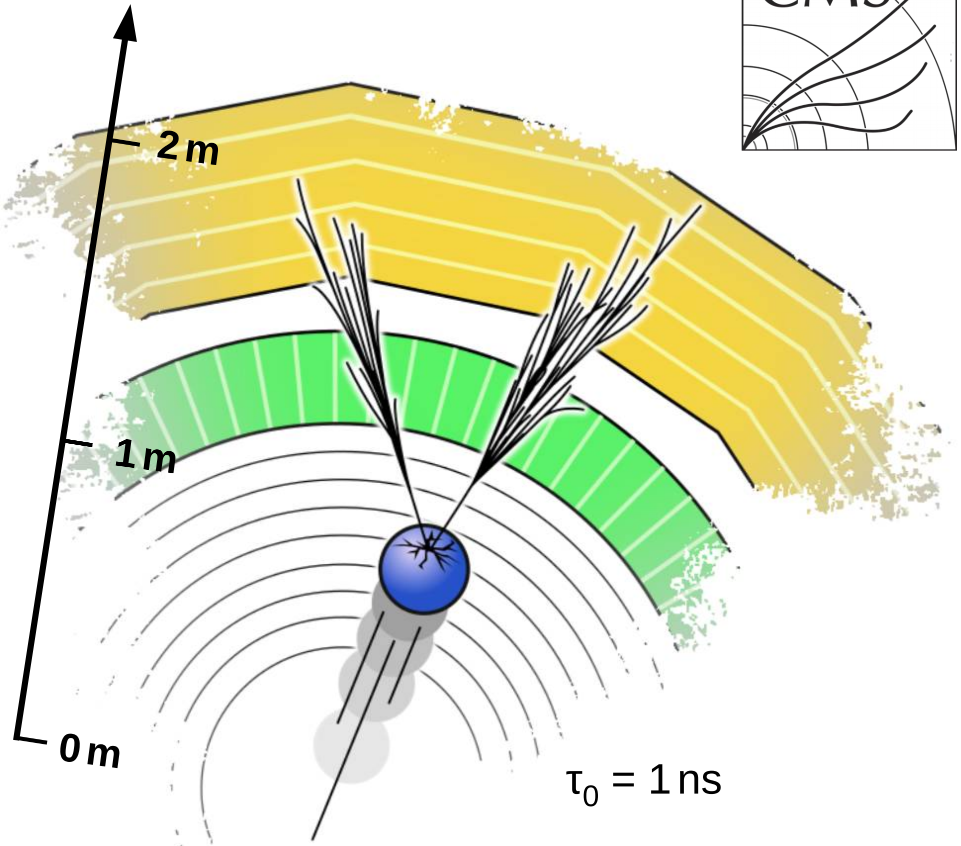
based on: → CMS Physics Analysis Summary, EXO-19-011, cds.cern.ch/record/2698267

Identification of new long-lived particle states using deep neural networks

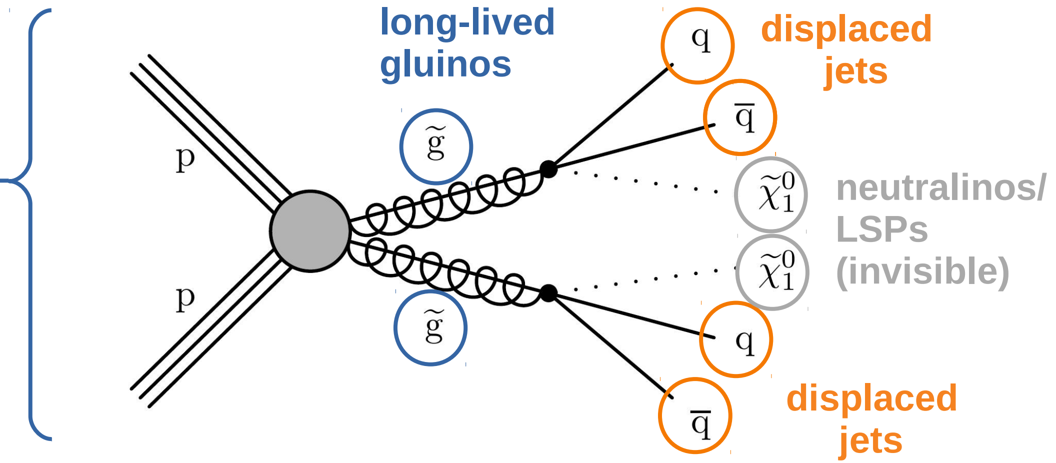


Imperial College London

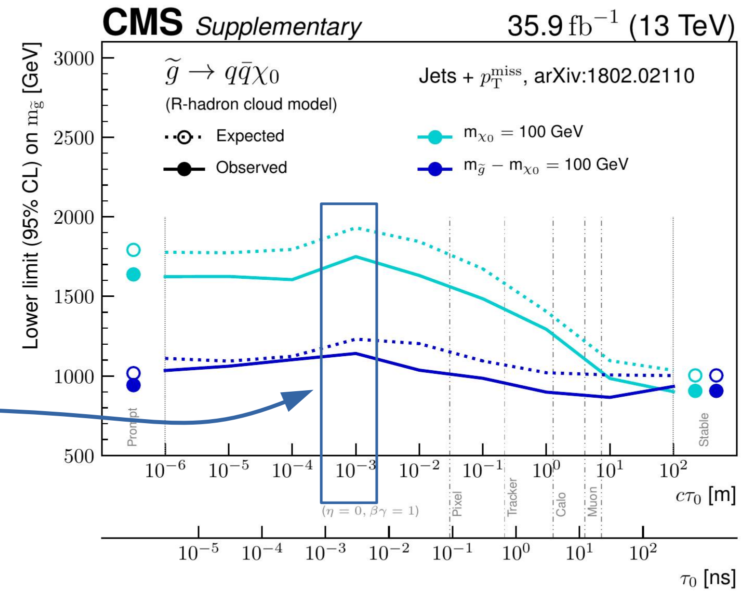
Matthias Komm
(for the CMS Collaboration)

Motivation

- long-lived particles
  - rich theoretical landscape: **split SUSY**, gauge-mediated SUSY breaking, R-parity violating SUSY, hidden valley, …
  - typically includes dark matter candidate
  - large parameter space:
    - proper decay length ($c\tau_0$)
      $\rightarrow$ $\mathcal{O}(10\ \mu\mathrm{m})\ldots\mathcal{O}(10\ \mathrm{m})$
    - gluino ($\tilde{g}$) mass & LSP ($\tilde{\chi}_1^0$) mass
      $\rightarrow$ mass difference controls p_T of jets



- existing search by CMS (JHEP 05 (2018) 025)
  - generic search for natural & split SUSY
    $\rightarrow$ sensitivity to LLPs through b-tagging ($c\tau_0 \approx 1\ \mathrm{mm}$)



