



LUNDS  
UNIVERSITET

# The DARKJETs project

HOW THE ERC CONTRIBUTES TO MAKING THE MOST OF LHC DATA

CATERINA DOGLIONI - LUND UNIVERSITY

@CatDogLund, she/her  
<http://www.hep.lu.se/staff/doglioni/>

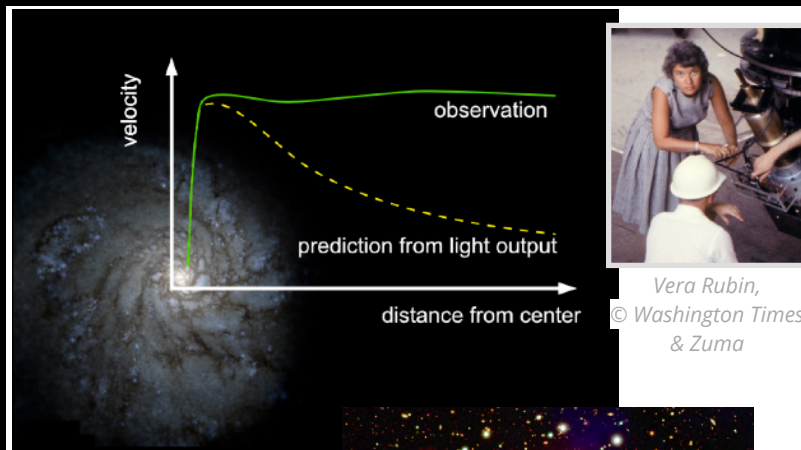






# DARK JETS

# DARK

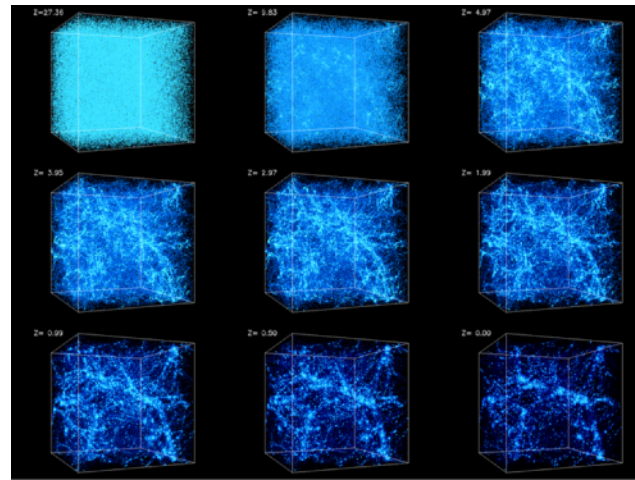


Vera Rubin,  
© Washington Times  
& Zuma

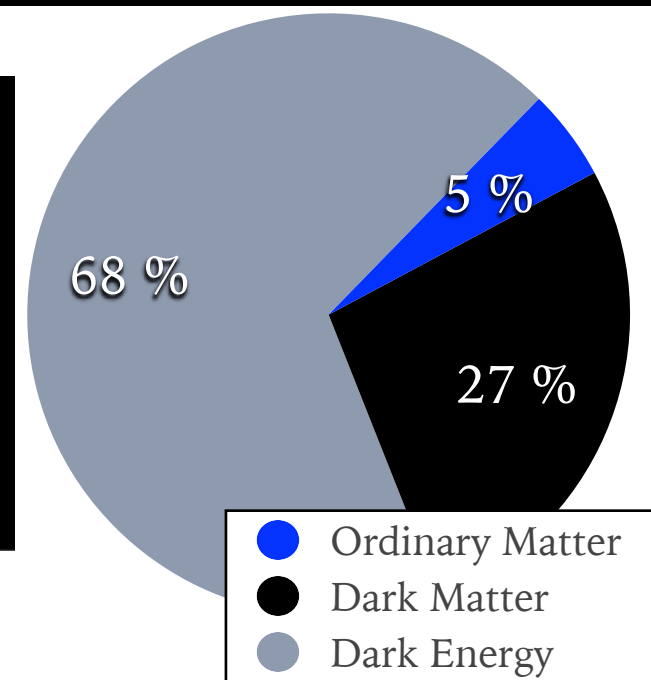
Wikipedia &  
Hopkins Research Group/Caltech



NASA/CXC/M. Weiss

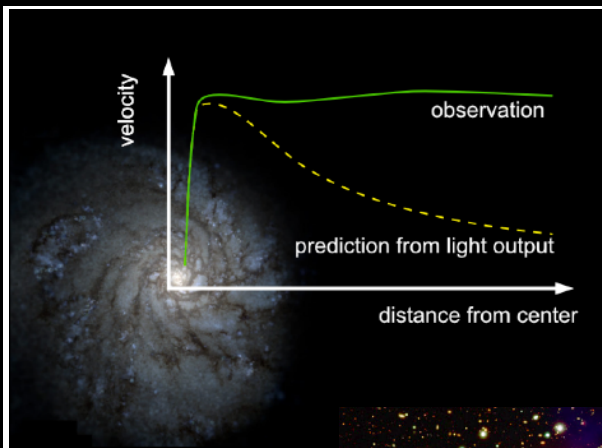


Simulations were performed at the National Center for  
Supercomputer Applications by A. Kravtsov and A. Klypin.

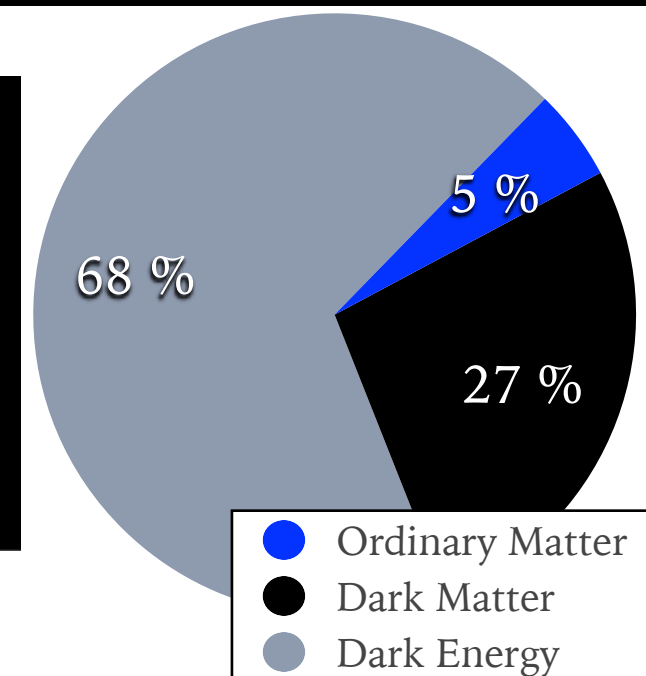
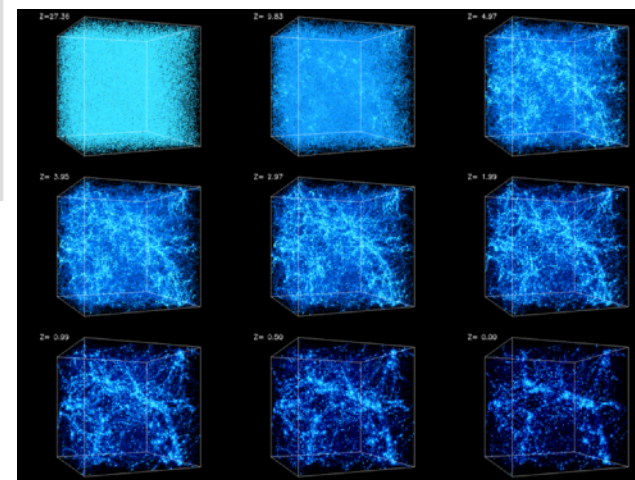


Empirical **problem** in the Standard Model (SM) of Particle Physics:  
no explanation for **Dark Matter**

# DARK



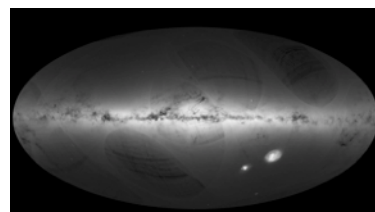
Wikipedia &  
Hopkins Research Group/Caltech



Empirical **problem** in the Standard Model (SM) of Particle Physics:  
[arXiv:0704.2276v1](https://arxiv.org/abs/0704.2276v1)  
no explanation for **Dark Matter**

A possible **solution**:  
invisible **Dark Matter particles** with weak SM interactions  
(Weakly Interacting Massive Particles)  
that can be detected by **a variety of experiments**

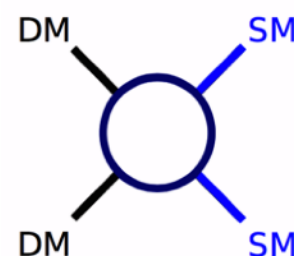
Credit: ESA/Gaia/DPAC.



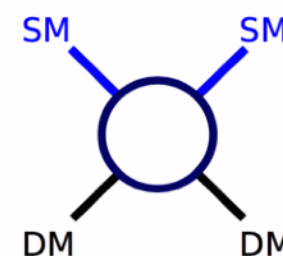
Astrophysics

Dark  
Matter

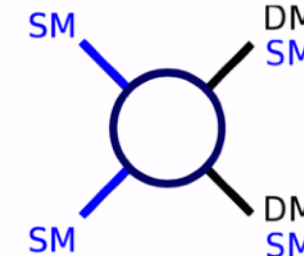
Standard  
Matter



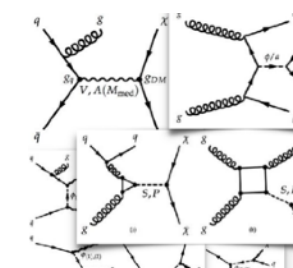
Indirect Detection



Direct Detection



Colliders & accelerators



Theory



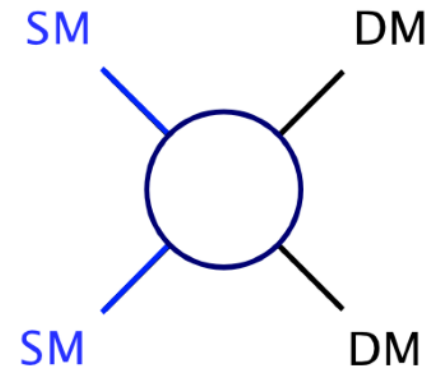
(If there are SM-DM interactions, like in WIMP models)

The LHC can **produce invisible particles**, in controlled conditions



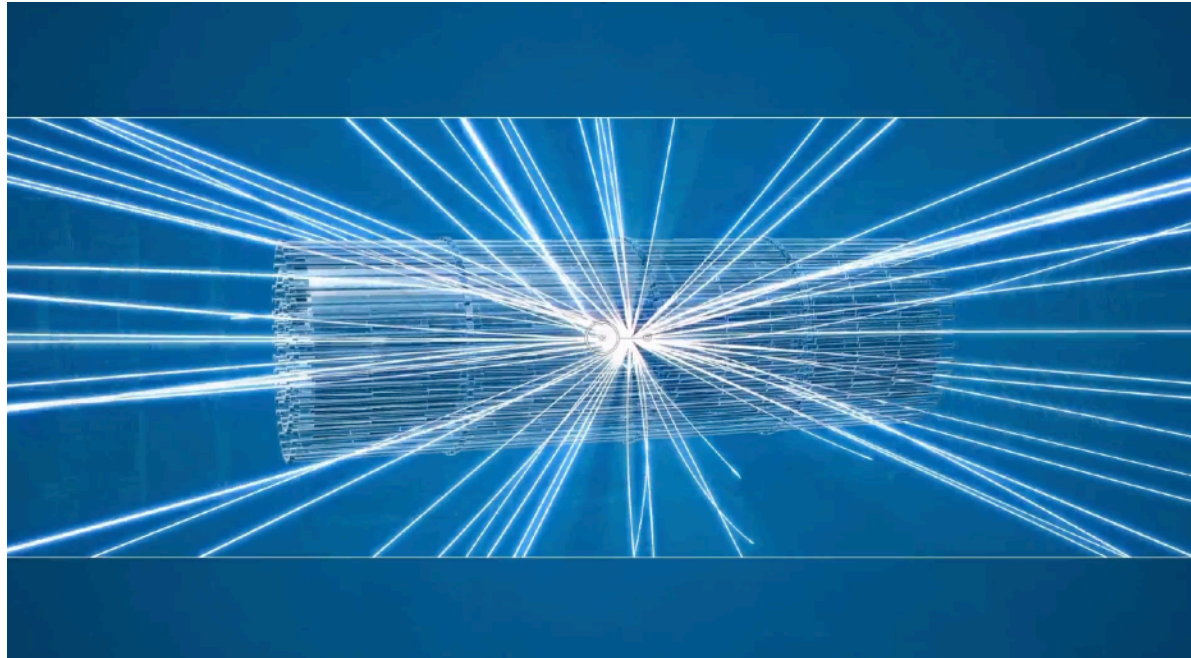
Standard  
Matter

Dark  
Matter



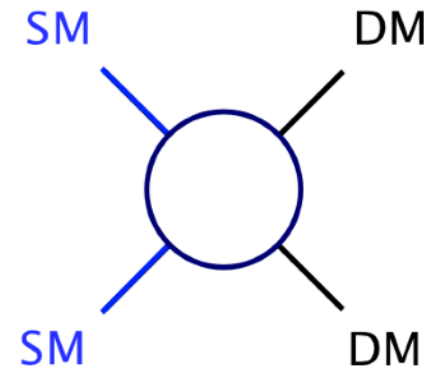
(If there are SM-DM interactions, like in WIMP models)

The LHC can **produce invisible particles**, in controlled conditions

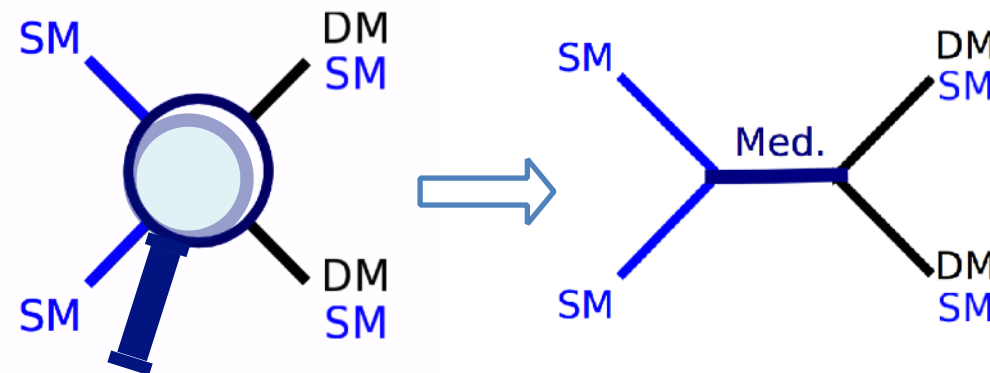


Standard  
Matter

Dark  
Matter



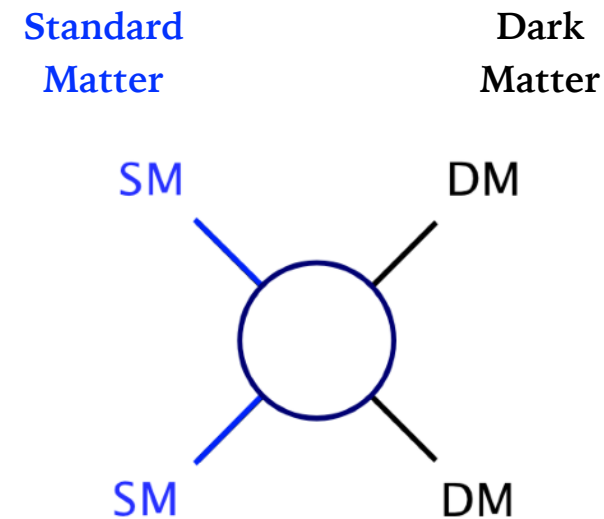
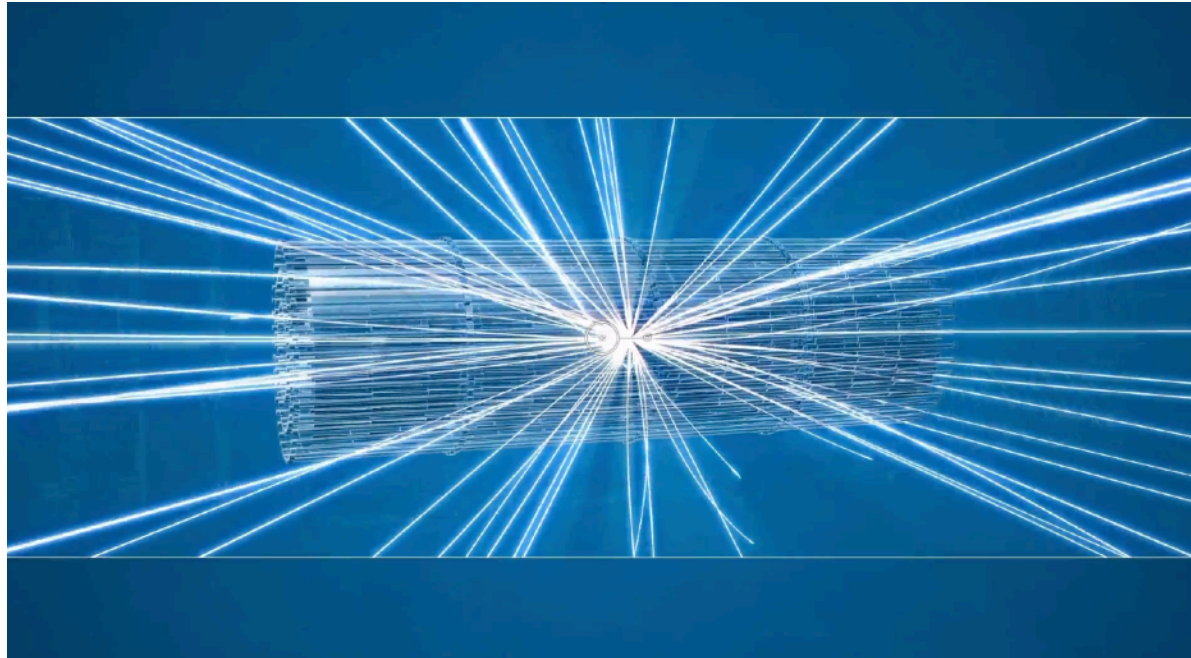
Experiments like ATLAS (and CMS) can also **probe the dark interaction** using the decays of a **mediator particle**



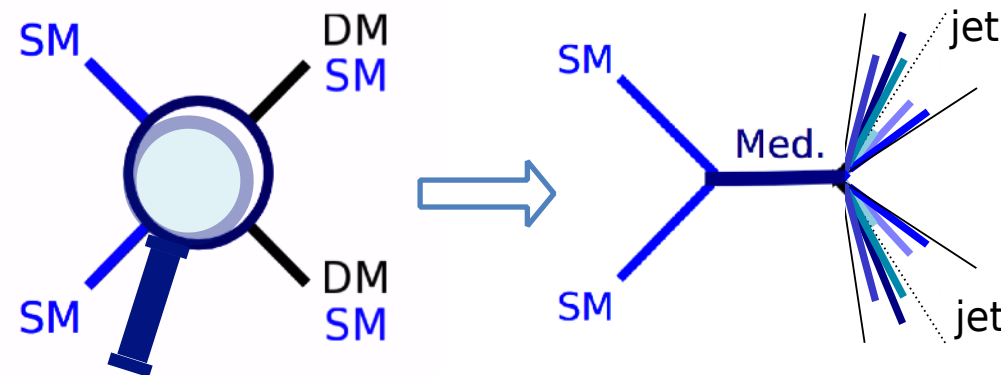


(If there are SM-DM interactions, like in WIMP models)

The LHC can **produce invisible particles**, in controlled conditions



Experiments like ATLAS (and CMS) can also **probe the dark interaction** using the decays of a **mediator particle**



# JETS

a **mediator particle** produced through interaction of quarks/gluons  
→ will also decay back into quarks/gluons

**DARKJETS**: searches for signs of dark matter  
in final states containing **jets of particles**

(originating from quarks and gluons)

# The Lund University DARKJETS team



Photo: Lena Björk Blixt.

