Peter Skands (Monash University — Melbourne Australia)

PYTHIA

for Flavour, Dark-Matter Physics, and Machine Learning

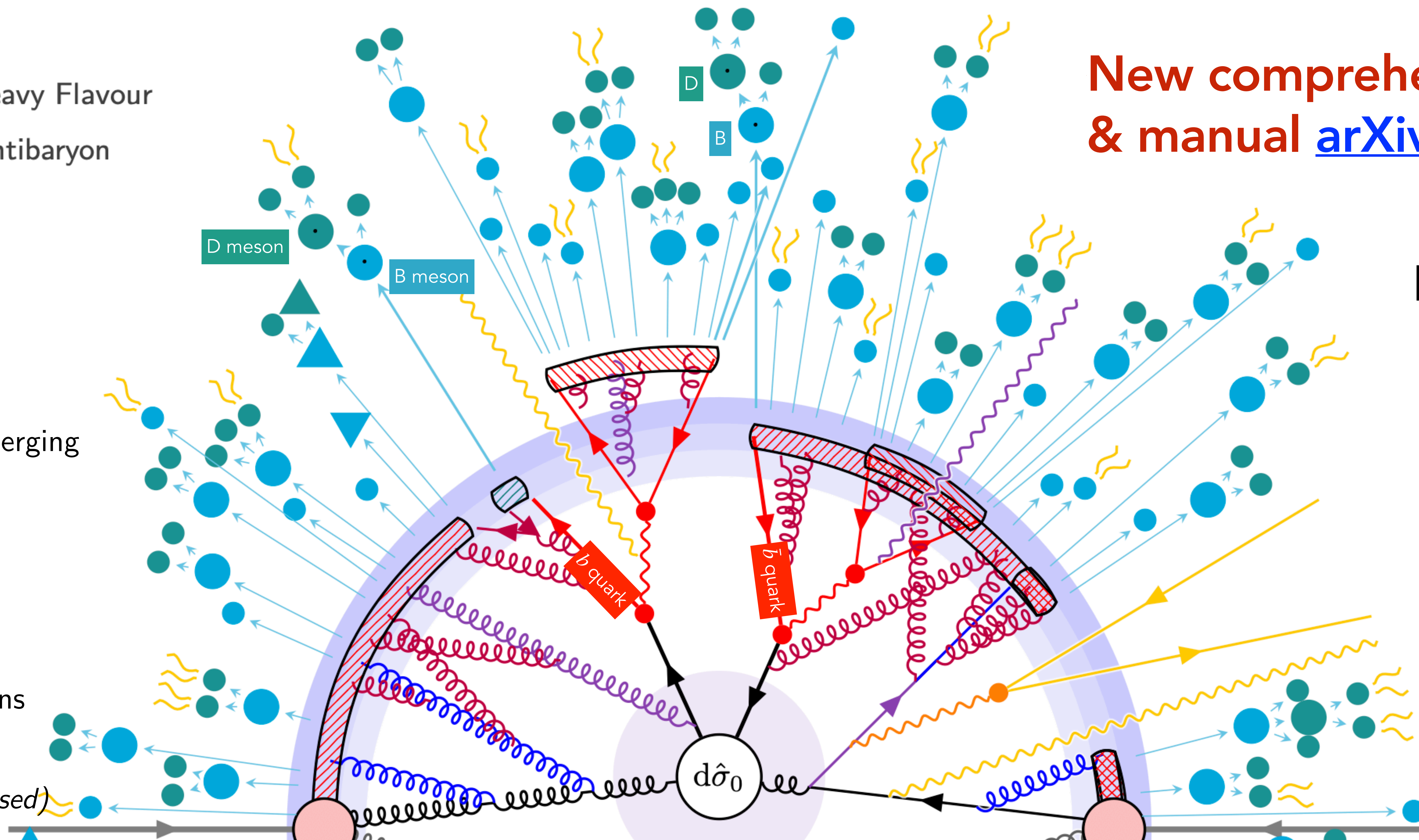
New comprehensive guide & manual [arXiv:2203.11601](https://arxiv.org/abs/2203.11601)

● Meson ⊙ Heavy Flavour
▲ Baryon ▼ Antibaryon

Parton Level

- Hard Interaction
- Resonance Decays
- MECs, Matching & Merging
- FSR
- ISR*
- QED
- Weak Showers
- Hard Onium
- Multiparton Interactions
- Beam Remnants*

(: incoming lines are crossed)*



Hadron Level

- Beam Remnants*
- Strings
- Clusters
- Colour Reconnections
- String Interactions
- Bose-Einstein & Fermi-Dirac
- Primary Hadrons
- Secondary Hadrons
- Hadronic Reinteractions
- QED in Hadron Decays

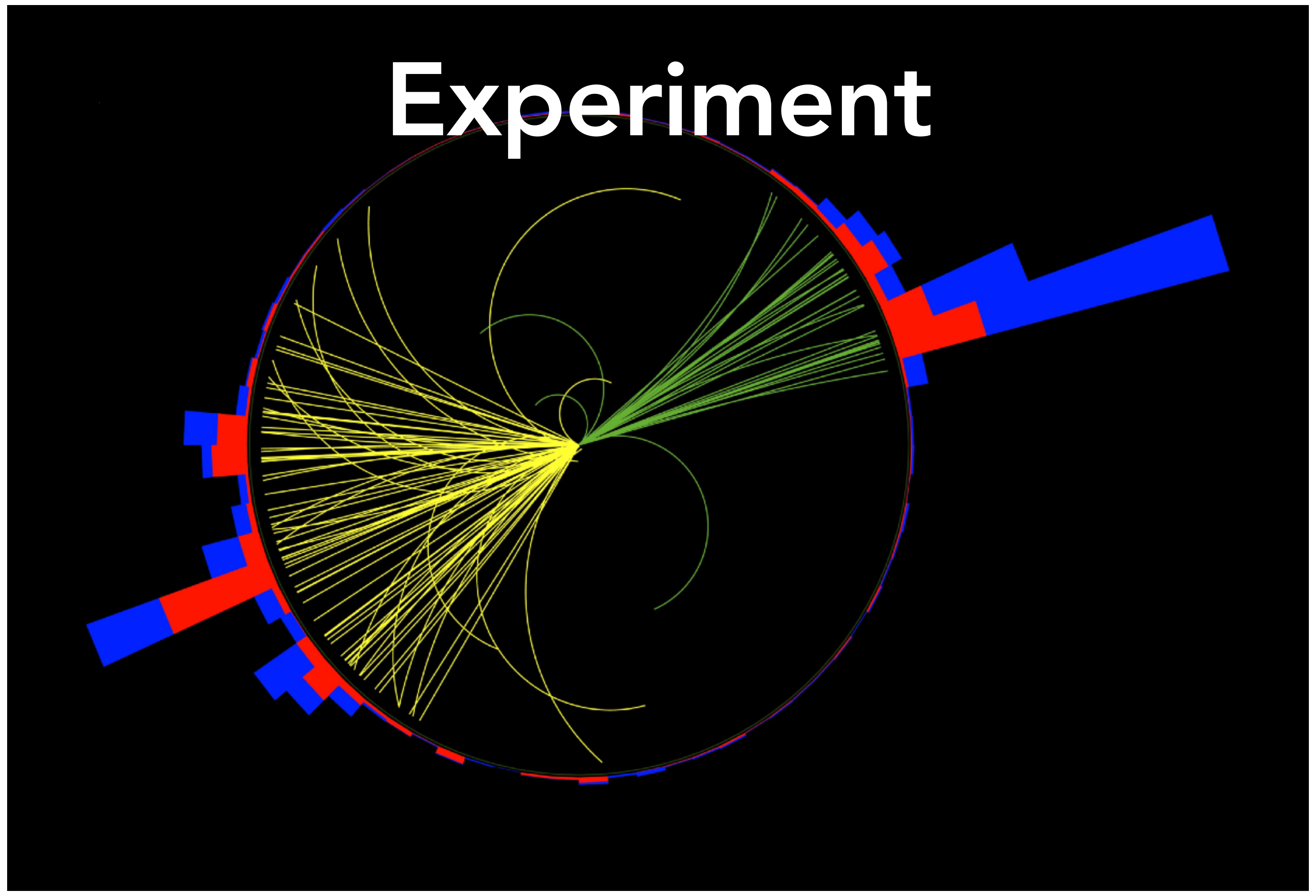
(: incoming lines are crossed)*

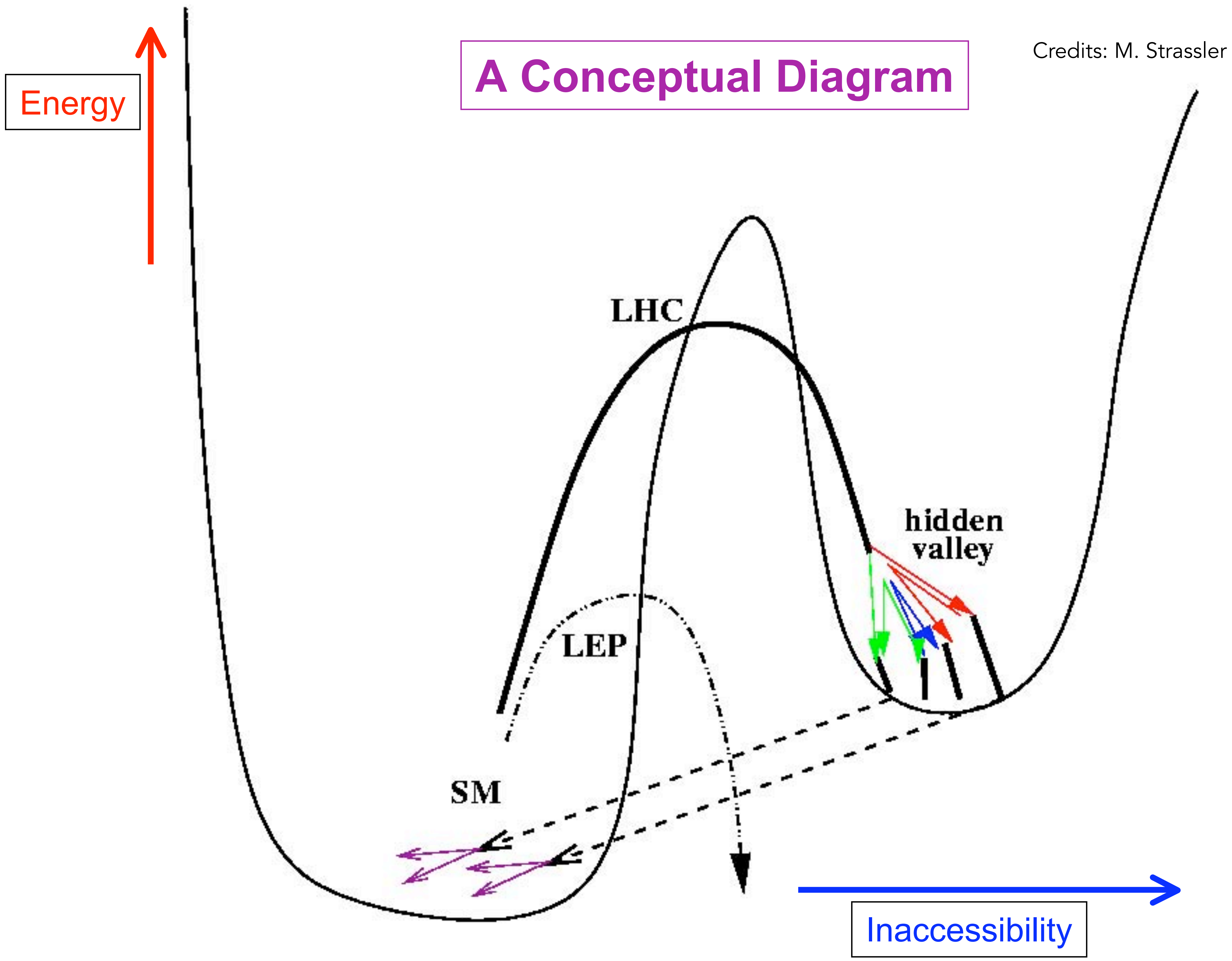
Model Building:

$$\mathscr{L}_{SM} \quad (+ \ \mathscr{L}_{BSM} \ ? \ )$$

Fundamental parameters

Fields, Symmetries, Couplings, Masses, …





Example: SM + Dark Sector (Hidden Valley)

From Theory to Observables (and back again)

Model Building:

$$\mathscr{L}_{SM} \quad (+ \ \mathscr{L}_{BSM}\ ?\ )$$

Fundamental parameters

Fields, Symmetries, Couplings, Masses, …

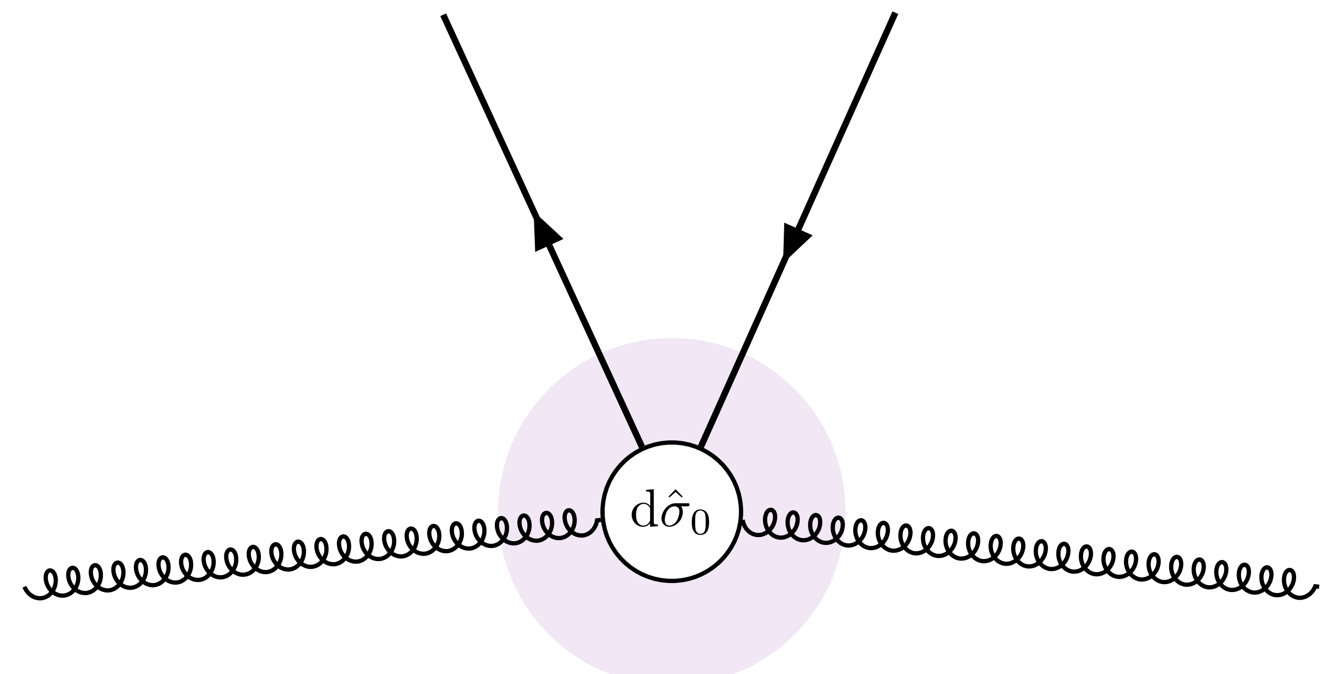
Can we measure this?

Phenomenology:

Compute observables:

$$\mathrm{d}\sigma_{AB\to X_1\ldots X_n},\ \mathrm{d}\Gamma_{A\to X_1\ldots X_n},\ \ldots$$

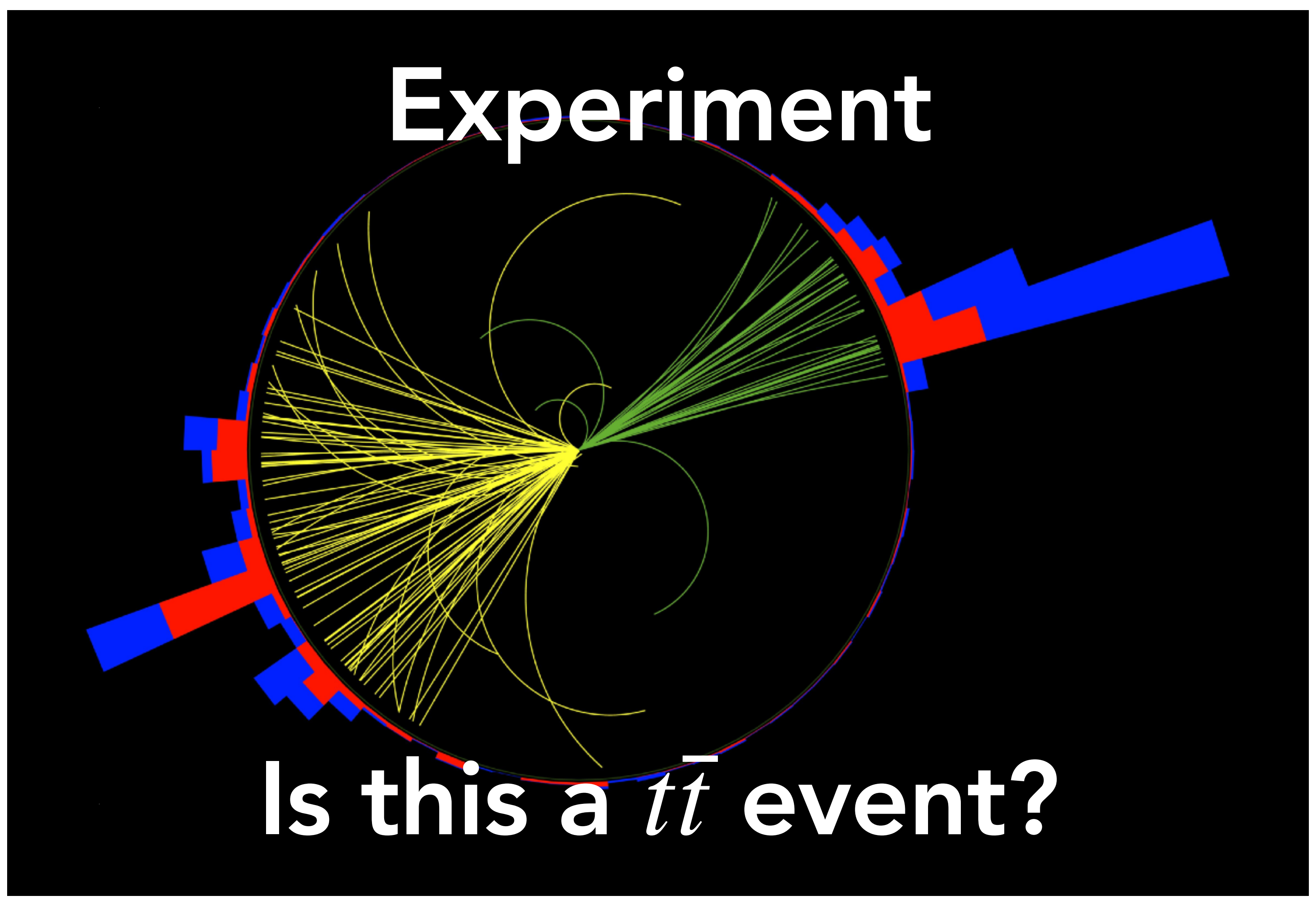
Example: top quarks



"Predict" probability distributions differentially in relevant phase spaces

➡ **This is a $t\bar{t}$ "event"**

Can we measure this?



