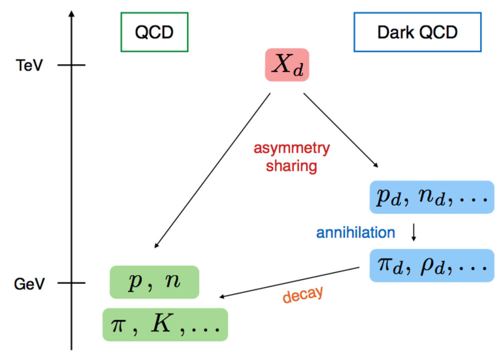
Search for New Physics with Emerging Jets in CMS

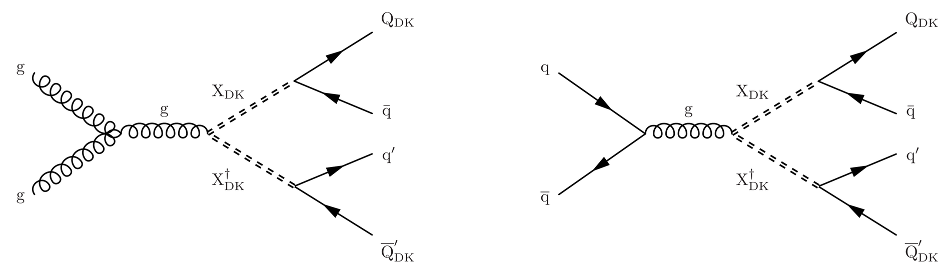
Yongbin Feng on behalf of CMS Collaboration
University of Maryland, College Park

Introduction

- **Dark QCD**: [arXiv:1306.4676]
  - One class of BSM models that has confining properties similar to QCD, but is not charged under the forces of the standard model
  - Provide proton-equivalent dark matter candidate
  - Naturally explain the observed similar mass densities of baryonic matter and dark matter



- **Emerging Jets model**: [arXiv:1502.05409]



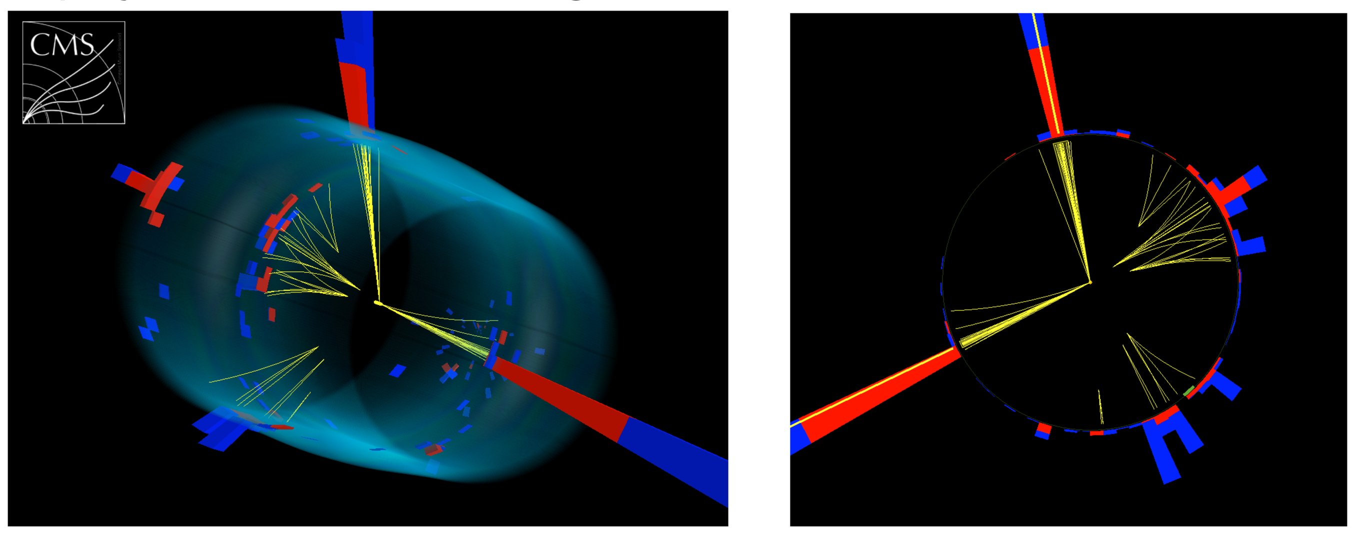
  - The complex scalar mediator X_{DK} can be pair produced via gluon fusion or quark-antiquark annihilation, and it decays to a down quark and a dark quark Q_{DK}
  - Dark hadrons will be produced from the dark quark, and then decay to SM particles via the mediator X_{DK}, producing many displaced tracks and vertices, and significant missing transverse momentum for models with long lifetime dark mesons (∼meter)

