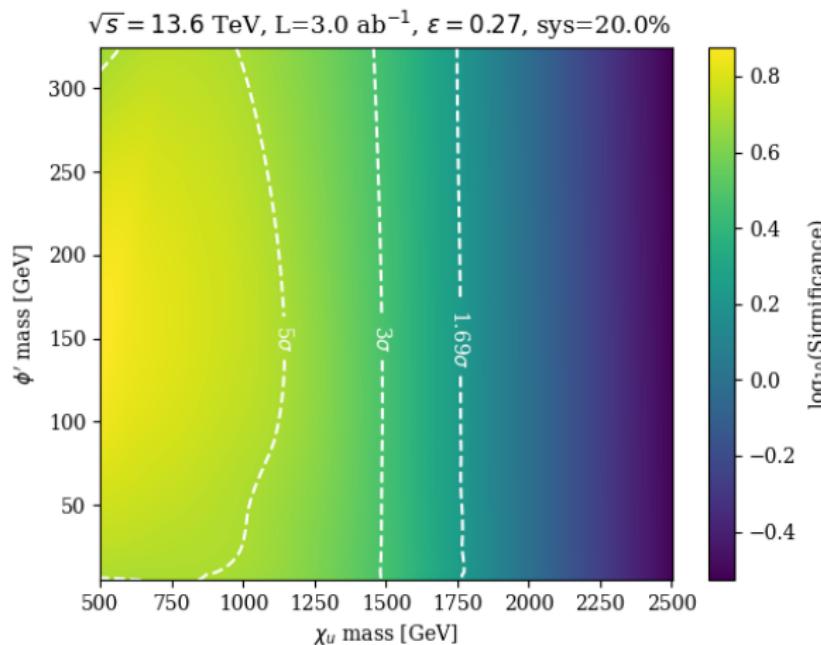


Fig.: Signal significance for the high luminosity LHC era, considering with  $3000 \text{ fb}^{-1}$  of collected data.

# $U_1$ Leptoquark Model

# The vector leptoquark $U_1$ model

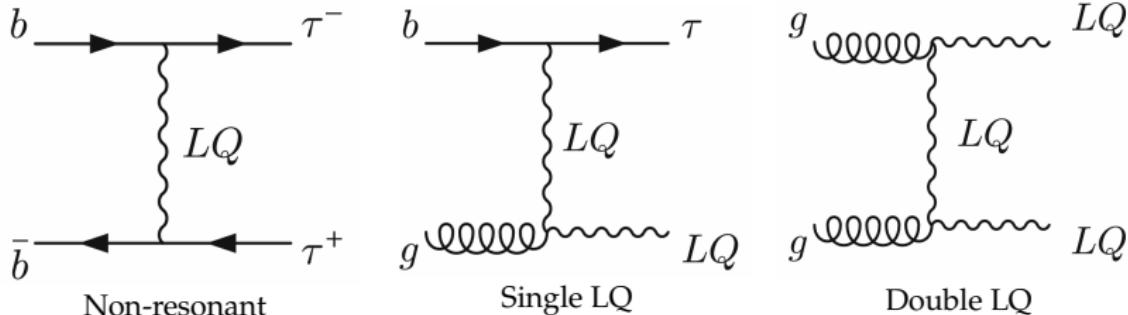
A leptoquark is defined as a particle with a vertex that mix vectors and quarks.

If  $U_1$  is a vector leptoquark that preserves the chirality on the vertex, we expect an interaction term like

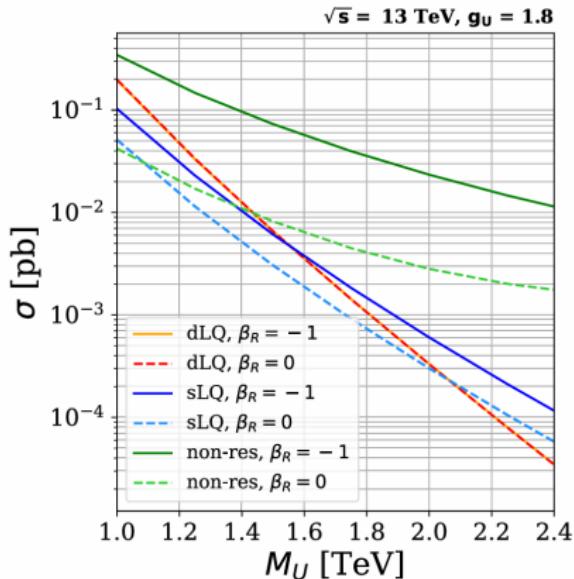
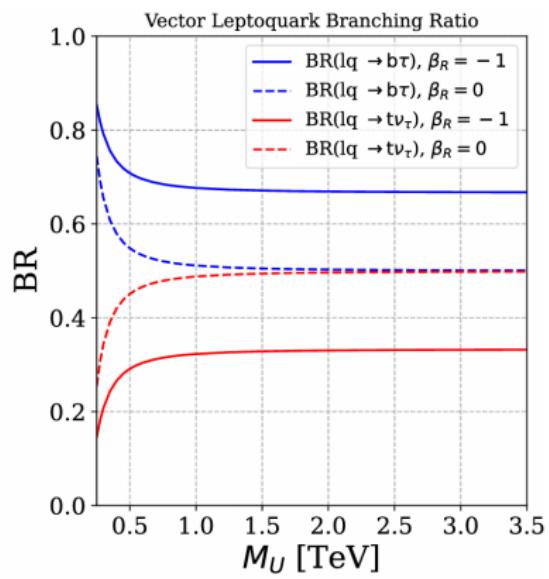
$$\sim \beta_L U_1^\mu \bar{q}_L \gamma_\mu \ell_L,$$