- idea: enhance sensitivity with generic displaced jet tagger

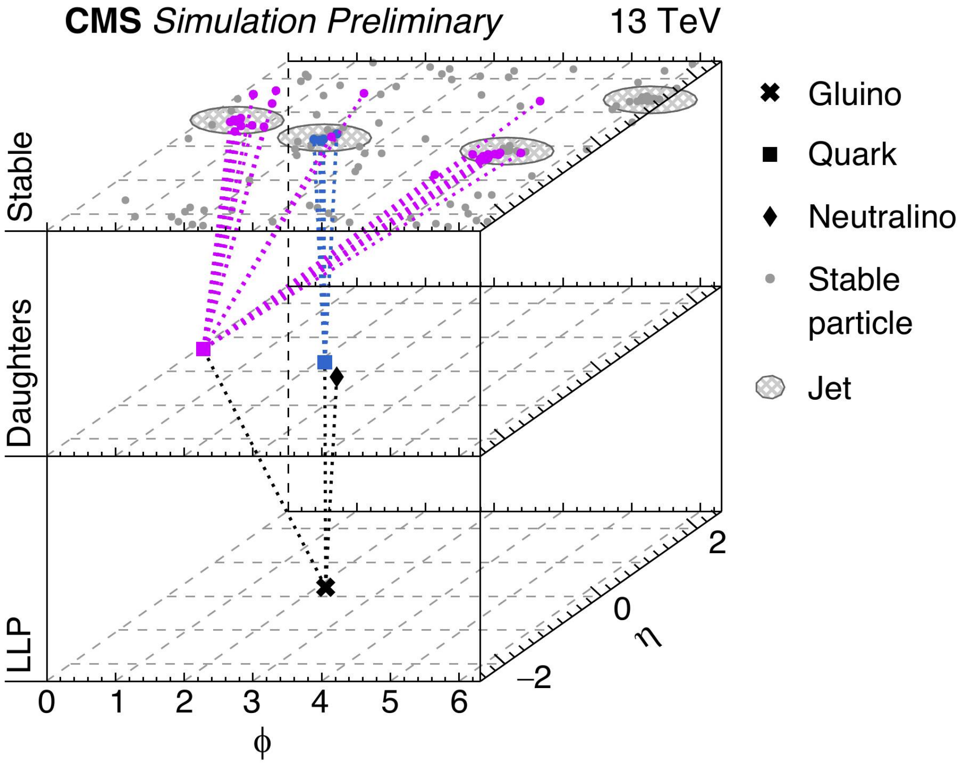
Labelling "displaced" jets

- problem: no definition in literature to be exploited
  - initial idea: "ghost" tagging as used for b, c jets
  - strong interactions between displaced quarks at the gluino decay vertex
    → ghost tagging cannot account for non-pointing jets or multiple jets from one parton

- solution
  - define jet momentum fraction of generator-level jet carried by clustered particles j per vertex v

$$f_v(\text{jet}) = \frac{\left(\sum_j \vec{p}_j \,\middle|\, j \in \text{vertex } v\right) \cdot \vec{p}_{\text{jet}}}{p_{\text{jet}}^2}, \quad f_v(\text{jet}) \in [0;\,1]$$

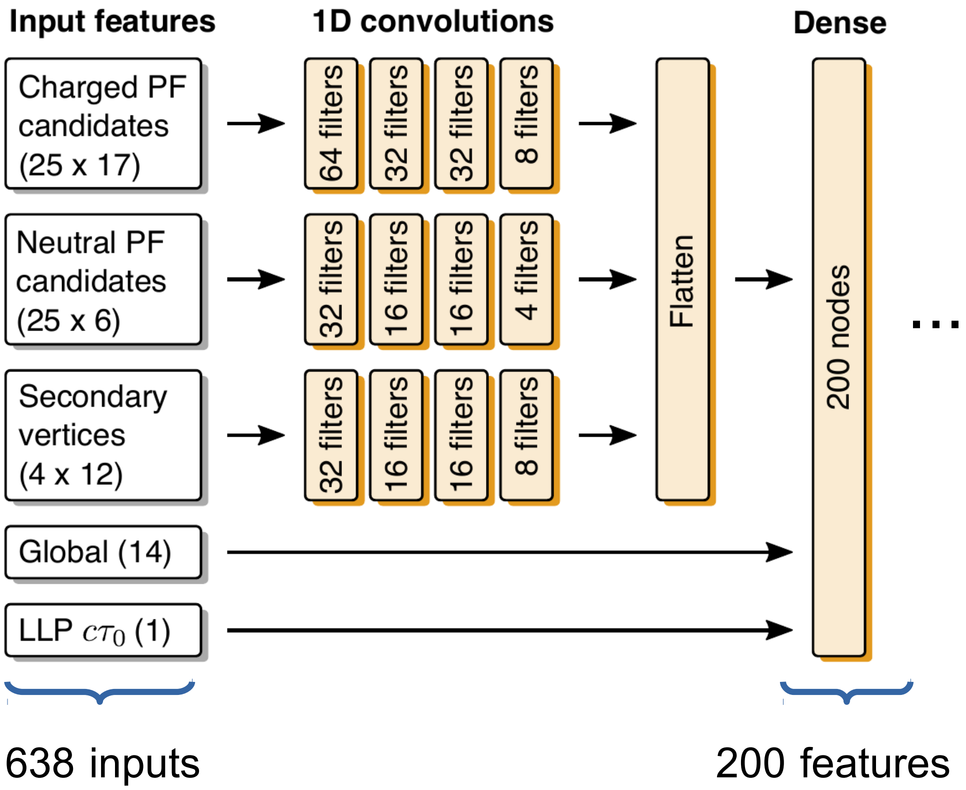
→ label jets 'LLP' where $f_v = \max$



Neural network architecture

- inspired by CMS DeepJet algorithm (latest b-tagging algorithm)
- parametrized network since importance of features changes with lifetime
- trained using jets from multijet, $\mathrm{t\bar{t}}$ & split SUSY samples to predict jet class: uds, g, b, c, LLP

➢ feature extraction
- 1d convolutions with kernel size of 1 → compresses features per constituent
- result combined with global features & lifetime → 200 highly discriminating features from 638 inputs