[LU Press Release](#)

In the picture from the left:

- *Postdocs:*

**William Kalderon** (now at Brookhaven National Lab), **Jannik Geisen** (superimposed)

- *PhD students:*

**Eric Corrigan, Eva Hansen** (remote on photo-day)

- *Master's student, now PhD:*

**Alexander Ekman**

+ Lund Master's students:

**Herjuno Nindhito**

now PhD student at the University of Geneva

**Prim Pasuwan, Isabelle John**

now PhD students at Stockholm University

**Zhiying Li**

now PhD student at the University of Oxford

**Sebastian Murk**

now PhD student at the University of Macquarie

**Erik Wulff**

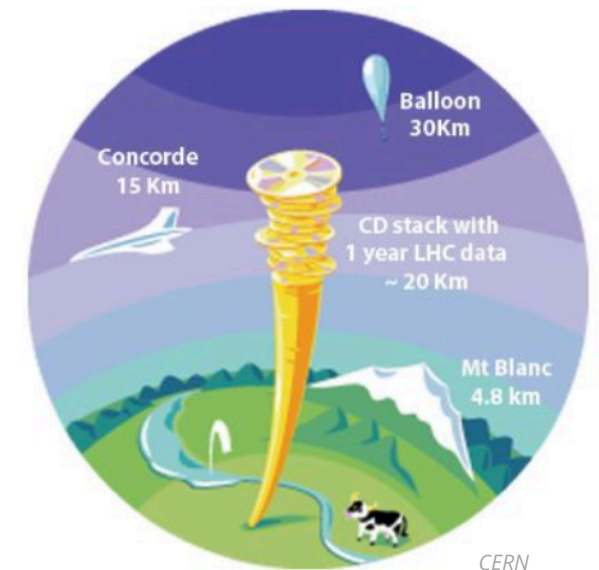
now Deep Learning engineer at Axis communication

[\[+ 12 Bachelor's students, see website\]](#)



# Our “Big Science” problem to solve: *too much data*

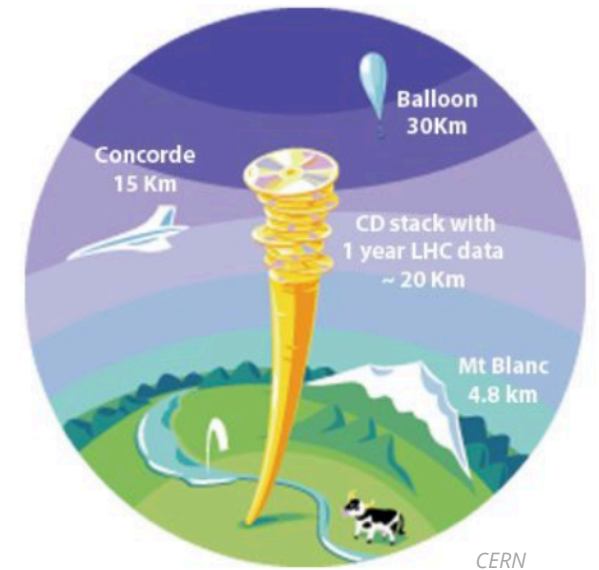
- The **dark matter signals** we are looking for are **rare**  
→ need enormous amount of collisions to produce them
- Their **backgrounds** look the same and are **much larger**
- **Problem:** recording all LHC data takes 400000 PB/year [\[Ref\]](#)
  - up to 30 million proton-proton collisions/second (MHz)
  - ~ 1-1.5 MB/data per collision event, including raw data



after selection of “interesting” data

# Our “Big Science” problem to solve: *too much data*

- The **dark matter signals** we are looking for are **rare**  
→ need enormous amount of collisions to produce them
- Their **backgrounds** look the same and are **much larger**
- **Problem:** recording all LHC data takes 400000 PB/year [\[Ref\]](#)
  - up to 30 million proton-proton collisions/second (MHz)
  - ~ 1-1.5 MB/data per collision event, including raw data



LHC experiments need to select “interesting” events (=trigger)  
in real-time (milli/microseconds)

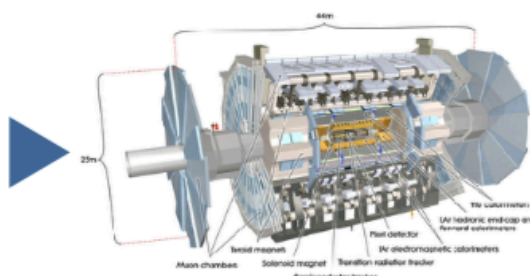
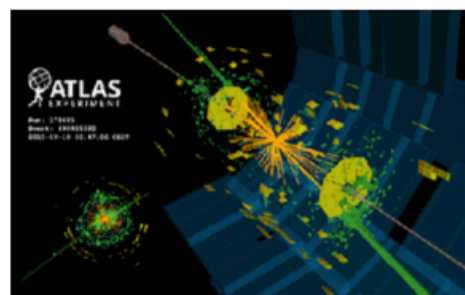
after selection of “interesting” data

Collisions at ~30 MHz  
(~1 MB of info each)

Hardware trigger  
outputs ~100 kHz

Software trigger  
outputs ~1 kHz

Online    Offline



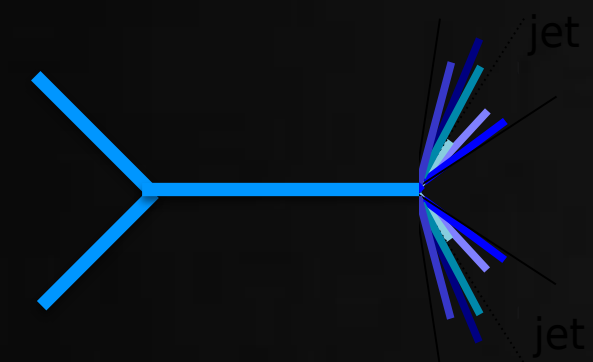
Event selection  
(trigger)

Object  
reconstruction  
and calibration

Data analysis



# Are we missing rare dark matter processes?



DM mediators

Events selected by the trigger

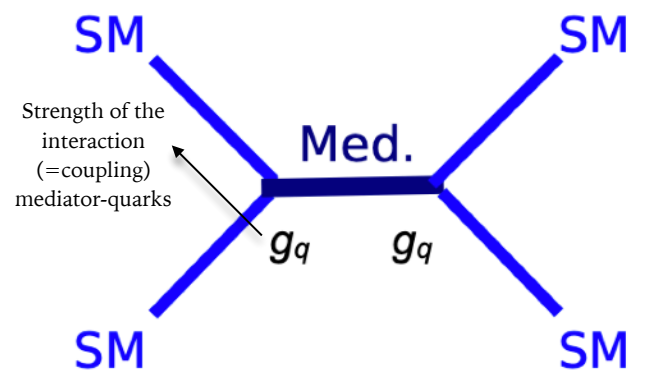
# Challenges for DM mediator searches (and more)

Selecting interesting events works for most of the LHC physics program...

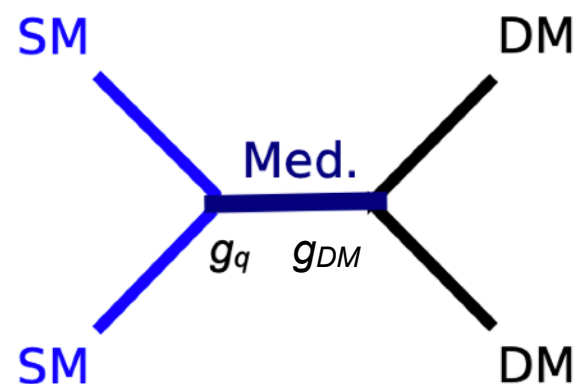
...but it is **not optimal** for rare processes with high-rate backgrounds:

we cannot record and store all data, and trigger **discards both background and signal**

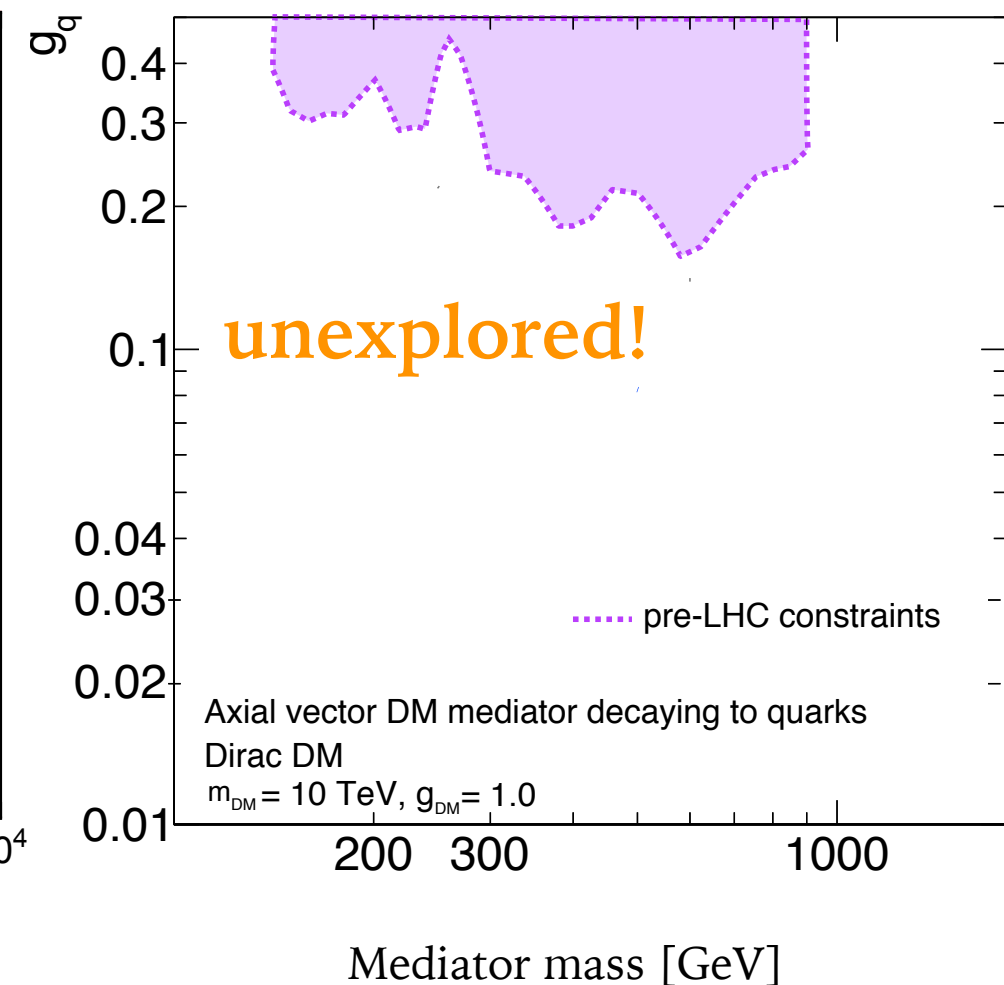
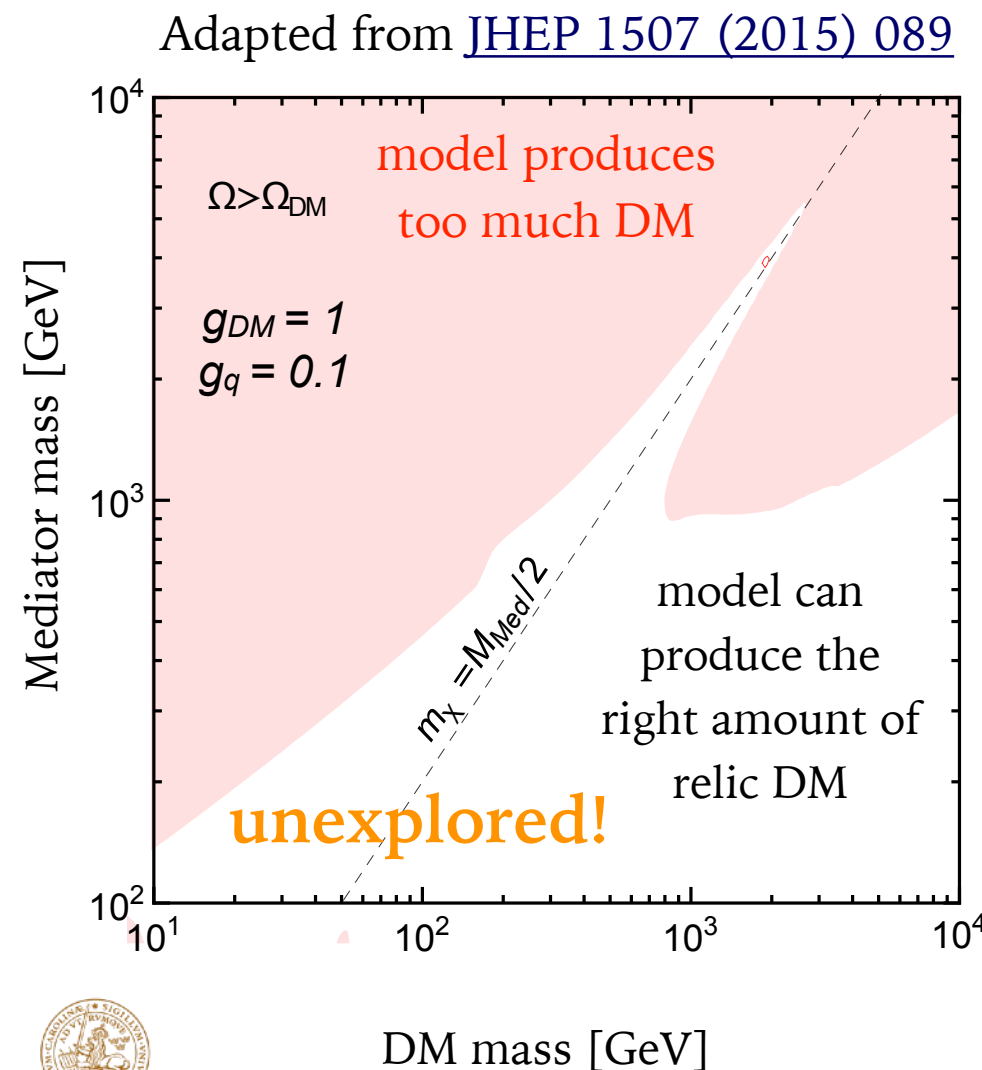
**This prevented us from being sensitive to low-mass DM mediators decaying into jets**



