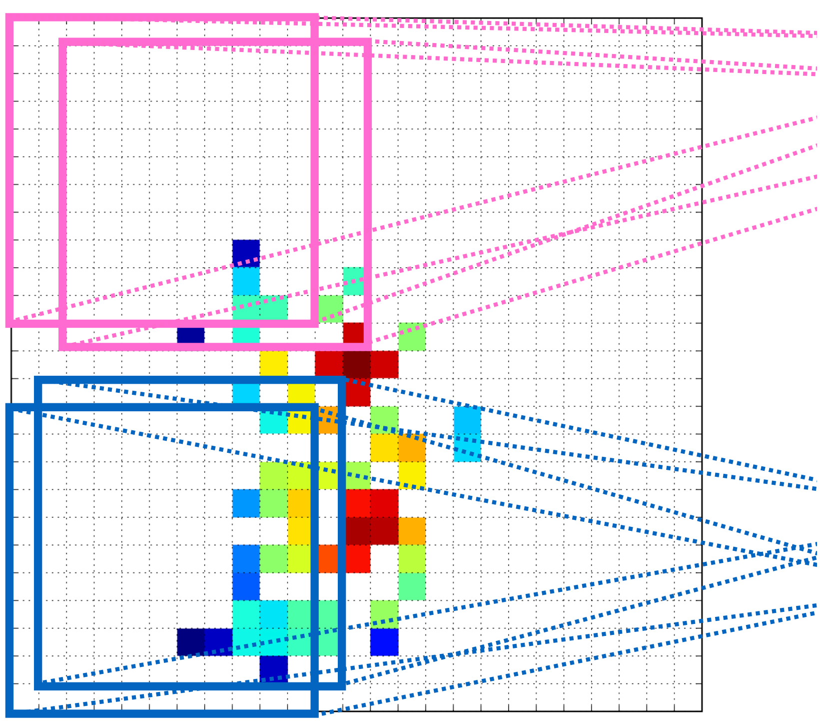
Jet Substructure and Machine Learning

Benjamin Nachman

Lawrence Berkeley National Laboratory

cern.ch/bnachman
bpnachman@lbl.gov

@bpnachman bnachman



BERKELEY LAB

BERKELEY EXPERIMENTAL PARTICLE PHYSICS

HENPIC
September 3, 2020

Physics Program of Jet Substructure

(1) (Beyond the) Standard model parameters

α_s (including running), m_{top}, EFTs, Higgs self coupling, …

(2) Unique tests of fundamental physics, including unique probes of high energy / collective behavior of the strong force.

interference & entanglement, dead cone, …

(3) Direct searches for new particles

Final states with boosted bosons, top quarks, …

(4) General-purpose Monte Carlo generator development and tuning

higher-order corrections, empower other measurements / searches, …

Physics Program of Jet Substructure

3

(1) (Beyond the) Standard model parameters

α_s (including running), m_{top}, EFTs, Higgs self coupling, …

(2) Un… probes …e.

Today, I'll give an example from each of these bullet points and then talk a bit about new tools that will help bring us into the future.

Final states with boosted bosons, top quarks, …

(4) General-purpose Monte Carlo generator development and tuning

higher-order corrections, empower other measurements / searches, …

(1) Towards SM parameters

Grooming makes *pp* jets “look like” e^+e^- jets.

Particular grooming algorithms (soft drop / modified mass drop) have desirable properties to make the above statement **quantitative**.

M. Dasgupta, A. Fregoso, S. Marzani, G. Salam, JHEP 09 (2013) 029

A. Larkoski, S. Marzani, G. Soyez, J. Thaler, JHEP 1405 (2014) 146

This makes observables on softdropped jets amenable to precision calculations for the ~**first time at a *pp* collider.**

precision = beyond LL (e.g. Pythia)

This is particularly important because JSS observables are dominated by **resummation** and **not fixed-order**!