Complex final states, backgrounds; efficiencies, calibrations, …

➡ Measured result + uncertainties

Adding Detail: QCD Showers

Model Building:

$$\mathscr{L}_{SM} \quad (+ \; \mathscr{L}_{BSM} \, ? \,)$$

Fundamental parameters

Fields, Symmetries, Couplings, Masses, …

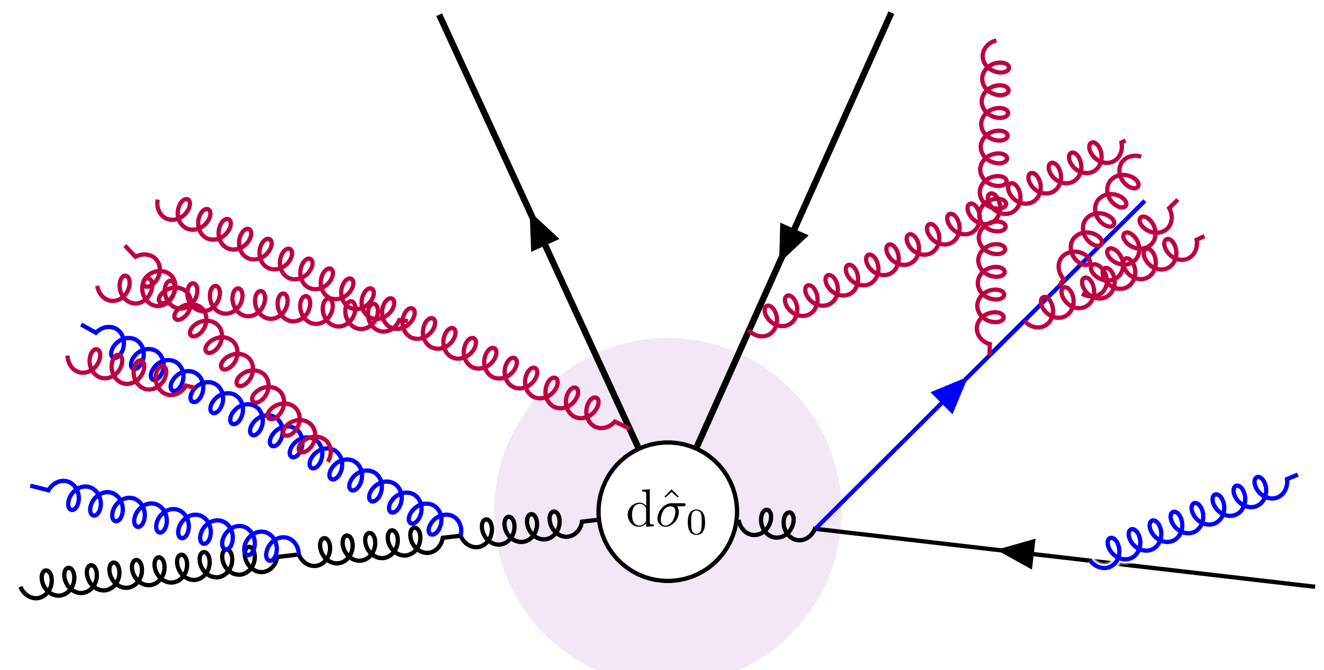
Can we measure this?

Phenomenology:

Compute observables:

$$\mathrm{d}\sigma_{AB\to X_1\ldots X_n}\,,\;\; \mathrm{d}\Gamma_{A\to X_1\ldots X_n}\,,\;\ldots$$

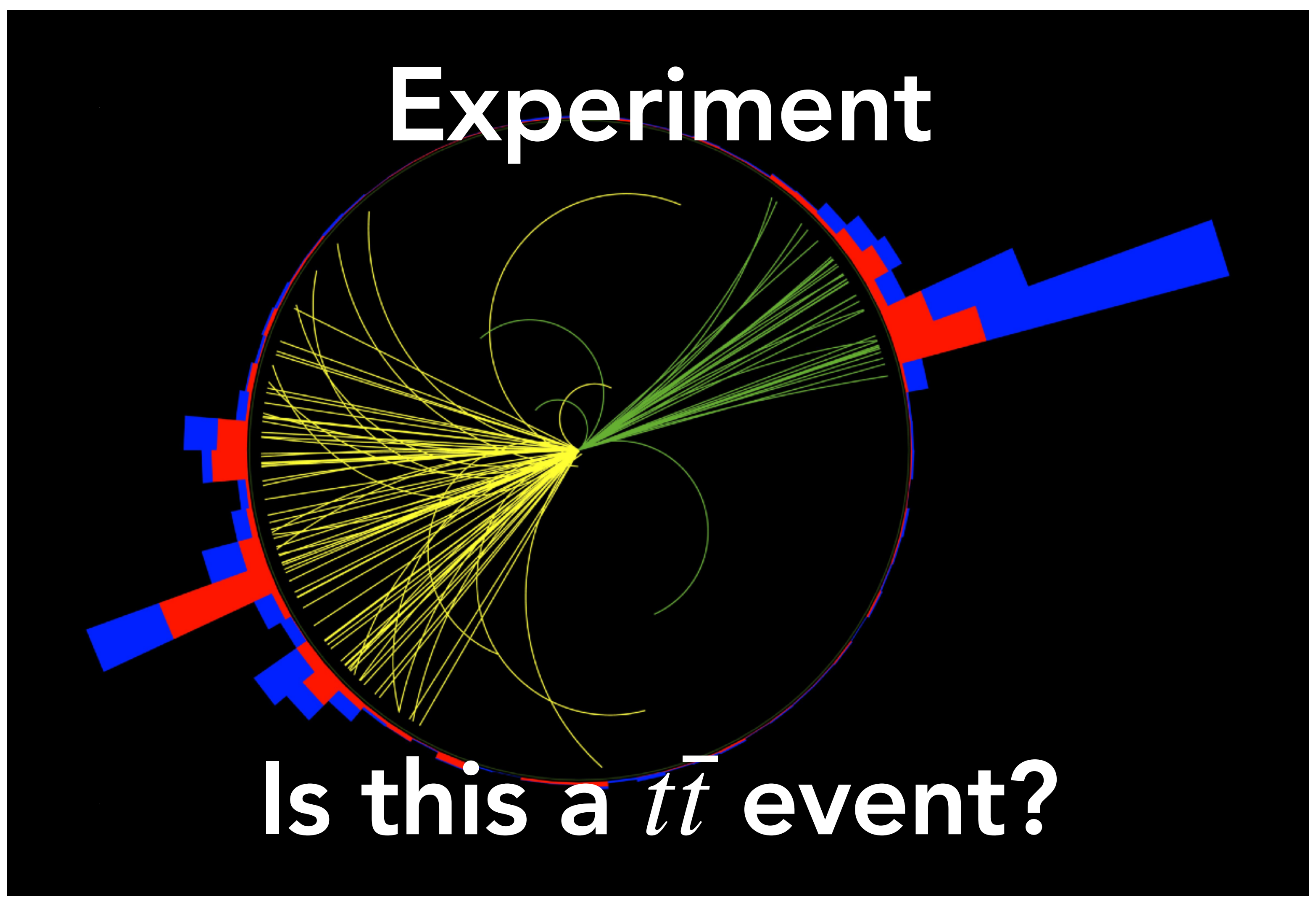
Example: top quarks



"Predict" probability distributions differentially in relevant phase spaces

➡ **This is a $t\bar{t}$ "event"**

Can we measure this?



Complex final states, backgrounds; efficiencies, calibrations, …

➡ Measured result + uncertainties

Adding Detail: Resonance Decays

Model Building:

$$\mathscr{L}_{SM} \quad (+ \ \mathscr{L}_{BSM} \ ? \ )$$

Fundamental parameters

Fields, Symmetries, Couplings, Masses, …

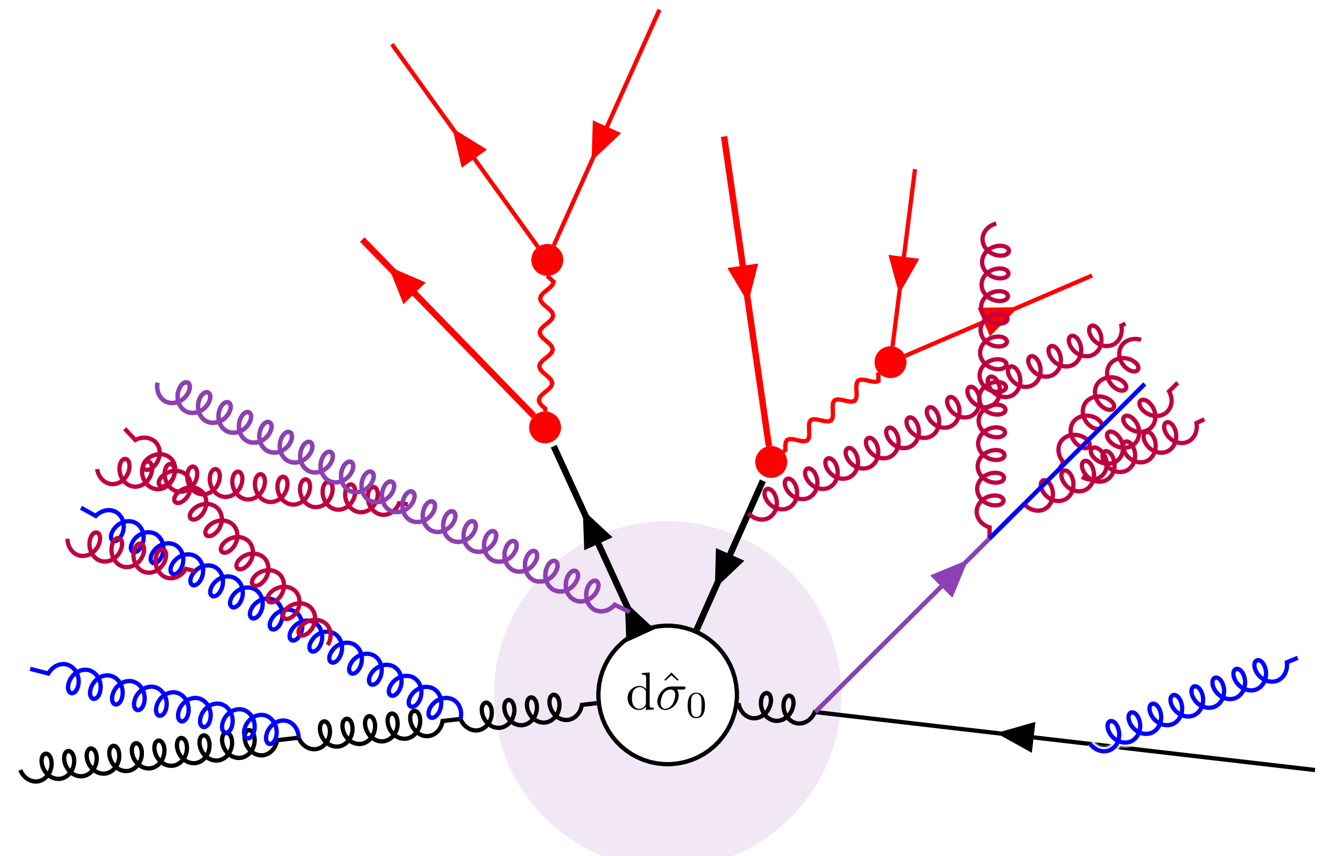
Can we measure this?

Phenomenology:

Compute observables:

$$\mathrm{d}\sigma_{AB\to X_1\ldots X_n}, \ \mathrm{d}\Gamma_{A\to X_1\ldots X_n}, \ \ldots$$

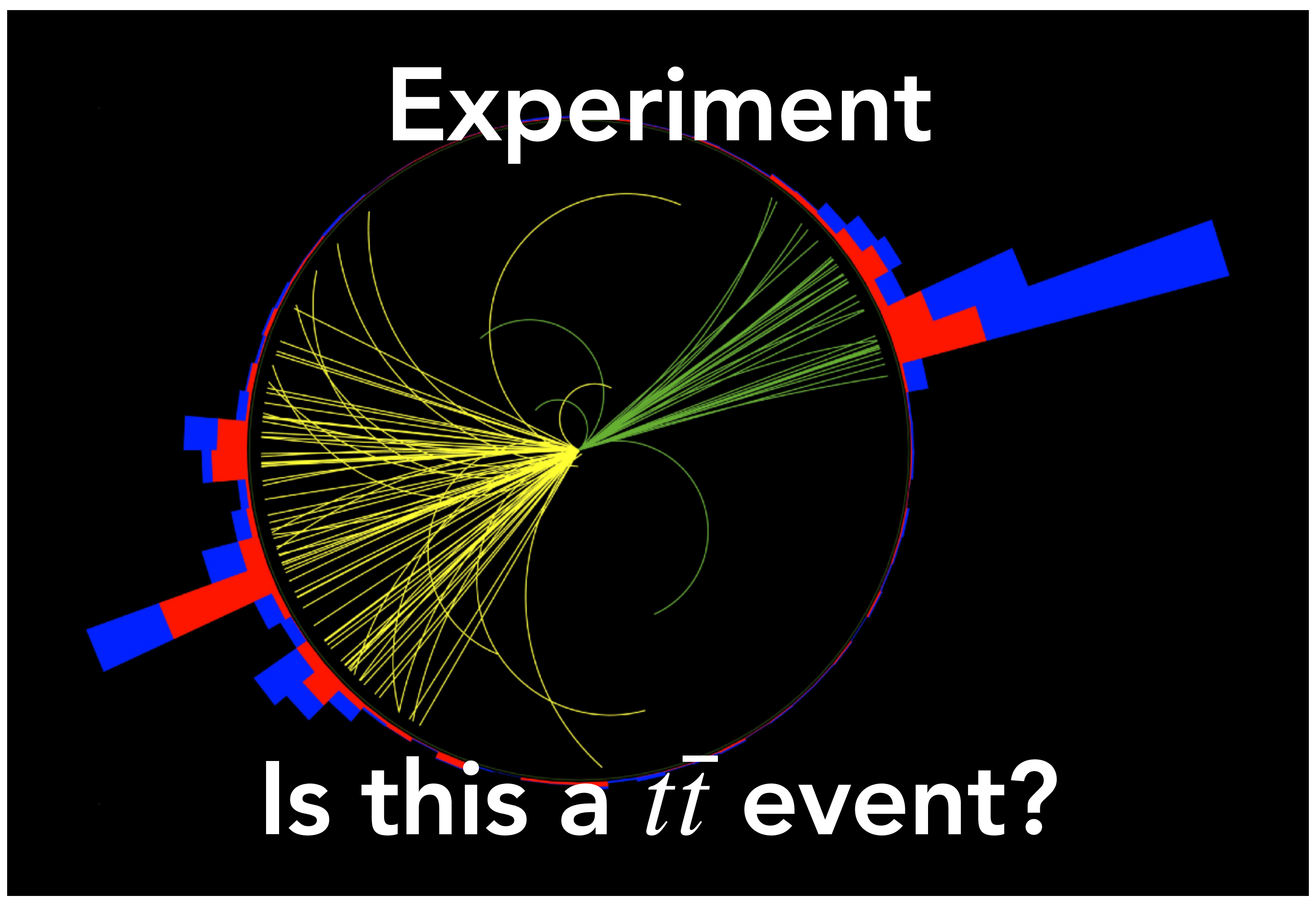
Example: top quarks



"Predict" probability distributions differentially in relevant phase spaces

➡ **This is a $t\bar{t}$ "event"**

Can we measure this?



Complex final states, backgrounds; efficiencies, calibrations, …

➡ Measured result + uncertainties

Adding Detail: **Showers in Decays**

Model Building:

$$\mathscr{L}_{SM} \quad (+ \ \mathscr{L}_{BSM} \, ? \,)$$

Fundamental parameters

Fields, Symmetries, Couplings, Masses, …

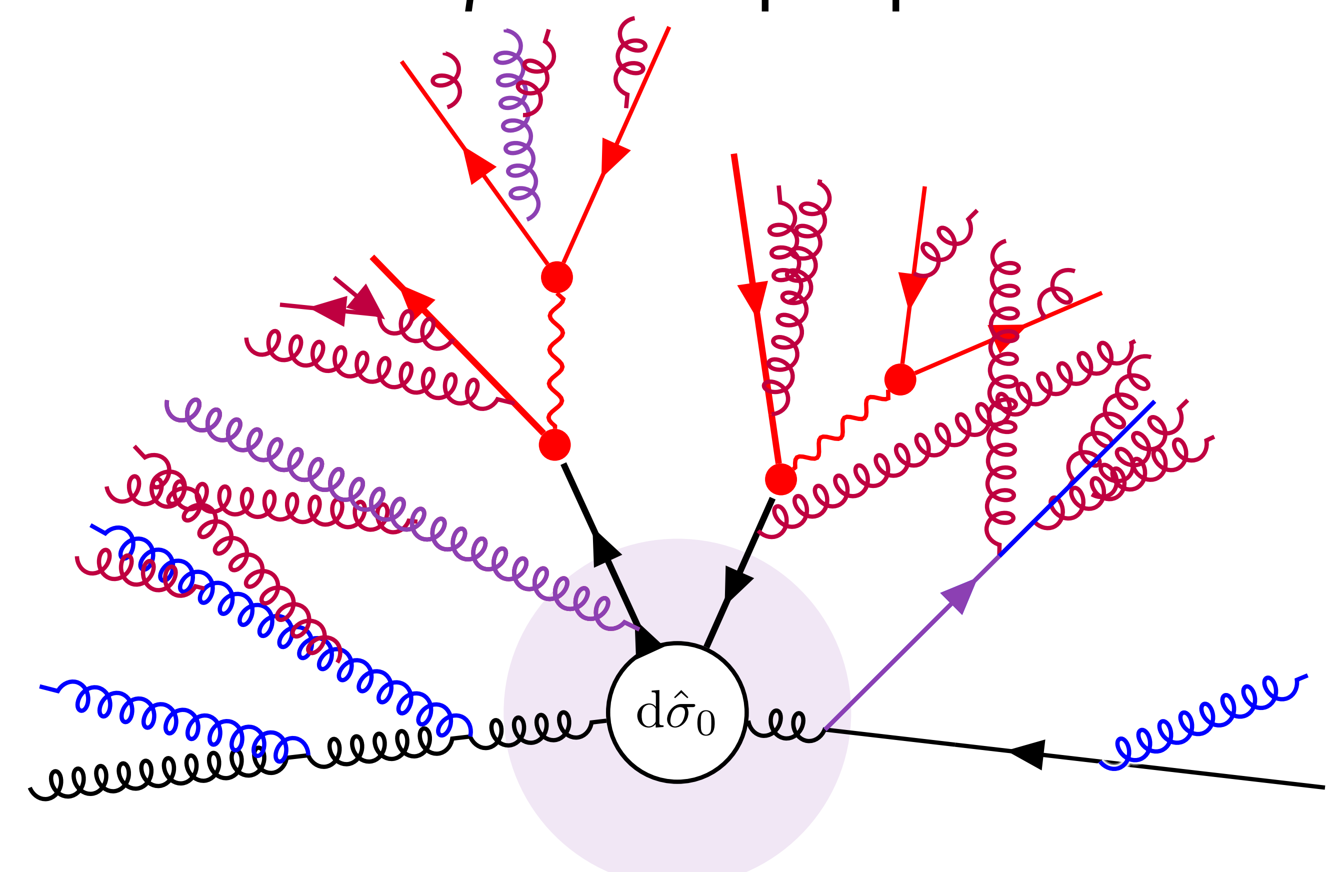
Can we measure this?