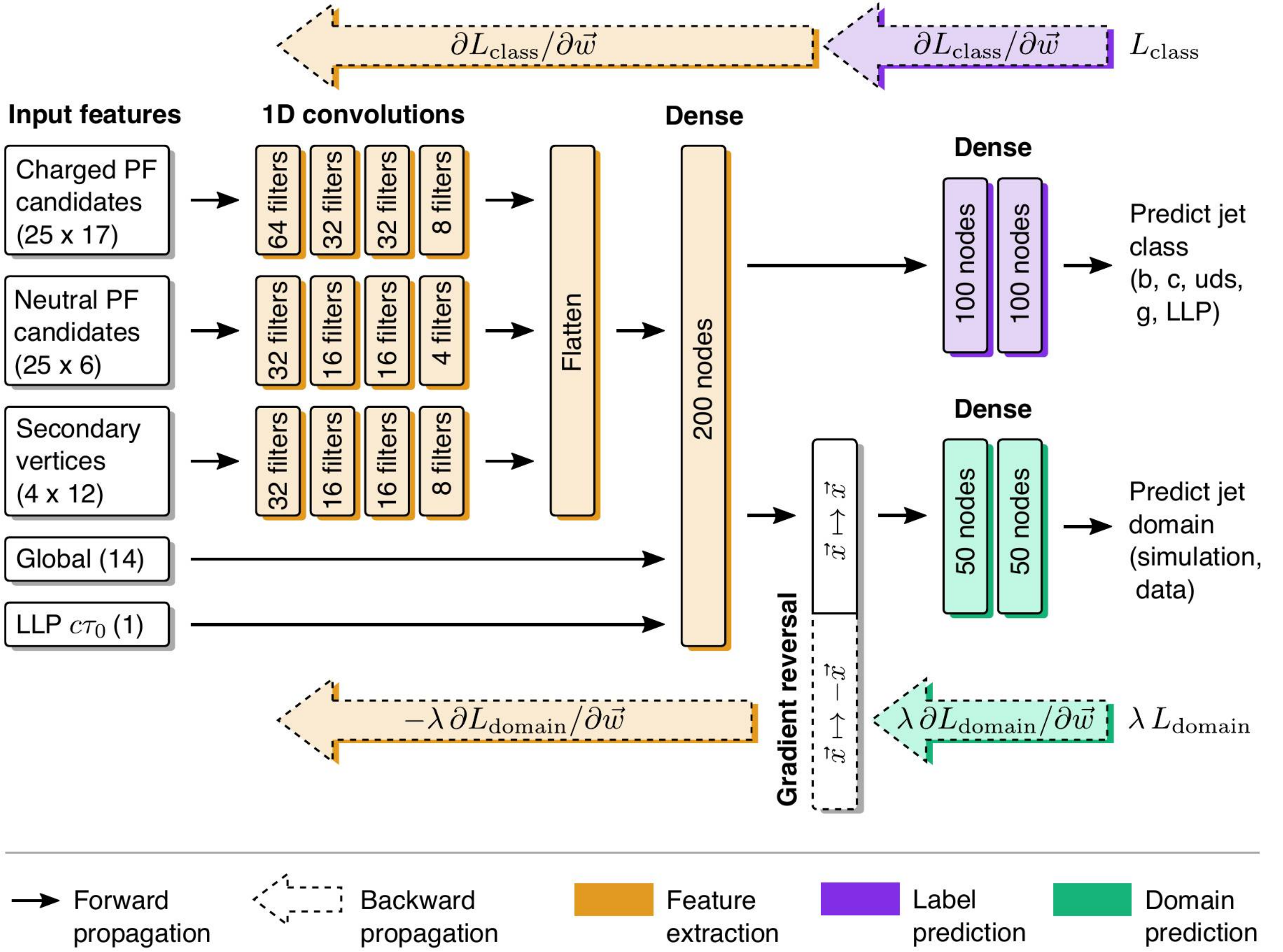


Domain adaptation

- apply domain adaptation by backpropagation to improve agreement between data/MC in control region (1505.07818)
- 200 extracted features are used to predict jet class & domain
- the summed loss is minimized: $L_{\mathrm{class}} + \lambda L_{\mathrm{domain}}$ (λ= hyperparameter)
- gradient reversal layer leads to maximization of weights wrt. domain loss in feature extraction layers

  → extracted features invariant wrt. domain; i.e. expect similar distribution & performance