and these allows a similar interaction term for the right handed currents

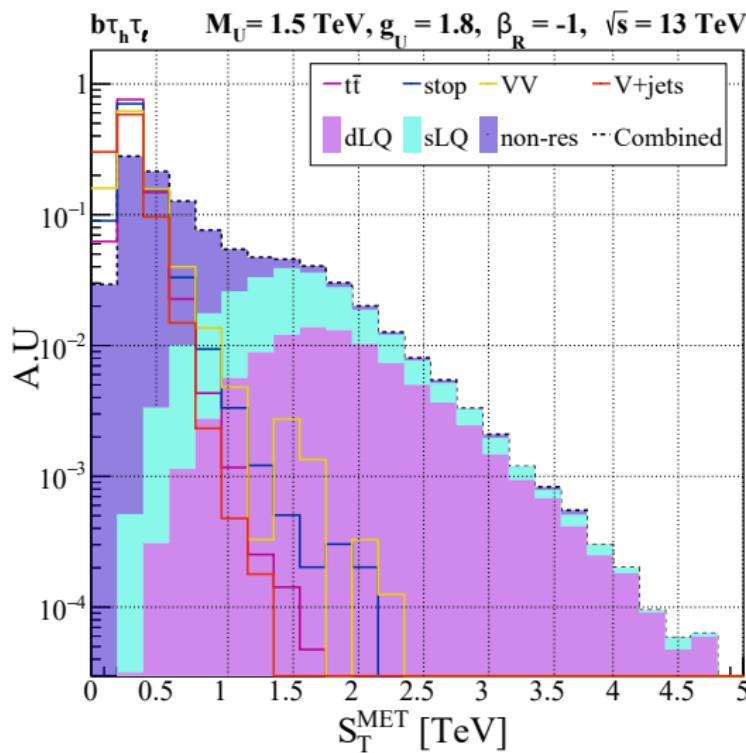
$$\sim \beta_R U_1^\mu \bar{d}_R \gamma_\mu e_R.$$



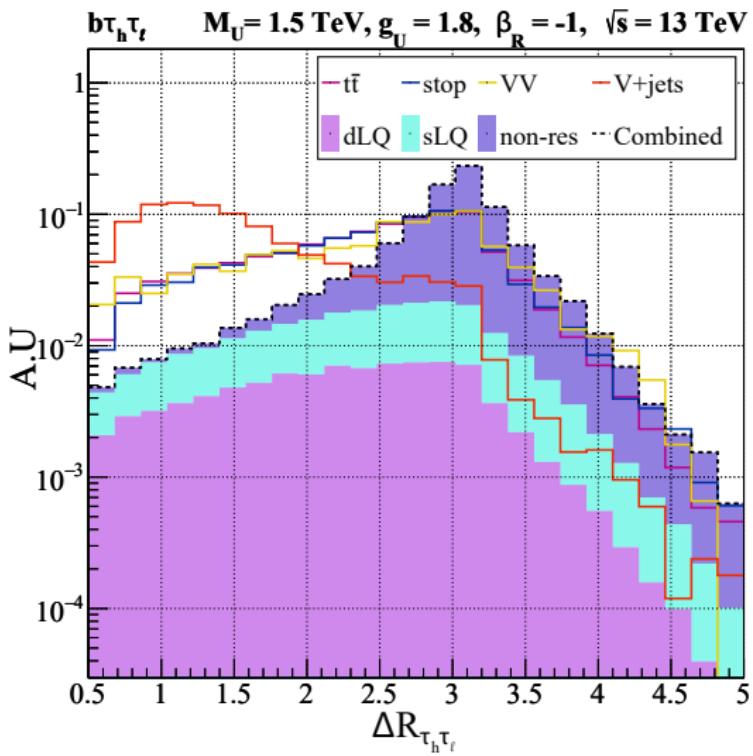
# Leptoquark Production at pp Colliders



$$S_T^{\text{met}} = \text{met} + \sum_i |p_T^i|$$



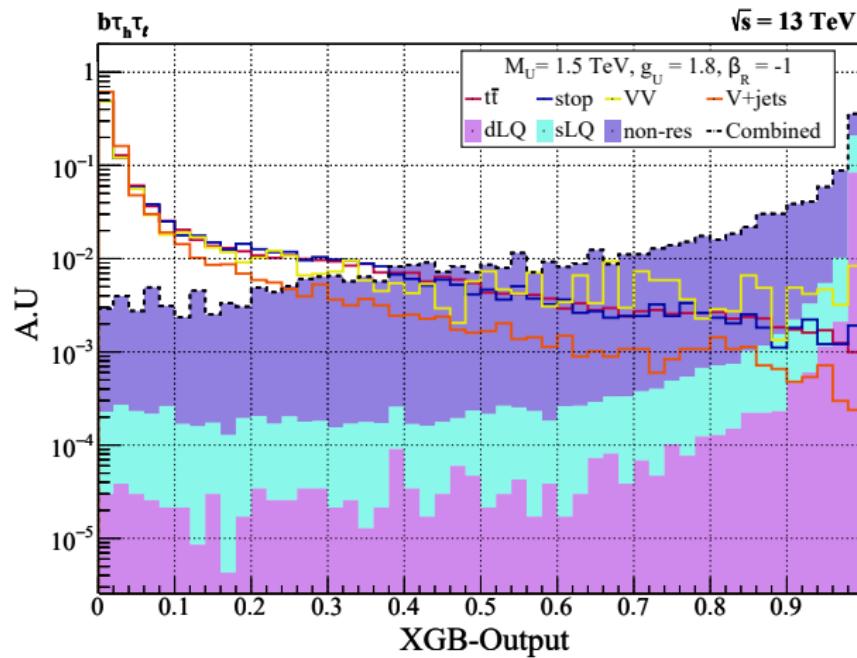
$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$$