Visible mediator decays



Invisible mediator decays

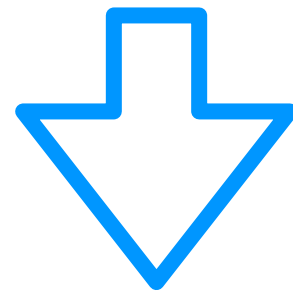




# The main technical idea behind DARKJETS

## Asynchronous data analysis

First record and store data, then reconstruct/analyze it



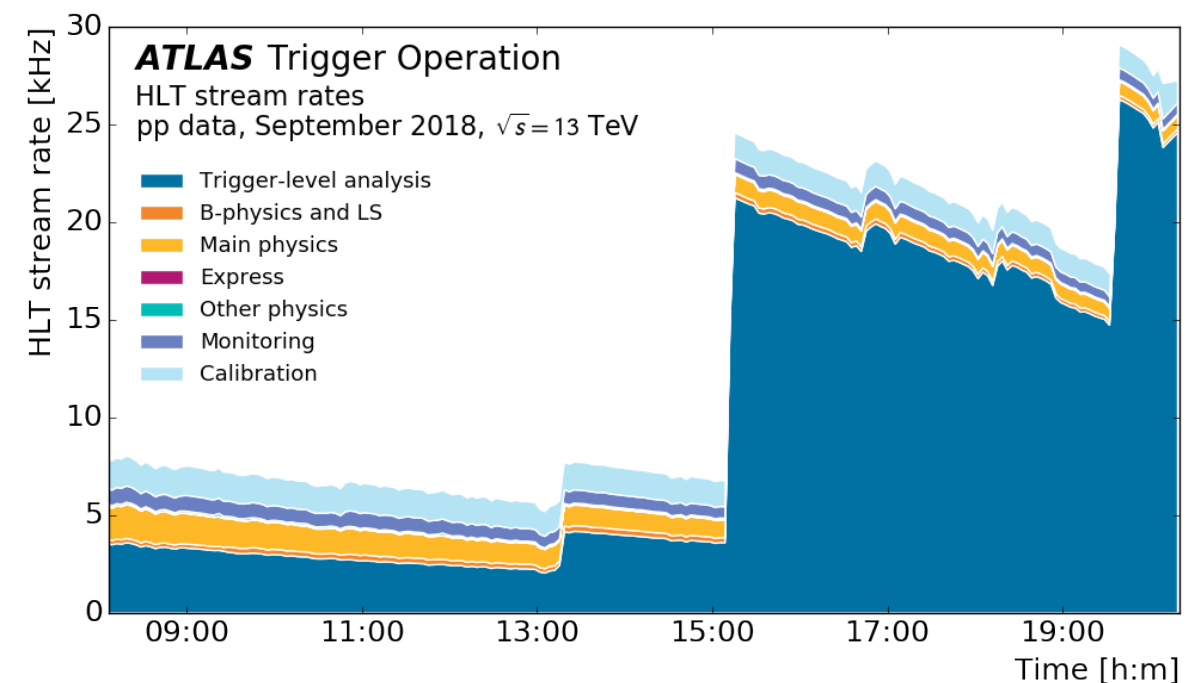
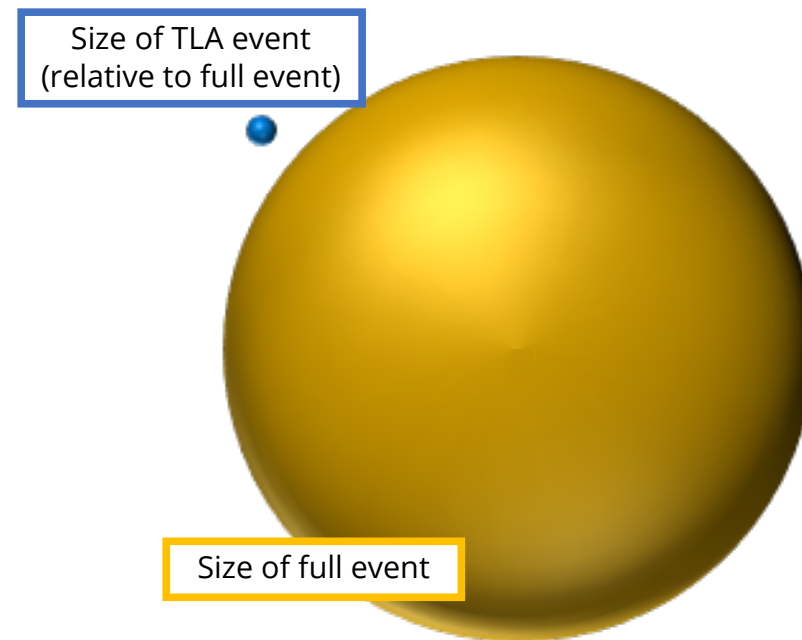
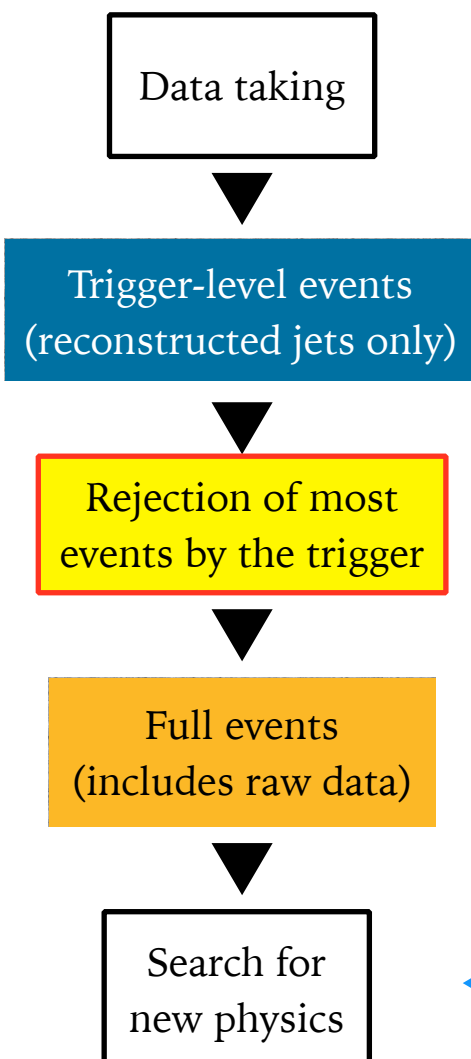
## Real-time data analysis

Reconstruct/analyse data as soon as it is read out  
so that only (**smaller**) final-state information  
needs to be stored

**CMS:** [Data Scouting](#), **LHCb:** [Turbo stream](#)

# ATLAS implementation: Trigger Level Analysis (TLA)\*

Much smaller event size  $\longrightarrow$  orders of magnitude more data can be recorded



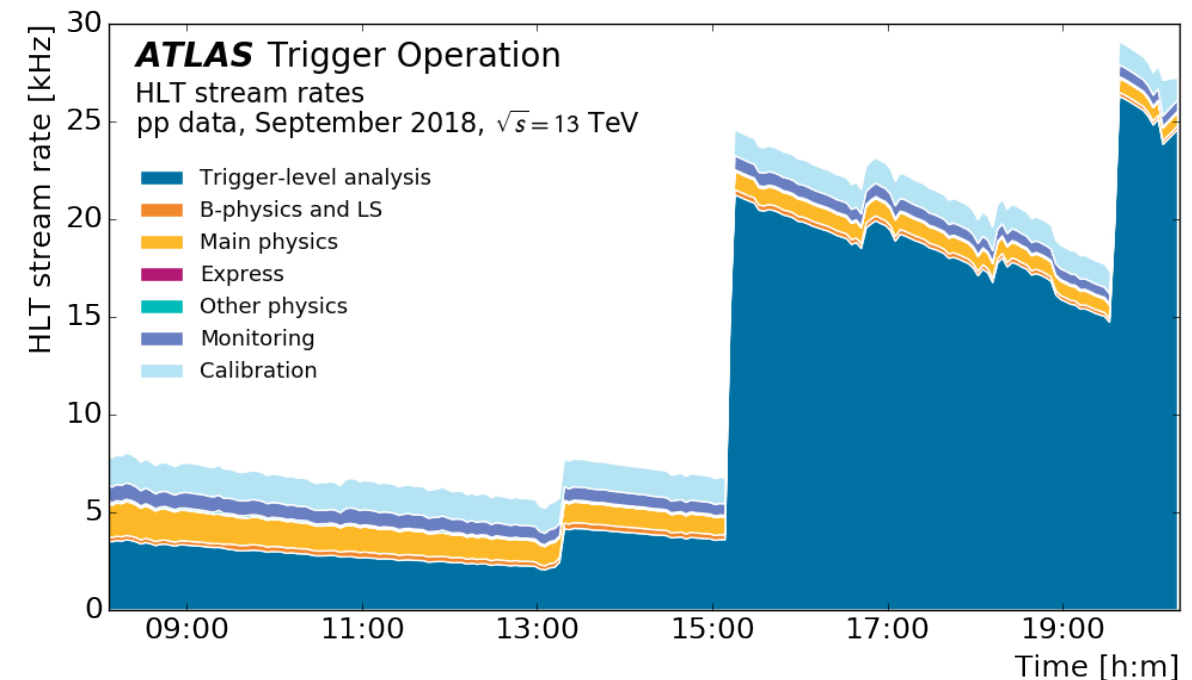
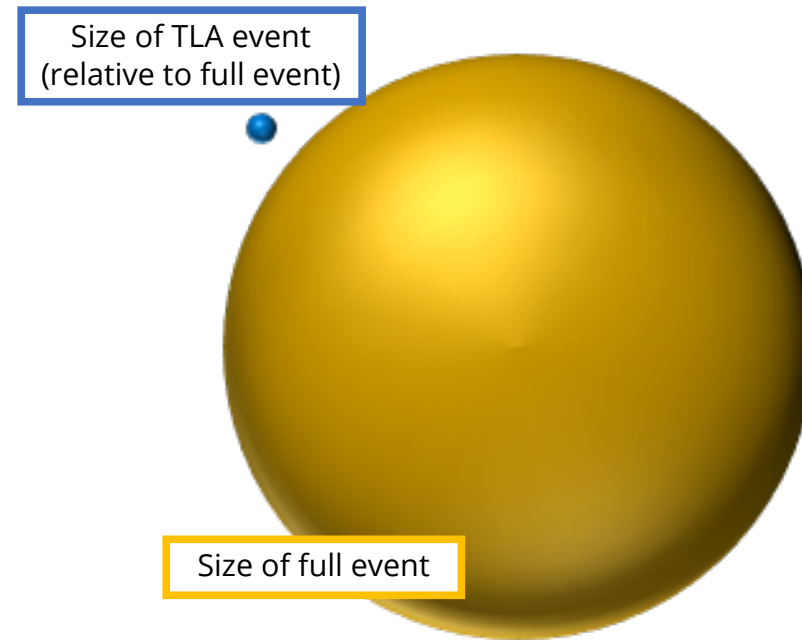
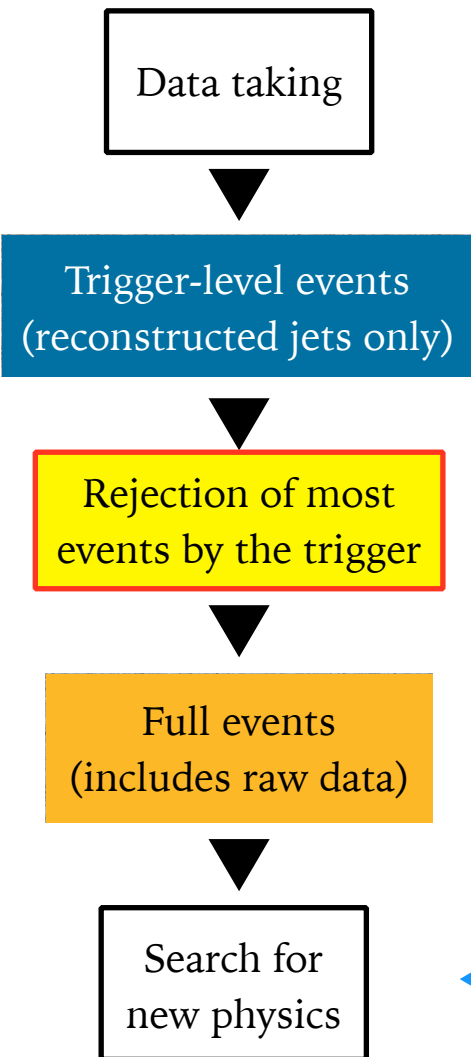
<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/TriggerOperationPublicResults>

- Challenges solved by **DARKJETS** team & collaborators:
  - software**: technical implementation and large-scale deployment
  - performance**: is a reduced data format “good enough” for a discovery”?
  - statistical analysis**: how to deal with unprecedented amounts of data?



# ATLAS implementation: Trigger Level Analysis (TLA)\*

Much smaller event size  $\longrightarrow$  orders of magnitude more data can be recorded



<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/TriggerOperationPublicResults>

## Challenges solved by DARKJETS team & collaborators:

- **software:** technical implementation
- **performance:** is a reduced data rate
- **statistical analysis:** how to combine data

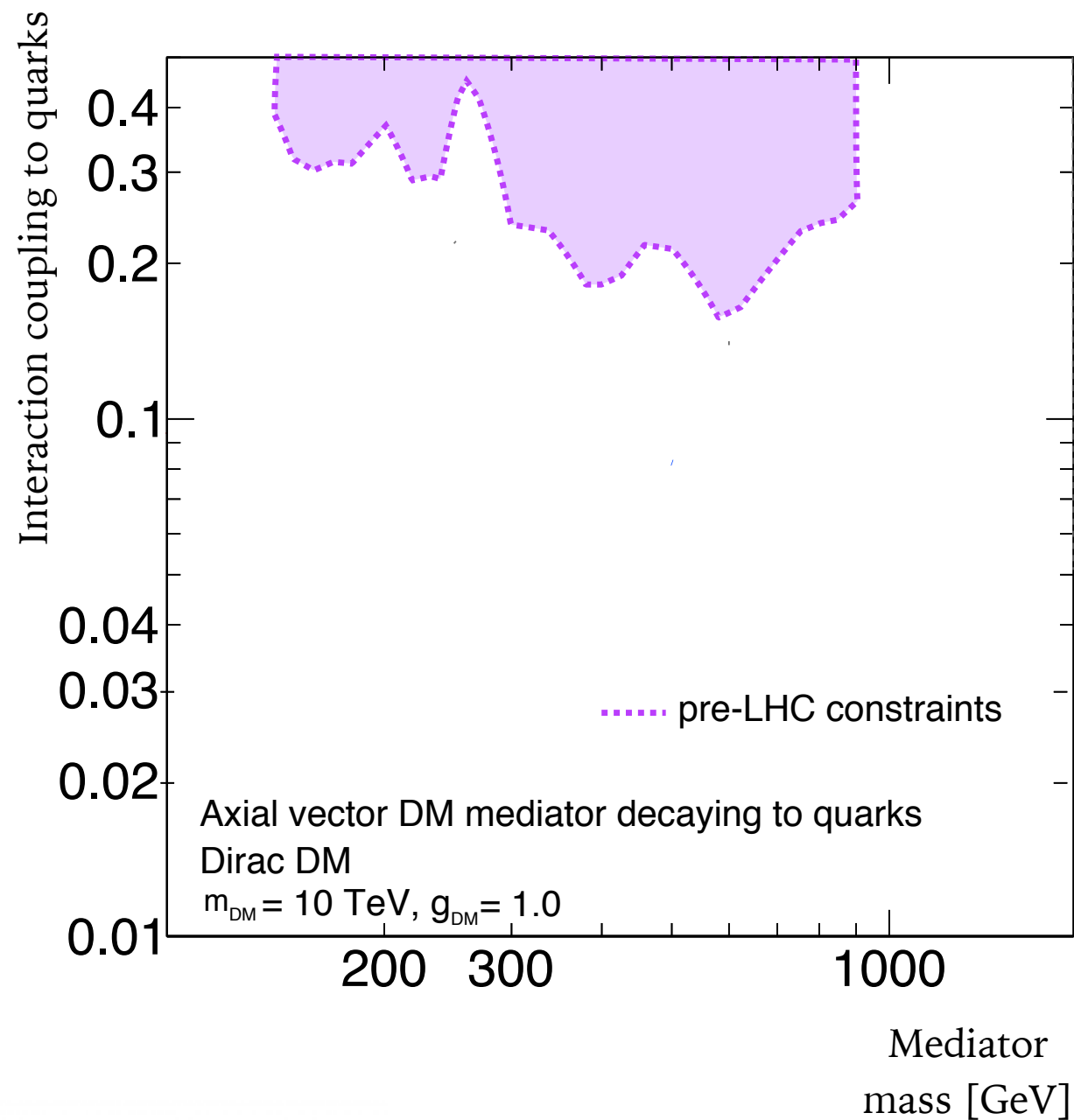
Heidelberg    Ohio State  
Geneva    Oregon  
Buenos Aires    CERN