Phenomenology:

Compute observables:

$$d\sigma_{AB\to X_1\ldots X_n}, \ d\Gamma_{A\to X_1\ldots X_n}, \ \ldots$$

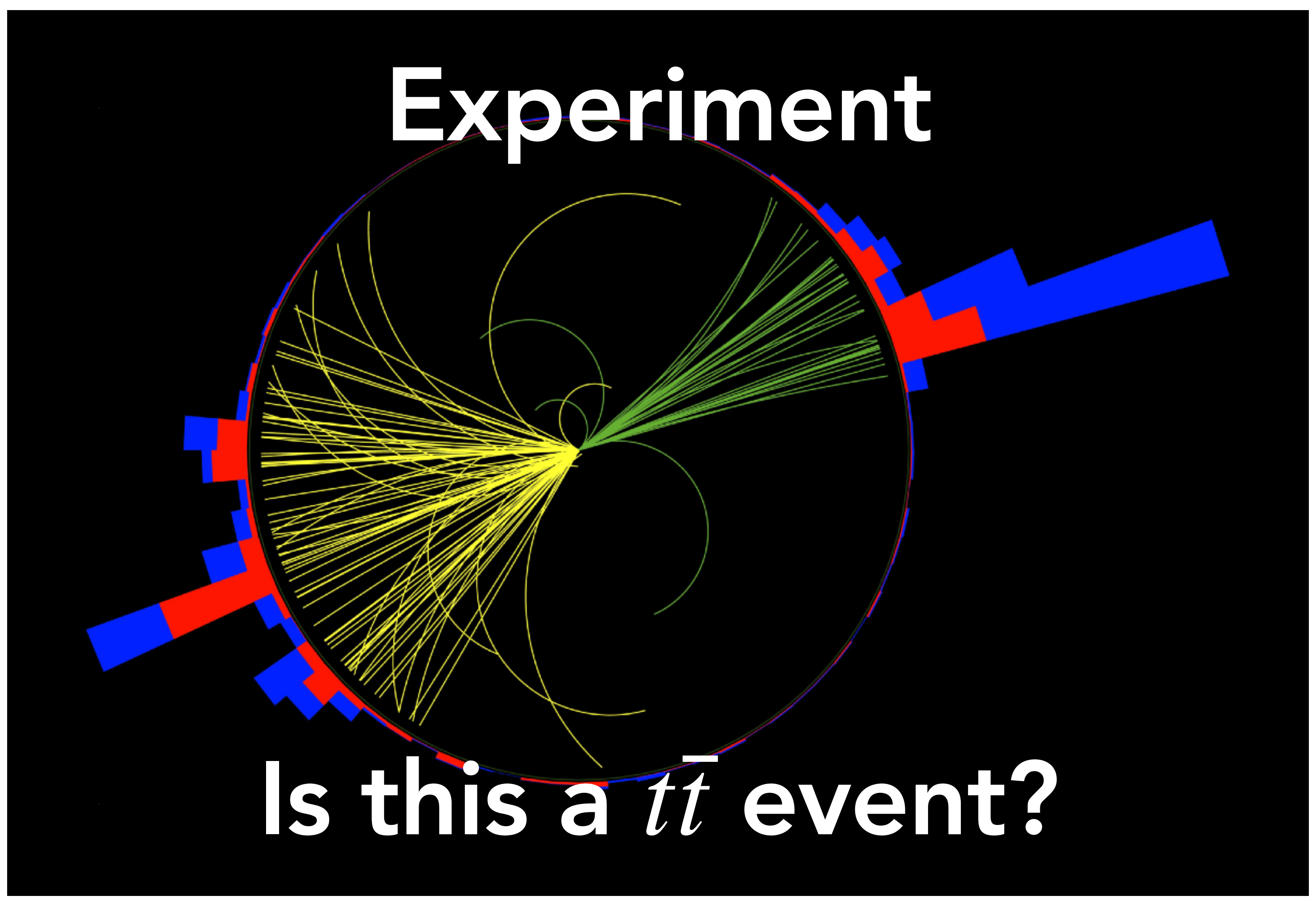
Example: top quarks



"Predict" probability distributions differentially in relevant phase spaces

➡ **This is a $t\bar{t}$ "event"**

Can we measure this?



Complex final states, backgrounds; efficiencies, calibrations, …
➡ Measured result + uncertainties

Adding Detail: QED and Weak Showers

Model Building:

$$\mathscr{L}_{SM} \quad (+ \; \mathscr{L}_{BSM} \; ? \;)$$

Fundamental parameters

Fields, Symmetries, Couplings, Masses, …

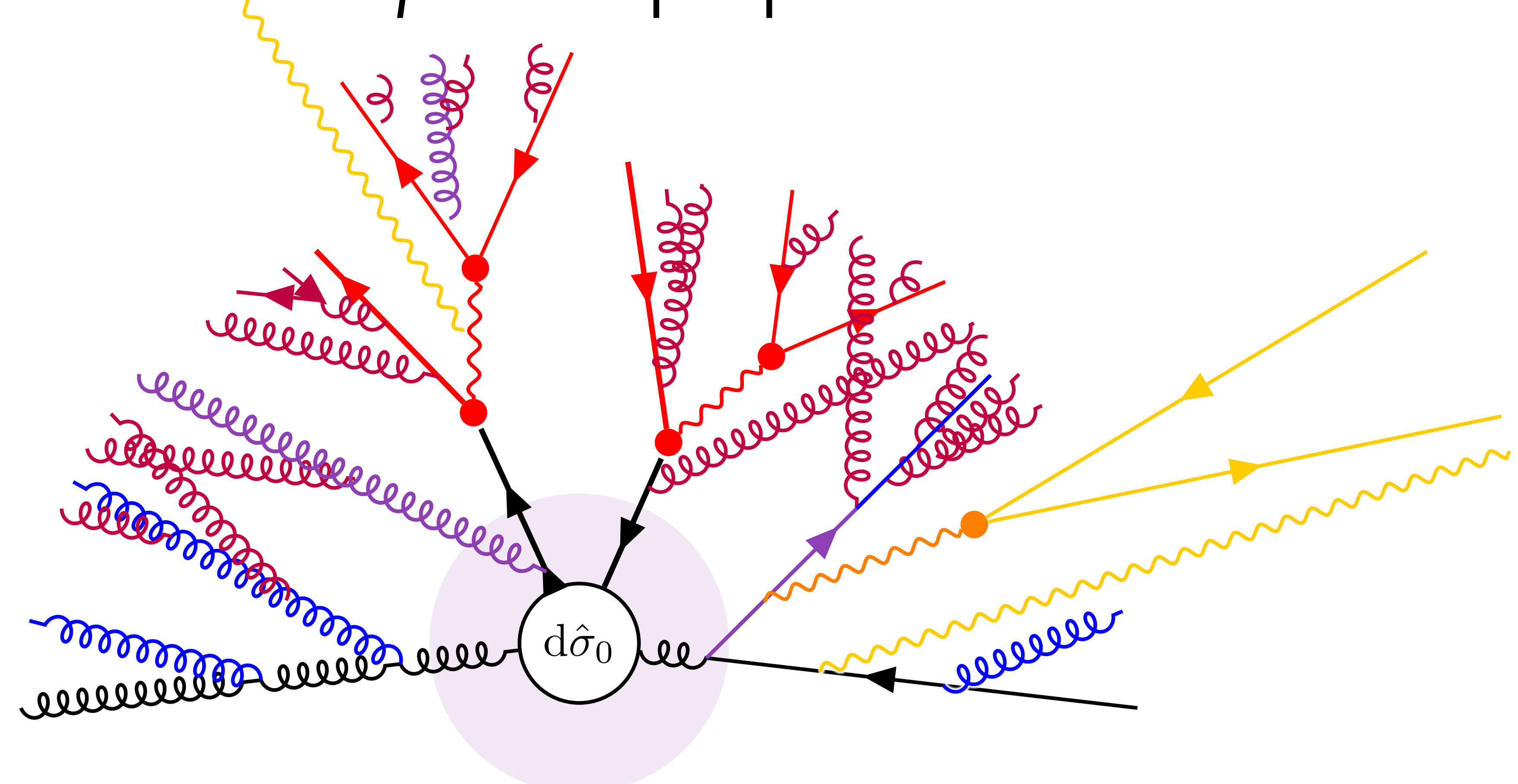
Can we measure this?

Phenomenology:

Compute observables:

$$\mathrm{d}\sigma_{AB\to X_1\ldots X_n}, \;\; \mathrm{d}\Gamma_{A\to X_1\ldots X_n}, \;\ldots$$

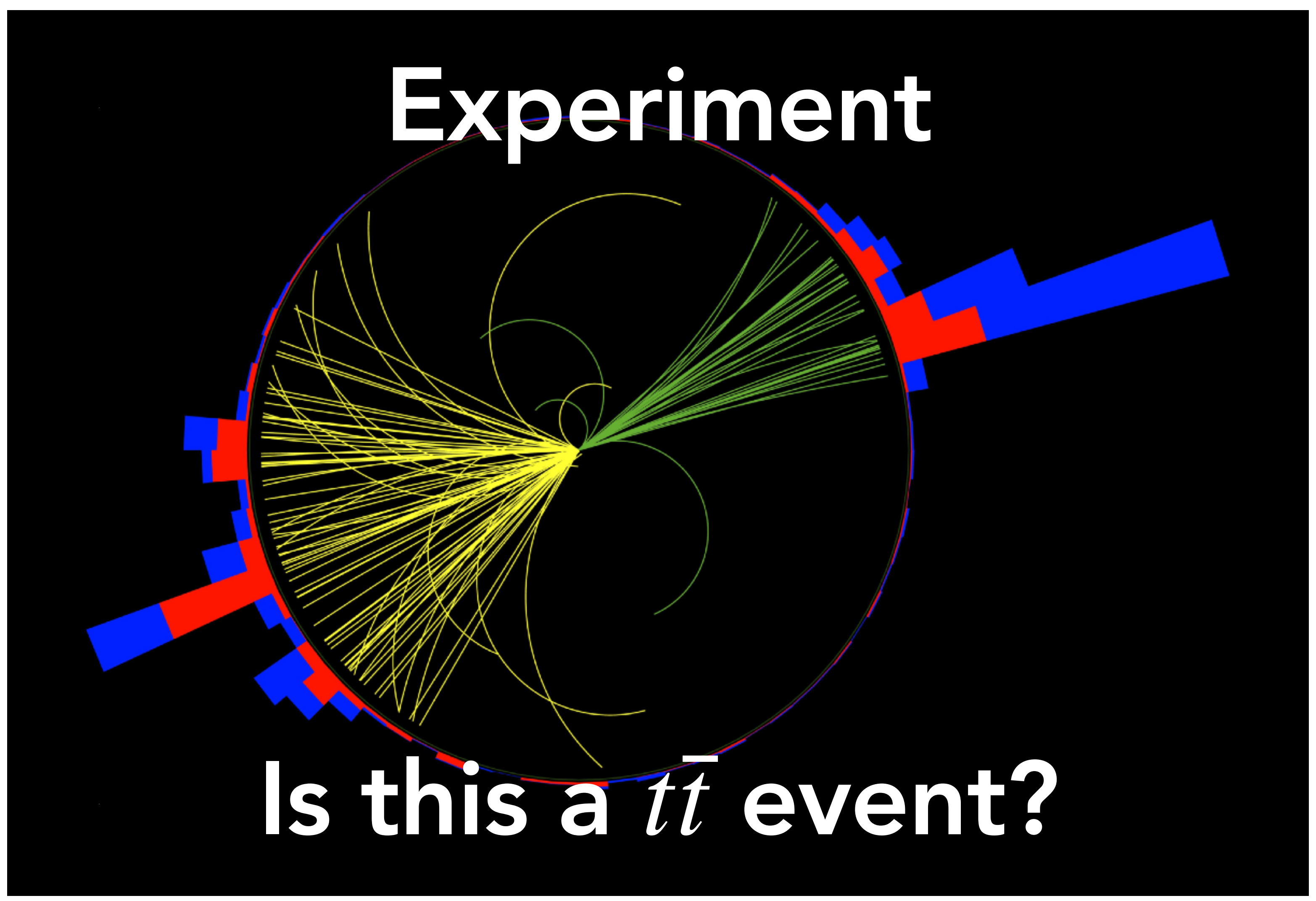
Example: top quarks



"Predict" probability distributions differentially in relevant phase spaces

➡ **This is a $t\bar{t}$ "event"**

Can we measure this?



Complex final states, backgrounds; efficiencies, calibrations, …
➡ Measured result + uncertainties

Model Building:

$$\mathscr{L}_{SM} \quad (+ \ \mathscr{L}_{BSM} \ ? \ )$$

Fundamental parameters

Fields, Symmetries, Couplings, Masses, …

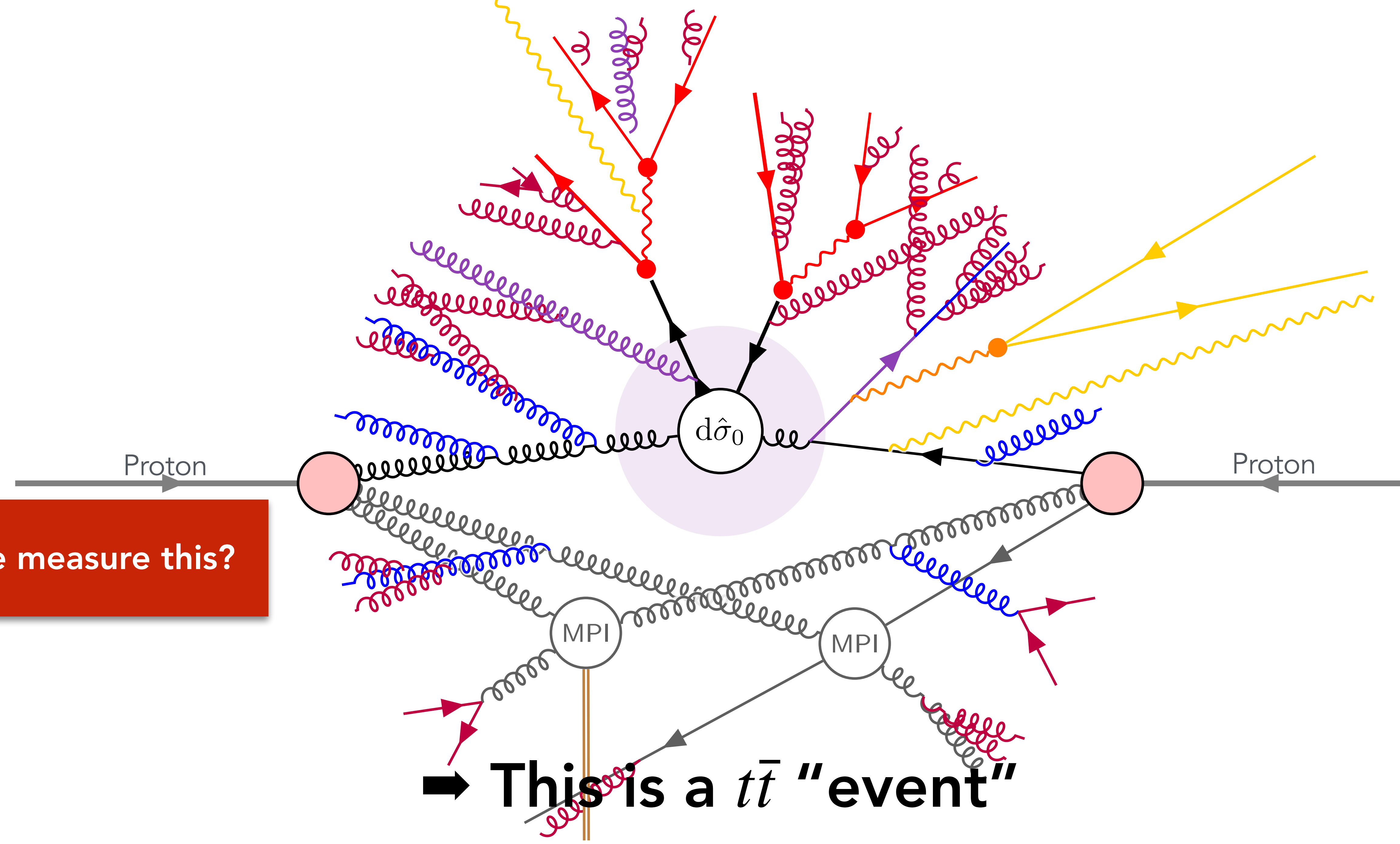
Can we measure this?

Phenomenology:

Compute observables:

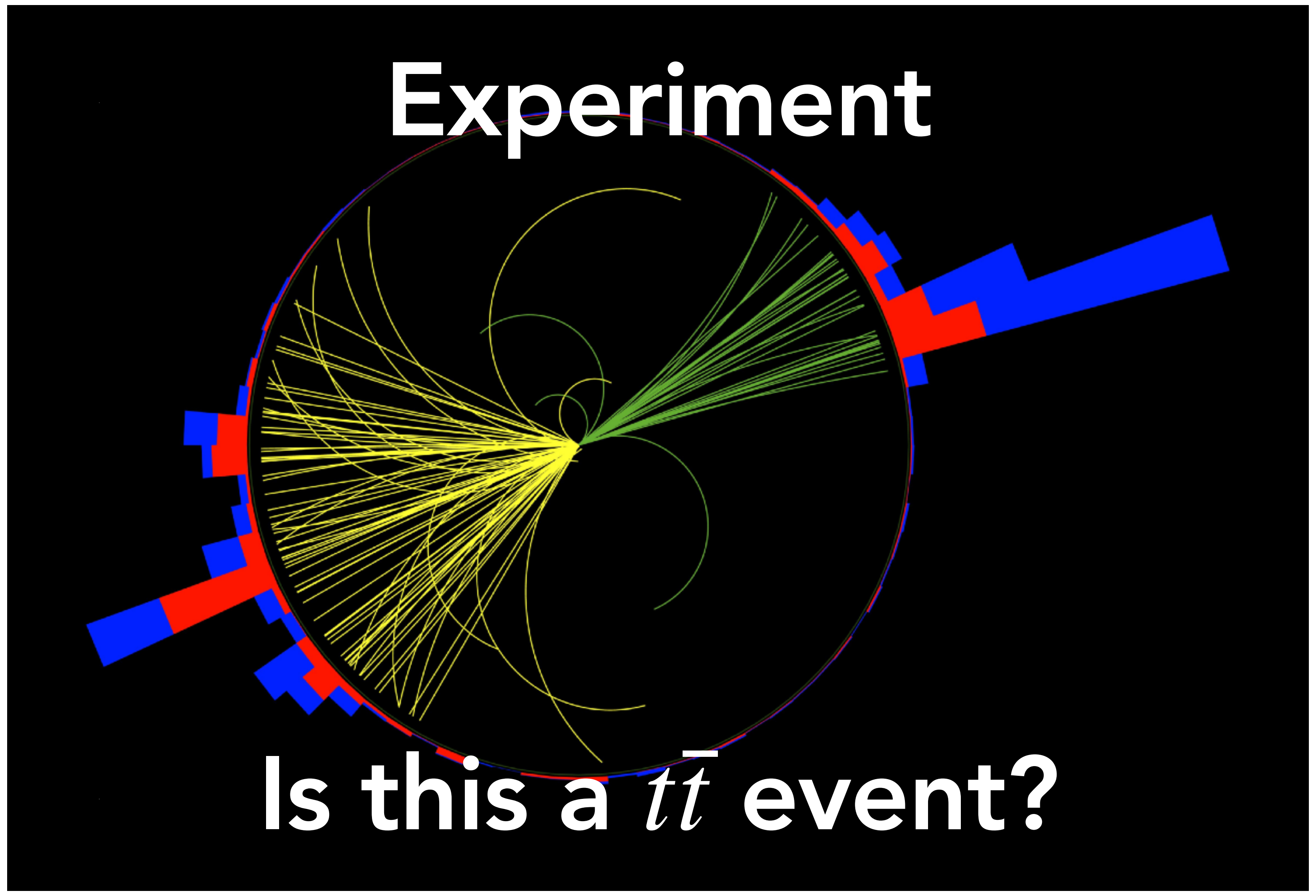
$$\mathrm{d}\sigma_{AB\to X_1\ldots X_n}, \ \mathrm{d}\Gamma_{A\to X_1\ldots X_n}, \ \ldots$$

Example: top quarks



➡ **This is a $t\bar{t}$ "event"**

Can we measure this?