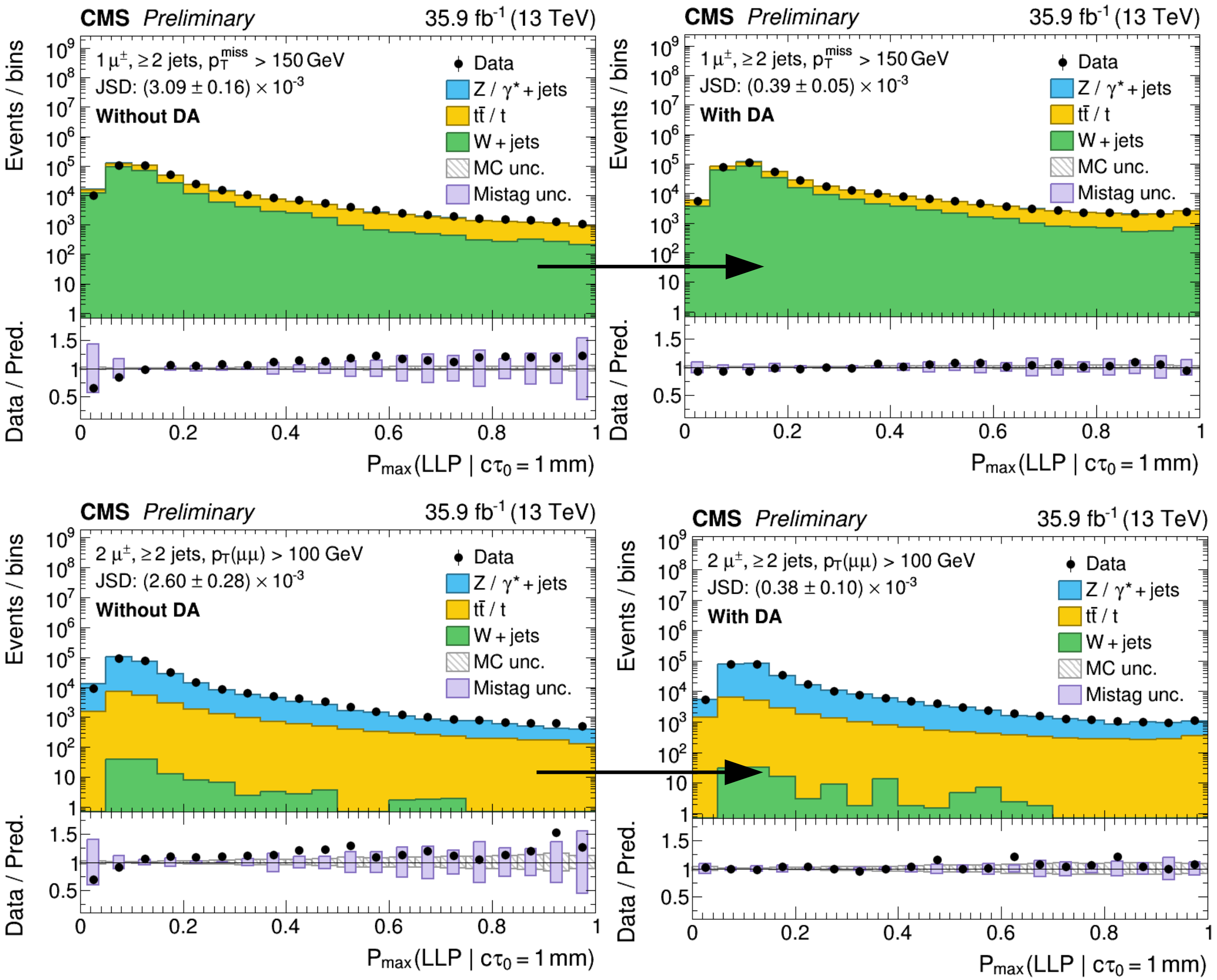
full architecture



Validation

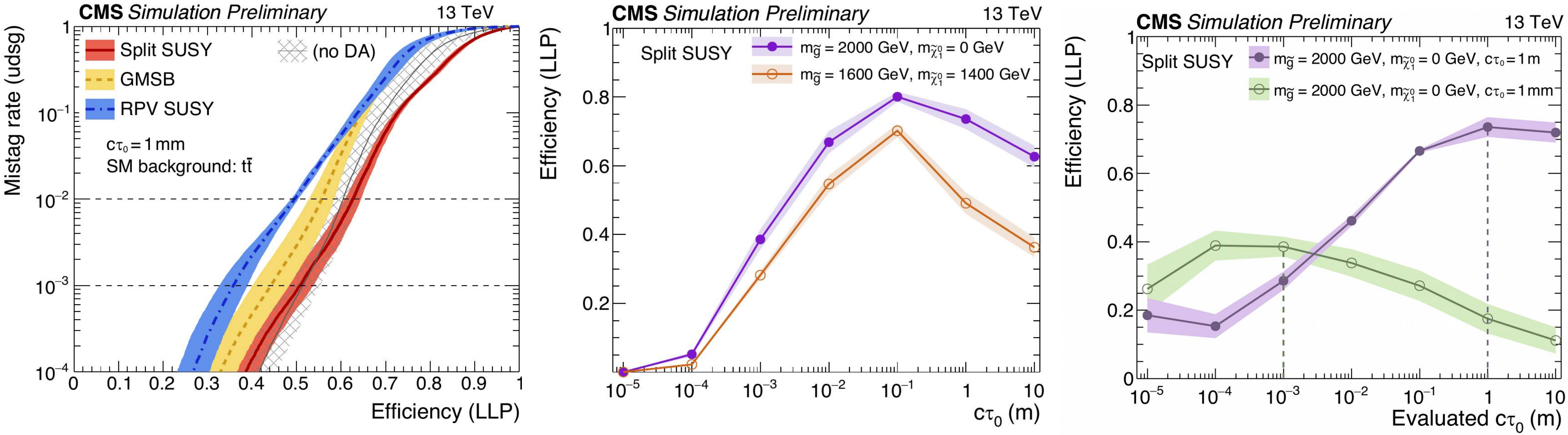
- domain adaptation uses simulated and real data jets from single muon control region
- improvement validated in dimuon control region
- deviations up to ±50% w/o DA reduced to ±10%

- mistag uncertainty derived in bins of $P(LLP|c\tau_0)$ from independent control region to cover for residual differences



Performance

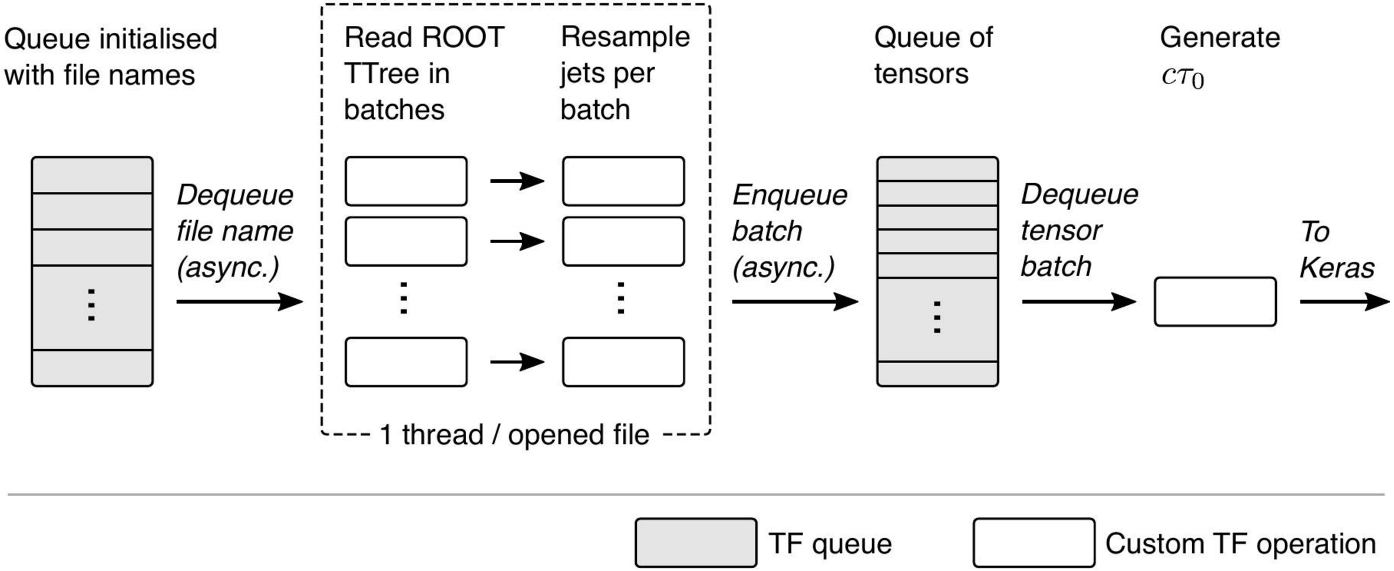
ROC curve $c\tau_0 = 1\ \mathrm{mm}$ | eff. @ fixed bkg. rate of 0.01% | evaluate for wrong $c\tau_0$



- good performance for GMSB ($\tilde{g} \to \tilde{G}g$) & RPV ($\tilde{t} \to b\ell$) models despite that the tagger was trained only with split SUSY sample
- lower discrimination power for $\mathcal{O}(10\ \mu\mathrm{m})$ lifetimes (~ within primary vertex resolution)
- evaluating at wrong lifetime results in degradation of performance
  → potential for estimating the lifetime of an unknown signal in data

Technical implementation

- training performed using keras & tensorflow packages
- developed custom preprocessing pipeline build on top of tensorflow (v1) queue system