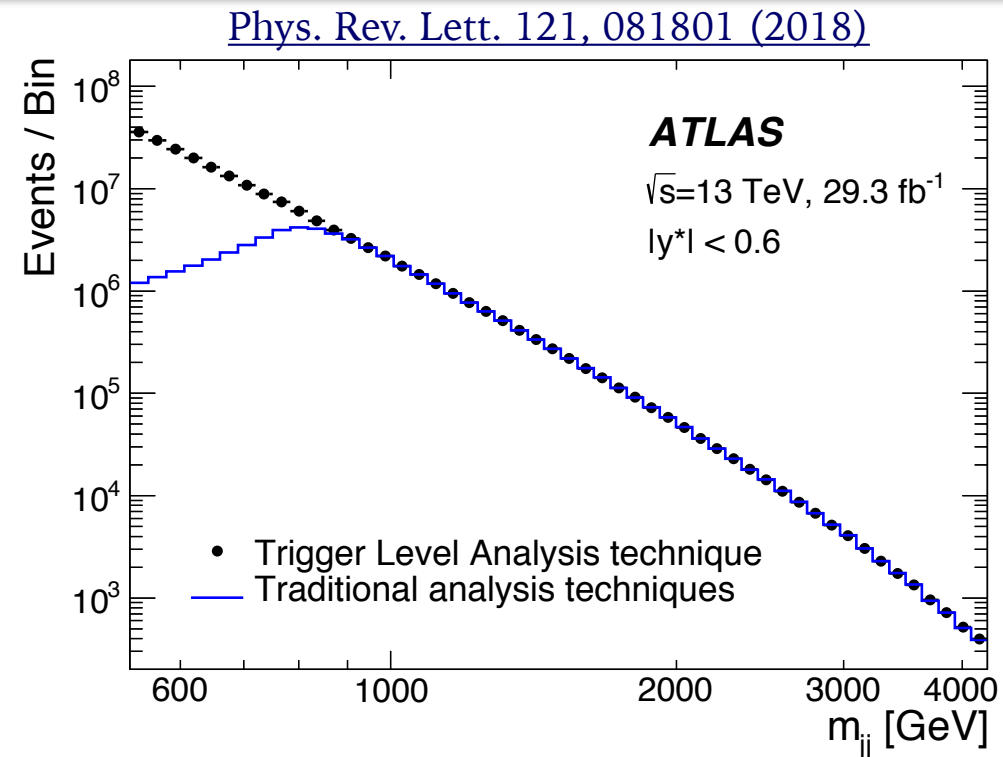
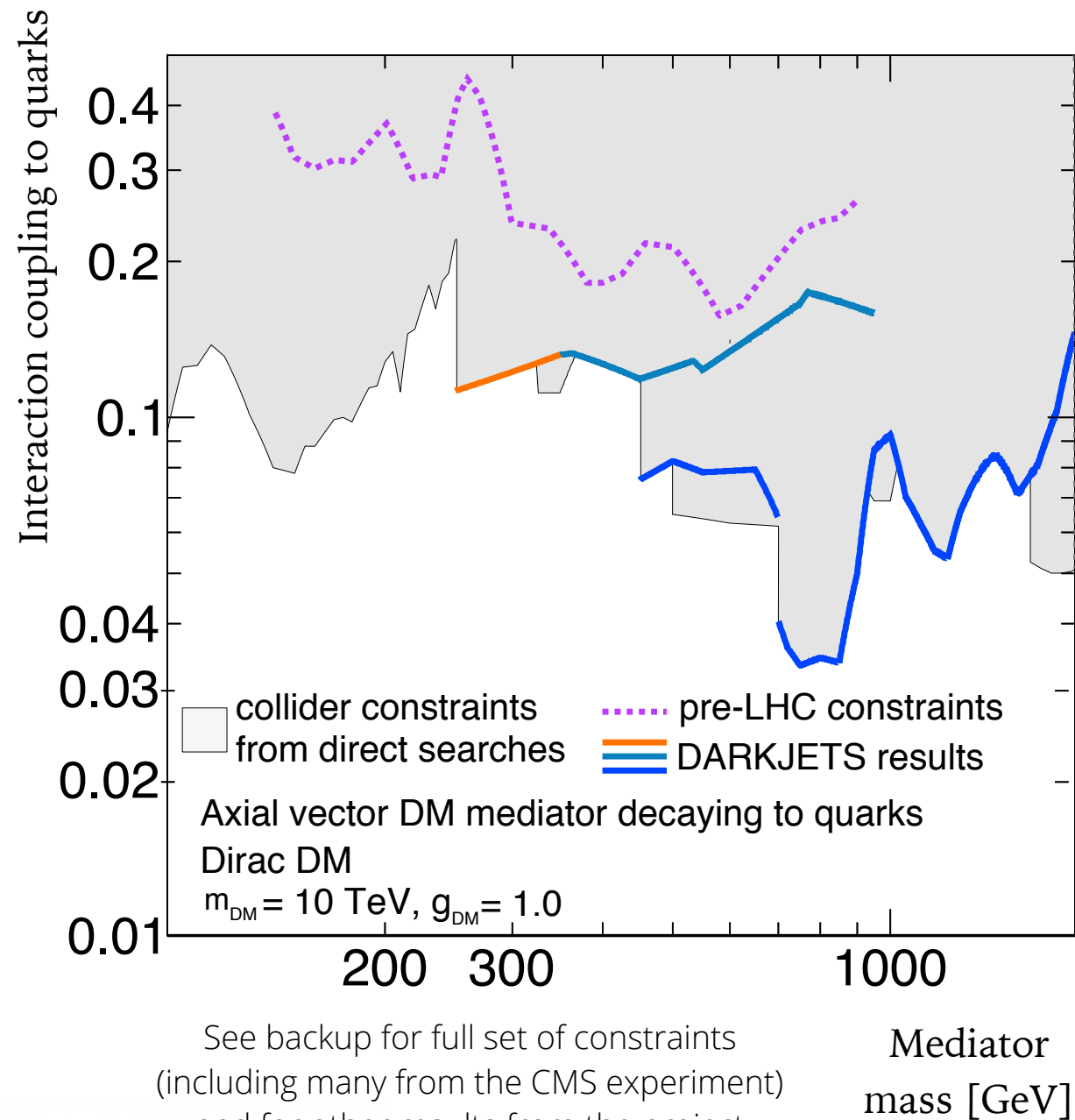
Caterina Doglio

\* Trigger Level Analysis is a Three Letter Algorithm (TLA)

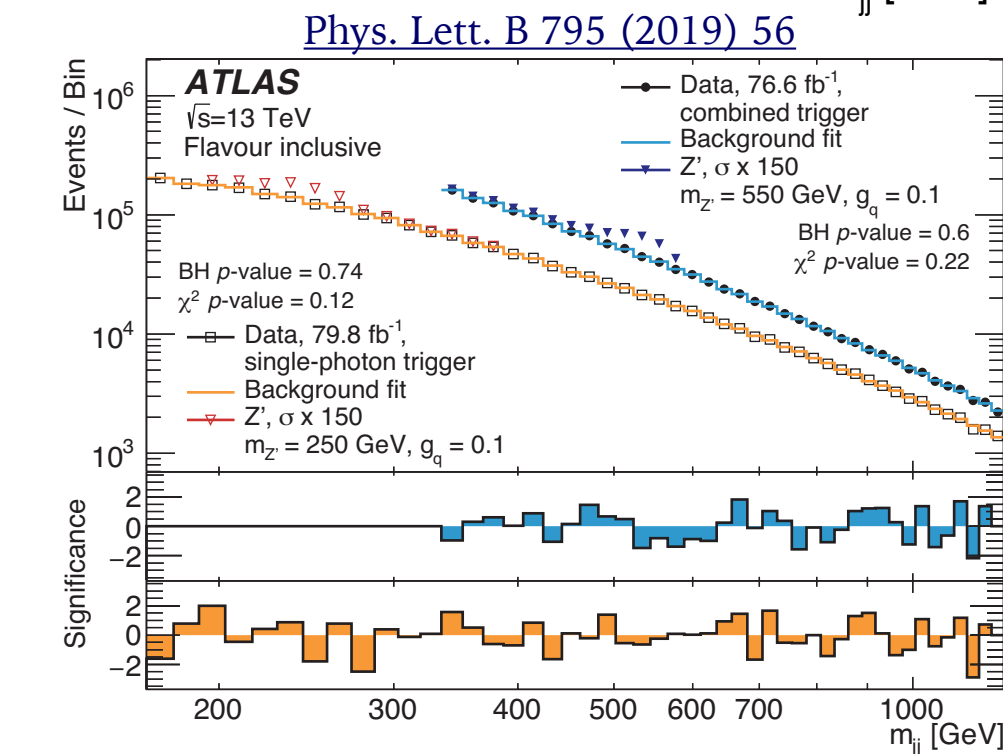
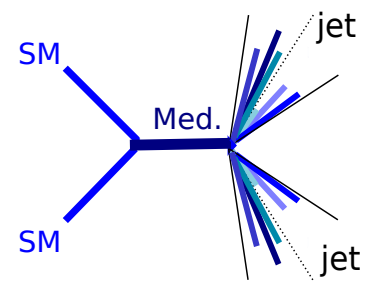
# (a subset of) Physics results from DARKJETS



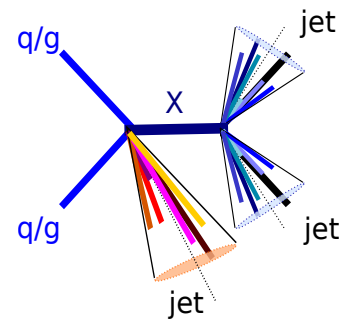
# (a subset of) Physics results from DARKJETS



**TLA technique:**  
Make the event size smaller



**Dijet+ISR signature:**  
Reduce the background





# A global view of WIMP dark matter

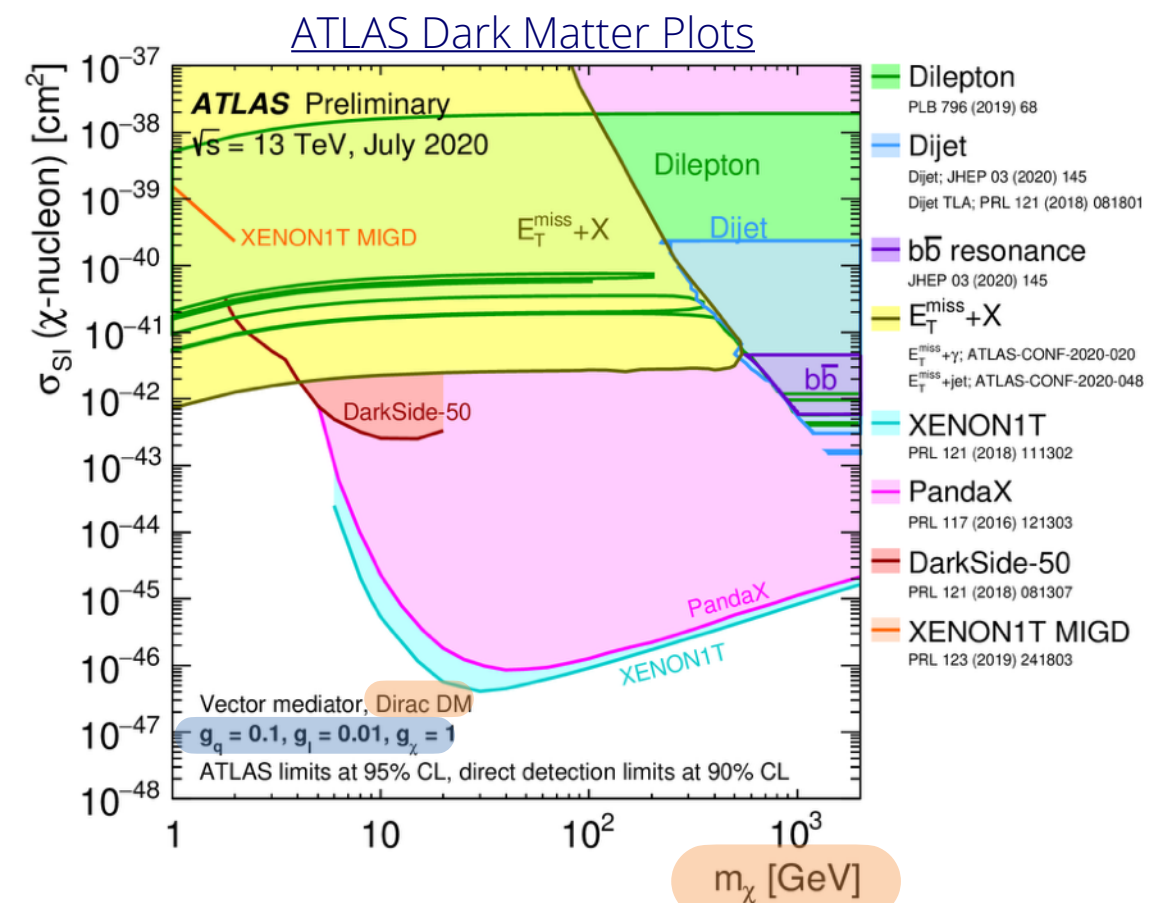
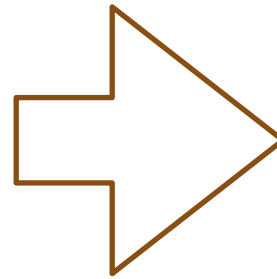
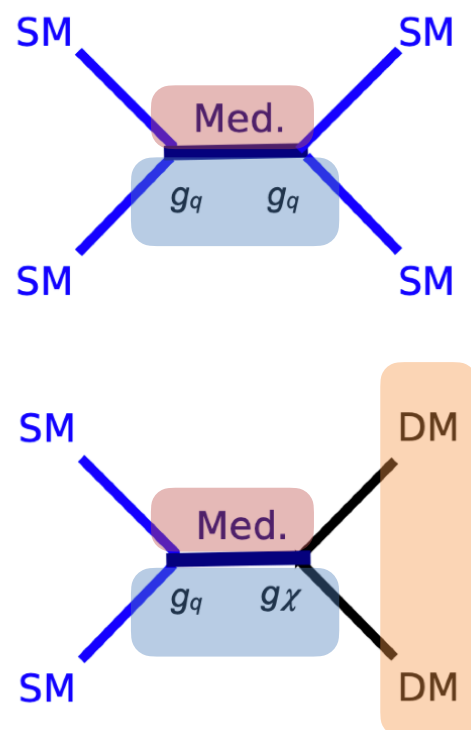
How do we compare results of different experiments  
~~in the most model independent way possible?~~

European Strategy Update  
“Big Question”

Comparisons are possible only in the context of a model  
Essential to **fully specify model/parameters** and **be aware of limitations**

LHC Dark Matter Working Group

<https://arxiv.org/abs/1603.04156>

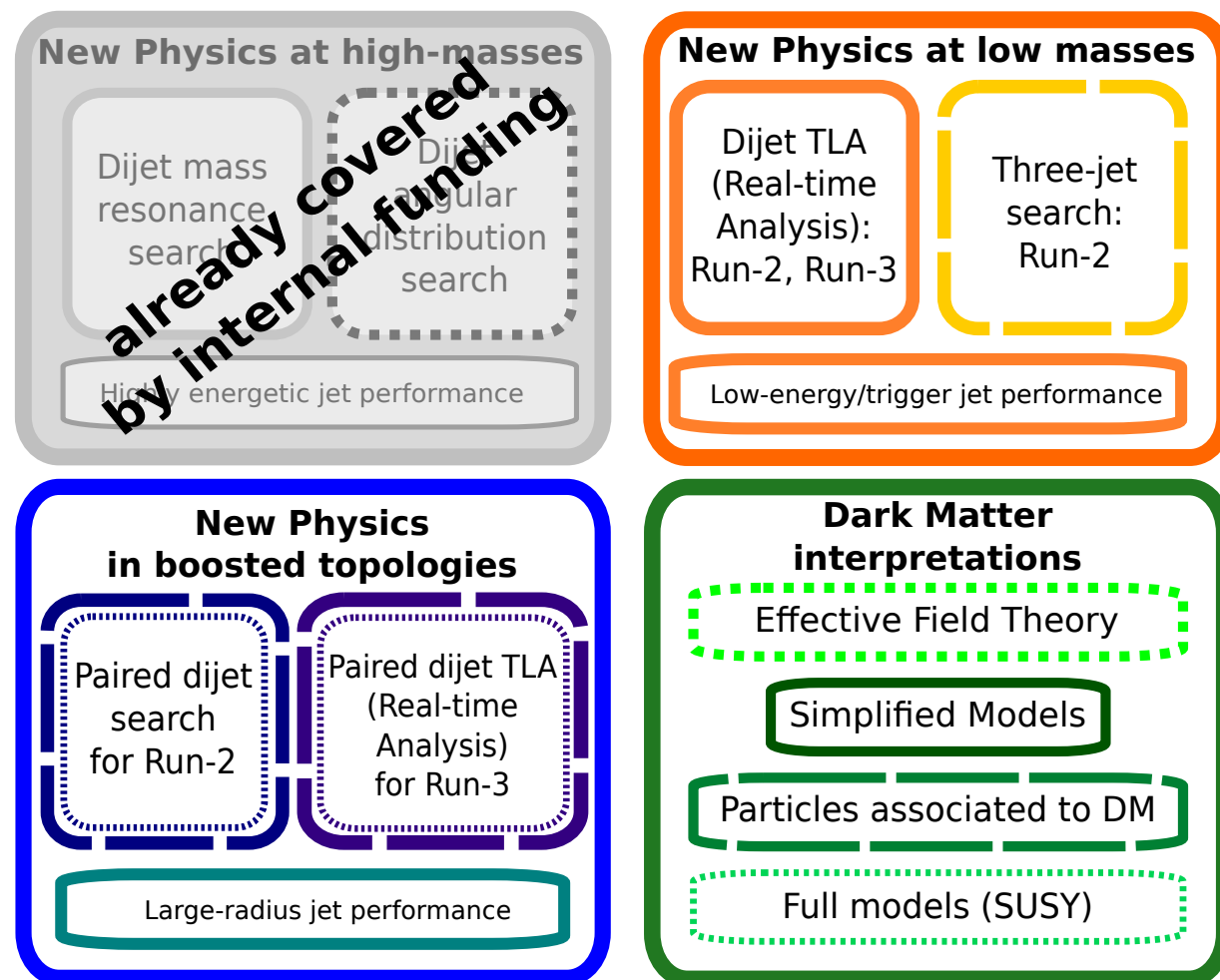


Complementarity of colliders with direct (indirect) detection  
performed **within the chosen benchmark models & parameters**

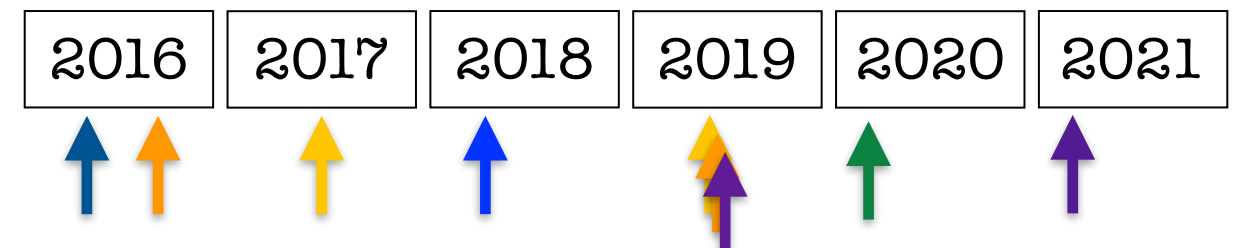
...more discussions with direct/indirect detection ongoing/to come...