Complex final states, backgrounds; efficiencies, calibrations, …

➡ Measured result + uncertainties

Adding Detail: Confinement

Model Building:

$$\mathscr{L}_{SM} \quad \left(+ \ \mathscr{L}_{BSM} \, ? \right)$$

Fundamental parameters

Fields, Symmetries, Couplings, Masses, …

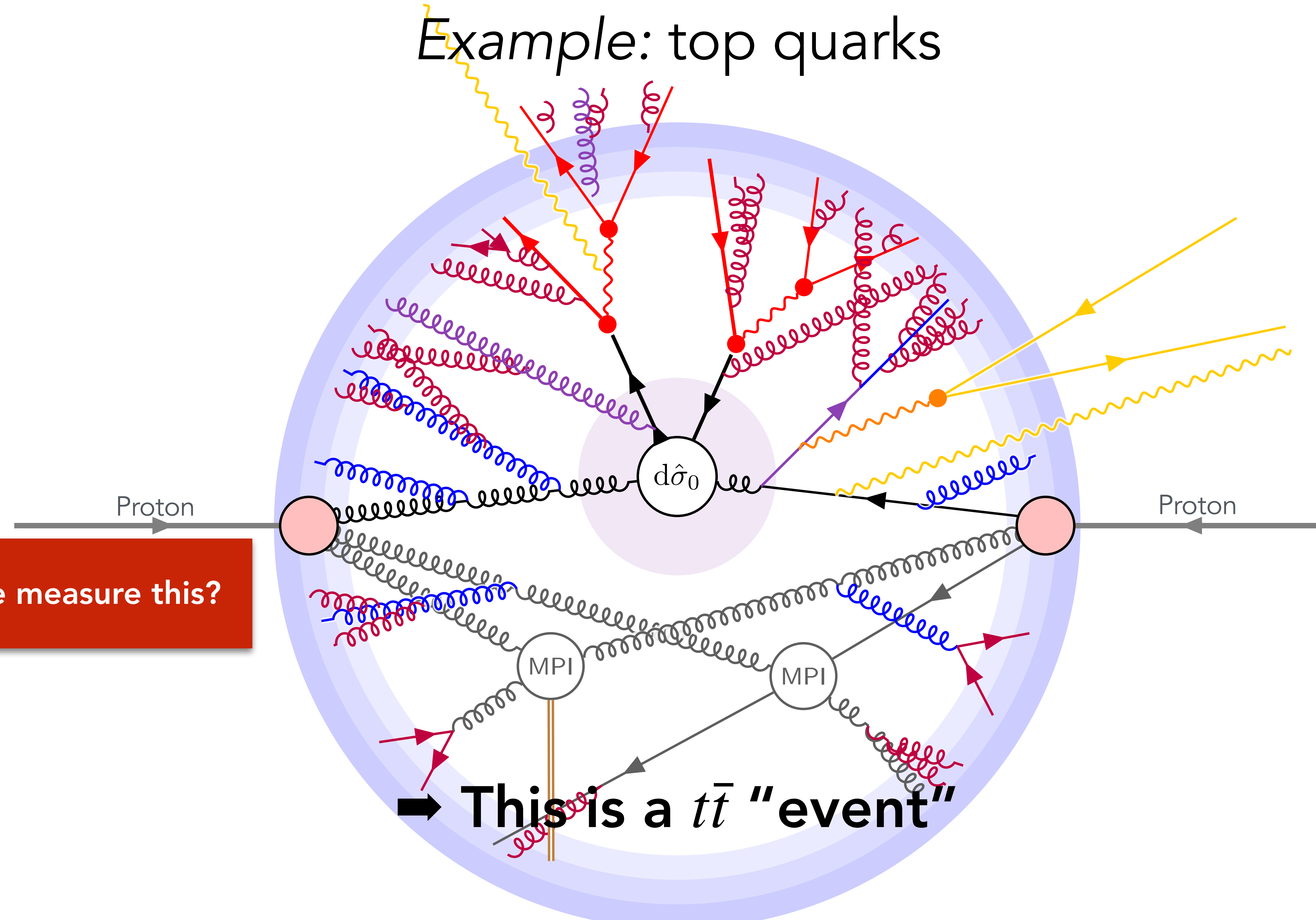
Can we measure this?

Phenomenology:

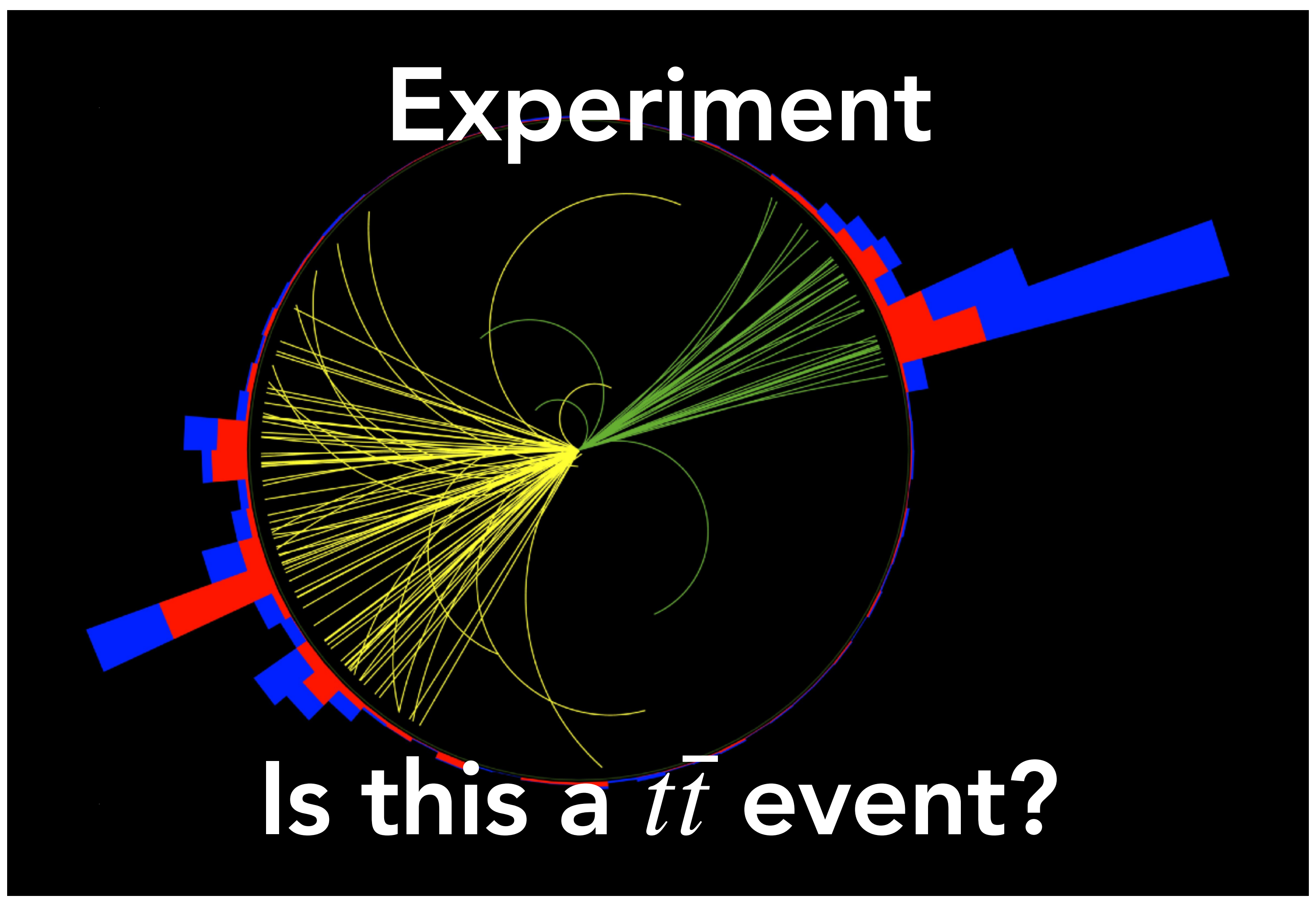
Compute observables:

$$d\sigma_{AB\to X_1\ldots X_n}, \quad d\Gamma_{A\to X_1\ldots X_n}, \ \ldots$$

Example: top quarks



Can we measure this?



Complex final states, backgrounds; efficiencies, calibrations, …

➡ Measured result + uncertainties

Adding Detail: **Strings** (= our model of confinement)

Model Building:

$$\mathscr{L}_{SM} \quad (+ \; \mathscr{L}_{BSM} \, ? \,)$$

Fundamental parameters

Fields, Symmetries, Couplings, Masses, …

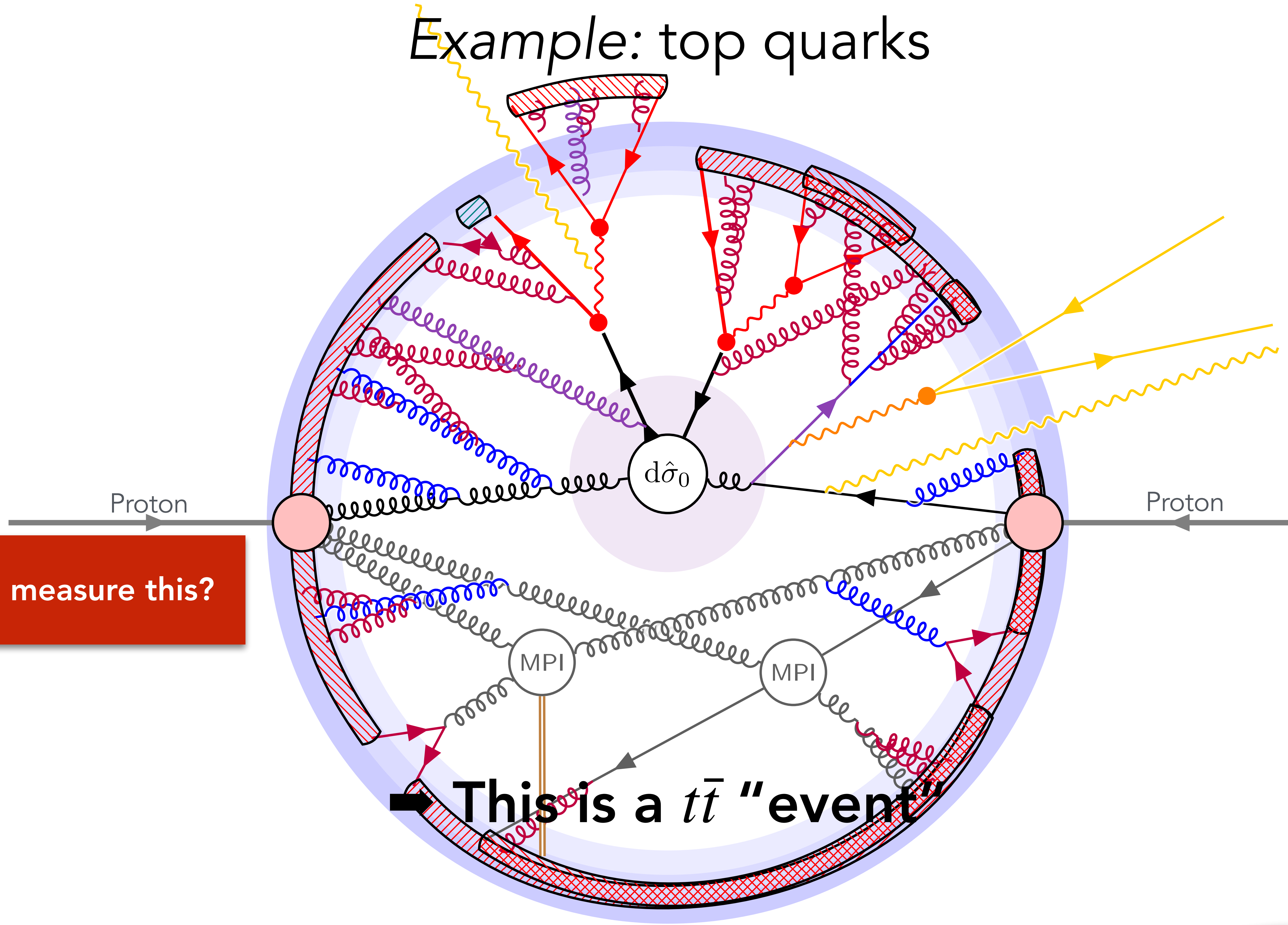
Can we measure this?

Phenomenology:

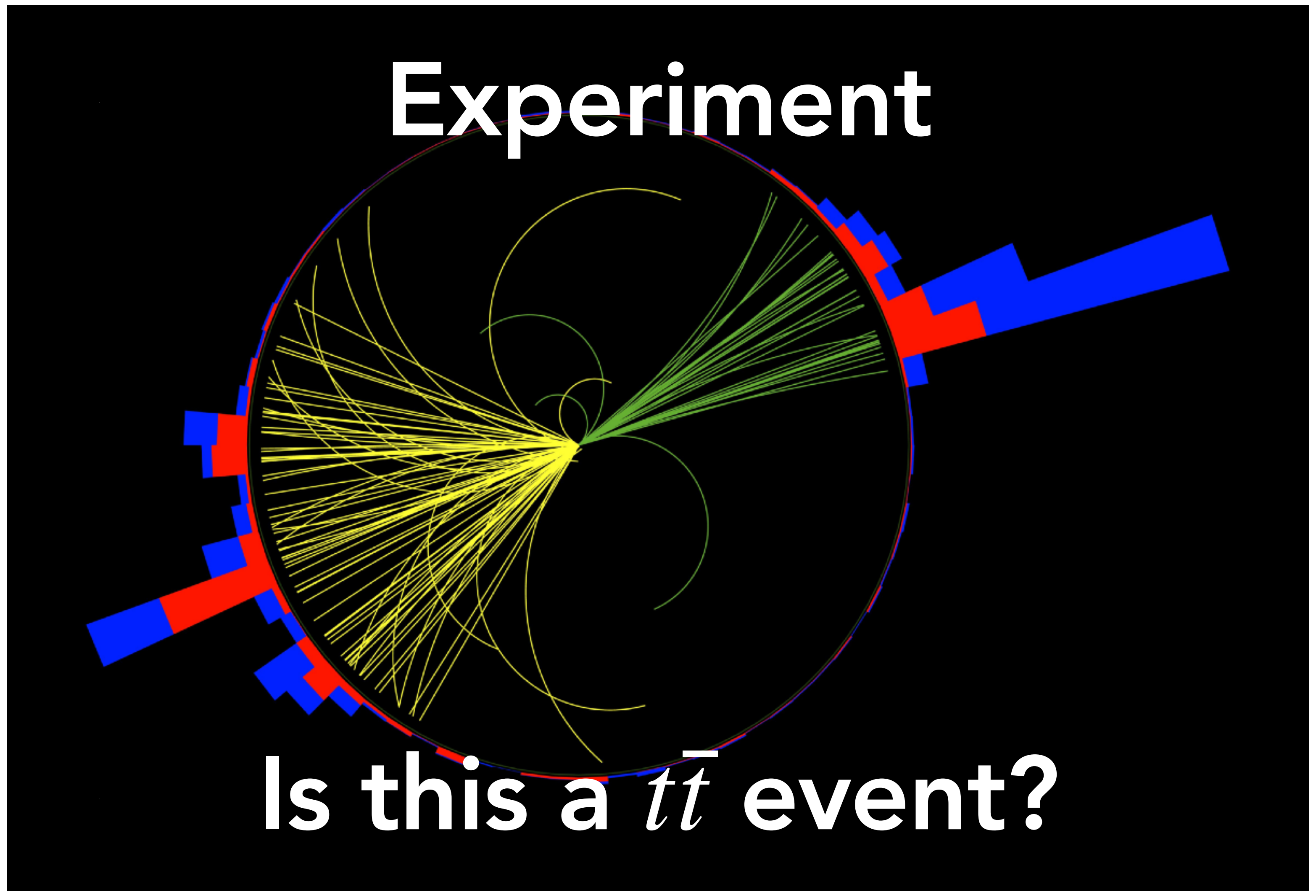
Compute observables:

$$\mathrm{d}\sigma_{AB \to X_1 \dots X_n} \, , \; \mathrm{d}\Gamma_{A \to X_1 \dots X_n} \, , \; \dots$$

Example: top quarks



Can we measure this?