- data is directly read & preprocessed from ROOT TTree asynchronously in CPU threads
- jets are resampled on-the-fly to achieve same p_T, η distribution for all jet classes
- a fake lifetime is generated for background jets by sampling from signal $c\tau_0$ distribution per batch

→ a demo will be released soon as well

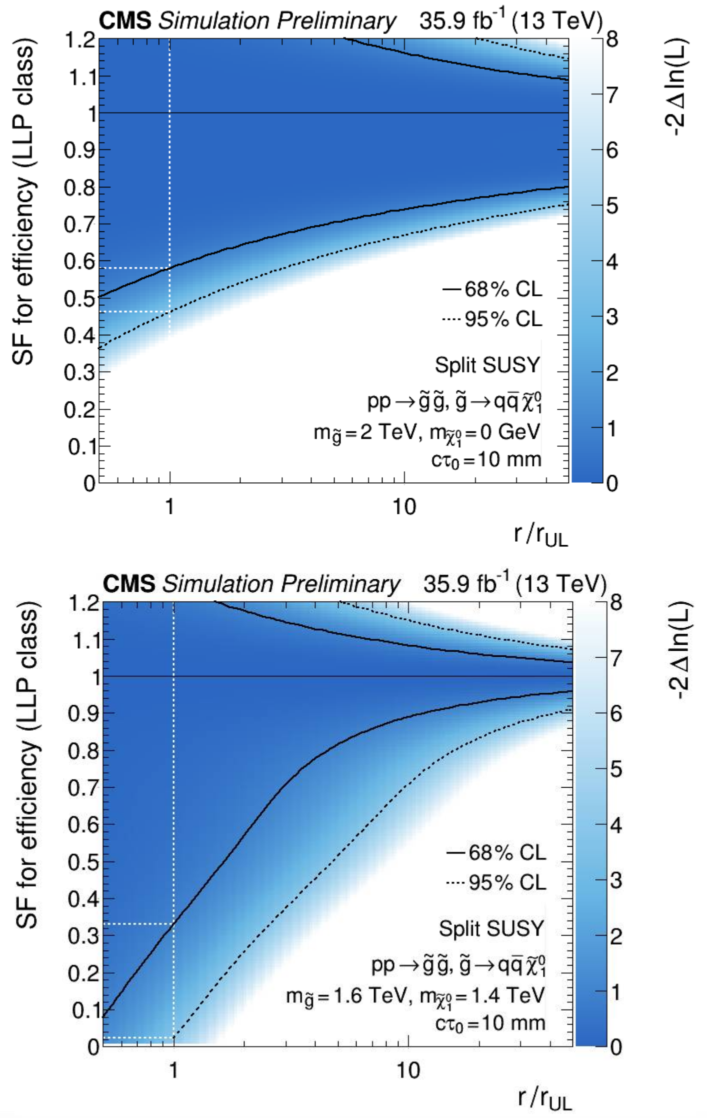
Showcase search for split SUSY

- strategy
  - select events with at least 3 jets ($p_{\mathrm{T}} > 30\ \mathrm{GeV},\ |\eta| < 2.4$) $H_{\mathrm{T}}^{\mathrm{miss}} > 300\ \mathrm{GeV}$, $H_{\mathrm{T}}^{\mathrm{miss}}/p_{\mathrm{T}}^{\mathrm{miss}} < 1.25$, veto $\mathrm{e}^{\pm}/\mu^{\pm}$
  - classify events depending on $H_{\mathrm{T}}, \#\mathrm{jets}, \#\mathrm{tags}$

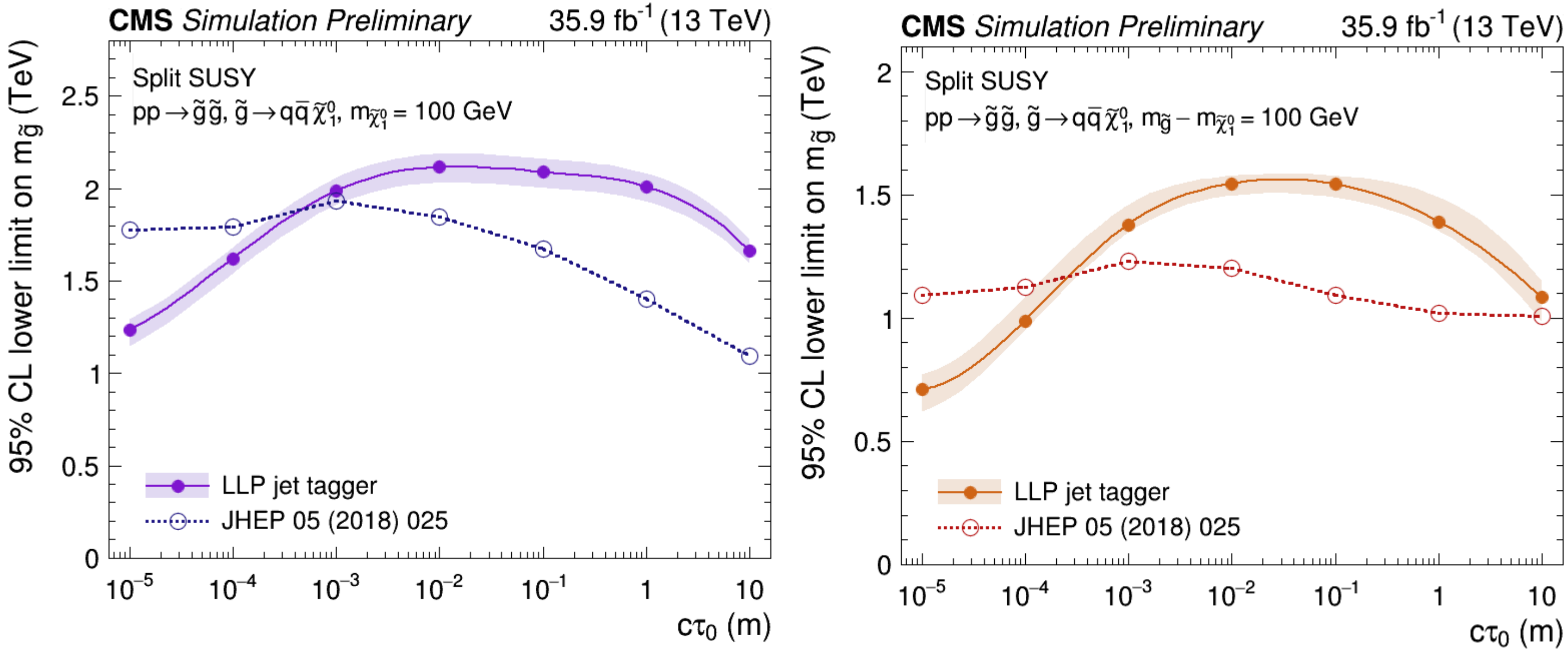
- signal efficiency
  - differences in signal efficiency between data/MC a priori unknown
  - idea: incorporate unknown signal efficiency as nuisance parameter in statistical model through event weight

$$w = \left(\frac{1-\mathrm{SF}\,\epsilon_{\mathrm{MC}}}{1-\epsilon_{\mathrm{MC}}}\right)^{(N_{\mathrm{jet}}-N_{\mathrm{tag}})} \times \mathrm{SF}^{N_{\mathrm{tag}}}$$