Signal Simulations

- Three free parameters: dark pion mass $m_{\pi_{DK}}$, its proper decay length $c\tau_{\pi_{DK}}$ and the mediator mass $m_{X_{DK}}$. This analysis generate different signal events samples:

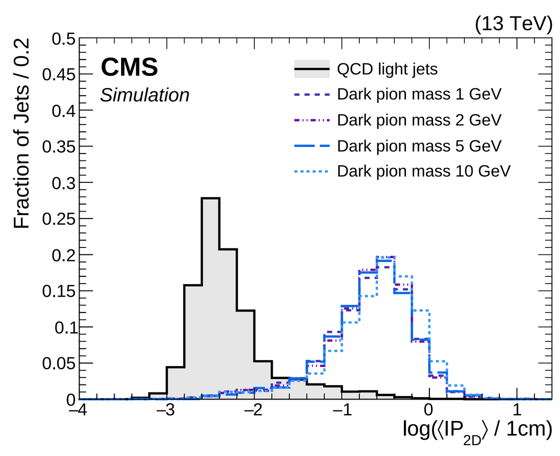
Signal model parameters	List of values
$m_{X_{DK}}$ [GeV]	400, 600, 800, 1000, 1250, 1500, 2000
$m_{\pi_{DK}}$ [GeV]	1, 2, 5, 10
$c\tau_{\pi_{DK}}$ [mm]	1, 2, 5, 25, 45, 60, 100, 150, 225, 300, 500, 1000

- Event display of one simulated signal event:



Event Selections

- Trigger: scalar p_T sum of all hadronic jets above 900 GeV
- Physics Objects:
  - Four leading jets reconstructed with anti-k_T algorithm ($\Delta R = 0.4$), with $|\eta| < 2.0$ and $p_T > 100$ GeV
  - High purity tracks with $p_T > 1$ GeV
- Emerging Jet tagging variables
  - Median of unsigned transverse impact parameters $\langle IP_{2D}\rangle$: large $\langle IP_{2D}\rangle$ for emerging jets

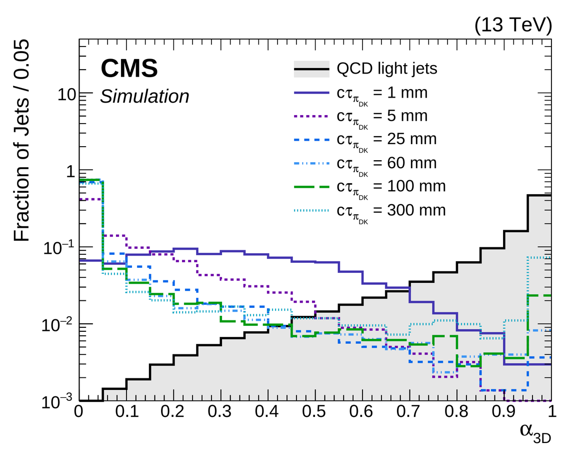


  - Define a new variable D_N as

$$D_N = \sqrt{\left[\frac{z_{PV} - z_{trk}}{0.01cm}\right]^2 + [IP_{sig}]^2}$$

where z_{PV} and z_{trk} are the z position of the primary vertex and of the track at its distance of closest approach to the PV. And a variable α_{3D} based on D_N:

$$\alpha_{3D} = \frac{p_T(\text{tracks with } D_N < \text{CUT})}{p_T(\text{all tracks})}$$