Complex final states, backgrounds; efficiencies, calibrations, …

➡ Measured result + uncertainties

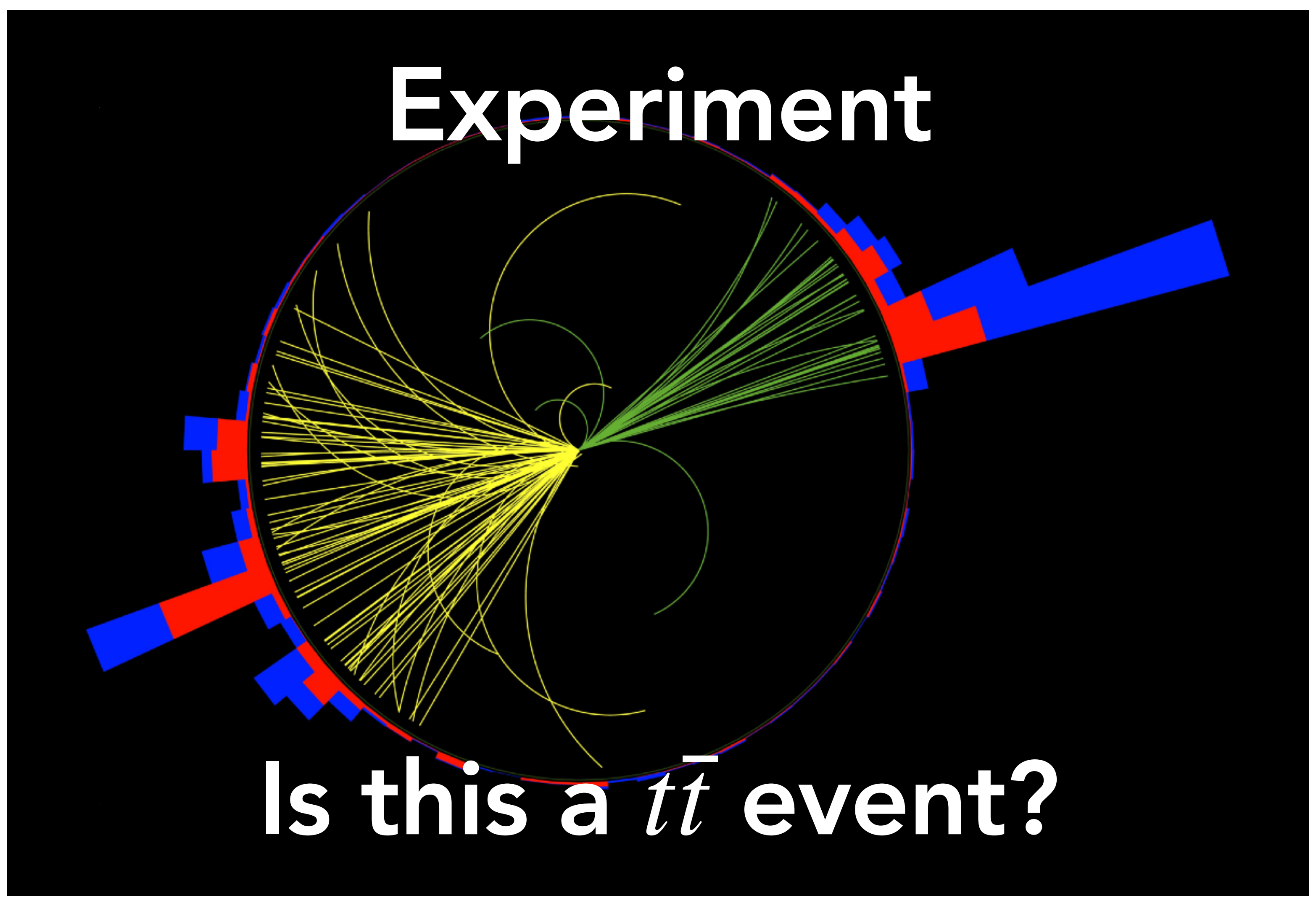
Adding Detail: Hadrons + Decays

Model Building:

$$\mathscr{L}_{SM} \quad (+ \ \mathscr{L}_{BSM} \, ? \,)$$

Fundamental parameters

Fields, Symmetries, Couplings, Masses, …



Complex final states, backgrounds; efficiencies, calibrations, …

➡ Measured result + uncertainties

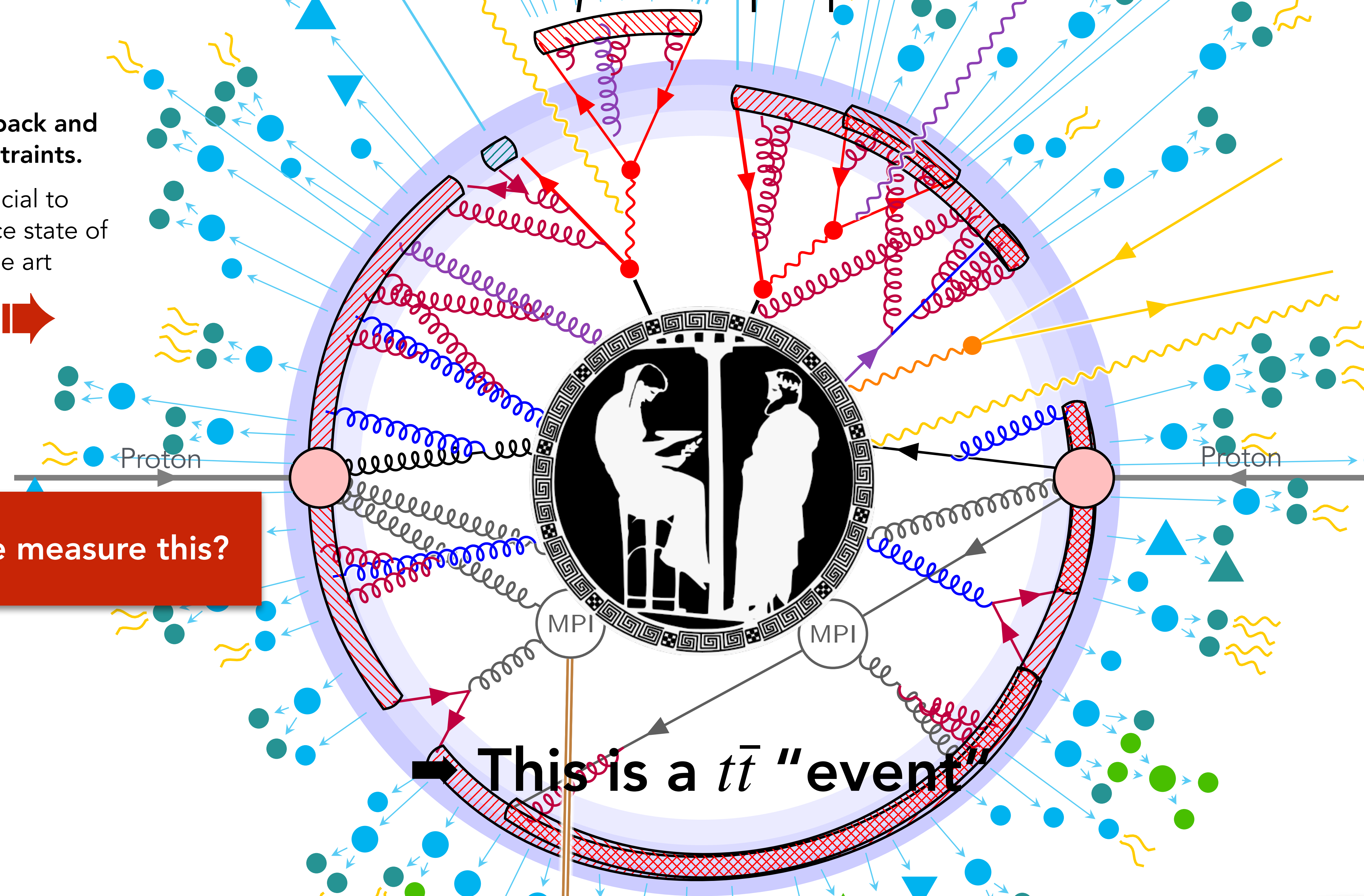
Can we measure this?

Phenomenology:

Compute observables:

$$d\sigma_{AB\to X_1\ldots X_n}, \quad d\Gamma_{A\to X_1\ldots X_n}, \ \ldots$$

Example: top quarks



Feedback and constraints.

Crucial to advance state of the art

Can we measure this?

(+ Detector Simulation, not touched on here)

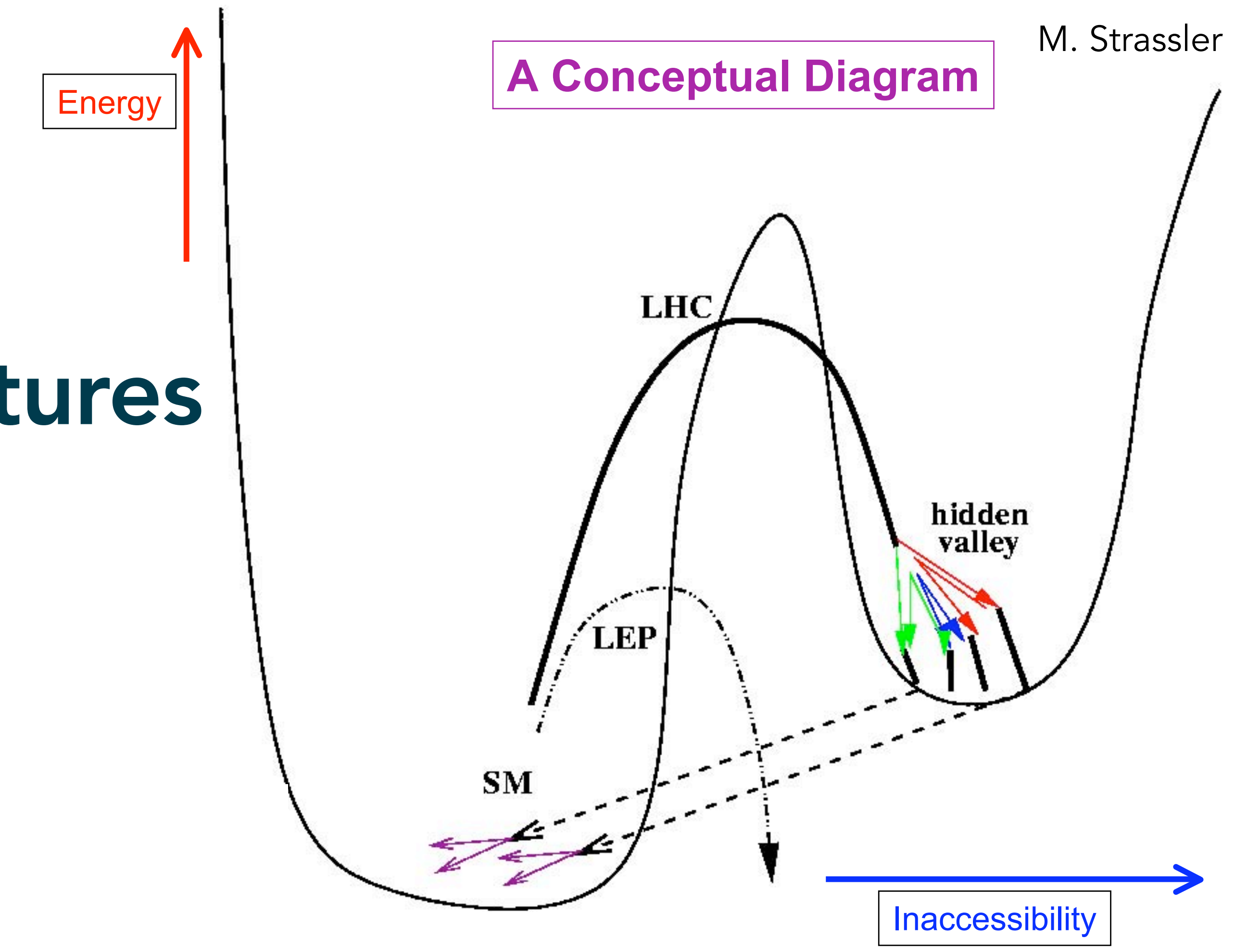
New Physics & Dark Matter 1: Hidden-Valley Scenarios

Hidden Valley ~ Dark Sector ~ Secluded Sector

New sector decoupled from SM at low energy

General framework in PYTHIA ➤ **study variety of exp signatures**

- **Abelian U(1)** (broken or unbroken $\Longrightarrow$ massive or massless γ_V)
- Or **Non-Abelian SU(N)** (unbroken, $N^2 - 1$ massless g_V)
- Key parameters: # of colours, # of valley-quarks q_V, masses



Three alternative production mechanisms in PYTHIA 8:

1. massive Z': $\mathrm{q}\overline{\mathrm{q}} \to \mathrm{Z}' \to \mathrm{q}_v\overline{\mathrm{q}}_v$,
2. kinetic mixing: $\mathrm{q}\overline{\mathrm{q}} \to \gamma \to \gamma_v \to \mathrm{q}_v\overline{\mathrm{q}}_v$,
3. massive F_v charged under both SM and hidden group, so e.g. $\mathrm{gg} \to \mathrm{F}_v\overline{\mathrm{F}}_v$. Subsequent decay $\mathrm{F}_v \to \mathrm{f}\mathrm{q}_v$. F_v spin either 0, 1/2 or 1 and matching q_v either 1/2 or 0.
4. (No Higgs portal, but doable. Qualitatively similar to Z'.)

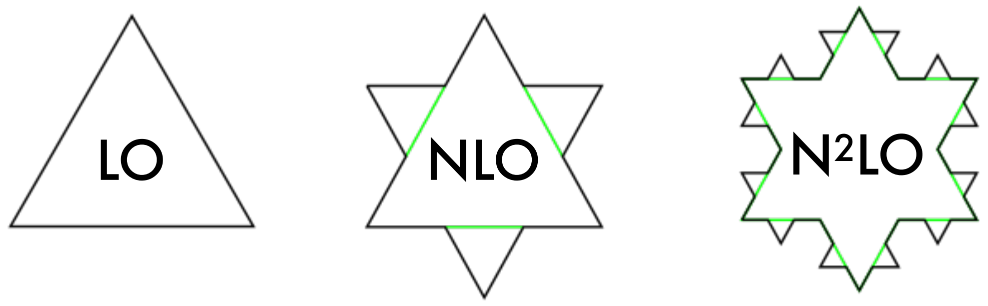
PYTHIA HTML Manual:

▼ Process Selection

Standard-Model Parameters
Couplings and Scales
Hard Processes
-- Hard QCD
-- Electroweak
-- Onia
-- Top
-- Fourth Generation
-- Higgs
-- SUSY
-- New Gauge Bosons
-- Left-Right Symmetry
-- Leptoquark
-- Compositeness
▶ Hidden Valleys
-- Extra Dimensions
-- Dark Matter

Based on Carloni & Sjöstrand 1006.2911, Carloni, Rathsman, Sjöstrand 1102.3795

Dark Showers



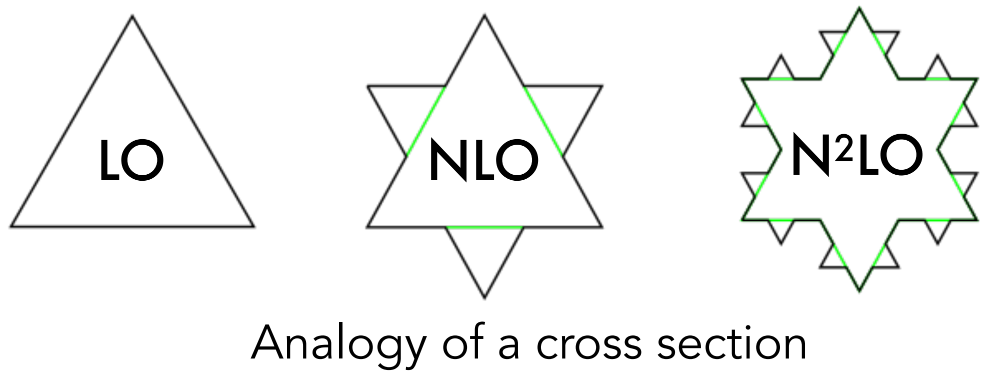
Analogy of a cross section

Accelerated Gauge Charges → Bremsstrahlung

In QFT: driven by propagator denominators (IR singularities)

- Pole structure is universal, same for QED/QCD/…
- Numerators depend on spins of radiators and of emitted quanta

All-orders quasi-fractal structure captured by Parton Showers

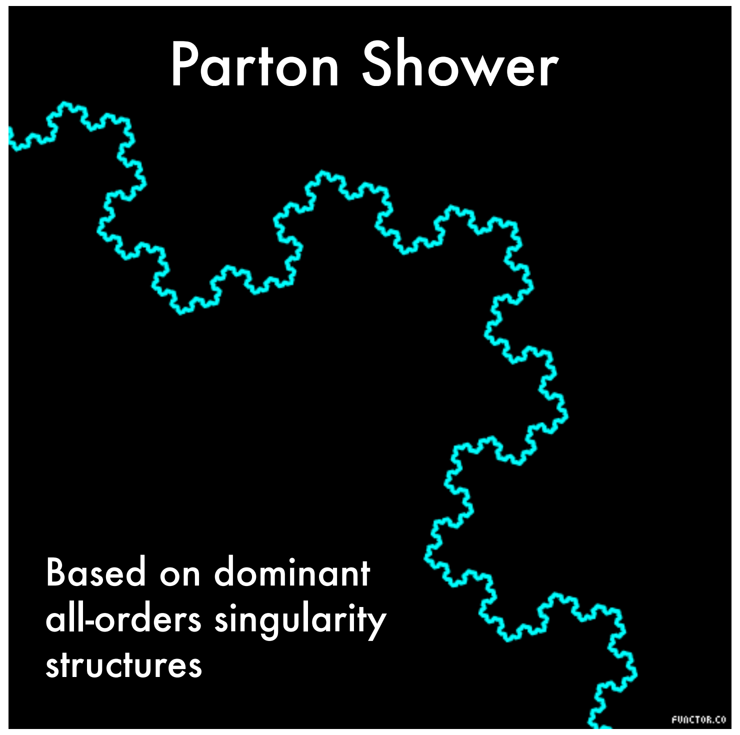


Accelerated Gauge Charges → Bremsstrahlung

In QFT: driven by propagator denominators (IR singularities)

- Pole structure is universal, same for QED/QCD/…
- Numerators depend on spins of radiators and of emitted quanta

All-orders quasi-fractal structure captured by Parton Showers



Valley-gluons/photons have IR (soft/collinear) singularities

Collinear limits (emission parallel to radiator): DGLAP kernels

Soft limits (emission of low-energy quanta): soft-eikonal factors

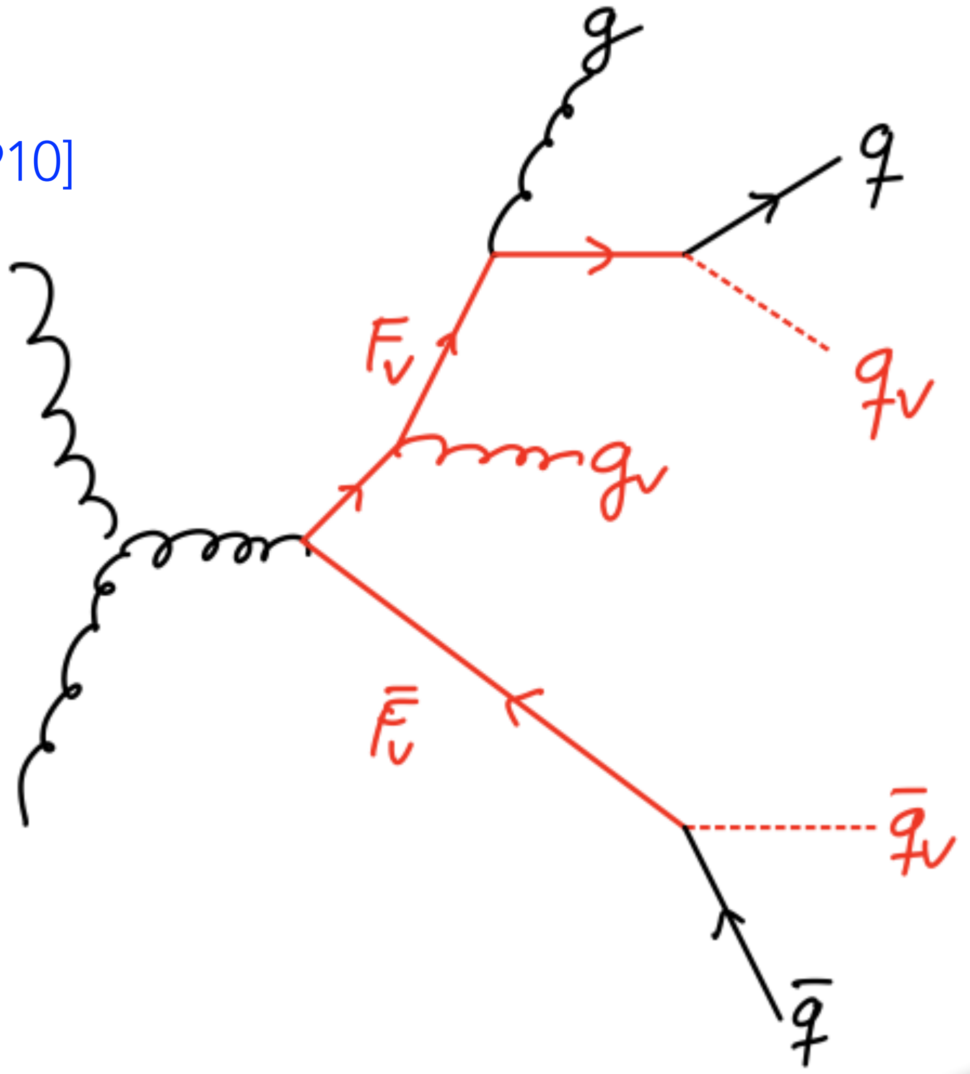
For massive radiators: collinear ones dampened [eg **ALICE**, *Nature* 605 (2022) 7910]

For broken U(1) $\Longrightarrow$ massive $\gamma_V \Longrightarrow$ soft ones dampened

Pythia: a single interleaved SM ⊕ HV shower evolution

Invisible sector emissions ➡ **recoil effects** in visible sector

Carloni & Sjöstrand 1006.2911



Note: also nice poster (& papers!) by E. Bernreuther; **Also:** Snowmass Summary on Dark Showers: 2203.09503

(Continued…): Dark-Sector **Hadronization + Decays**

Dark-sector particles may remain invisible, or:

Carloni & Sjöstrand 1006.2911, Carloni, Rathsman, Sjöstrand 1102.3795

Broken U(1) $\Longrightarrow$ Radiated γ_V's decay back, $\gamma_V \to \gamma \to f\bar{f}$

BRs as photon ➡ lepton pairs!