# DARKJETS: the plan

Plan in grant proposal:

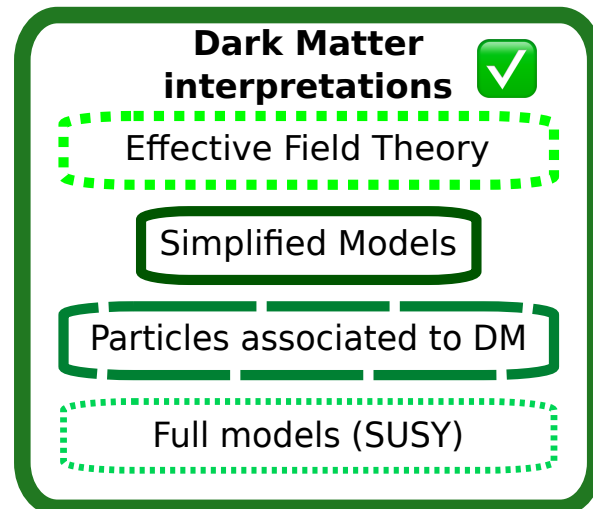
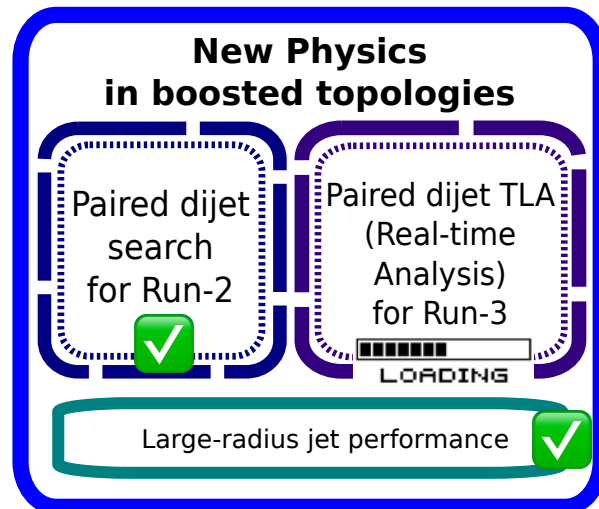
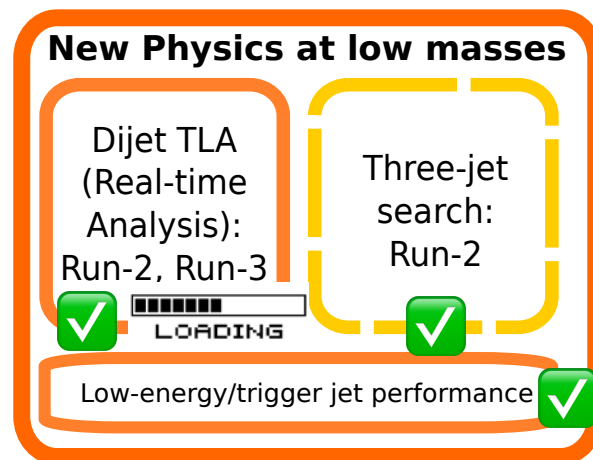
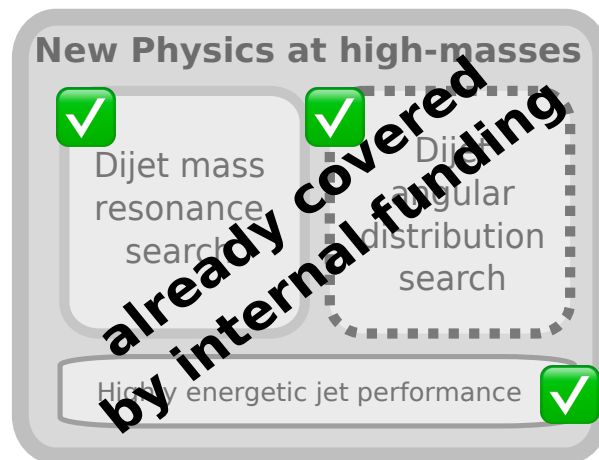


arrow = publication in the original plan



# DARKJETS: the outcomes

[Covered in this talk]  
Eric Corrigan's thesis + Eva Hansen & Alexander Ekman's contributions + Will Kalderon's focus



**Extra:** Paired-fat-jet analysis for non-WIMP DM models

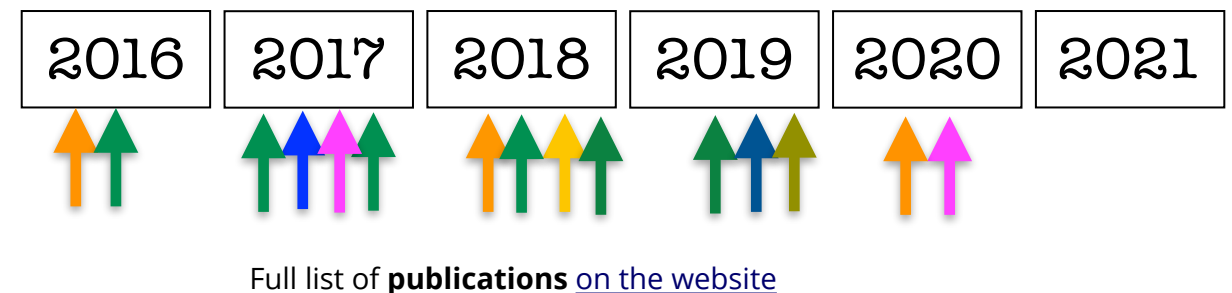
**Extra:** synergies with other DM searches  
[European Strategy of particle physics](#)  
[Snowmass process \(US\)](#)

Not covered in this talk (see backup)  
Eva Hansen's thesis + Jannik Geisen's focus

**Extra:** software and computing for DM and real-time analysis within the [HEP Software Foundation](#)



arrow = publication in a given year (not including conference papers)



## Dissemination:

- Not everything goes according to plan or is in your hands (especially in a large collaboration)
  - Still, **reached all project milestones**
- Organised, chaired, ~~cancelled~~ **conferences**
- A number of **conference talks** given by team members (see [Eva Hansen's ICHEP talk](#))

## Communication & outreach:

- We contributed to various newsletters related to DM/computing/particle physics, CERN Courier, ATLAS features [[see website](#)]
- We are responsible for the [Lund IPPOG Masterclasses](#) (including Day of Women and Girls in Science)



# What next for/after DARKJETS?

DARKJETS research questions related to  
*finding signs of **dark matter** in **jet** final states at the LHC*

after ~4 years: many **answers**,  
but also many **new directions** for future research

Further exploration of the  
electroweak scale ( $\sim 100$  GeV)

Non-WIMP searches  
with non-standard jet signatures

DM @ colliders complementarity with  
accelerator experiments & astrophysics

Upgrade ATLAS trigger for next LHC run

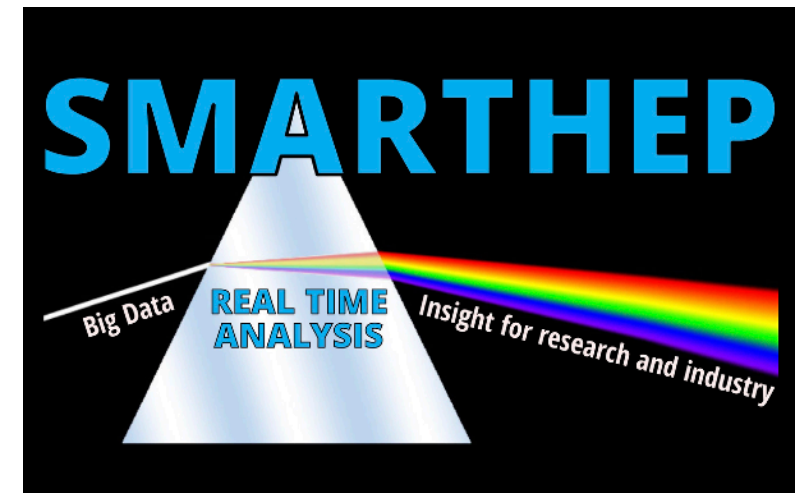
Start from jet TLA prototype  
→ make real-time analysis widely usable  
for searches (and measurements)

Sustainability and reusability of LHC/DM  
analyses, in terms of data and pipelines

# Collaborations and interdisciplinary projects

Ongoing **collaborations** related to DARKJETS:

- [SMARTHEP](#) network (submitted as ETN)
  - gathering real-time analysis experts from all main LHC experiments (6 ERC grantees)
- [REALTIME](#) interdisciplinary Advanced Study Group
  - members from engineering, IT, maths, astrophysics, law, social sciences
- [HELIOS Lund-Hamburg research school](#) on intelligent instrumentation in physics
  - similar “knowledge environments” of university + labs (ESS, European xFEL, DESY + CERN...)



Other **networking opportunities** within Horizon2020/Horizon Europe, some examples for prospective grantees [[link](#)]:

- ERC Visiting Fellow program [[Czech link](#)]
- Implementing arrangements with non-EU countries



# Conclusions and acknowledgements

## What we did in DARKJETS so far:

- Advanced real-time analysis methods at ATLAS
  - Helping pave the way for doing *more physics with fewer resources* at upgraded LHC experiments
- Helped define a framework for DM searches at the LHC
  - Many collaborative efforts started through the [LHC Dark Matter Working Group](#)
- Looked for DM mediators in a challenging but promising region
  - No signs yet, but we keep on looking, in this and different directions
- Trained postdocs and students on DM physics & big science tools
  - Career development includes links with other disciplines and industry

## An ERC grant is a **unique opportunity** to:

- Pursue the **research questions** you find most interesting (and fun)
- Train a group of **talented researchers** and work with them to answer those questions
- Build **networks** that will last throughout your career

You can (and should)  
**think big**, about  
science that wouldn't  
happen without you /  
without this grant

Disclaimer: I didn't go into more  
"unsolicited grant-advice" but feel  
free to ask in the Q&A!

## This grant / work would not have been possible without:

- Lund research services (especially Sophie Hyden Picasso and Karin Langborger)
- Theory & experimental collaborators / friends
  - and everyone who advised/mentored me along the way
- Lund colleagues (not only from HEP)
- Many others not mentioned here...

Thanks!





# Backup slides



**LUNDS**  
UNIVERSITET

For more information, see recent (virtual) seminars:  
[Seminar at UK STFC Particle Physics Experiment \(RAL\)](#)  
[Virtual MIT Colloquium](#)  
[plus more on the website]

# Video: LHC collisions and data selection

**CERN-VIDEO-2015-024-001**



# High energy physics collaborations (a fraction of)

Grant-related question:

**ATLAS** How to find CD (or other ERC grantees) in such large experiments?



**CMS**



**LHCb**



**ALICE**





# The Standard Model is not enough [arXiv:0704.2276v1](https://arxiv.org/abs/0704.2276v1)

Example of **aesthetical problem**:

- Measured Higgs boson mass is “fine-tuned”

Example of **empirical problem**:

- Dark Matter not described in the Standard Model

The **Standard Model** may be only a **low-energy approximation** of a more complete theory

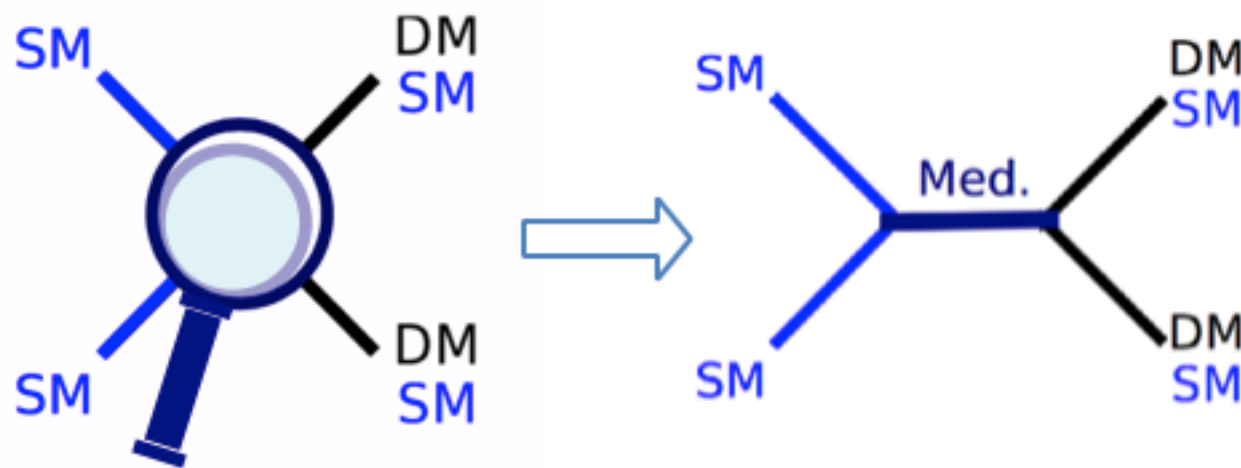
- **Many different theories** can explain those shortcomings
  - None of these theories is yet favored by data
  - Very different detector signatures
    - Some of them buried in **high-rate backgrounds**
    - Some of them **rare but very unusual**

Making the most of the data: **enabling discoveries**  
ensuring these events are recorded and analyzed

# Dark Matter mediators at the LHC

If there's a force other than gravity, there's a **mediator**,  
and the LHC could **detect** it:

*simplified models* are popular Run-2 LHC search benchmarks



Physics of the Dark  
Universe  
Volume 27, January 2020, 100371



Dark Matter benchmark models for  
early LHC Run-2 Searches: Report of the  
ATLAS/CMS Dark Matter Forum

## Dark Matter Forum & Working Group

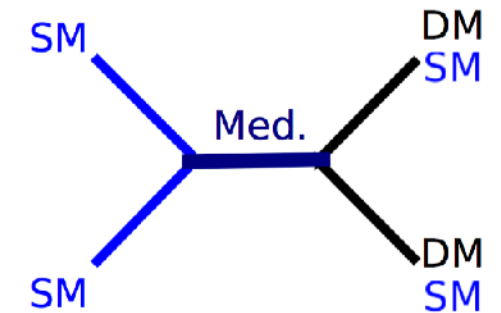
<https://lpsc.web.cern.ch/content/lhc-dm-wg-dark-matter-searches-lhc>  
[Phys. Dark Univ. 26 \(2019\) 100371](#) & references within



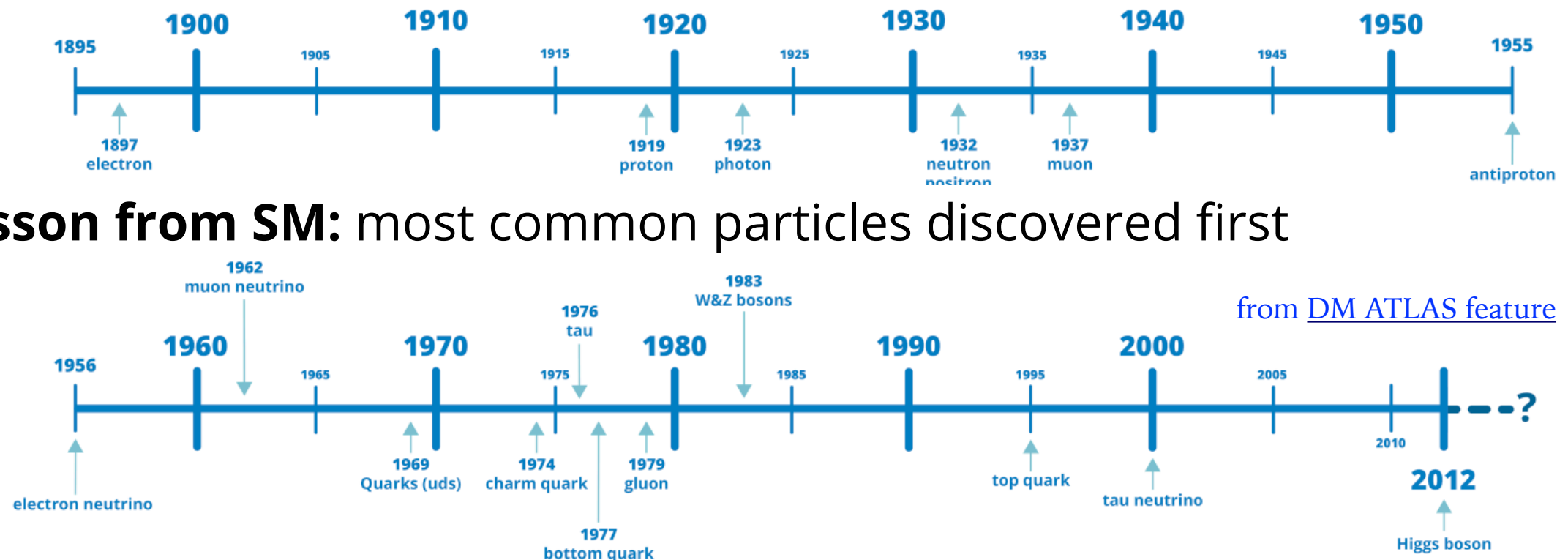
# Choice of benchmarks

<https://abstrusegoose.com/406>

*“Why should we choose/believe the simplest models?”*  
*“Do we think DM is all made of a single WIMP model?”*  
 (not really...see dark sectors!)