→ scale factor (SF) can be constrained in-situ with the chosen categorization of events



Expected limits on $pp \to \tilde{g}\tilde{g},\ \tilde{g} \to q\bar{q}\chi_1^0$



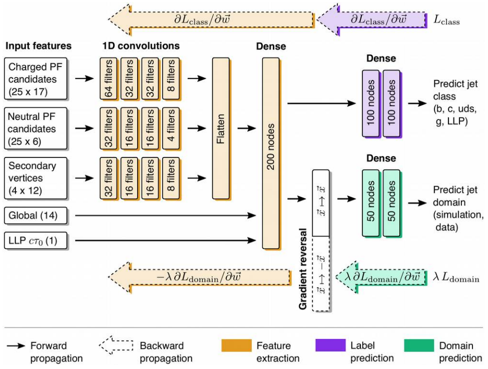
- competitive (expected) limits obtained using simple binning scheme
- largest uncertainty from finite simulation sample size
- clear gain for lifetimes $c\tau_0 \geq 1\mathrm{mm}$ over previous search based on b-tagging
- less sensitive at lower lifetimes since event kinematics were not (yet) explored

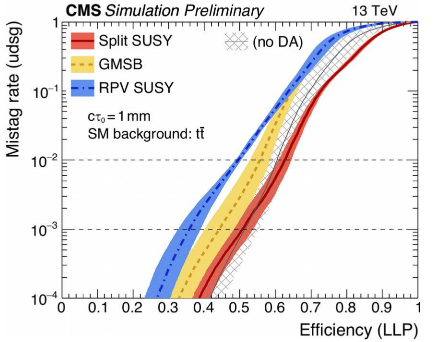
Summary

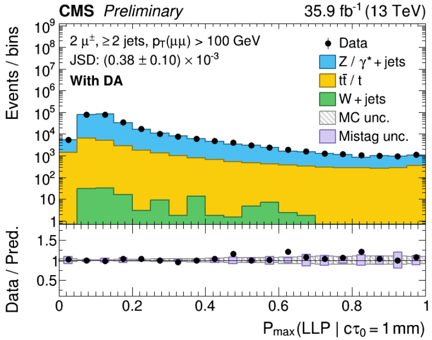
- jet tagger for generic displaced jets
  - displaced jets definition
  - parametrized neural network
  - domain adaptation to improve data/MC modeling in control regions
  - good performance also for models not in training
  - custom input pipeline for preprocessing which reads ROOT TTree directly
- showcase application
  - signal model: split SUSY
  - simple event categorization ($H_\mathrm{T}, \#\mathrm{jets}, \#\mathrm{tags}$)
  - in-situ constraint of unknown signal efficiency
  - competitive expected limits obtained for $c\tau_0 \geq 1\mathrm{mm}$

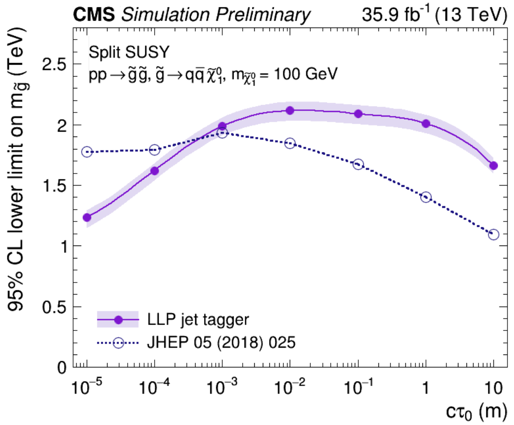
- further information

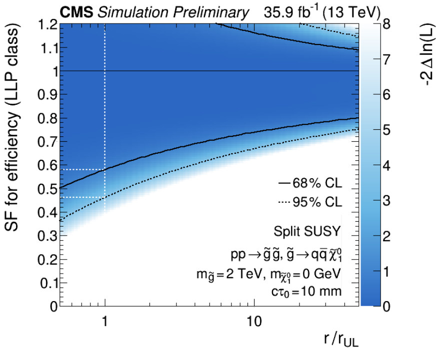
  → CMS Physics Analysis Summary, EXO-19-011, cds.cern.ch/record/2698267