The Soft Drop Procedure

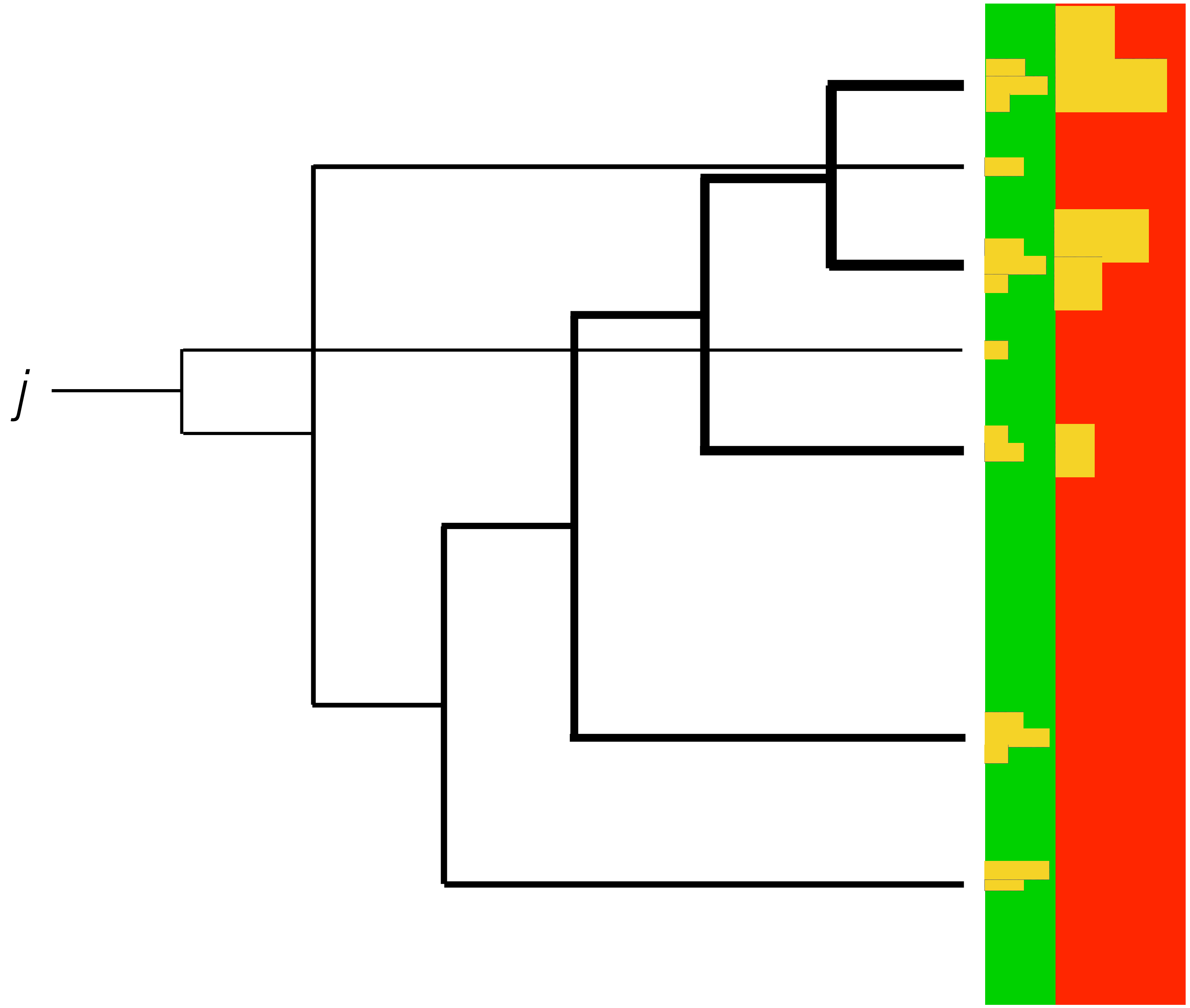
Take a jet clustered with e.g. anti-k_t

Re-cluster it with C/A

Traverse the clustering tree backwards

If a branch point satisfies the soft drop condition, stop.