# The optimized observable

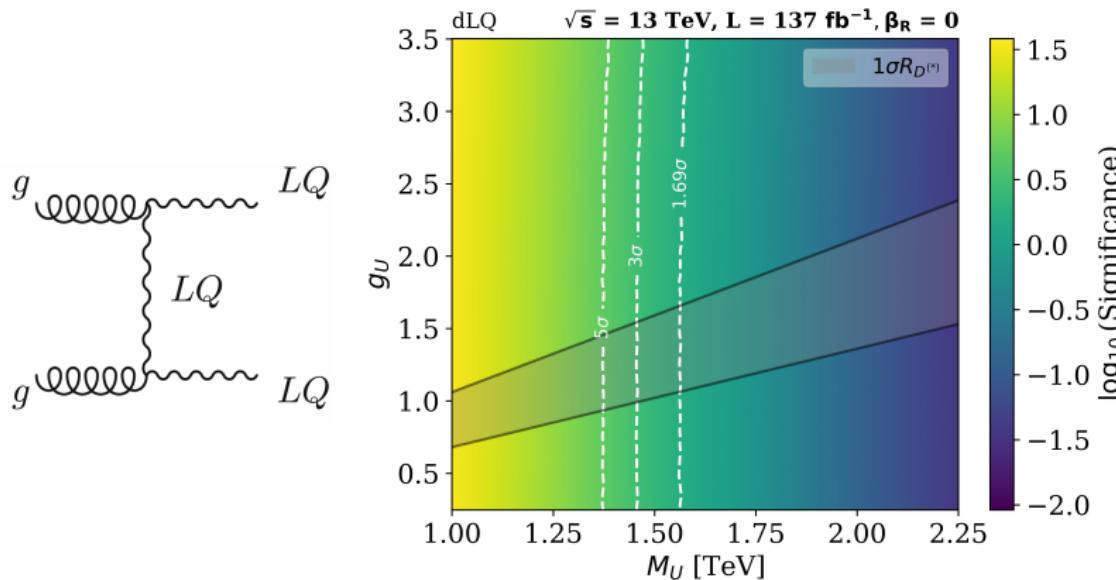
XGB-output

We can evaluate a score for the signal and background events using the discriminator algorithm.



# Double Leptoquark Production

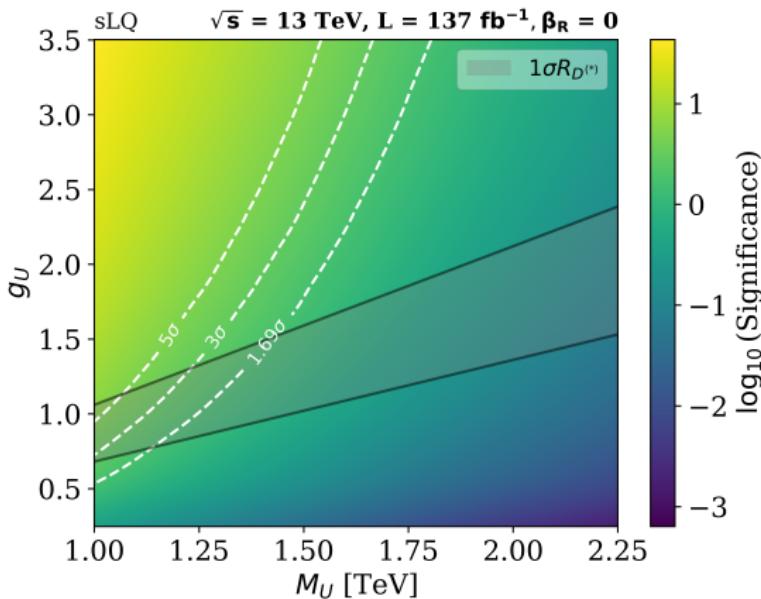
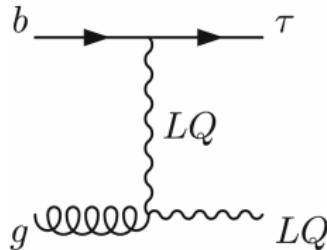
The Sensitivity Reach / only left-handed currents



Double leptoquark production is sensitive to the leptoquark mass, its production depends only on the QCD coupling constant and the available energy.