SU(N) $\Longrightarrow$ **full hidden-sector string fragmentation**

Up to 8 different q_V

Many different valley mesons & baryons (PYTHIA 8.307) NEW

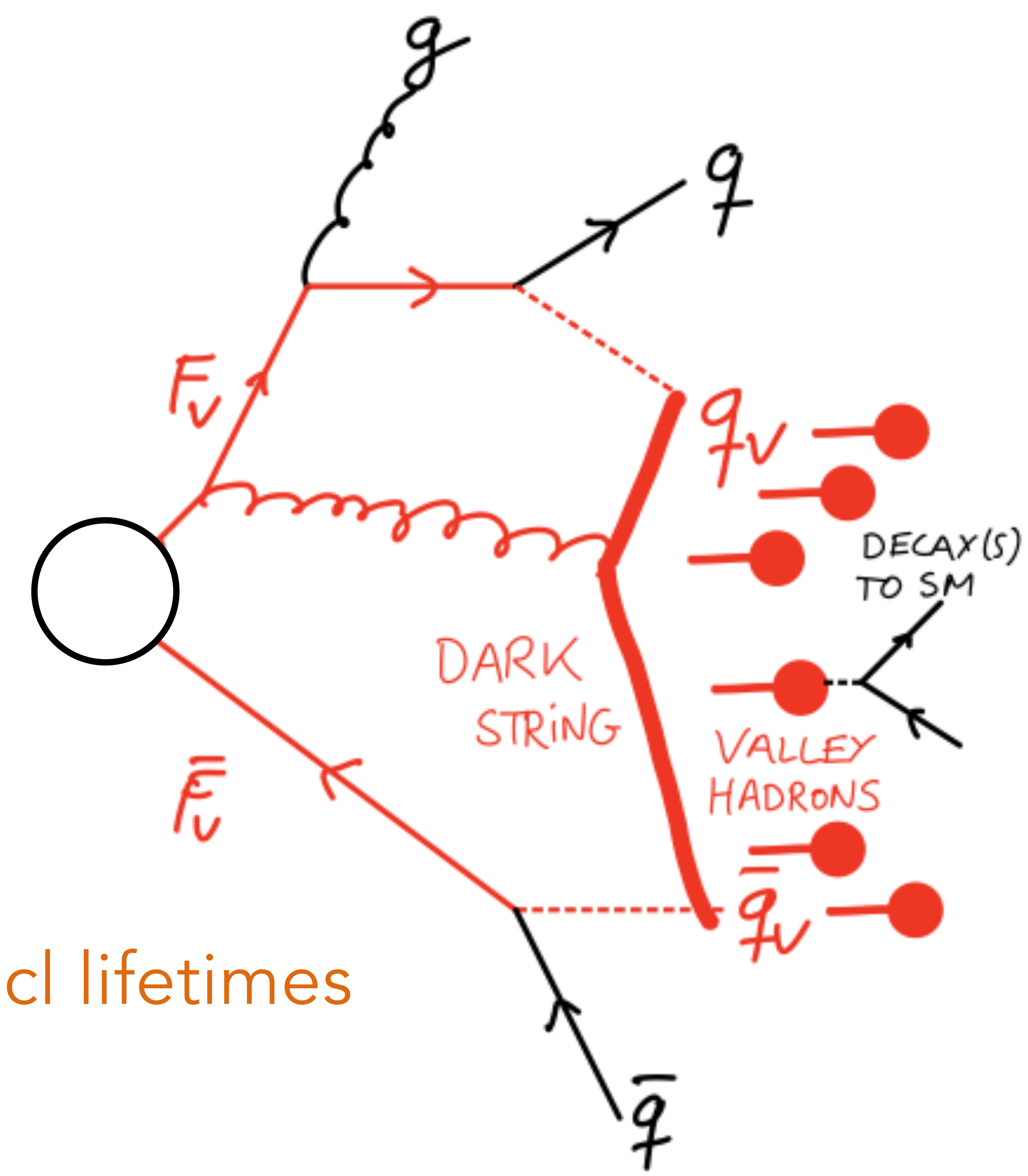
(Baryons mainly for N=3; some applicability for N≥4; so far no dedicated model for N=2)

Flavour-diagonal valley mesons can decay back to SM

(Here assuming mediator(s) conserve "valley flavour")

Fraction = $1/n_{q_V}$

Can set masses of different flavour mesons + decay tables incl lifetimes



↳ **Displaced vertices (LLPs)**, by adjusting lifetimes

➤ Unusual signatures: displaced leptons; semi-visible jets, emerging jets, … (e.g., Faucet et al., 2208.10062) (e.g., Schwaller et al., 1502.05409)

***(Also Note:** ∃ experimental **repo for HNL production** using τ decays in Pythia:)*

https://gitlab.com/hnls/pythia – Main Contact: Phil Ilten

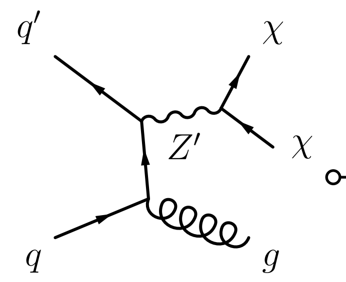
General DM Capabilities of PYTHIA 1: Production at Colliders

Many existing processes/models can be used straight away

General 2HDM model (couplings between scalars and SM can be set by hand)

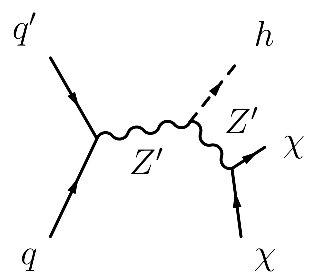
Recycle SUSY machinery for colour/EW charged mediators ("t-channel" models)

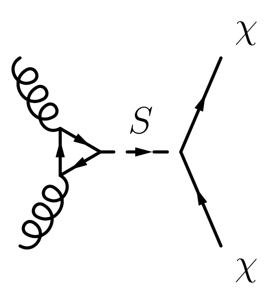
Set decay by hand: e.g. $\tilde{\ell} \to \ell\tilde{\chi}$ (for fermionic DM), $W + \tilde{\nu}$ (scalar DM), etc.



New resonances W', Z'; can decay into new particles (not just SM)

Set of Generic Dark Matter Processes





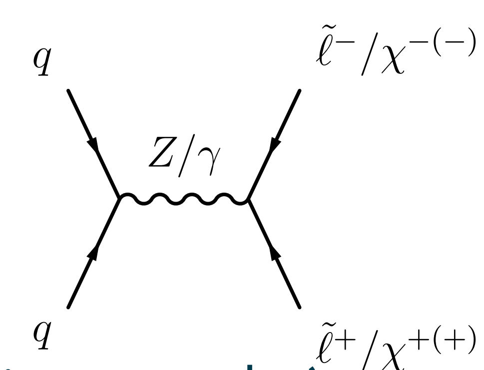
→ (Axial-) vector resonance (+ jet) includes kinetic mixing

→ Associated production of Z' with SM Higgs (for mono-Higgs)

→ (Pseudo-) scalar resonance (+ jet)

Also: Drell-Yan production of new mediators $f\bar{f} \to Z/\gamma^* \to \chi\chi$

E.g. SU(2) N-plet fermions or scalar with U(1) charge.



(Prompt or **long-lived**: displaced vertex; displaced leptons; disappearing tracks)

New particles, decays, hard processes can be added in user code

Using PYTHIA's **ParticleData** scheme and/or SLHA scheme

Masses and lifetimes of mediators can be set by user (e.g., **long-lived signatures**)

Semi-internal processes & resonance decays inheriting from existing ones

Or: **Les Houches Event Files** (LHEF, e.g., from MadGraph)

PYTHIA HTML Manual:

▼ Process Selection

Standard-Model Parameters
Couplings and Scales
Hard Processes
-- Hard QCD
-- Electroweak
-- Onia
-- Top
-- Fourth Generation
-- Higgs
-- SUSY
-- New Gauge Bosons
-- Left-Right Symmetry
-- Leptoquark
-- Compositeness
-- Hidden Valleys
-- Extra Dimensions
-- Dark Matter

▼ Interfaces to External Hard Processes

Les Houches Accord (LHA)
Les Houches Event Files
Semi-Internal Processes
MadGraph5 Processes
HelacOnia Processes
Alpgen Event Interface

▶ Parton Showers

▶ Matching and Merging

▼ Particles & Decays

Particle Data
Mass-dependent Hadron Widths
Resonance Decays
Hadron & Tau Decays
Semi-Internal Resonances
External Decays

Based on N. Desai, *Collider signatures for dark matter and long-lived particles with Pythia 8*, 1807.04240

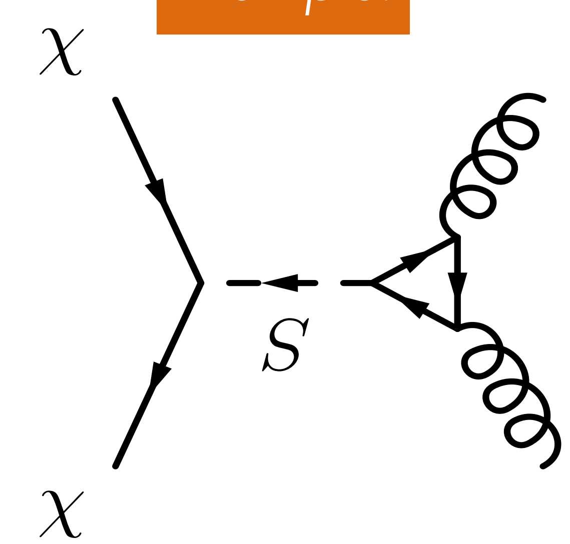
Dark Matter Physics 2: **Annihilation Spectra**

Annihilation spectra for indirect-detection experiments

`main07.cc`: dummy production process + user-defined decay table

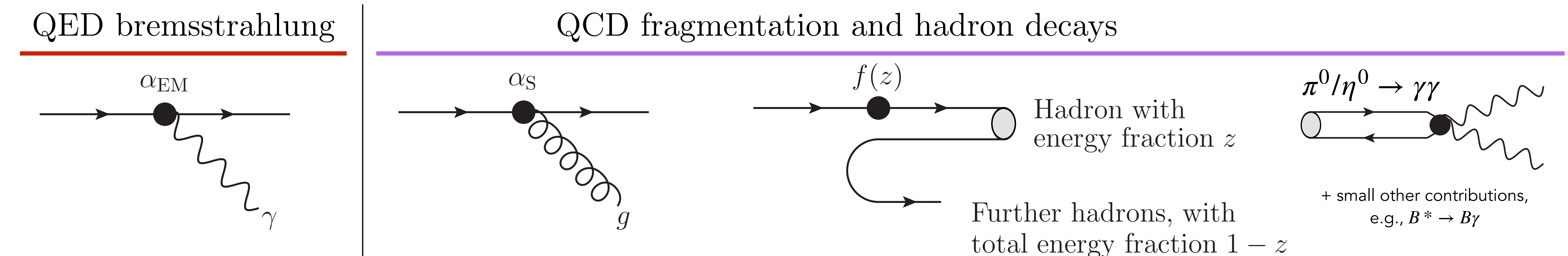
For example: spectrum for DM + DM → $g\,g$ obtained by adding dummy resonance with mass $2m_{\rm DM}$ decaying into two gluons.

Example:



PYTHIA: full fragmentation modelling

For example, for gamma-ray spectra from annihilation to final states with quarks:



Dominates at high x_γ	Photons from $\pi^0 \to \gamma\gamma$ dominate bulk (and peak) of spectra

Main/best constraints: $\pi^\pm$ spectra in e^+e^- annihilation (by sospin); + a few more: mainly **LEP**

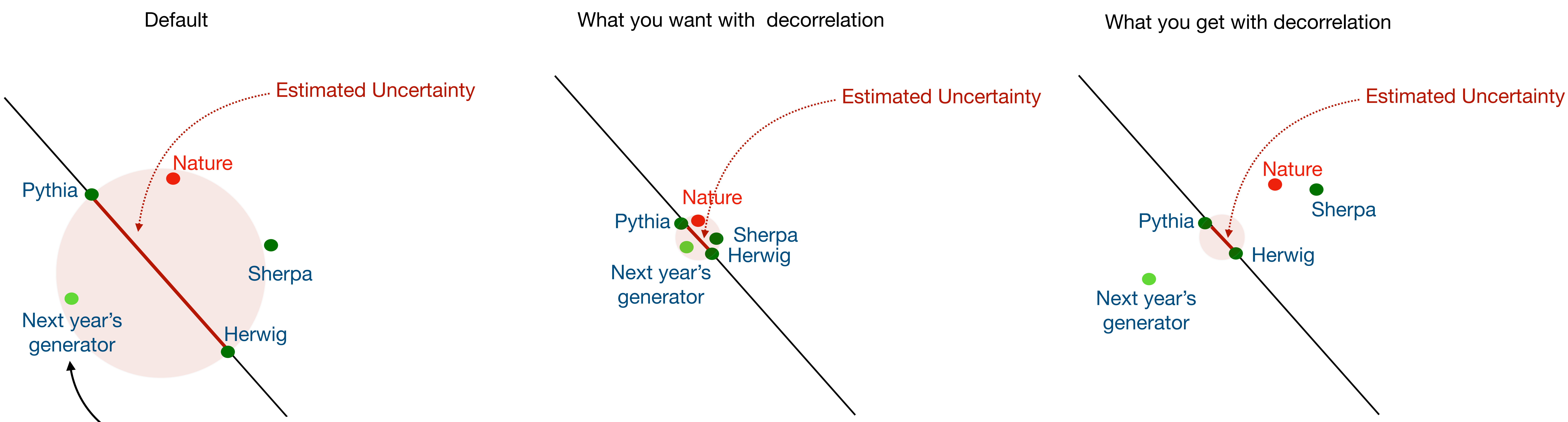
(+ you also want to know sources & ranges of QED bremsstrahlung in the modelling)

See, e.g., PS et al., "The Monash Tune", 1404.5630, and S. Amoroso et al., 1812.07424

Uncertainties (e.g., on DM Annihilation Spectra)

Borrowed slide from A Ghosh, Tuesday ↔ Machine Learning of/for Theory Models

EPJC:s10052.022.10012.w: **Aishik Ghosh**, Benjamin Nachman



*Personal Comment: I would kind of hope next year's generator would be **closer** to Nature, not further from it…*

Instruction to ML: "Please shrink Pythia vs Herwig difference"

Model will learn to fool you !

ML methods don't often generalise the way you would hope

17

Another Aspect of the Problem: Pythia, Herwig, Sherpa all tuned to ~ same data

No guarantee that they span the experimental uncertainties (similar issue as of old with PDFs)

Based on A. Jueid et al., 1812.07424 (gamma rays, *eg for GCE*) and 2202.11546 (antiprotons, *eg for AMS*) + **last week: 2303.11363 (all)**

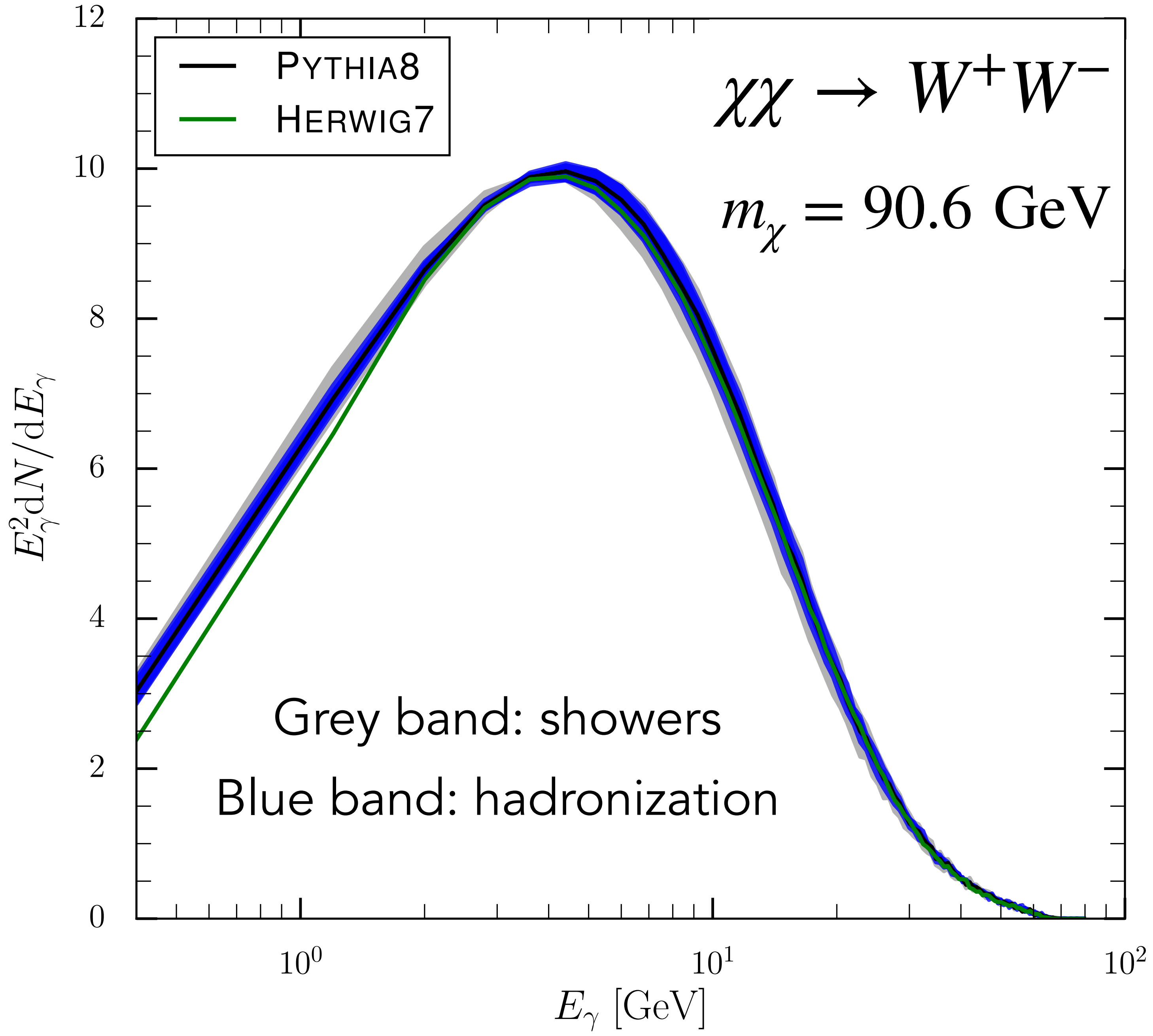
QCD uncertainties on Dark-Matter Annihilation Spectra