Otherwise remove the softer branch and continue down the harder branch.



clusters hardest radiation first

The Soft Drop Procedure

Take a jet clustered with e.g. anti-k_t

↓

Re-cluster it with C/A

↓

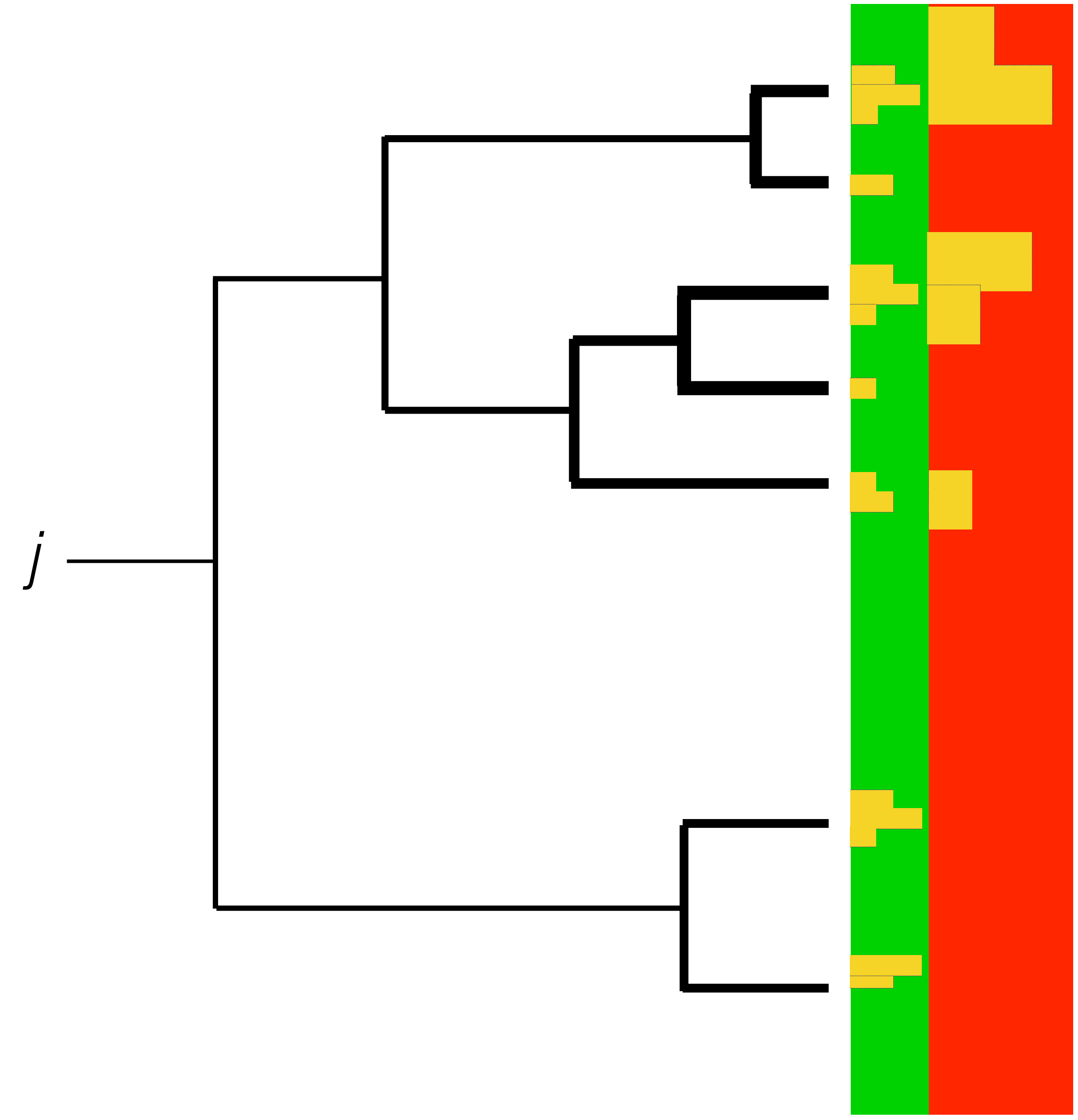
Traverse the clustering tree backwards

↓

If a branch point satisfies the soft drop condition, stop.