# Single leptoquark production

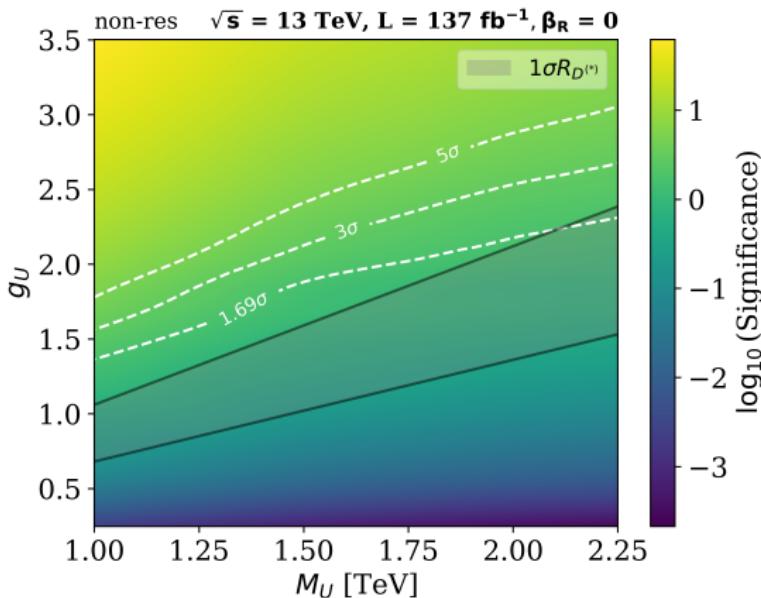
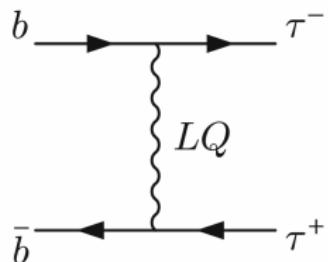
The Sensitivity Reach / only left-handed currents



Single leptoquark production is sensitive to both, mass and couplings. It contributes to the regions of high coupling constants at higher masses than double leptoquark production.

# Non-resonant Production

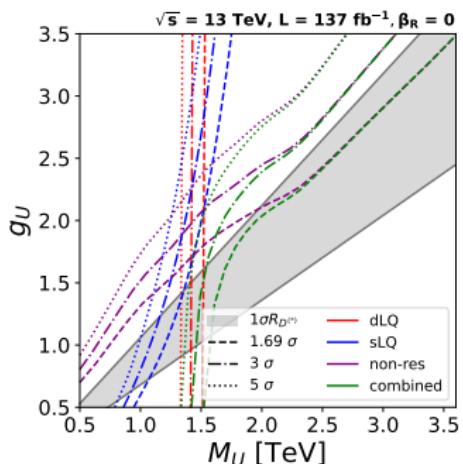
The Sensitivity Reach / only left-handed currents



Non-resonant production is highly dependent on the couplings, so it dominates the regions of high coupling constants at all masses.

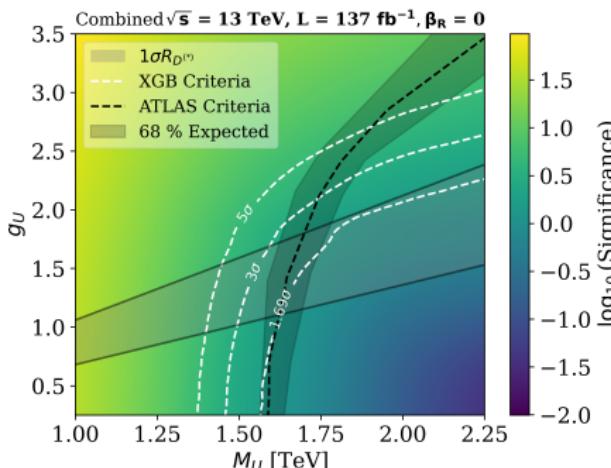
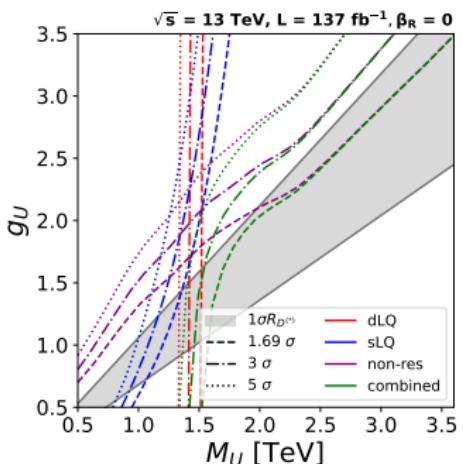
# Combined Sensitivity Reach