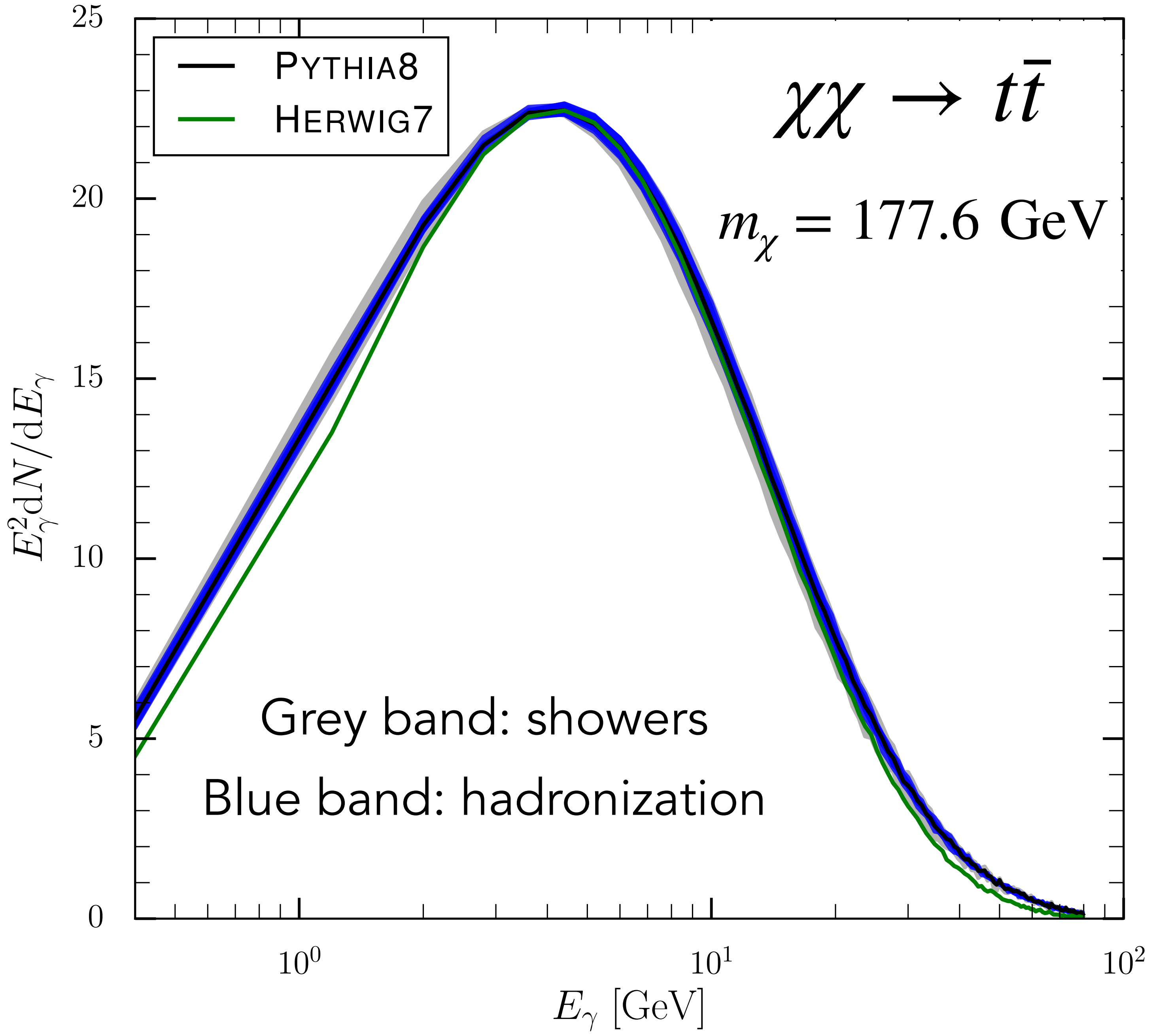
Compare different generators? **Problem:** all tuned to ~ same data

No guarantee that they span the experimental uncertainties (similar issue as of old with PDFs)

Instead, did **parametric refittings** of constraining data within PYTHIA's modelling

Photon spectra





Same done for antiprotons, positrons, antineutrinos

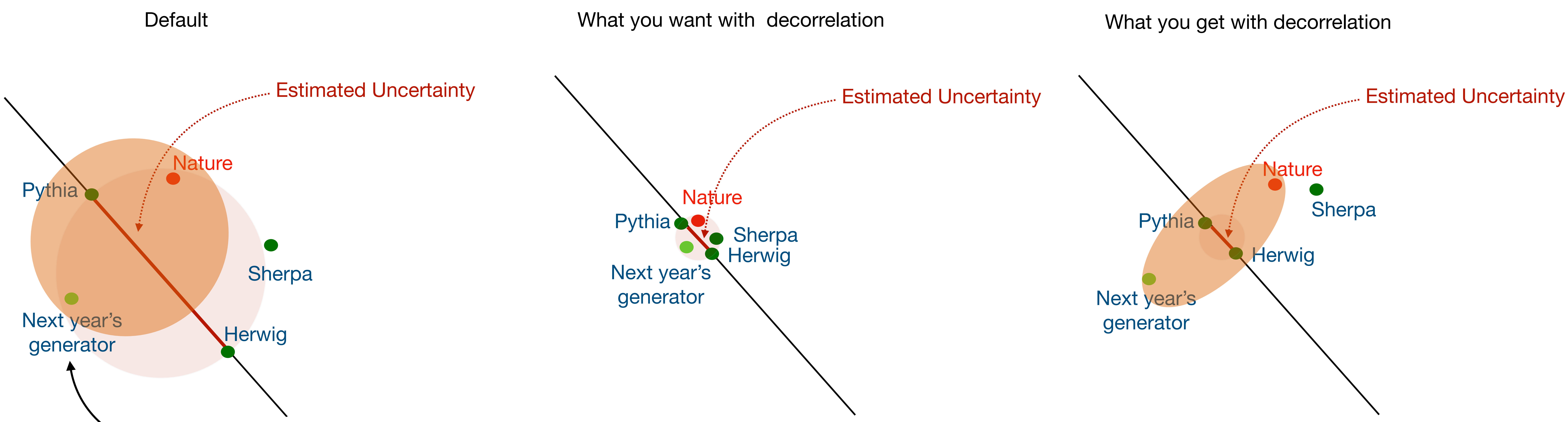
Tables with uncertainties available on request. Also the spanning tune parameters of course.

Main Contact: adil.jueid@gmail.com

Uncertainties (e.g., on DM Annihilation Spectra)

Borrowed slide from A Ghosh, Tuesday ↔ Machine Learning of/for Theory Models

EPJC:s10052.022.10012.w: **Aishik Ghosh**, Benjamin Nachman



Personal Comment: I would kind of hope next year's generator would be ***closer*** *to Nature, not further from it…*

Instruction to ML: "Please shrink Pythia vs Herwig difference" Model will learn to fool you !

ML methods don't often generalise the way you would hope

17

Another Aspect of the Problem: Pythia, Herwig, Sherpa all tuned to ~ same data

No guarantee that they span the experimental uncertainties (similar issue as of old with PDFs)

Cosmic-Ray Air Showers

Single incident particle ➡ billions of final-state particles (forget about GEANT).
Recently started a collaboration with CORSIKA 8 fast/optimised air-shower tracker

New: PythiaCR [Based on Sjöstrand + Utheim, **2005.05658** & **2108.03481**]

Provide **hadron-air cross sections** ⊕ **perform collisions** ⊕ simulate **hadron decays**

(Air ~ ^{14}N + ^{16}O; currently also ^{40}Ar, ^{208}Pb; few hours of manual labour to add more)

Cosmic-ray "beams" are **heterogenous** and **not mono-energetic**:

Achieved by initialising multiple beams in energy grids + rapid beam switching

CR (re-)interactions "fixed-target"; can probe **low CM energies** (by HEP standards)

Standard (collider) Pythia only applies for $\sqrt{s} > 10\,\mathrm{GeV}$

New **extensive** **low-energy (re)interaction model**s

➡ **Arbitrary hadron-hadron collisions at low E, and arbitrary hadron-p/n at any energy)**

Extend to **hadron-nucleus** using nuclear-geometry part of ANGANTYR

So far limited comparisons with data - interested in feedback

A positive technical note: native C++ simplifies CORSIKA 8 - PYTHIA 8 interfacing

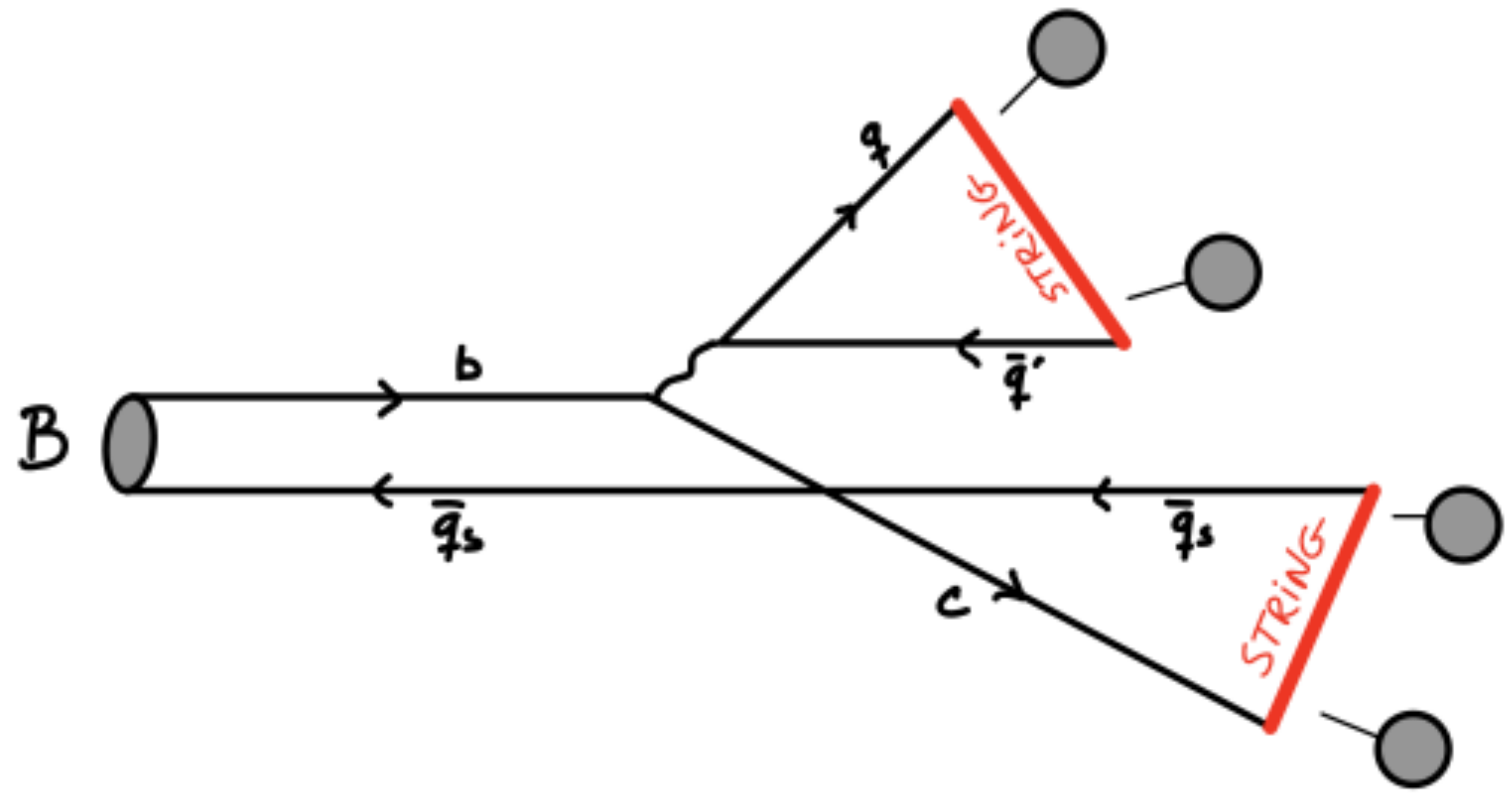
See also M. Reininghaus et al. *Pythia 8 as hadronic interaction model in air shower simulations*, **2303.02792**

Future Research Directions relevant to **Flavour Physics**

1. State of the art for B decays: PYTHIA + EVTGEN (M. Kreps, Warwick U)

PDG far from complete. Also need differential distributions: ME or simply flat phase space

+ **Many** modes modelled via $b \to c \otimes$ string fragmentation



B mesons: 1/3 of all decays

Λ_b baryons: 1/2 of all decays

New grant with Warwick U (LHCb):
"Beautiful Strings"
incl investigating these decays

2. Electromagnetic Corrections (QED FSR) in B hadron decays:

Energetic photons affect mass, q^2 reconstructions *(e.g., $B \to \mu^+\mu^-K^+\pi^- + \gamma$)*

HERWIG and SHERPA have dedicated hadron-level QED (YFS) showers Schonherr & Krauss 0810.5071 Hamilton & Richardson 0603034

For PYTHIA, external PHOTOS lib: Fortran, static variables: multi-threading bottleneck

➡ **VINCIA** has novel (& unique) **multipole QED showers** & **interleaved cascade decays**

PS & R. Verheyen 2002.04939 H. Brooks, PS, R. Verheyen 2108.10786

Come to Australia

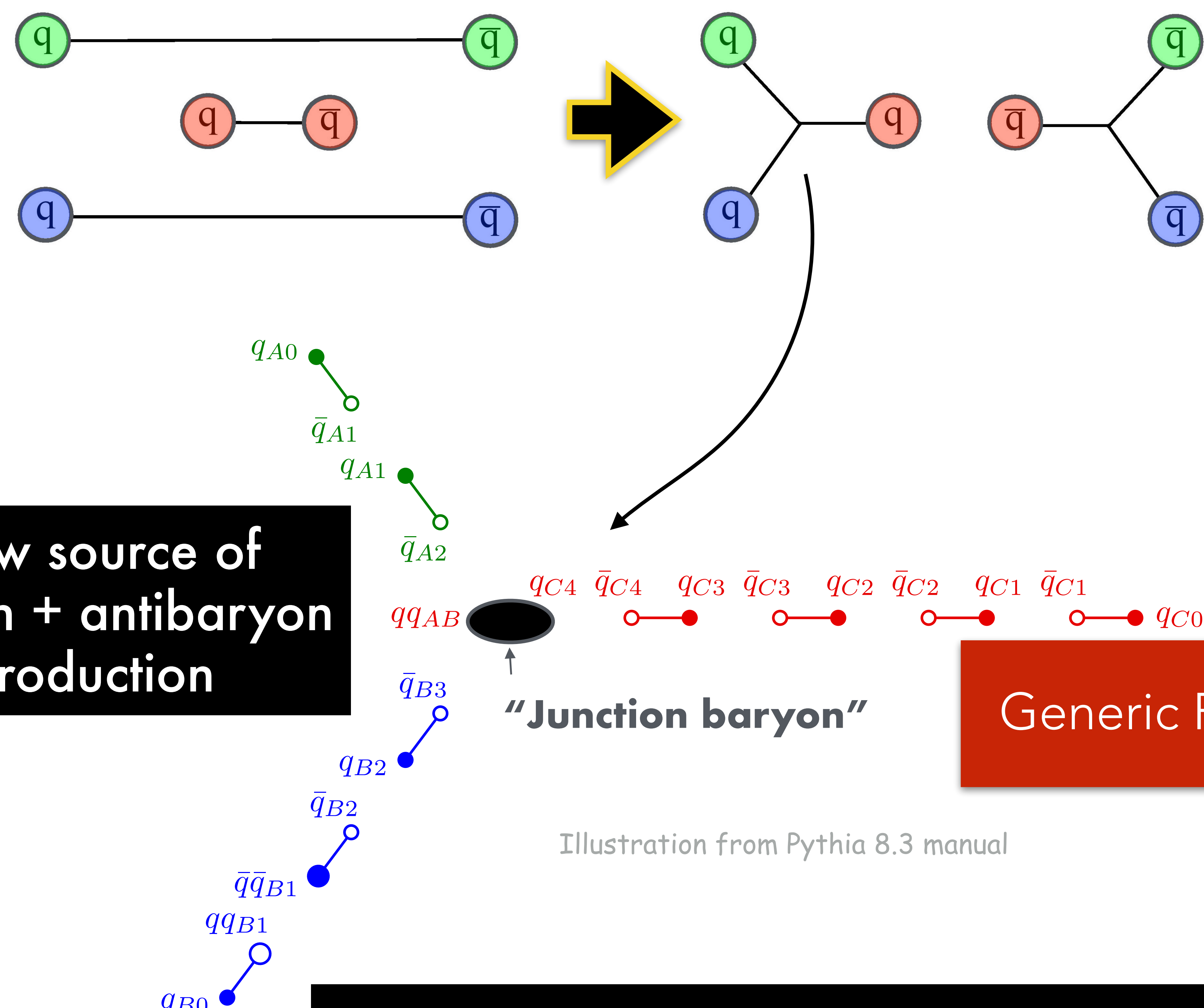
Poster Submissions still Open

Australia's Top 5 Most Dangerous Animals:
1. Horses (7.7 fatalities / yr)
2. Cows (3.3)
3. Dogs (2.7)
4. Kangaroos (1.8)
5. Bees (1.6) (tied with sharks)

Unique feature of SU(3): Y-Shaped 3-String "Junctions" (Low-mass limit = baryons)

Christiansen & PS (2015): Toy model of "Colour reconnections" based on stochastic sampling of SU(3) group probabilities: allows for random (re)connections in colour space, guided by string area law

For example:



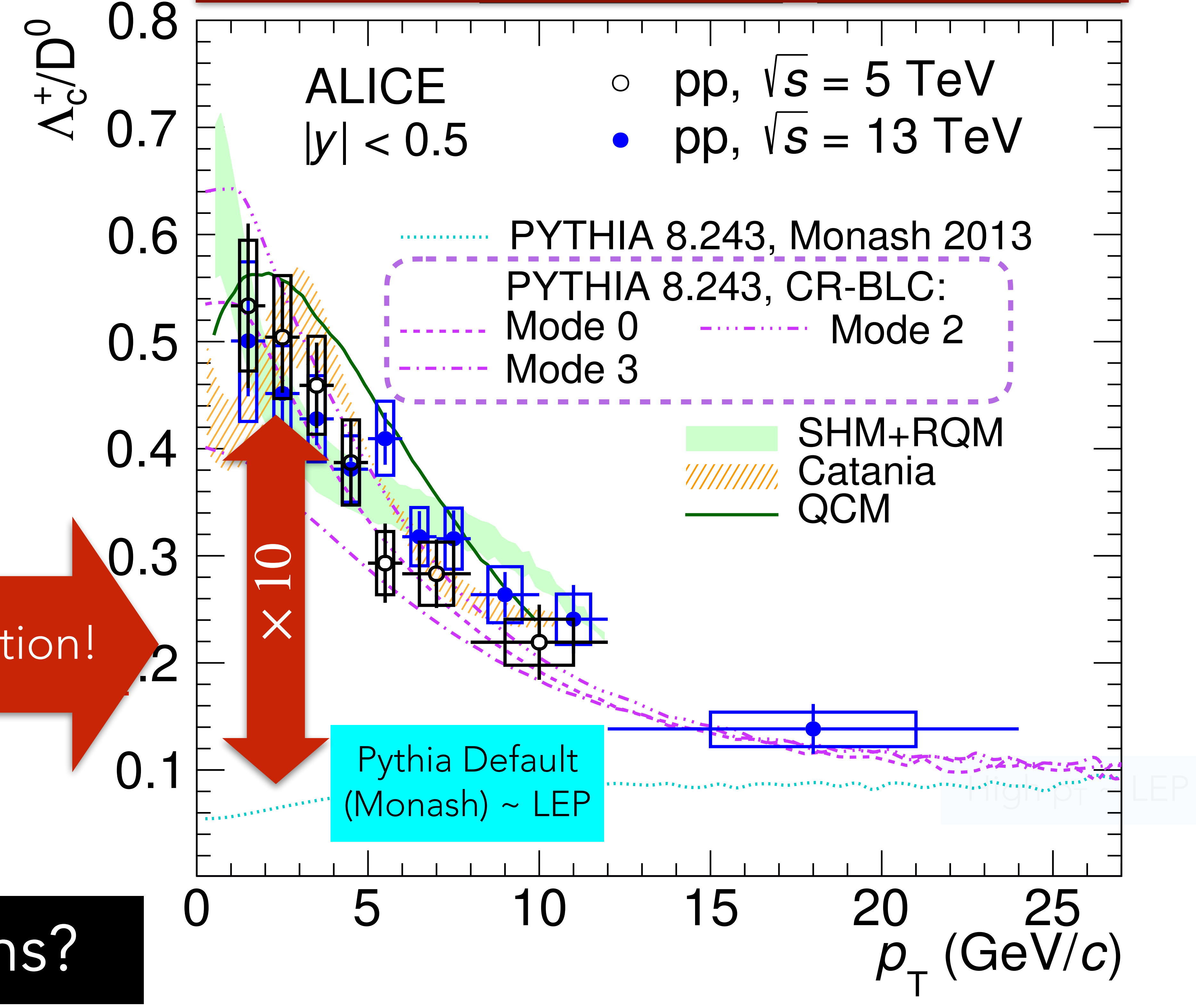
New source of baryon + antibaryon production

Illustration from Pythia 8.3 manual

Cool Discovery! String Junctions?