↓

Otherwise remove the softer branch and continue down the harder branch.



clusters closest radiation first

The Soft Drop Procedure

Take a jet clustered with e.g. anti-k_t

↓

Re-cluster it with C/A

↓

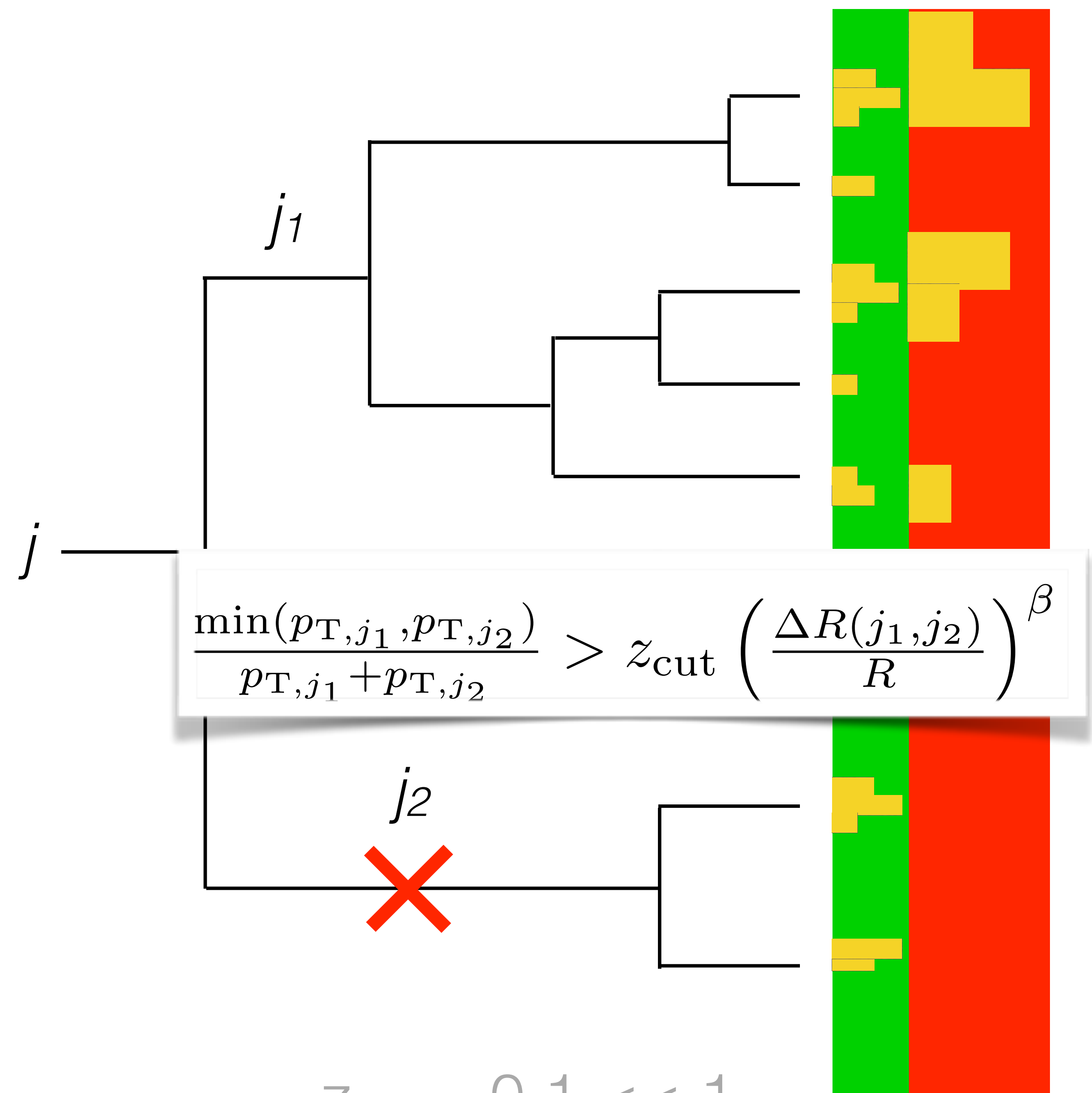
Traverse the clustering tree backwards

↓

If a branch point satisfies the soft drop condition, stop.