The Sensitivity Reach / only left-handed currents



# Combined Sensitivity Reach

The Sensitivity Reach / only left-handed currents

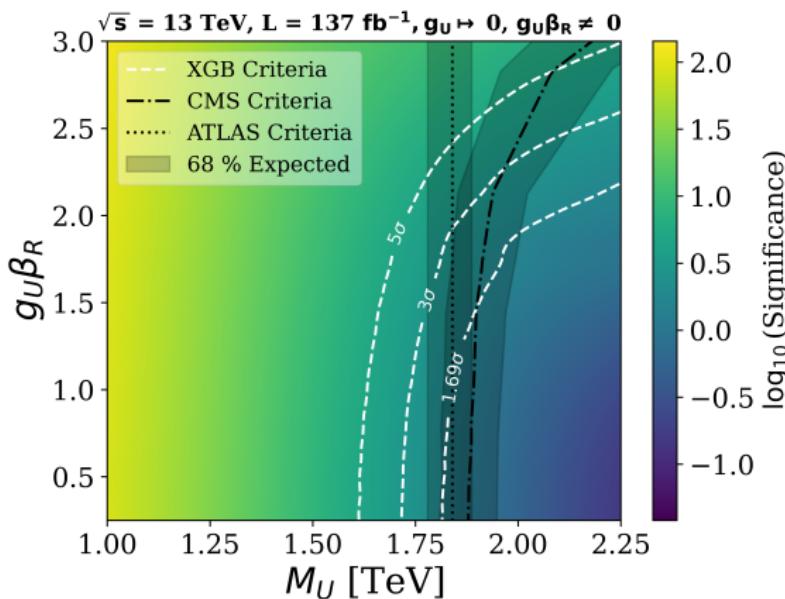


The sensitivity of all signal production processes combined compares our expected exclusion region with the latest one from the ATLAS Collaboration [ArXiv:2305.15962], but we suggest better sensitivity for high coupling constants.

## Combined Sensitivity Reach

# Combined Sensitivity Reach

The Sensitivity Reach / only right-handed currents



The case  $BR(lq \rightarrow b\tau) = 1$  corresponds to the only right-handed currents coupling. The sensitivity compared with the latest one from the CMS [2308.07826] and ATLAS Collaborations [2303.01294], again we suggest better sensitivity for high coupling constants.

# $Z'$ Interferences

# The need of a $Z'$ boson in gauge $U_1$ models

If  $U_1$  has a gauge origin, we could rewrite the interaction term in the covariant derivative as

$$\psi_L^{\text{SM}} = \begin{pmatrix} q_{Lr} \\ q_{Lg} \\ q_{Lb} \\ \ell_L \end{pmatrix} \implies \mathcal{L}_{\text{int}} \sim U_{1\alpha}^\mu \bar{\psi}_L^{\text{SM}} \gamma_\mu T_+^\alpha \psi_L^{\text{SM}} + \text{h.c.}, \quad T_+^\alpha = \begin{pmatrix} 0 & 0 & 0 & \delta_{r\alpha} \\ 0 & 0 & 0 & \delta_{g\alpha} \\ 0 & 0 & 0 & \delta_{b\alpha} \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

we have six generators  $T_\pm^\alpha$  with closure relation and projecting into a color singlet operator:

$$\sum_\alpha [T_+^\alpha, T_-^\alpha] = 3T_{B-L} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix}.$$

So, the gauge group with this leptoquark must include a  $U(1)_{B-L}$  symmetry (The right-handed currents also have a similar interaction term).

# The need of a $Z'$ boson in gauge $U_1$ models

If  $U_1$  has a gauge origin, we could rewrite the interaction term in the covariant derivative as

$$\psi_L^{\text{SM}} = \begin{pmatrix} q_{Lr} \\ q_{Lg} \\ q_{Lb} \\ \ell_L \end{pmatrix} \implies \mathcal{L}_{\text{int}} \sim U_{1\alpha}^\mu \bar{\psi}_L^{\text{SM}} \gamma_\mu T_+^\alpha \psi_L^{\text{SM}} + \text{h.c.}, \quad T_+^\alpha = \begin{pmatrix} 0 & 0 & 0 & \delta_{r\alpha} \\ 0 & 0 & 0 & \delta_{g\alpha} \\ 0 & 0 & 0 & \delta_{b\alpha} \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

we have six generators  $T_\pm^\alpha$  with closure relation and projecting into a color singlet operator:

$$\sum_\alpha [T_+^\alpha, T_-^\alpha] = 3T_{B-L} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix}.$$

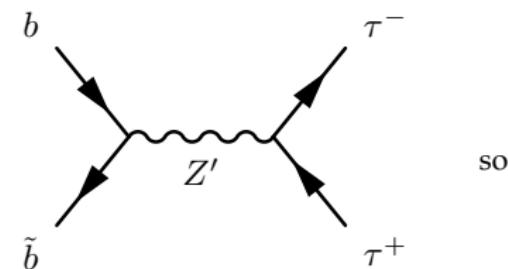
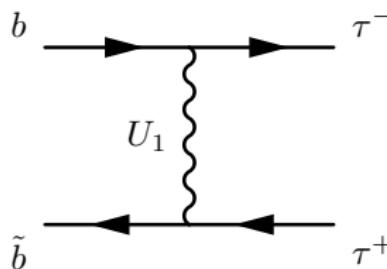
So, the gauge group with this leptoquark must include a  $U(1)_{B-L}$  symmetry (The right-handed currents also have a similar interaction term).