## Key particle discoveries



- **Lesson from SM:** most common particles discovered first
- Even simple models can encapsulate **relevant experimental characteristics** representing wider classes of theories

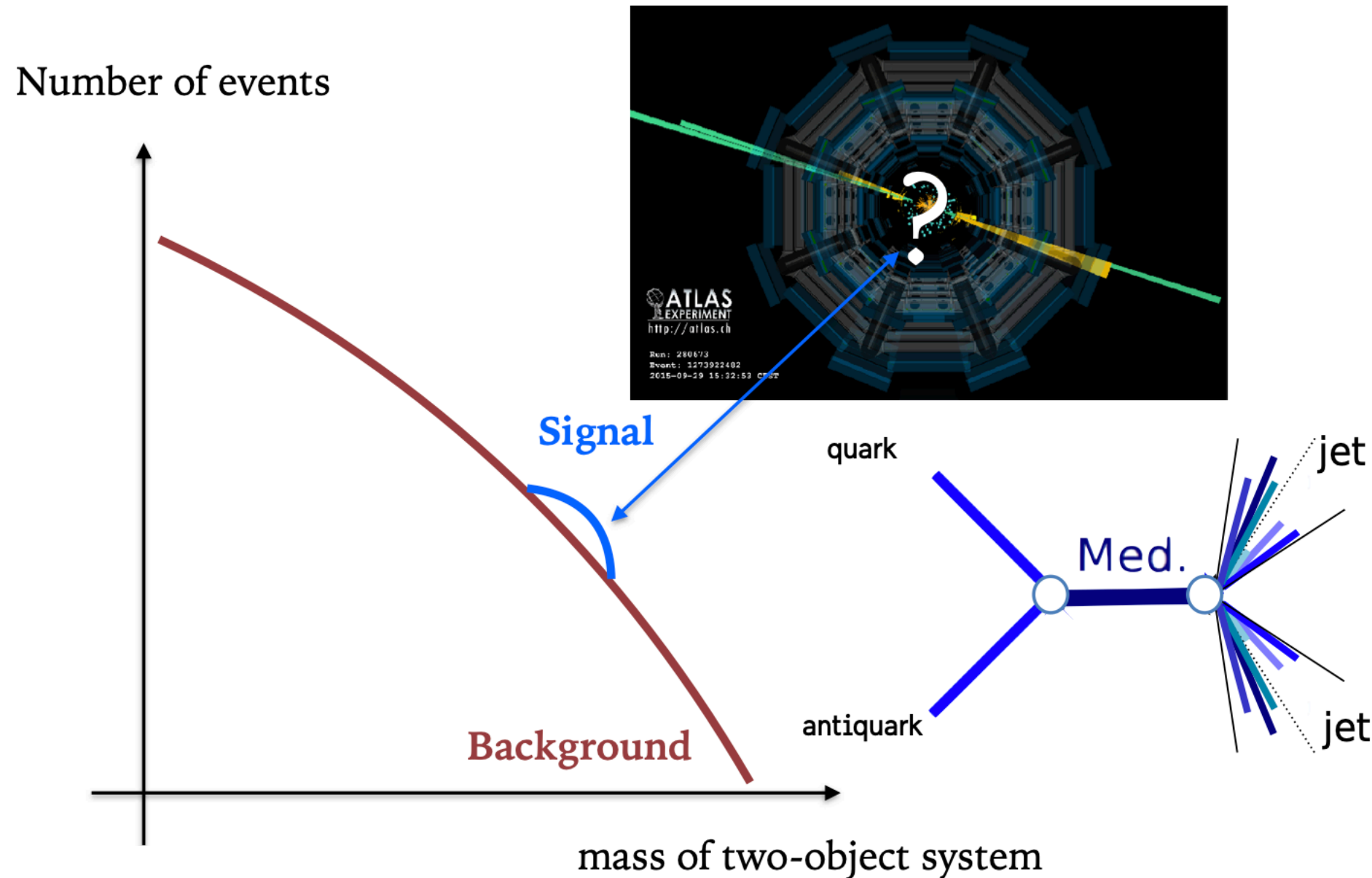
as long as we are aware that they can be more rare than what we choose as example





# How would new particles manifest?

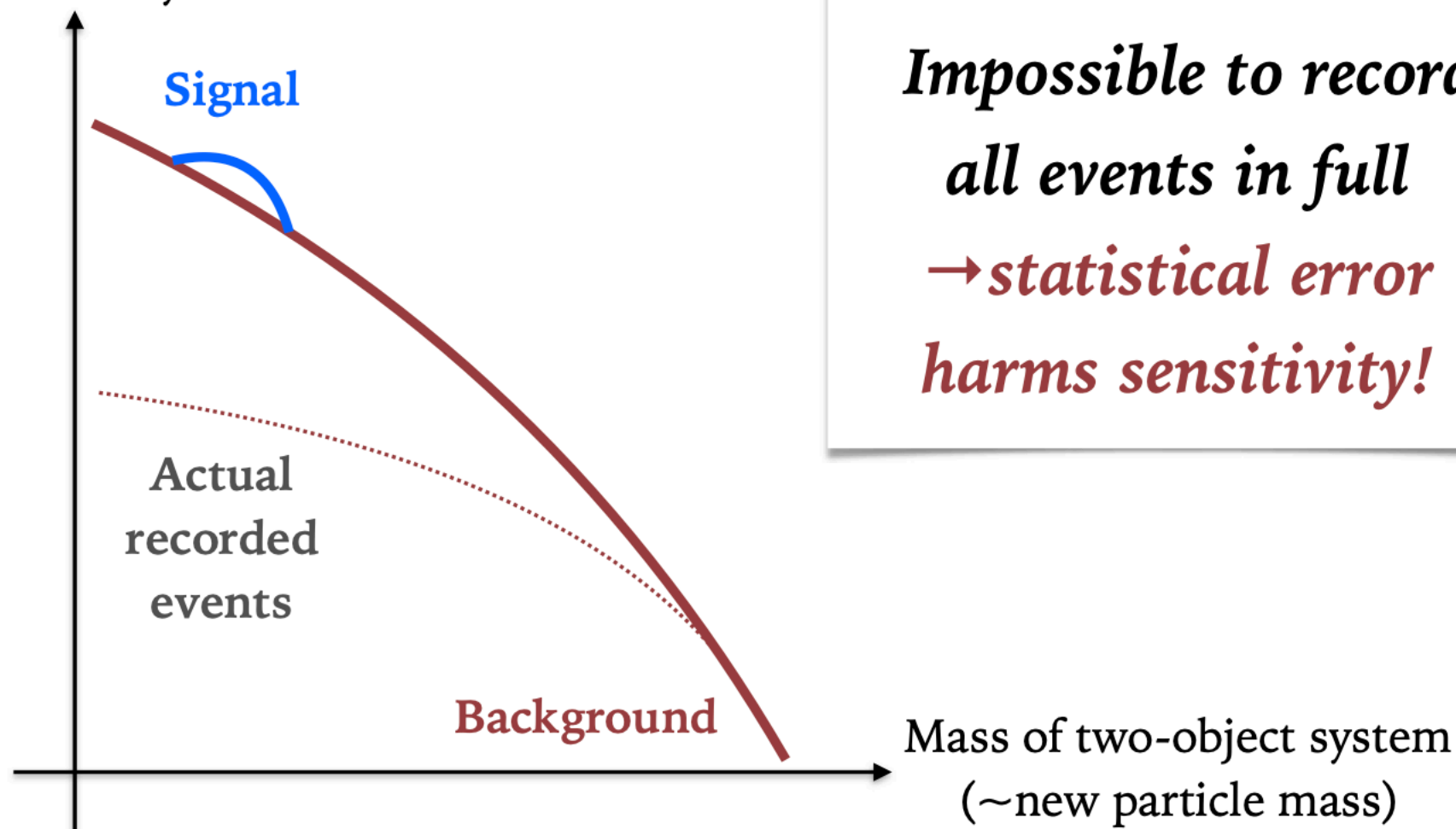
**New particles:** resonant excess (bump) over Standard Model background



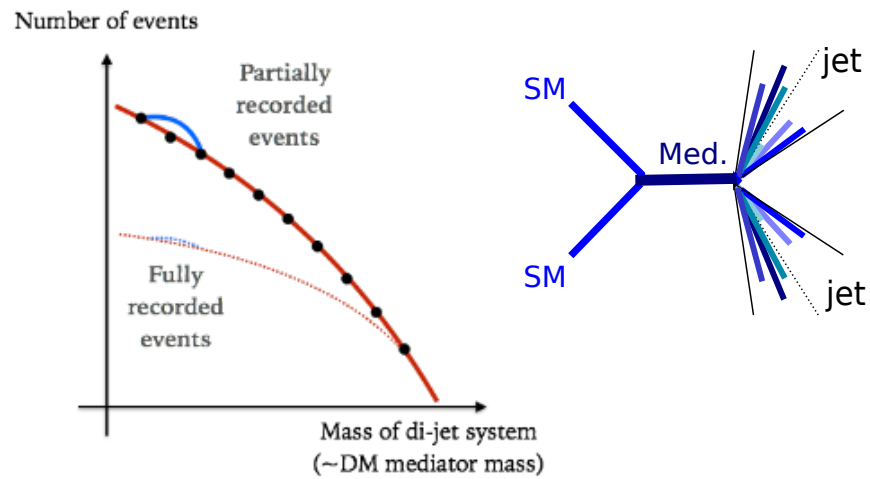
# Challenges for new particle searches

**Main challenge for resonance searches:** large backgrounds  
and signal that looks very much like background

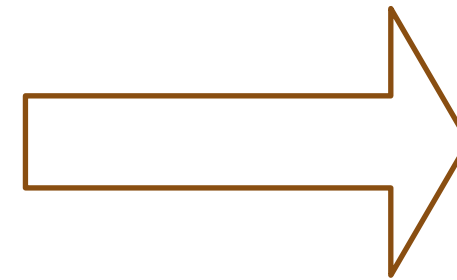
Number of events  
produced by the LHC



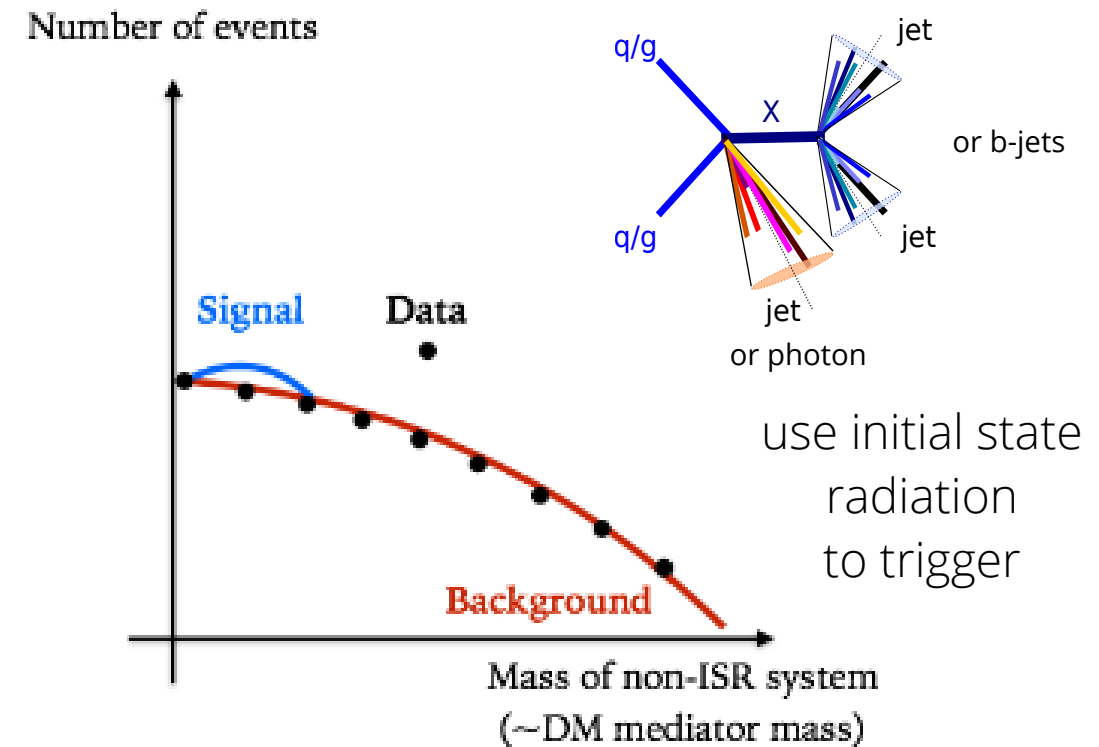
# Results from dijet+ISR search



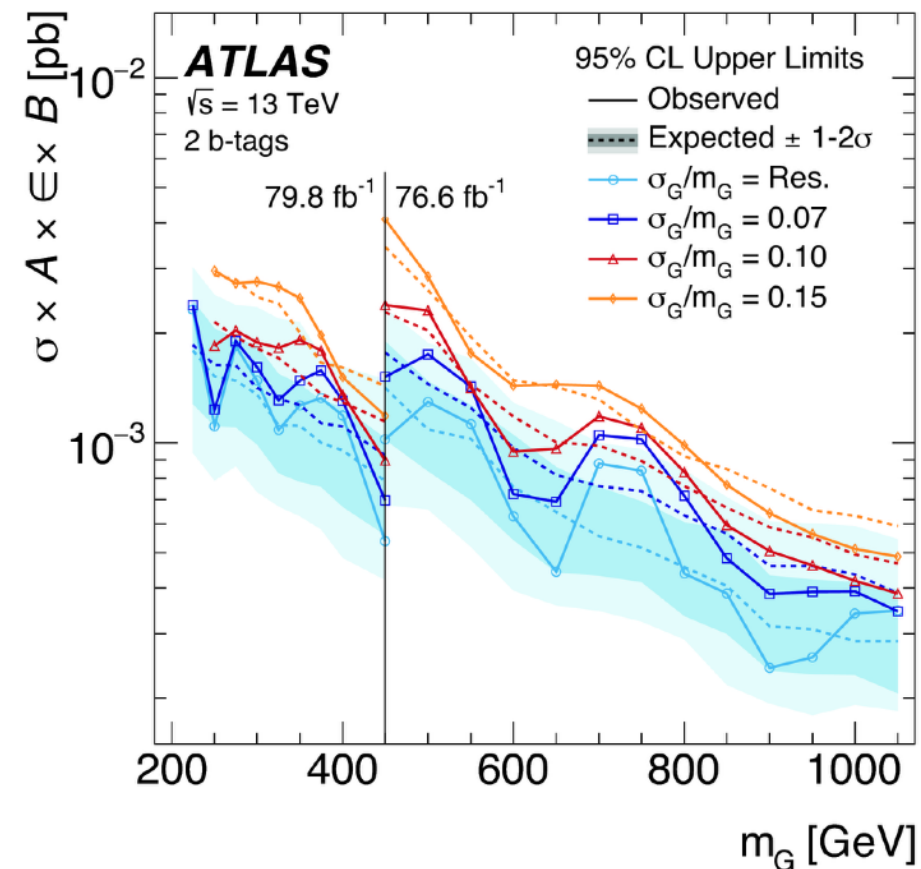
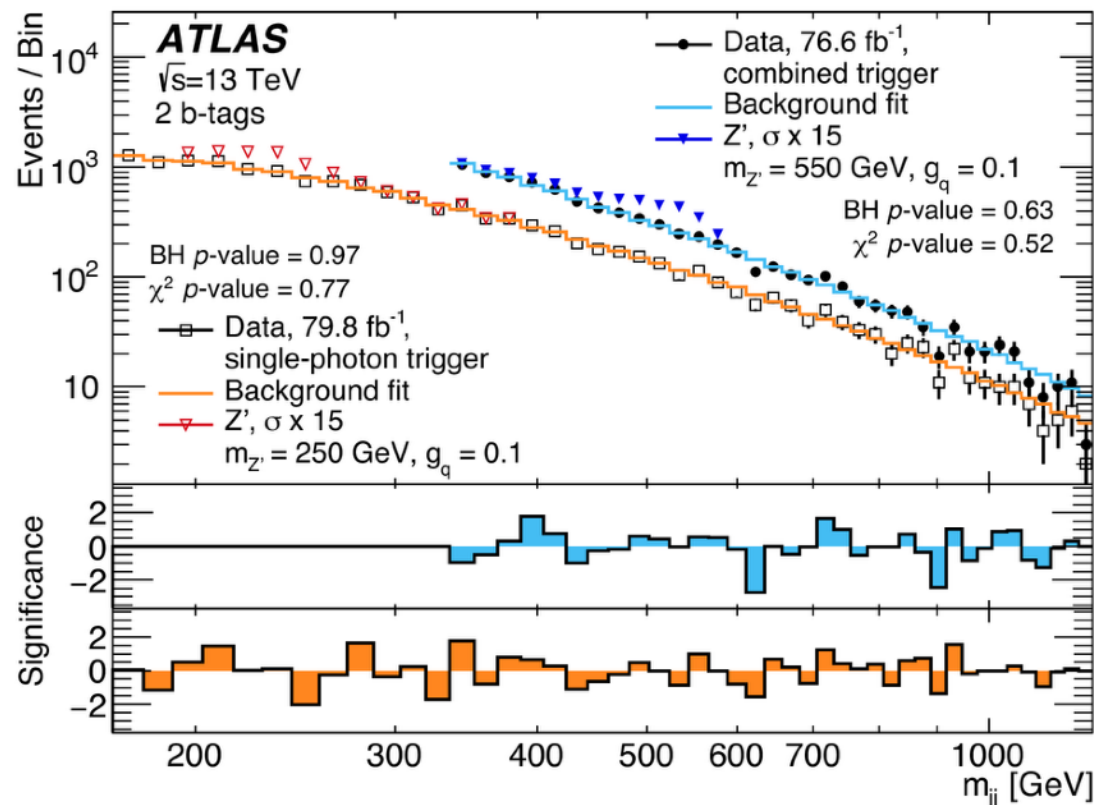
TLA restores full sensitivity, but range still limited by L1 trigger



reduce backgrounds to look for mediators with lower masses



[Phys. Lett. B 795 \(2019\) 56](#)