Generic Prediction!

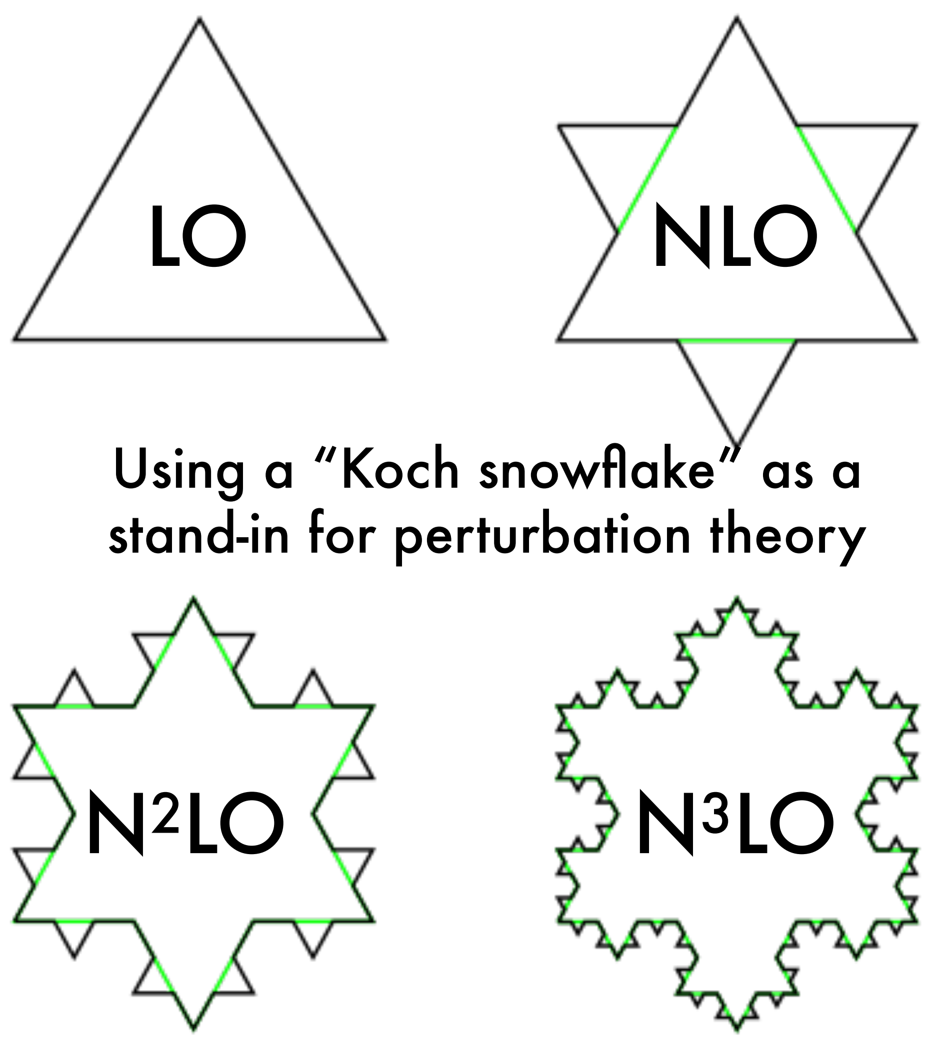
ALICE 2021: huge Λ_c^+ rates



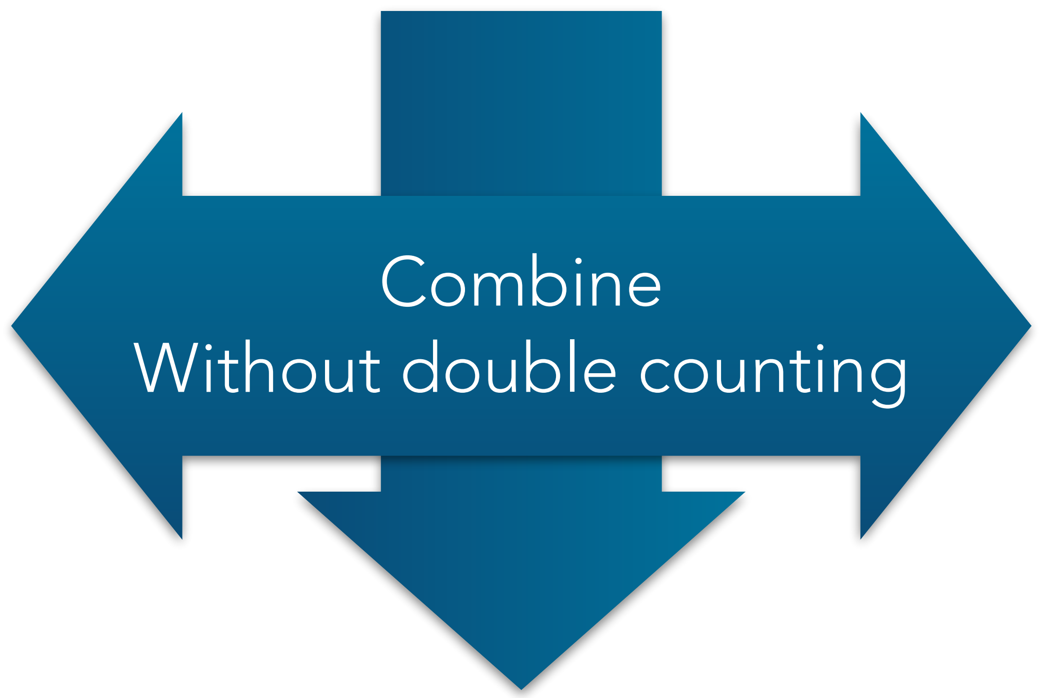
Note: Same appears to be the case for Λ_b

Perturbative QCD: Matching & Merging

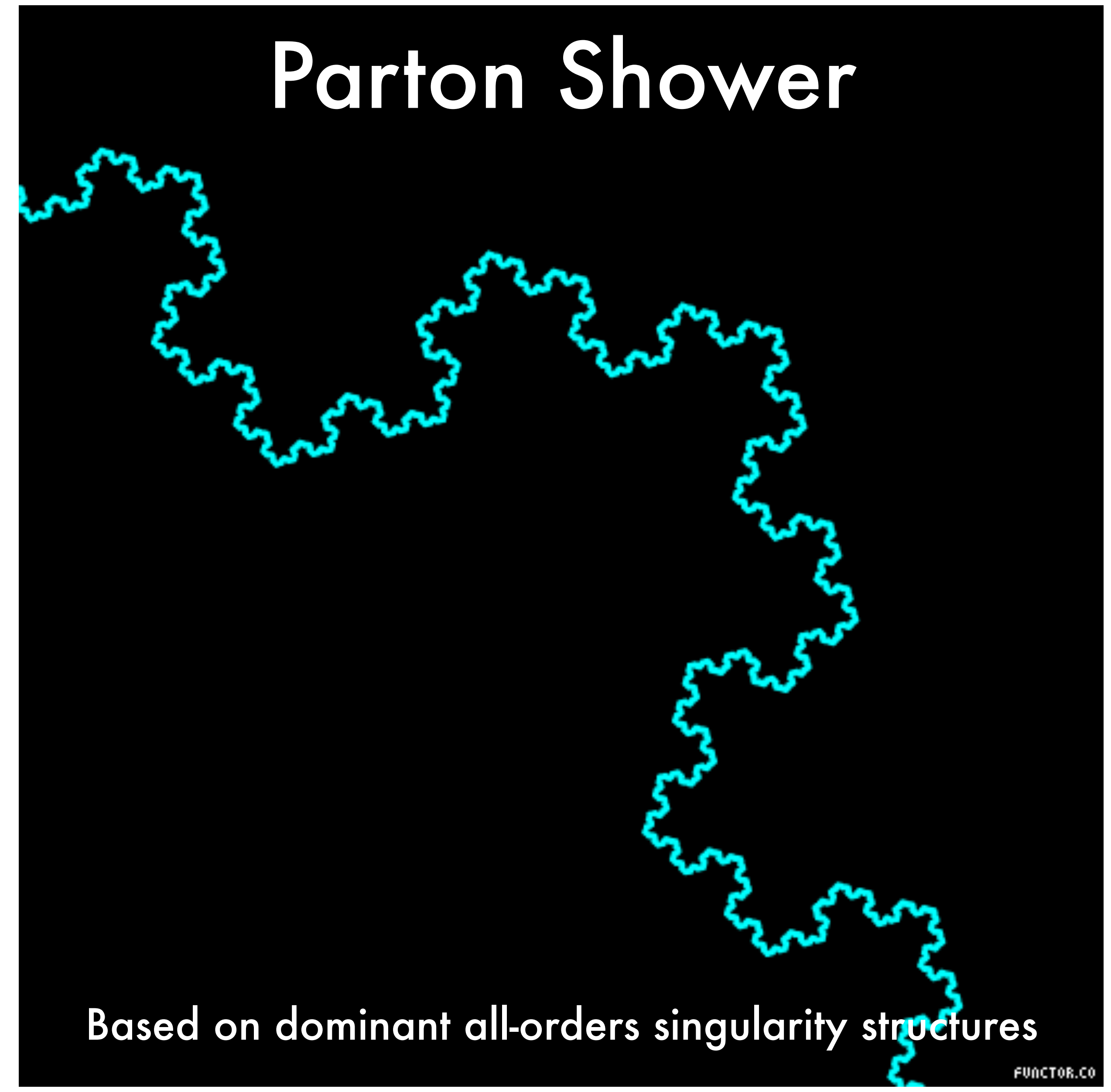
Simplified analogy of a "cross section":



Include more orders



Improve accuracy & detail



PYTHIA HTML Manual:

▼ Matching and Merging

Matching and Merging
-- POWHEG Matching
-- aMC@NLO Matching
-- MLM Jet Merging
-- CKKW-L Merging
-- MESS Merging (Vincia)
-- UMEPS Merging
-- NLO Merging
Merging Variations
Implement New Schemes

Combinations well established for first few orders

Fixed-order QCD for the first few orders; then shower "takes over"

- **1st-order LO matrix-element corrections** (MECs) (Sjöstrand et al., 80s)
  - ➤ standard in PYTHIA for 2 → 1 processes and SM resonance decays
- **1st-order NLO matched calculations** (MC@NLO, POWHEG '00s)
- **Multiple LO merged calculations** (CKKW & Lönnblad, '00s + more recent)

Note: PYTHIA also has many example programs showing how to do this in practice: main80.cc - main89.cc

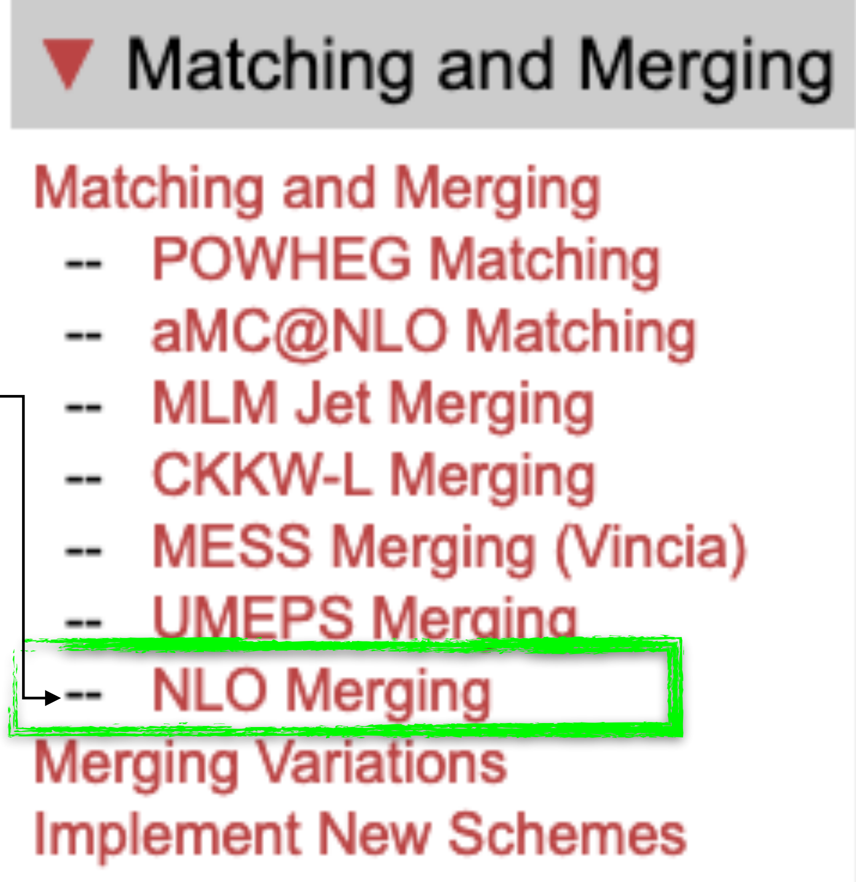
Perturbative QCD: The Cutting Edge

State of the art (for LHC phenomenology right now):

Merging several NLO + PS matched calculations

➤ NL3, UNLOPS (POWHEG-based), FxFx (MC@NLO-based), …

main88.cc



Next Step: NNLO + PS matching

Conceptual issue:

LL PS ≠ full NNLO singularity structure

GENEVA & **MINNLOPS**:

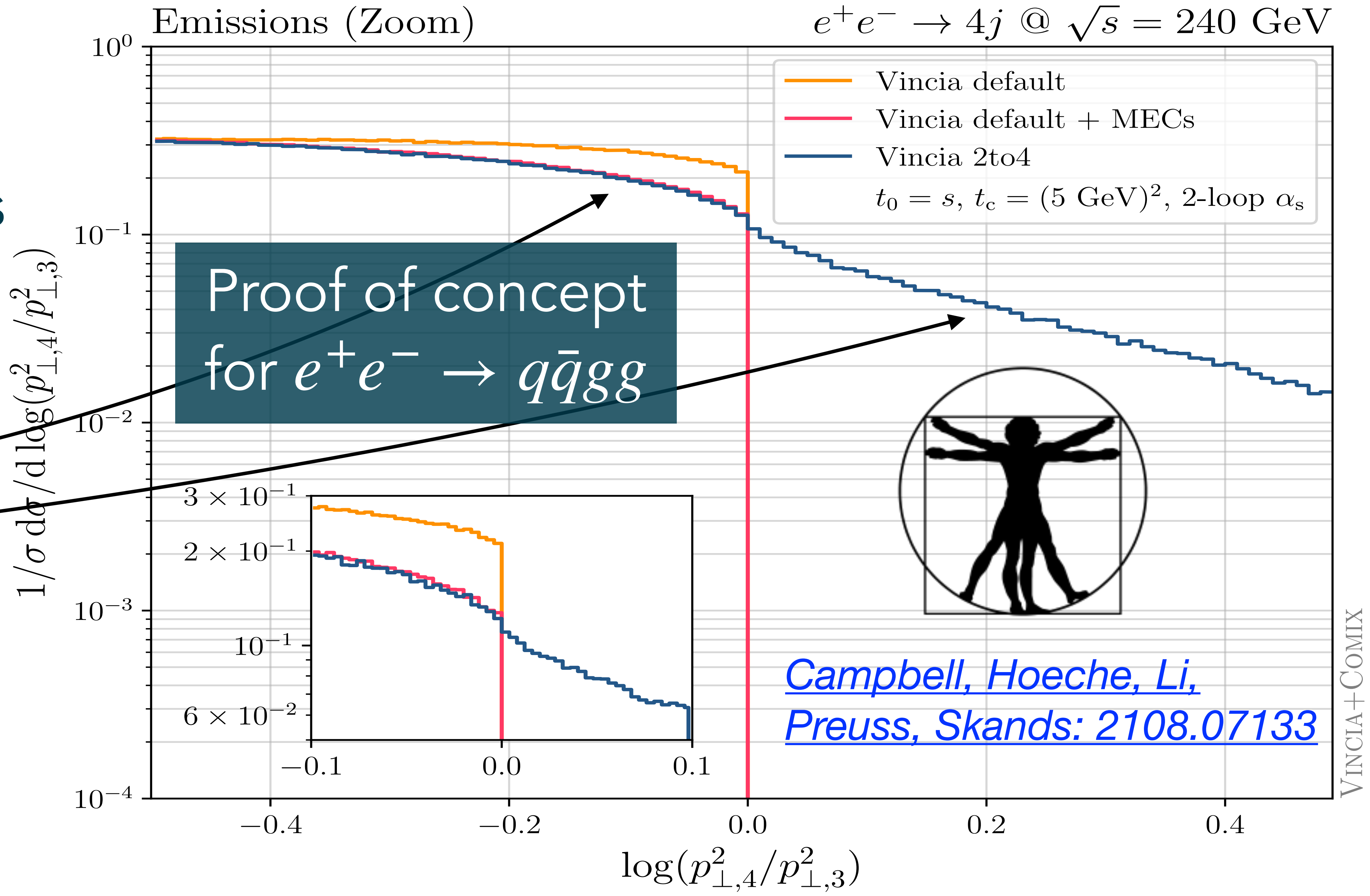
Supply **analytical** NNLL Sudakov factors

VINCIA (arXiv:2108.07133):

New type of shower: interleaves iterated **ordinary** $2 \to 3$ **branchings** with new **"direct"** $2 \to 4$ **ones**

Does describe the full (differential) NNLO singularity structure!

(Interesting also for **NNLL showers**)



Proof of concept for $e^+e^- \to q\bar{q}gg$

Campbell, Hoeche, Li, Preuss, Skands: 2108.07133