The interaction terms for the  $Z'$  boson have the form

$$\begin{aligned} \mathcal{L}_{\text{int}} &\sim Z'_\mu \left( \bar{\psi}_L^{\text{SM}} \gamma^\mu (3T_{B-L}) \psi_L^{\text{SM}} \right) \\ &\sim Z'_\mu \left( \bar{q}_L \gamma^\mu q_L - 3\bar{\ell}_L \gamma^\mu \ell_L \right). \end{aligned}$$

# Interference with a $Z'$ vector boson

Non-Resonant Production (leptoquarks)    Resonant Production (neutral bosons)



$$\mathcal{M}_{U_1} \sim \frac{1}{t - m_{U_1}^2 + im_{U_1}\Gamma_{U_1}}, \quad (4)$$

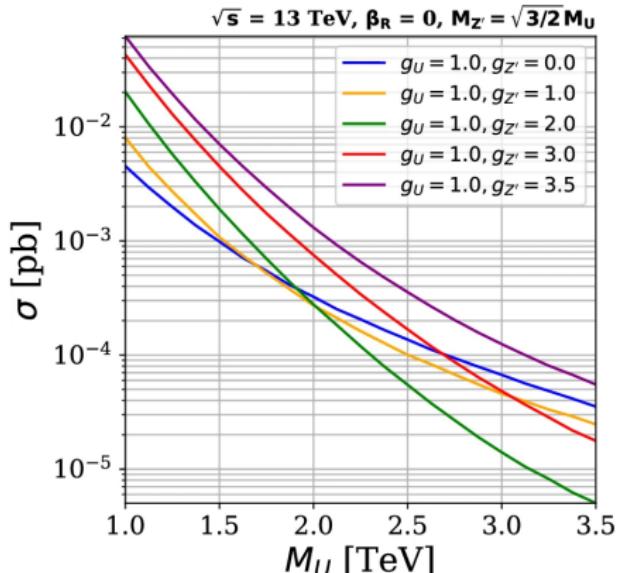
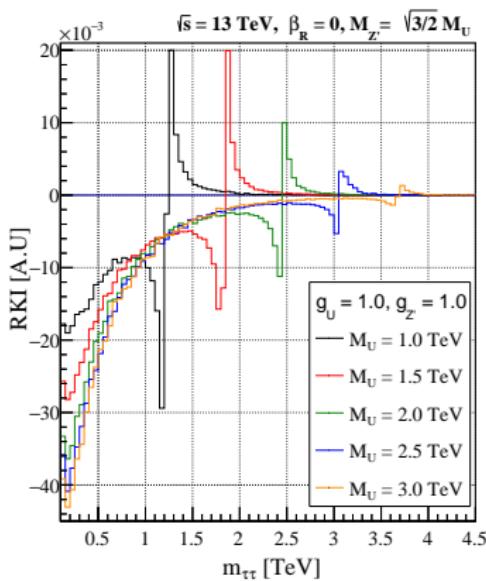
$$\mathcal{M}_{Z'} \sim \frac{1}{s - m_{Z'}^2 + im_{Z'}\Gamma_{Z'}}, \quad (5)$$

the interference has the form

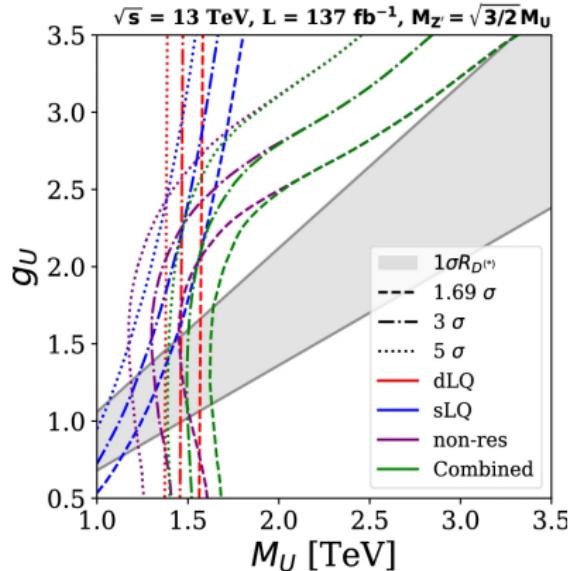
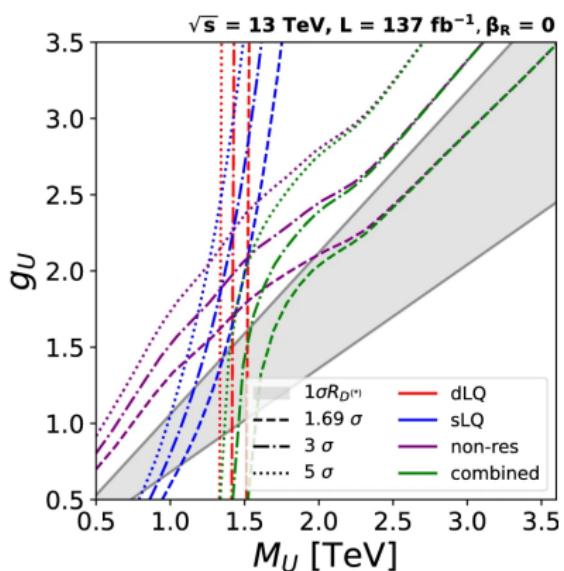
$$\frac{d}{dm} \left[ \sigma_{LQ+Z'} - \left( \sigma_{LQ} + \sigma_{Z'} \right) \right] \sim \frac{g_{z'} g_U}{s} \frac{m_{LQ} m_{Z'} \Gamma_{LQ} \Gamma_{Z'} - (t - m_{LQ}^2)(s - m_{Z'}^2)}{\left[ (t - m_{LQ}^2)^2 + m_{LQ}^2 \Gamma_{LQ}^2 \right] \left[ (s - m_{Z'}^2)^2 + m_{Z'}^2 \Gamma_{Z'}^2 \right]}.$$

Similary for Polarized final states.