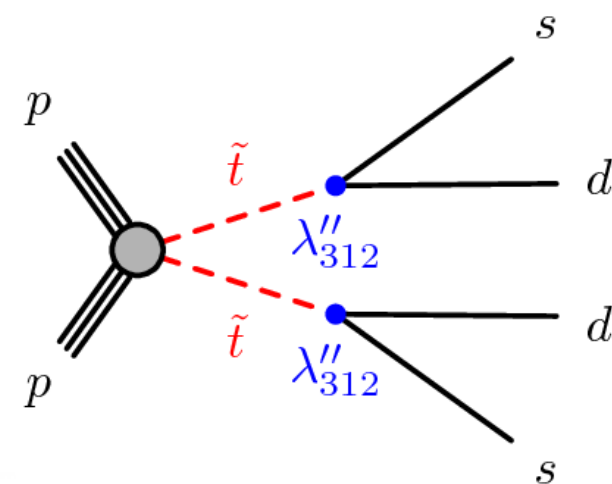
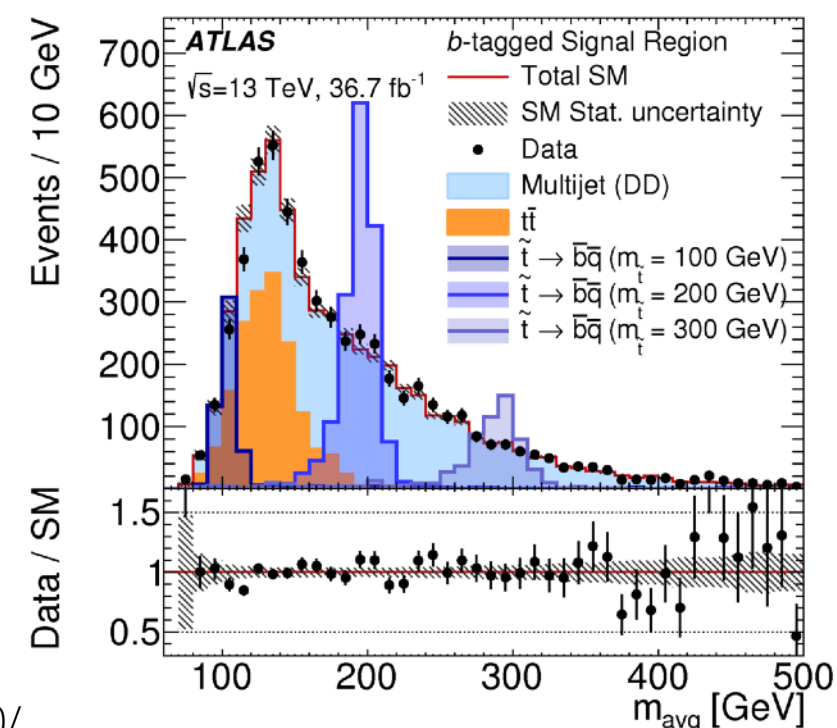
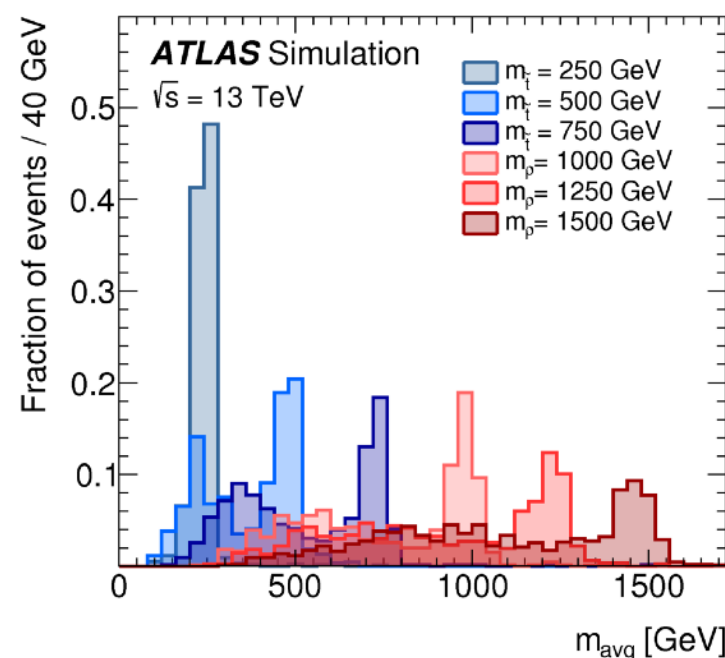
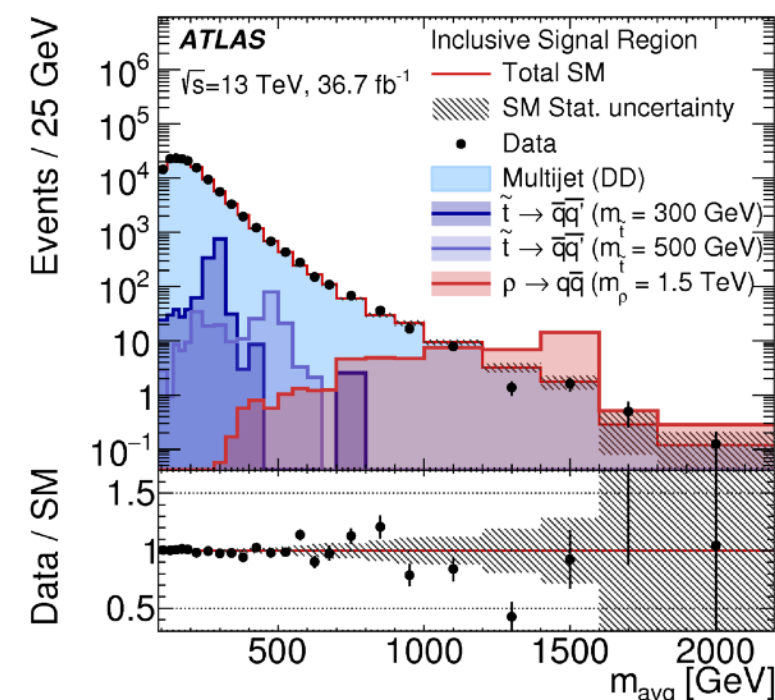




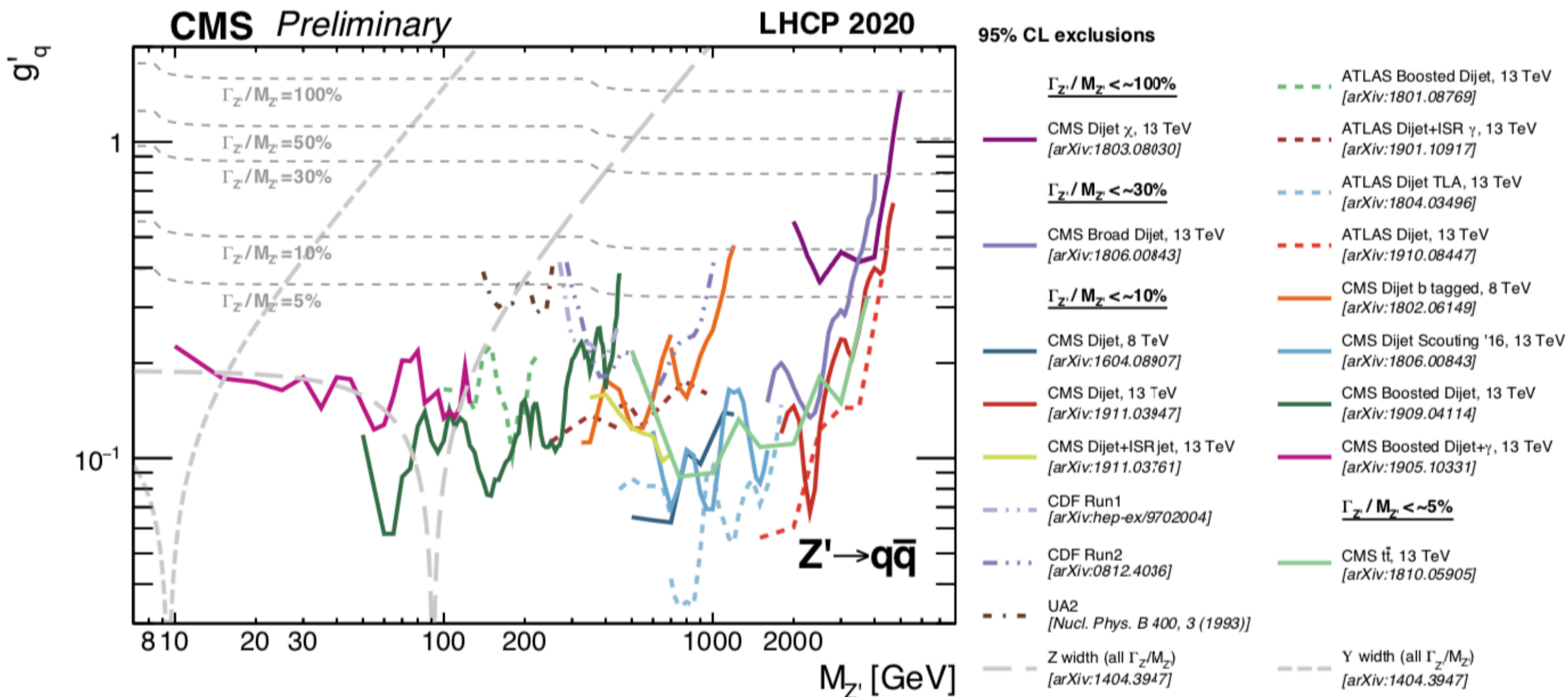
# Results from paired dijet search

[Eur. Phys. J. C 78 \(2018\) 250](#)

- For even lower masses:  
produce two new particles, each  
decaying in two jets
  - Not (yet) limited by triggers at  $\sim 100$  GeV
- Benchmark signal: (R-parity violating)  
supersymmetric top partner
- Connections to non-WIMP signals  
(e.g. *dark QCD*)
  - Developments ongoing...

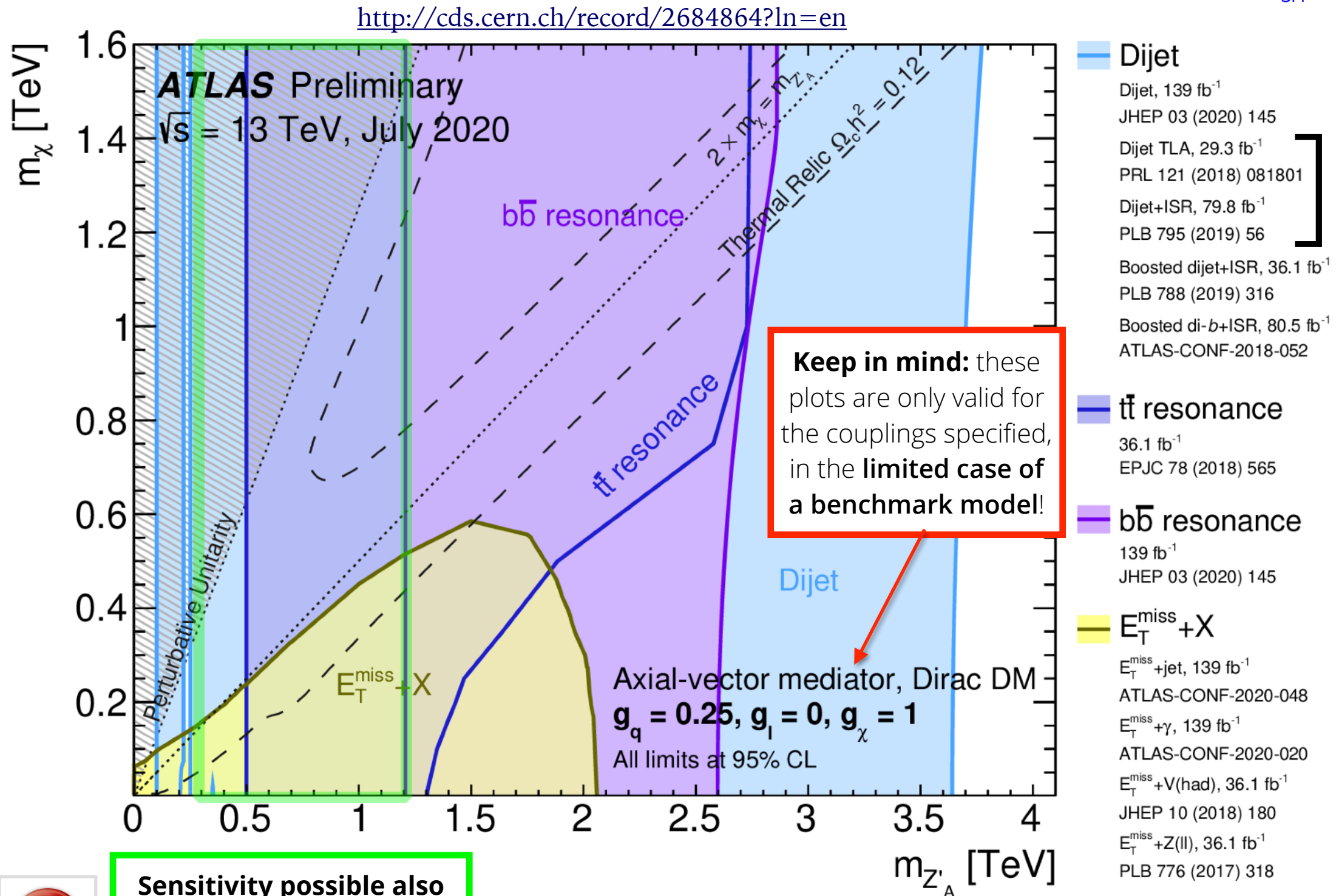
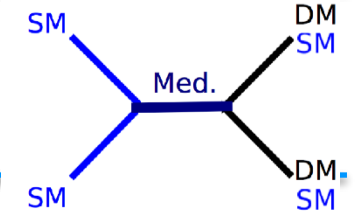


# Full set of mediator constraints from dijet searches



<https://twiki.cern.ch/twiki/pub/CMSPublic/SummaryPlotsEXO13TeV/>

# Visible/invisible searches for DM (mediators)



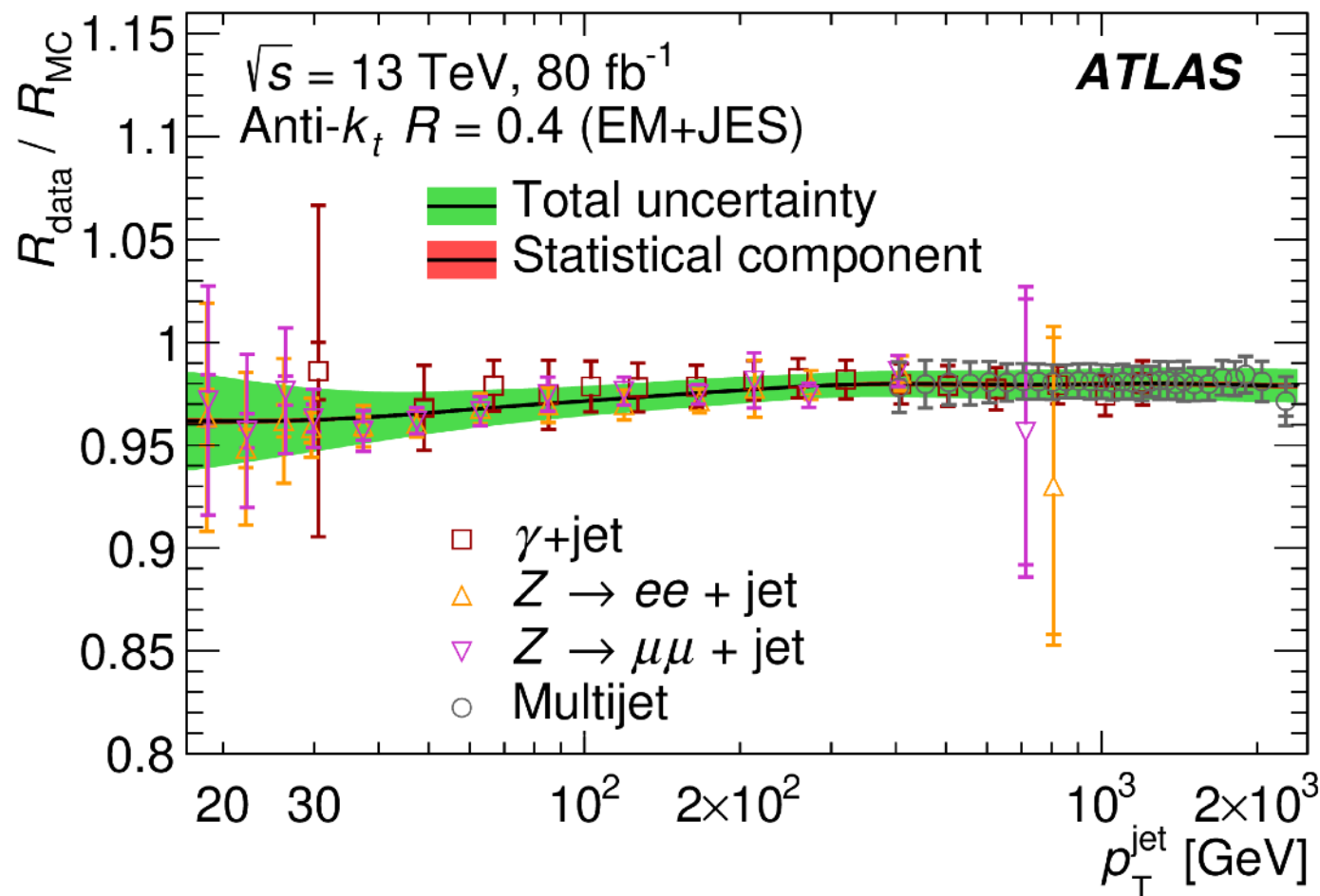
**Sensitivity possible also thanks to non-standard analysis workflows**