- Cuts are optimized based on $S/\sqrt{S + B + (0.1B)^2}$ and merged if the significance difference is less than 10%, resulting in a total of 7 different cut sets. One example:
  - Kinematic selections: event $H_T > 900$ GeV, four leading jets with $p_T > 225$, 100, 100, 100 GeV
  - At least two jets passing emerging selections: $D_N < 4.0$, $\langle IP_{2D}\rangle > 0.05$ cm, and $\alpha_{3D} < 0.25$

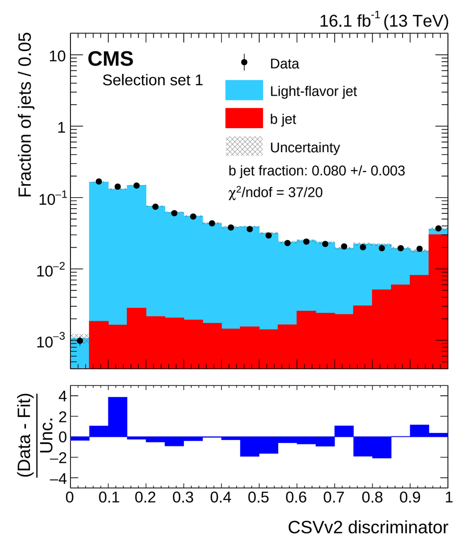
Background Estimation Method

- Mis-identification (MisID) rate has a strong dependence on jet flavor: large MisID rate for b jets
- b jets are often pair produced in LHC, causing correlation for background prediction

Need to measure MisID rates for b jets ϵ_{fb} and light jets ϵ_{fl} separately

- Use events with a high p_T photon to measure MisID rate, in order to prevent signal contamination

(a). Prepare two samples with different b jet fractions, using tag-and-probe method:
  - One requires the event with an additional jet with b-tagging discriminator value smaller than 0.2, to suppress b jet fraction
  - The other requires an additional jet with b-tagging discriminator value bigger than 0.8, to increase b jet fraction

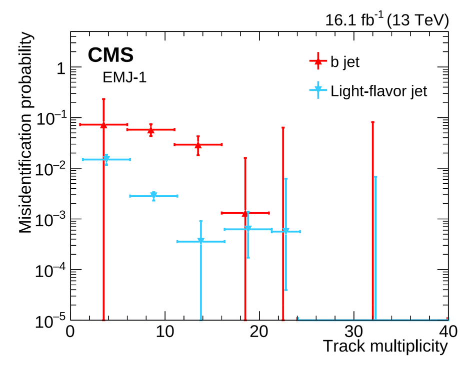


(b). Measure the MisID rate of these two samples, ϵ_{f1} and ϵ_{f2}

(c). Measure the b jet fractions of the two samples f_{b1} and f_{b2}, by fitting the b-tagging discriminator distribution with templates of b jets and light jets from MC

(d). Invert the matrix to get ϵ_{fb} and ϵ_{fl}:

$$\begin{pmatrix}\epsilon_{fb}\\ \epsilon_{fl}\end{pmatrix} = \begin{pmatrix}\frac{1-f_{b2}}{f_{b1}-f_{b2}} & \frac{-(1-f_{b1})}{f_{b1}-f_{b2}}\\ \frac{-f_{b2}}{f_{b1}-f_{b2}} & \frac{f_{b1}}{f_{b1}-f_{b2}}\end{pmatrix}\begin{pmatrix}\epsilon_{f1}\\ \epsilon_{f2}\end{pmatrix}$$