# Interference with a $Z'$ vector boson



# Effects on the Sensitivity reach



# Summary and Conclusions



# Backup Slides

Where the SM charges for the leptoquark, in the  $Y = 2(Q - T_3)$  convention, are

	$\bar{q}_L$	$\ell_L^j$	$\bar{q}_L \gamma_\mu \ell_L$	$U_1^\mu$
$U(1)$	$-1/3$	$-1$	$-4/3$	$+4/3$
$SU(2)$	$\bar{\mathbf{2}}$	$\mathbf{2}$	$\mathbf{1}$	$\mathbf{1}$
$SU(3)$	$\bar{\mathbf{3}}$	$\mathbf{1}$	$\mathbf{3}$	$\mathbf{3}$

Then, the leptoquark  $U_1 \sim (\mathbf{3}_C, \mathbf{1}_I, 4/3_Y)$ .

Where the SM charges for the leptoquark, in the  $Y = 2(Q - T_3)$  convention, are

	$\bar{q}_L$	$\ell_L^j$	$\bar{q}_L \gamma_\mu \ell_L$	$U_1^\mu$
$U(1)$	$-1/3$	$-1$	$-4/3$	$+4/3$
$SU(2)$	$\bar{\mathbf{2}}$	$\mathbf{2}$	$\mathbf{1}$	$\mathbf{1}$
$SU(3)$	$\bar{\mathbf{3}}$	$\mathbf{1}$	$\bar{\mathbf{3}}$	$\mathbf{3}$

Then, the leptoquark  $U_1 \sim (\mathbf{3}_C, \mathbf{1}_I, 4/3_Y)$ .

The full Lagrangian for the vector leptoquark is

$$\begin{aligned} \mathcal{L}_U = & -\frac{1}{2} U_{\mu\nu}^\dagger U^{\mu\nu} + M_U^2 U_\mu^\dagger U^\mu - i g_s U_\mu^\dagger T^a U_\nu G_a^{\mu\nu} - \frac{2i}{3} g' U_\mu^\dagger U_\nu B^{\mu\nu} \\ & + \frac{g_u}{\sqrt{2}} \left[ U_1^\mu \left( \beta_L^{ij} \bar{q}_L^i \gamma_\mu e_L^j + \beta_R^{ij} \bar{d}_R^i \gamma_\mu e_R^j \right) + \text{h.c.} \right] \end{aligned}$$

where  $U_{\mu\nu} = \mathcal{D}_\mu U_\nu - \mathcal{D}_\nu U_\mu$ ,  $\mathcal{D}_\mu = \partial_\mu - i g_s G_\mu^a T^a - i \frac{2}{3} g_Y B_\mu$ , and the couplings  $\beta_L$  and  $\beta_R$  are complex  $3 \times 3$  matrices in flavor space.

Where the SM charges for the leptoquark, in the  $Y = 2(Q - T_3)$  convention, are

	$\bar{q}_L$	$\ell_L^j$	$\bar{q}_L \gamma_\mu \ell_L$	$U_1^\mu$
$U(1)$	$-1/3$	$-1$	$-4/3$	$+4/3$
$SU(2)$	$\bar{\mathbf{2}}$	$\mathbf{2}$	$\mathbf{1}$	$\mathbf{1}$
$SU(3)$	$\bar{\mathbf{3}}$	$\mathbf{1}$	$\bar{\mathbf{3}}$	$\mathbf{3}$

Then, the leptoquark  $U_1 \sim (\mathbf{3}_C, \mathbf{1}_I, 4/3_Y)$ .

The full Lagrangian for the vector leptoquark is

$$\begin{aligned} \mathcal{L}_U = & -\frac{1}{2} U_{\mu\nu}^\dagger U^{\mu\nu} + M_U^2 U_\mu^\dagger U^\mu - i g_s U_\mu^\dagger T^a U_\nu G_a^{\mu\nu} - \frac{2i}{3} g' U_\mu^\dagger U_\nu B^{\mu\nu} \\ & + \frac{g_u}{\sqrt{2}} \left[ U_1^\mu \left( \beta_L^{ij} \bar{q}_L^i \gamma_\mu e_L^j + \beta_R^{ij} \bar{d}_R^i \gamma_\mu e_R^j \right) + \text{h.c.} \right] \end{aligned}$$

where  $U_{\mu\nu} = \mathcal{D}_\mu U_\nu - \mathcal{D}_\nu U_\mu$ ,  $\mathcal{D}_\mu = \partial_\mu - i g_s G_\mu^a T^a - i \frac{2}{3} g_Y B_\mu$ , and the couplings  $\beta_L$  and  $\beta_R$  are complex  $3 \times 3$  matrices in flavor space.

Constraints from  $\Delta F = 2$  and lepton flavor violation impose a hierarchy with dominant third generation couplings:

$$|\beta_L^{11}|, |\beta_L^{12}|, |\beta_L^{21}|, |\beta_L^{22}|, |\beta_L^{31}| \ll |\beta_L^{13}| \ll |\beta_L^{23}|, |\beta_L^{32}| \ll |\beta_R^{33}|, |\beta_L^{33}| = \mathcal{O}(1), \quad (6)$$

where  $\beta_R$  is diagonal.