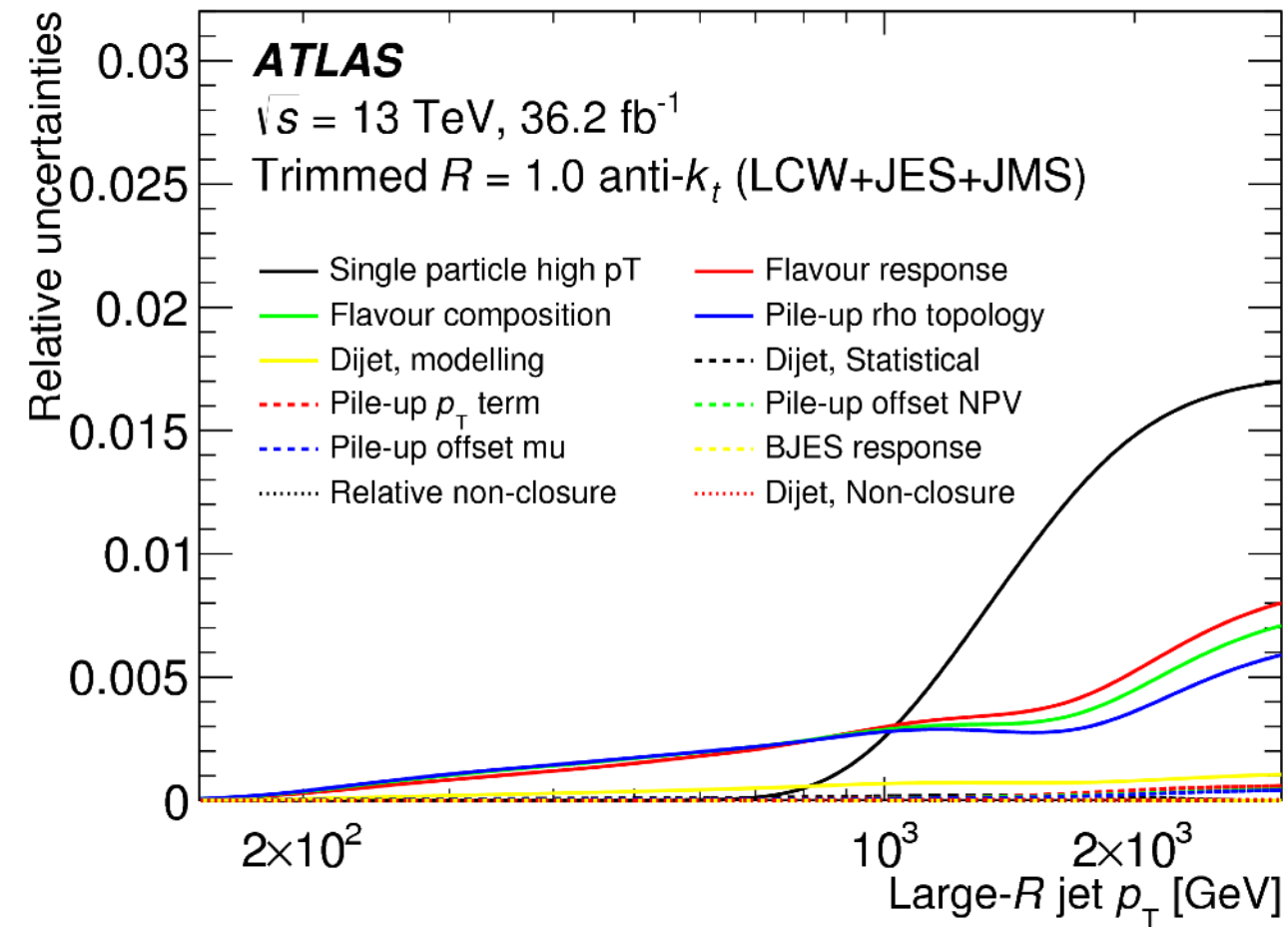
# Performance of small/large-R jets in ATLAS

[arXiv:2007.02645](https://arxiv.org/abs/2007.02645)



Final jet energy scale correction for small-R jets

[Eur. Phys. J. C 79 \(2019\) 135](https://arxiv.org/abs/1905.02645)



Jet energy scale uncertainty for large-R jets

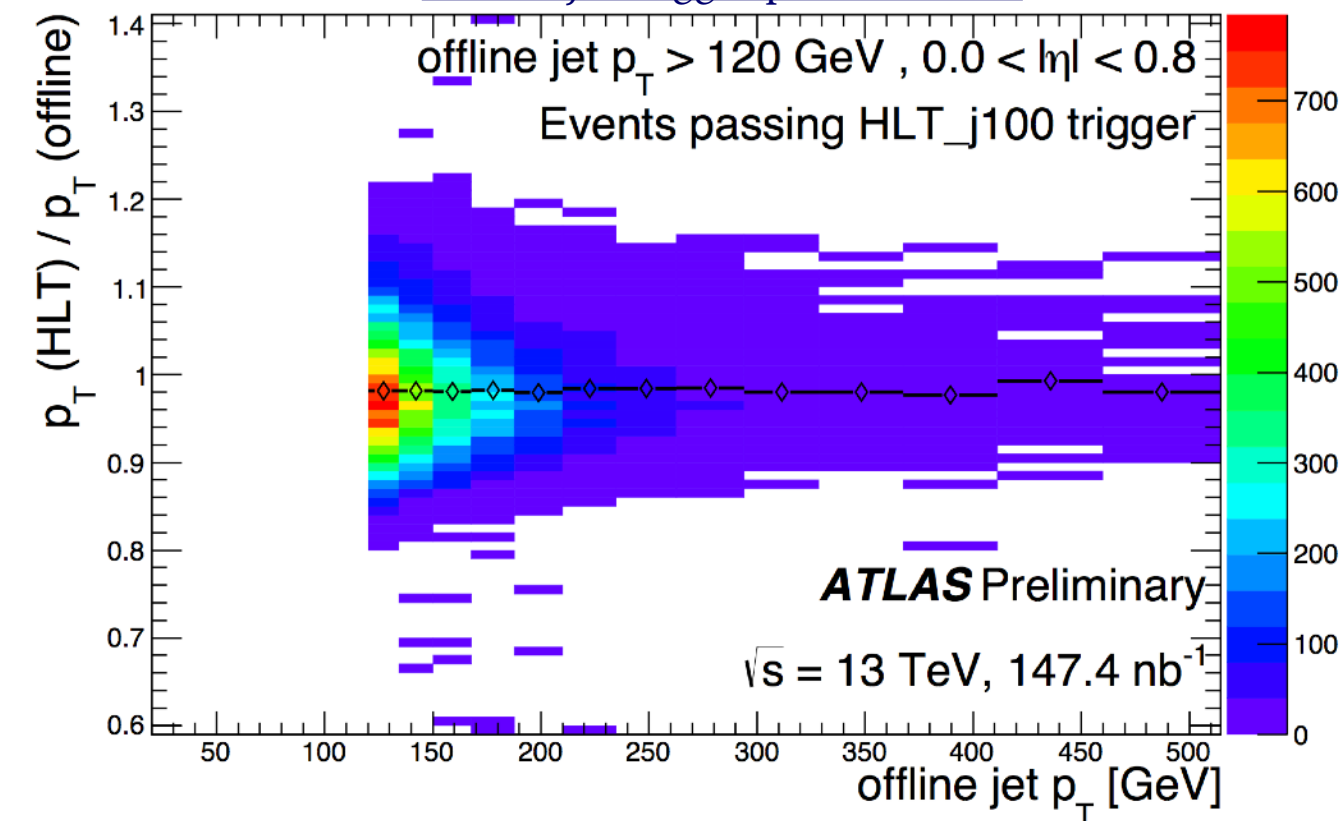
# Are trigger jets good enough?

...they have to be, to be able to detect very small signals!

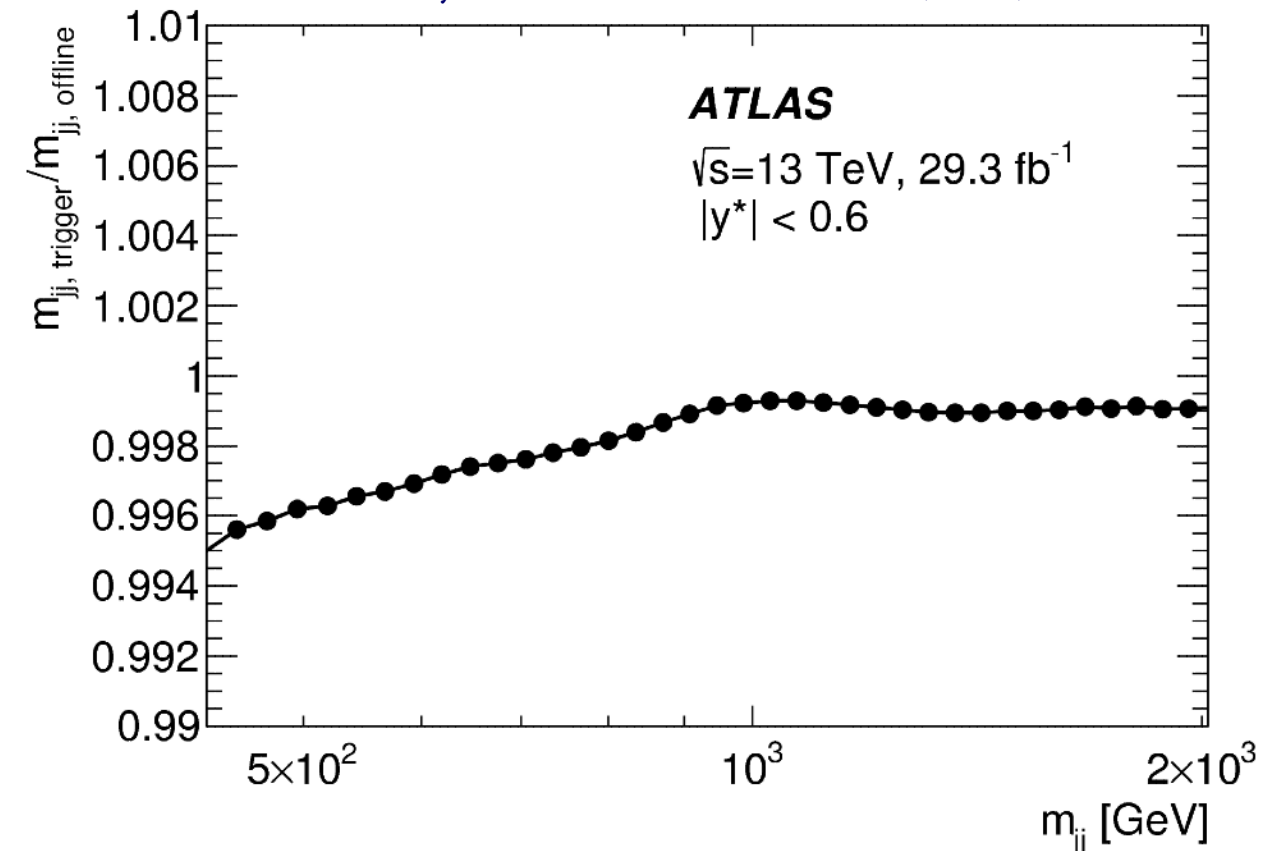
September 2015 (interview)

May 2017 (paper submission)

[ATLAS jet trigger public results](#)



[Phys. Rev. Lett. 121, 081801 \(2018\)](#)



(Note the change in y-axis scale)

# Possibilities for cross-talk in collaborative environments

- **High Energy Physics Software Foundation**

- Not only LHC experiments
- Forum for physicists with interest in software for HEP
- Instrumental for creation of [IRIS-HEP](#) NSF effort
- Working groups including [trigger & reconstruction](#)
  - [Website](#)
  - Mailing list: [hsf-forum@googlegroups.com](mailto:hsf-forum@googlegroups.com) (google group)



- **DarkMachines**

- Collective of astro/particle physicists interested in machine learning tools for dark matter
- Various (self-organized) efforts, including unsupervised searches
  - [Website](#)
  - [Subscribe to main mailing list](#)

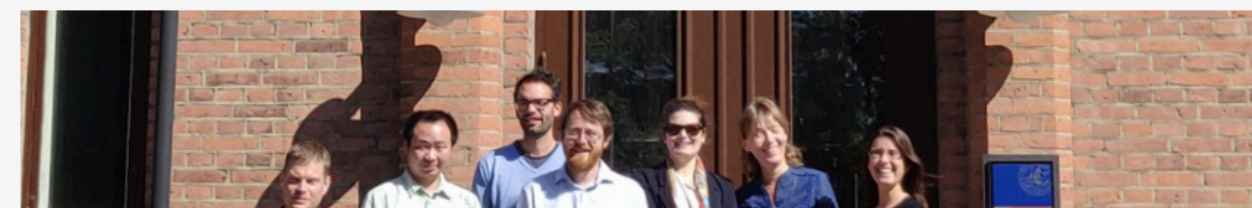
## About Dark Machines

Dark Machines is a research collective of physicists and data scientists. We are curious about the universe and want to answer cutting edge questions about Dark Matter with the most advanced techniques that data science provides us with.



# Real-time analysis (well) beyond HEP

Welcome to our REALTIME ASG blog! We're a group of researchers interested in real-time data acquisition and decision making as well as open data working as an Advanced Study Group 2019-2020 at the Pufendorf Institute for Advanced Studies.




- **Connections to astrophysics**
  - Real-time alerts for interesting events
- **Connections to accelerator beams / physics / engineering**
  - Beam steering, “triggering” for synchrotron and laser experiments
    - Sharing of technology, hybrid architectures (FPGA/GPU)
- **Connections to social sciences & law**
  - Real-time data deposition promotes open data and credibility of science
    - Implications for treatment and ownership of data treatment
    - Sharing of tools (versioning / data persistency)
-  about the [REALTIME](#) Advanced Study Group

open to discuss further collaborations!



# Real-time analysis in HEP and industry

- **Big data analysis / real-time analysis**

- Shared tools: machine learning, hybrid computing architectures
- Connections: IT, industry (e.g. Internet-of-Things)
  - data is mostly cheap to record
  - time-to-insight is the key metric
-  about the [SMARTHEP](#) and HELIOS networks

open to discuss further collaborations!

Innovative Training Networks (ITN)  
Call: H2020-MSCA-ITN-2020



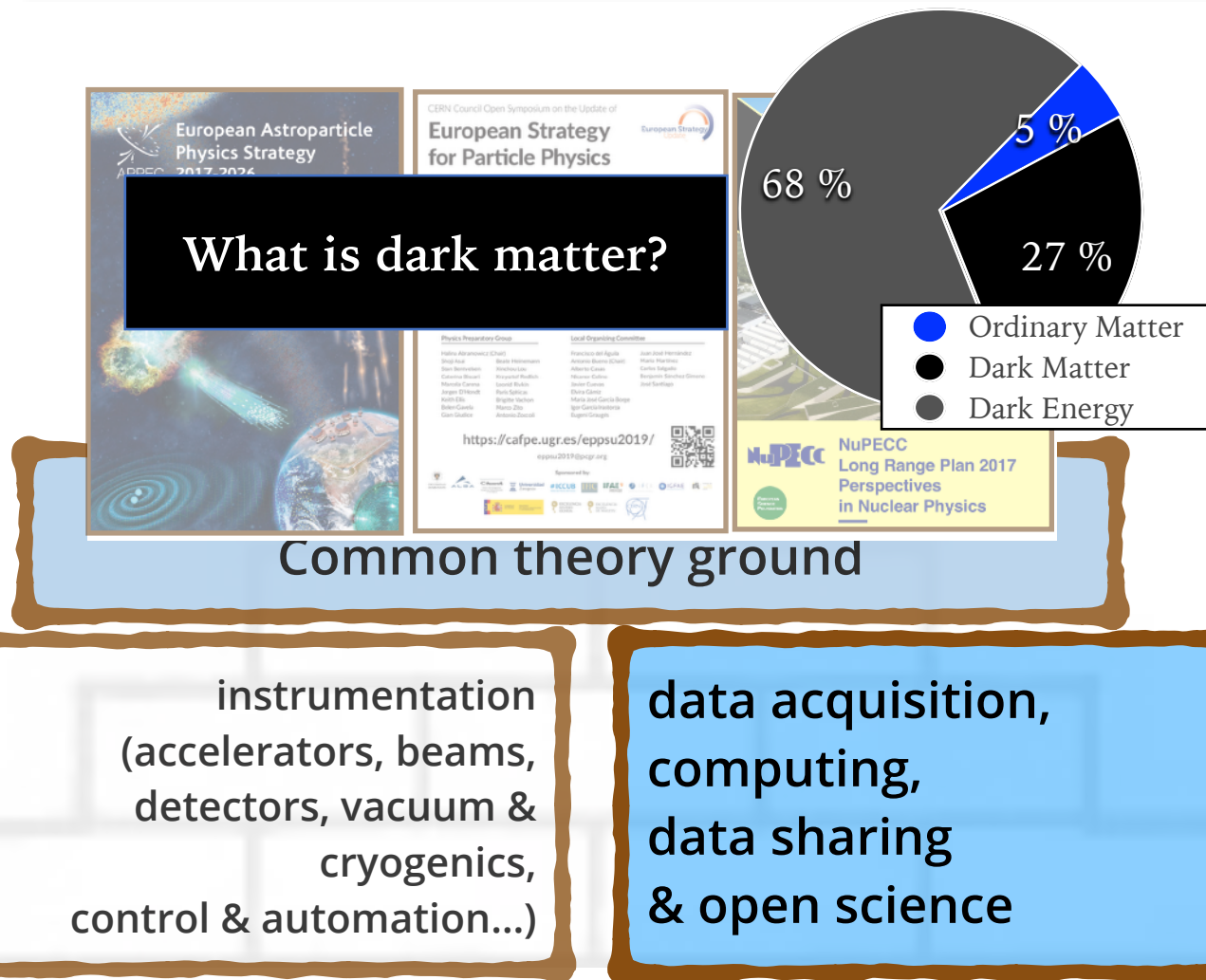
(proposal currently in "reserve list")

Synergies between **MA**chine learning, **Rea**l **T**ime analysis  
and **Hy**brid architectures for efficient **E**vent **P**rocessing and decision making  
**SMARTHEP**



(proposal funded as Helmholtz International Graduate School)

# Foundations needed to exploit synergies



Talk at EPS-HEP / ECFA session 2019, CERN EP Newsletter

Slides from European Strategy Update release, 19/06/2020



## 2020 Strategy Statements

### 5. Synergies with neighbouring fields

#### Particle and Astroparticle Physics

- Synergies exist at the level of infrastructure, detectors, computing, interaction models and physics goals (ex.: neutrinos, dark matter, cosmic rays and gravitational waves)
- The need to foster these synergies has been clearly identified in the national inputs



# Two ongoing projects focused on Dark Matter

## searches & interpretation

JENAS EoI: Initiative for Dark Matter in Europe and beyond: Towards facilitating communication and result sharing in the Dark Matter community (iDMEu)

<https://indico.cern.ch/event/869195/>  
[ESCAPE newsletter](#) [APPEC newsletter](#)

build a discussion platform and tools to facilitate collaboration of existing groups/ efforts on **dark matter searches** and **interpretation**



Common theory ground

instrumentation  
(accelerators, beams, detectors,  
vacuum & cryogenics,  
control & automation...)

data acquisition,  
software, computing,  
data sharing  
& open science



**Towards a Dark Matter  
Test Science Project**

[ESCAPE Progress Meeting, 2020](#)

compare **end-to-end analysis workflows** for WIMP searches, towards their implementation in a common **Software Catalogue** and as input to the design of the **European Open Science Cloud**

## software & data

More initiatives and links in backup slides