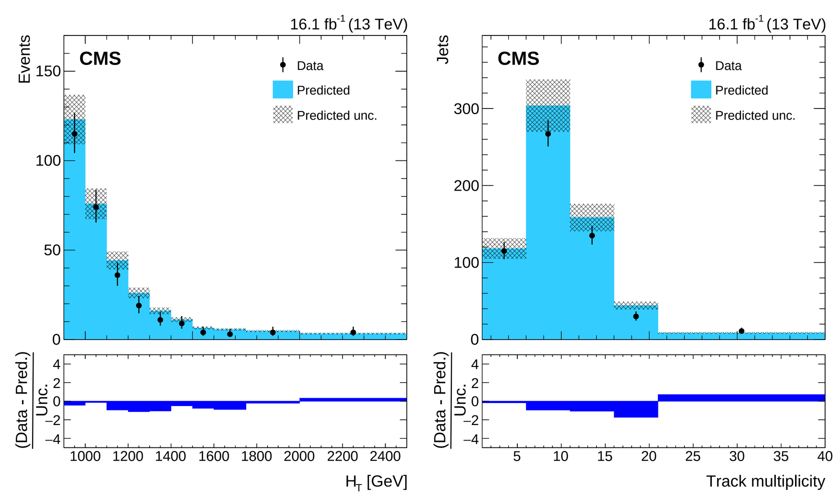


(e). Use the final MisID rate ϵ_f for the background prediction, calculated with

$$\epsilon_f = \epsilon_{fb} f_b + \epsilon_{fl}\,(1 - f_b)$$

where f_b is the b jet fraction in the control region, measured in the same way as above

Background Estimation Test

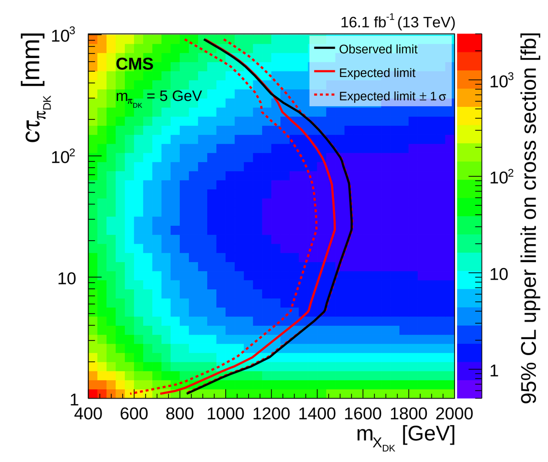


- Test the background estimation method in the SM QCD-enhanced region:
  - Data: 279
  - Predicted 317±35

Results and Limits

- Expected and observed event yields for each selection set. The benchmark column shows the expected event yield of the most sensitive signal model for each set

Cut No.	Benchmark	Expected	Observed
1	317 ± 41	168 ± 15 ± 5	131
2	2730 ± 380	31.8 ± 5.0 ± 1.4	47
3	4.34 ± 0.44	19.4 ± 7.0 ± 5.5	20
4	0.24 ± 0.04	22.5 ± 2.5 ± 1.5	16
5	336 ± 34	13.9 ± 1.9 ± 0.6	14
6	69.1 ± 7.1	9.4 ± 2.0 ± 0.3	11
7	0.37 ± 0.05	4.40 ± 0.84 ± 0.28	2



- Upper limits at 95% C.L. on the signal cross section with $m_{\pi_{DK}}$ of 5 GeV

Conclusion

- A search for events consistent with pair production of a heavy mediator particle that decays to a light quark and a new fermion called a Q_{DK} is performed, with 16.1 fb^{-1} 13 TeV data
- A new data-driven background prediction method is presented, which helps us handle the correlation and predict the background well
- The data are consistent with the expected contributions from SM background process. Limits are set on the dark pion proper decay length for mediator masses between 400 and 1250 GeV
- These are the first results from a dedicated search for new particles predicted by such a model

References: CMS-EXO-18-001 [arXiv:1810.10069]