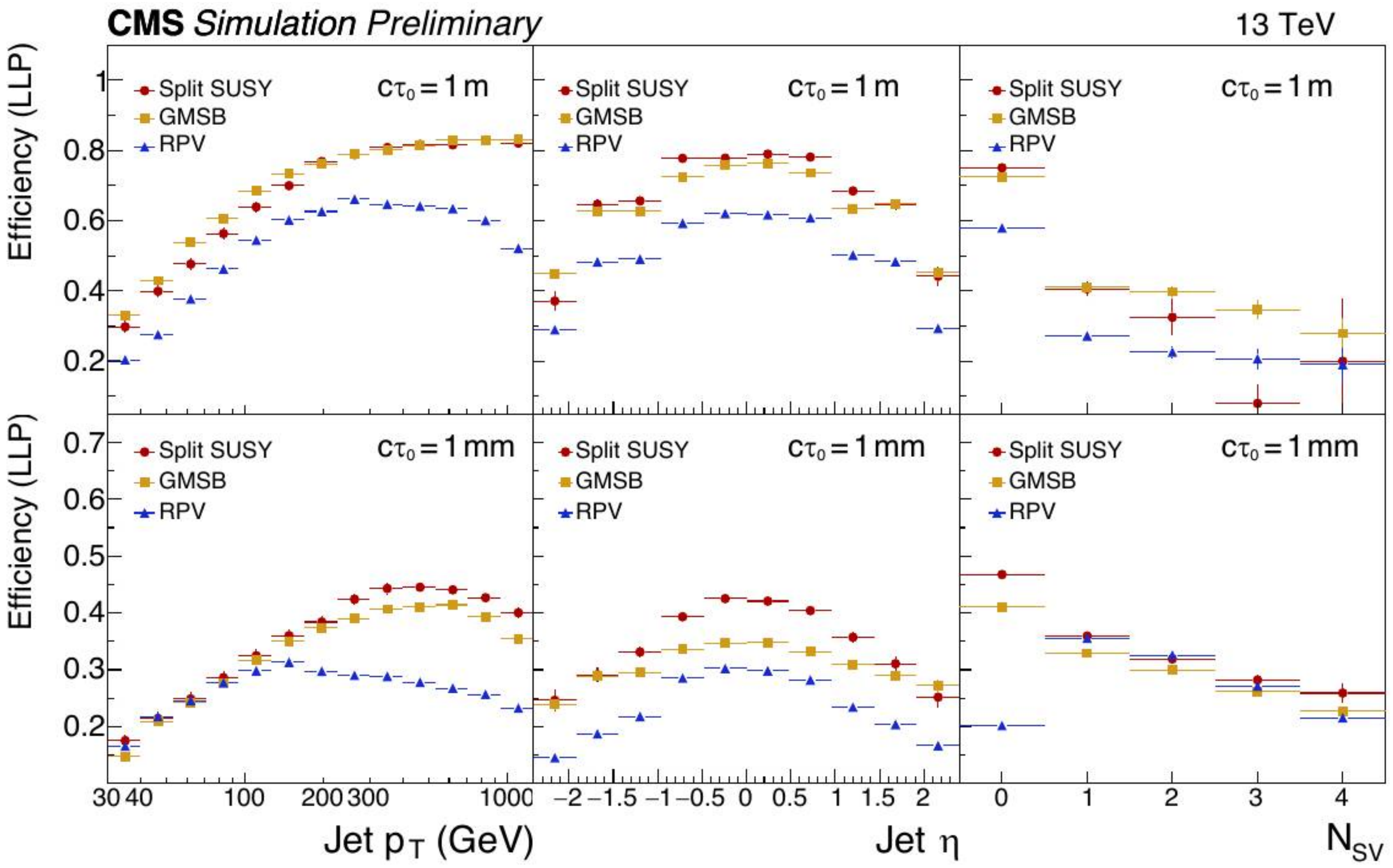






Backup

Performance

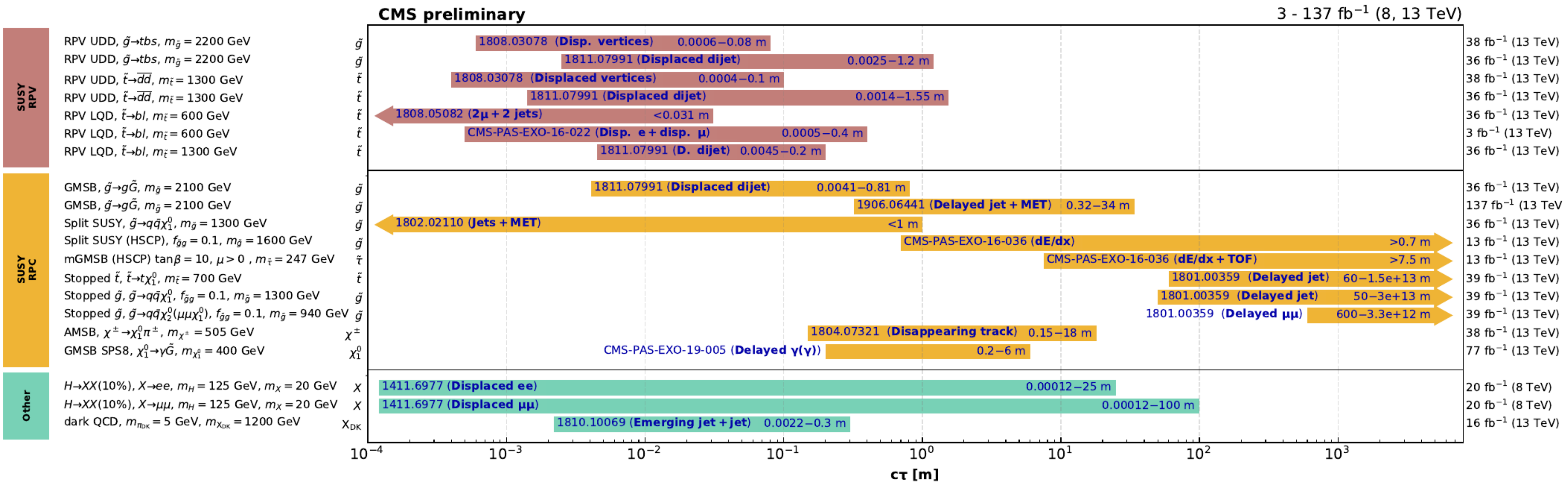


Yield table

H_T (GeV)	300–800	300–800	300–800	>800	>800	>800
(N_{jet}, N_{tag})	(3–4, $\geq$2)	(5, $\geq$2)	($\geq$6, $\geq$3)	(3–4, $\geq$2)	(5, $\geq$2)	($\geq$6, $\geq$3)
$Z^0(\rightarrow \nu\overline{\nu})$+jets	40.7 ± 39.2	6.5 ± 5.8	0.6 ± 0.4	3.3 ± 2.8	1.6 ± 1.2	0.1 ± 0.1
$W(\rightarrow \ell\nu)$+jets	56.3 ± 44.1	11.6 ± 5.1	1.5 ± 0.5	3.6 ± 2.5	1.2 ± 3.0	< 0.1
$t\bar{t}$	39.6 ± 36.1	17.9 ± 15.7	1.9 ± 1.1	2.1 ± 1.3	3.2 ± 2.4	3.0 ± 2.1
Single top	5.7 ± 5.2	2.6 ± 2.2	0.3 ± 0.2	0.6 ± 0.4	0.5 ± 0.3	0.4 ± 0.3
Total SM	142 ± 69	38.5 ± 17.6	4.3 ± 1.3	9.7 ± 4.0	6.5 ± 4.1	3.5 ± 2.5
Uncompressed	< 0.1	< 0.1	< 0.1	3.0 ± 2.9	3.8 ± 3.7	5.7 ± 5.5
Compressed	5.4 ± 5.0	4.2 ± 3.8	2.8 ± 2.5	1.1 ± 0.9	2.5 ± 2.2	4.5 ± 4.1

- signal scenarios:
  - uncompressed: $c\tau_0 = 1\ \mathrm{mm},\ m_{\tilde{g}} = 2\ \mathrm{TeV},\ m_{\tilde{\chi}_1^1} = 0\ \mathrm{TeV}$
  - compressed: $c\tau_0 = 1\ \mathrm{mm},\ m_{\tilde{g}} = 1.6\ \mathrm{TeV},\ m_{\tilde{\chi}_1^1} = 1.4\ \mathrm{TeV}$