# Two body scattering

CM-Frame

Consider the process

$$A(\vec{p}_1) + B(\vec{p}_2) \longrightarrow C(\vec{p}_3) + D(\vec{p}_4), \quad (7)$$



From the Golden Rule, the cross section is given by

$$\sigma = \frac{n!(2\pi)^4}{4\sqrt{(\vec{p}_1 \cdot \vec{p}_2)^2 - (m_1 m_2)^2}} \int |\mathcal{M}|^2 \delta^{(4)}(p_1 + p_2 - p_3 - p_4) \frac{d^3 \vec{p}_3}{(2\pi)^3 2E_3} \frac{d^3 \vec{p}_4}{(2\pi)^3 2E_4}. \quad (8)$$

But, in the CM frame,  $\vec{p}_1 + \vec{p}_2 = 0$ , where

$$\sqrt{(\vec{p}_1 \cdot \vec{p}_2)^2 - (m_1 m_2)^2} = E_1 E_2 |\vec{p}_1|, \quad (9)$$

$$\delta^{(4)}(p_1 + p_2 - p_3 - p_4) = \delta(E_1 + E_2 - E_3 - E_4) \delta^{(3)}(\vec{p}_3 + \vec{p}_4). \quad (10)$$

Thus

$$\sigma = \left(\frac{1}{8\pi}\right)^2 \frac{n!}{(E_1 E_2) |\vec{p}_1|} \int |\mathcal{M}|^2 \frac{\delta\left(E_1 + E_2 - \sqrt{\vec{p}_3^2 + m_3^2} - \sqrt{\vec{p}_4^2 + m_4^2}\right)}{\sqrt{\vec{p}_3^2 + m_3^2} \sqrt{\vec{p}_4^2 + m_4^2}} d\vec{p}_3 \quad (11)$$

Integrating over the radial part  $|\vec{p}_3|$ , we get

$$\sigma = \left(\frac{1}{8\pi}\right)^2 \frac{n! |\vec{p}_3|}{(E_1 + E_2)^2 |\vec{p}_1|} \int |\mathcal{M}|^2 d\Omega, \quad (12)$$

with

$$|\vec{p}_3| = \frac{1}{2} \frac{\sqrt{((E_1 + E_2)^2 - m_3^2 - m_4^2)^2 - 4m_3^2 m_4^2}}{E_1 + E_2}, \quad (13)$$

the outgoing momentum in the CM frame.

We prefer work with differential cross section as

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2} \frac{n!}{(E_1 + E_2)^2} \frac{|\vec{p}_3|}{|\vec{p}_1|} |\mathcal{M}|^2. \quad (14)$$

Defining  $\sqrt{s} = E_1 + E_2$ , we have

$$|\vec{p}_3| = \frac{1}{2} \frac{\sqrt{(s - m_3^2 - m_4^2)^2 - 4m_3^2 m_4^2}}{\sqrt{s}}, \quad |\vec{p}_1| = \frac{1}{2} \frac{\sqrt{(s - m_1^2 - m_2^2)^2 - 4m_1^2 m_2^2}}{\sqrt{s}}. \quad (15)$$

so the differential cross section is

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2} \frac{n!}{s} \sqrt{\frac{(s - (m_3 + m_4)^2)(s - (m_3 - m_4)^2)}{(s - (m_1 + m_2)^2)(s - (m_1 - m_2)^2)}} |\mathcal{M}|^2. \quad (16)$$

Note that at this point, we don't need to know the explicit form of the matrix element  $\mathcal{M}$ , so it is a generic result.

## Writting $t$ in terms of $s$ and $\theta$

In general, there are three Lorentz-invariant useful kinematical variables to describe the scattering process, known as Mandelstam variables:

$$\hat{s} = (p_1 + p_2)^2 = (p_3 + p_4)^2 = m_1^2 + m_2^2 + 2p_1^\mu p_{2\mu} = m_3^2 + m_4^2 + 2p_3^\mu p_{4\mu}, \quad (17)$$

$$\hat{t} = (p_1 - p_3)^2 = (p_2 - p_4)^2 = m_1^2 + m_3^2 - 2p_1^\mu p_{3\mu} = m_2^2 + m_4^2 - 2p_2^\mu p_{4\mu}, \quad (18)$$

$$\hat{u} = (p_1 - p_4)^2 = (p_2 - p_3)^2 = m_1^2 + m_4^2 - 2p_1^\mu p_{4\mu} = m_2^2 + m_3^2 - 2p_2^\mu p_{3\mu}. \quad (19)$$

In the CM-frame,  $\hat{s} = s = (E_1 + E_2)^2$ . If,  $m_3 = m_4$  and  $m_1 = m_2$  with  $E_1 = E_2 = E_3 = E_4 = E$ , we have

$$t = -(\vec{p}_1 - \vec{p}_3)^2 = -\vec{p}_1^2 - \vec{p}_3^2 + 2\vec{p}_1 \cdot \vec{p}_3 \quad (20)$$

where  $\vec{p}_1^2 = E^2 - m_1^2$  and  $\vec{p}_3^2 = E^2 - m_3^2$ .

So, in terms of  $s$ ,  $t$  could be written as

$$t = -2s + (m_1^2 + m_3^2) + 2\sqrt{(s/4 - m_1^2)(s/4 - m_3^2)} \cos \theta, \quad (21)$$

with  $\theta$  the scattering angle in the CM-frame.