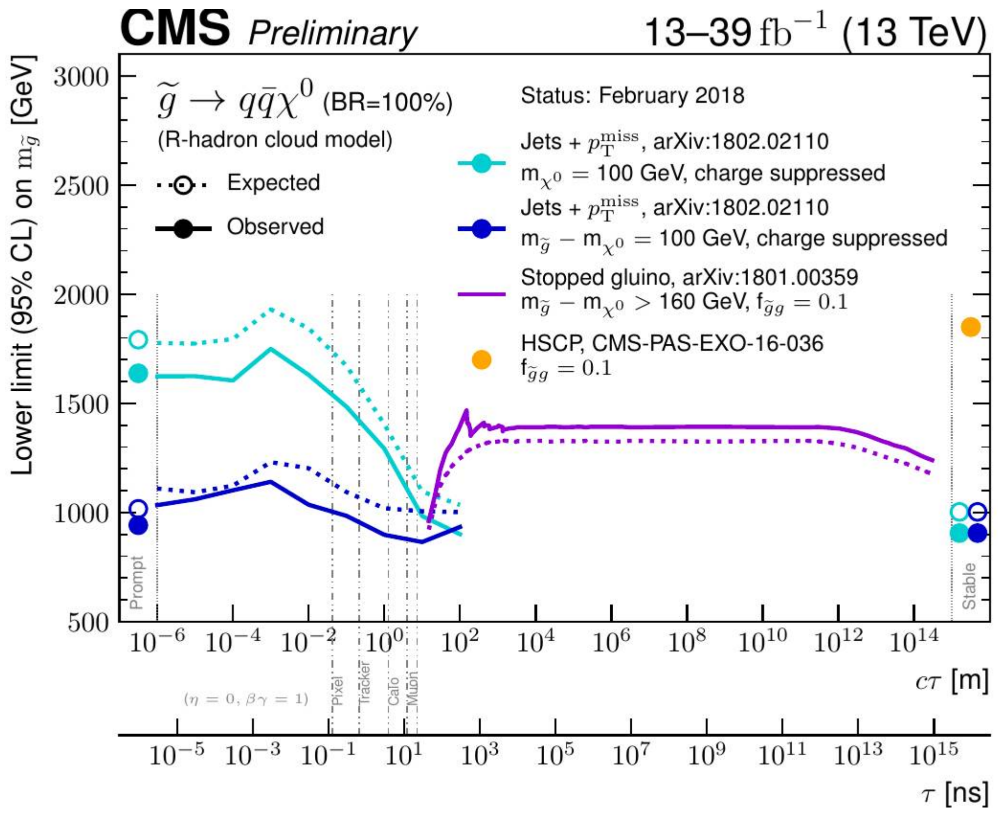
CMS limits overview

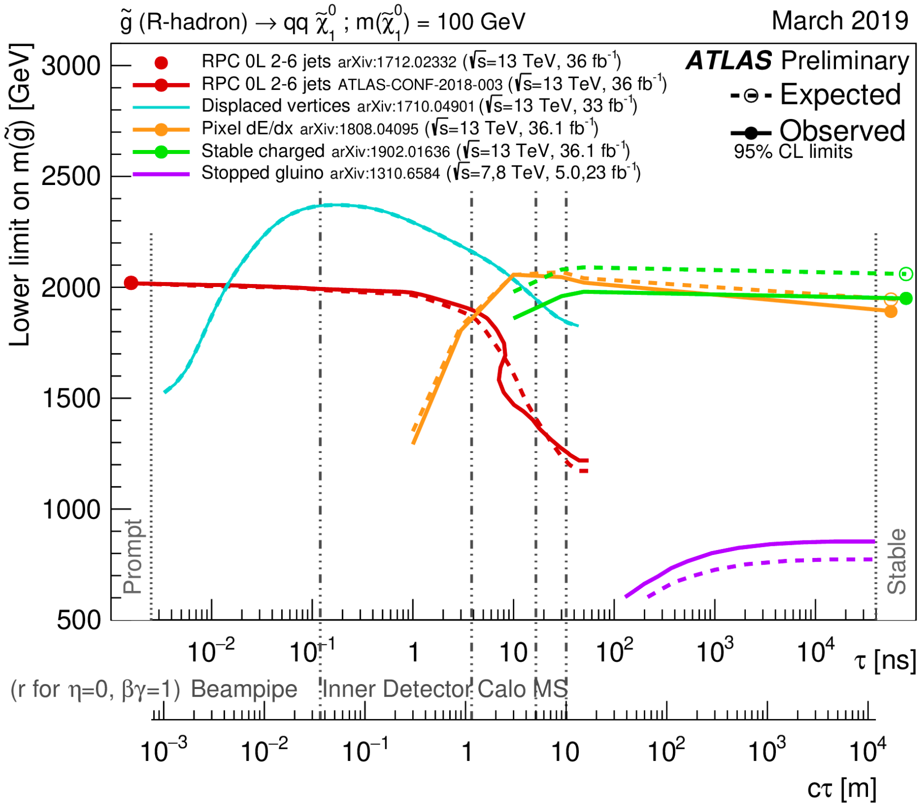


July 2019

Selection of observed exclusion limits at 95% C.L. (theory uncertainties are not included). The y-axis tick labels indicate the studied long-lived particle.

Limits


CMS Preliminary
13–39 fb^-1 (13 TeV)
g̃ → qq̄χ^0 (BR=100%)
(R-hadron cloud model)
Expected
Observed
Status: February 2018
Jets + p_T^miss, arXiv:1802.02110
m_χ0 = 100 GeV, charge suppressed
Jets + p_T^miss, arXiv:1802.02110
m_g̃ − m_χ0 = 100 GeV, charge suppressed
Stopped gluino, arXiv:1801.00359
m_g̃ − m_χ0 > 160 GeV, f_g̃g = 0.1
HSCP, CMS-PAS-EXO-16-036
f_g̃g = 0.1
Lower limit (95% CL) on m_g̃ [GeV]
cτ [m]
τ [ns]
Prompt
Stable
Pixel
Tracker
Calo
Muon
(η = 0, βγ = 1)


g̃ (R-hadron) → qq χ̃1^0 ; m(χ̃1^0) = 100 GeV
March 2019
ATLAS Preliminary
Expected
Observed
95% CL limits
RPC 0L 2-6 jets arXiv:1712.02332 (√s=13 TeV, 36 fb^-1)
RPC 0L 2-6 jets ATLAS-CONF-2018-003 (√s=13 TeV, 36 fb^-1)
Displaced vertices arXiv:1710.04901 (√s=13 TeV, 33 fb^-1)
Pixel dE/dx arXiv:1808.04095 (√s=13 TeV, 36.1 fb^-1)
Stable charged arXiv:1902.01636 (√s=13 TeV, 36.1 fb^-1)
Stopped gluino arXiv:1310.6584 (√s=7,8 TeV, 5.0,23 fb^-1)
Lower limit on m(g̃) [GeV]
Prompt
Stable
τ [ns]
(r for η=0, βγ=1) Beampipe Inner Detector Calo MS
cτ [m]