↓

Otherwise remove the softer branch and continue down the harder branch.



$$\frac{\min(p_{T,j_1}, p_{T,j_2})}{p_{T,j_1}+p_{T,j_2}} > z_{cut}\left(\frac{\Delta R(j_1,j_2)}{R}\right)^{\beta}$$

$z_{cut} = 0.1 << 1$

The Soft Drop Procedure

Take a jet clustered with e.g. anti-k_t

↓

Re-cluster it with C/A

↓

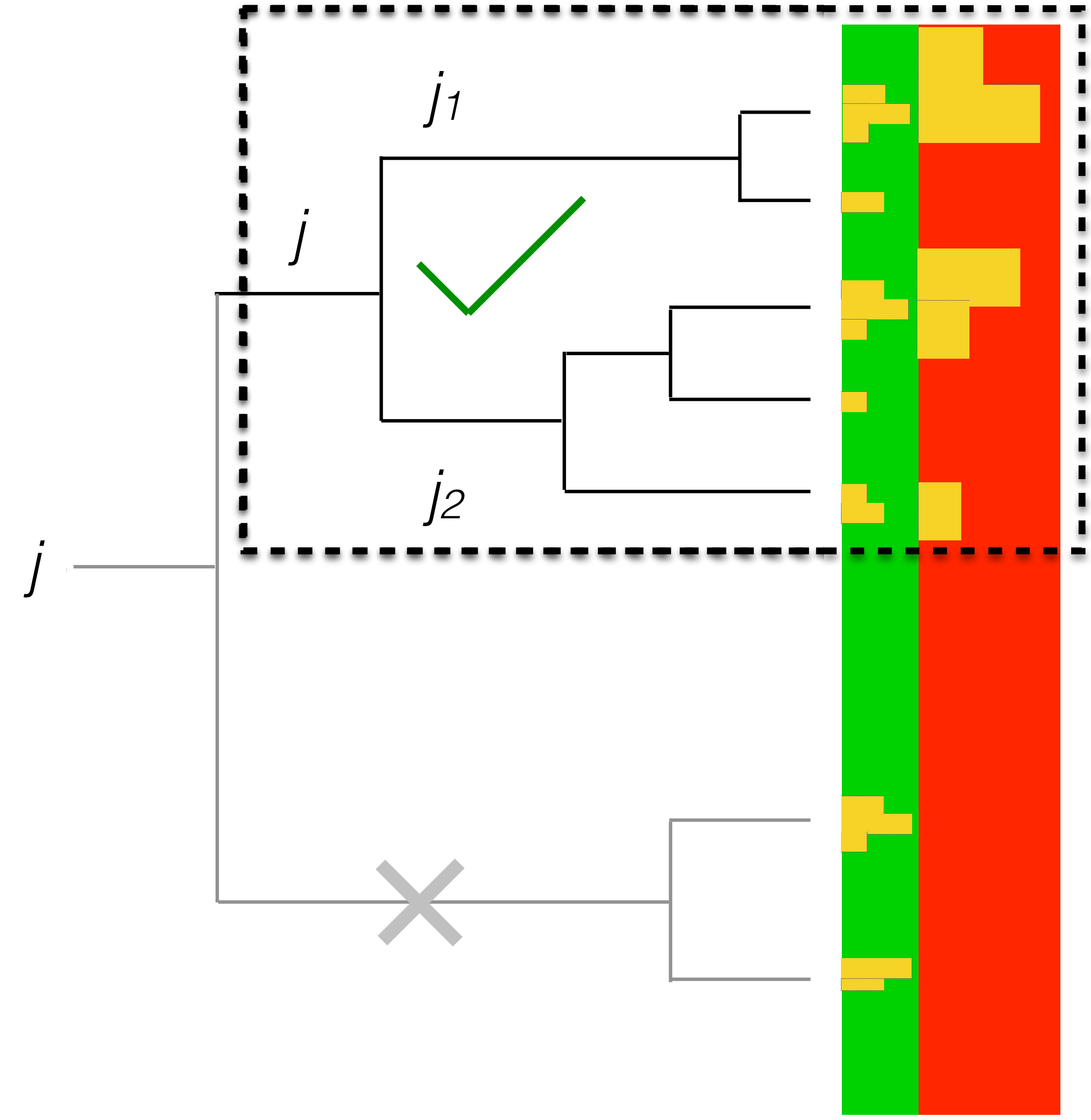
Traverse the clustering tree backwards

↓

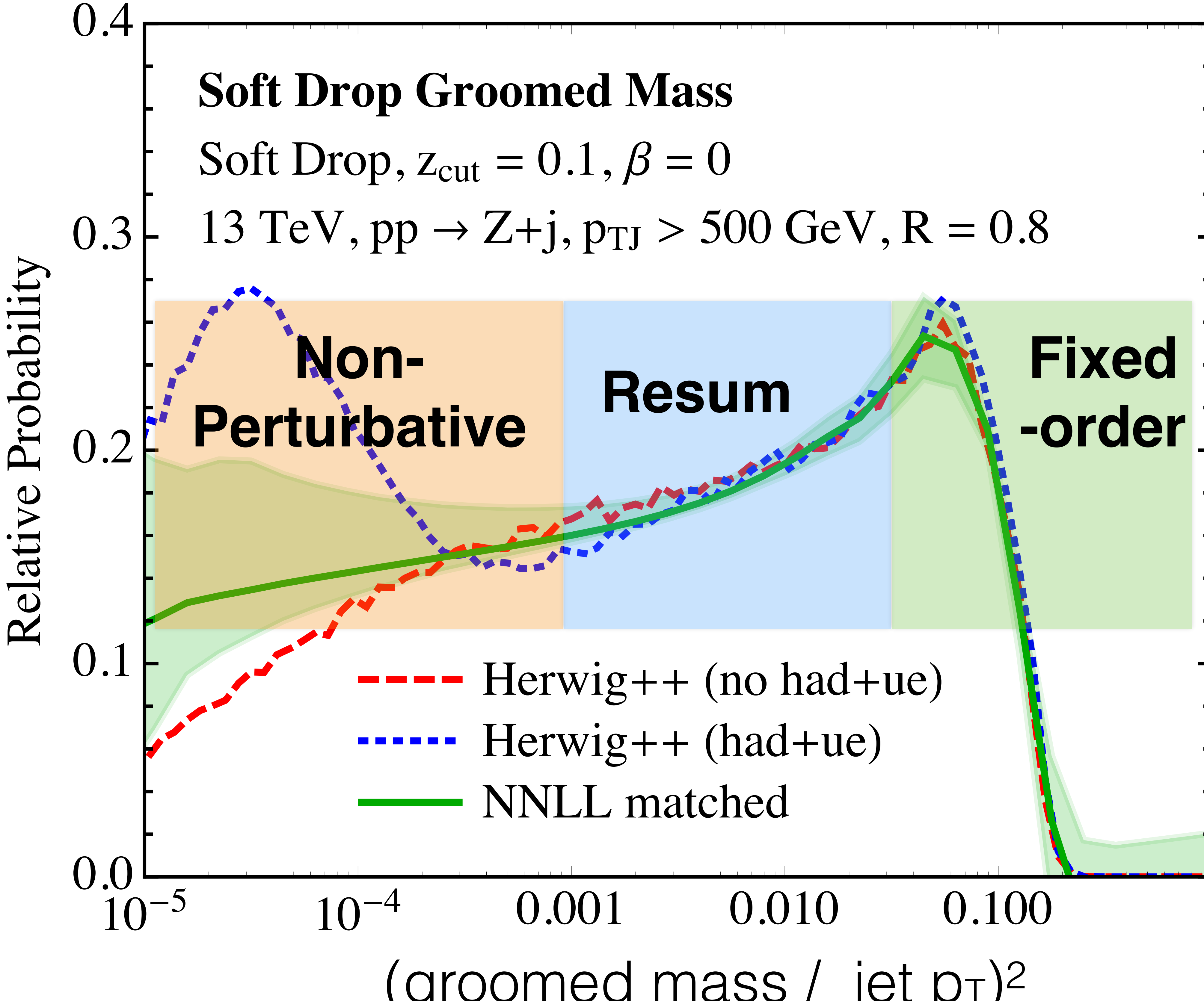
If a branch point satisfies the soft drop condition, stop.

↓

Otherwise remove the softer branch and continue down the harder branch.



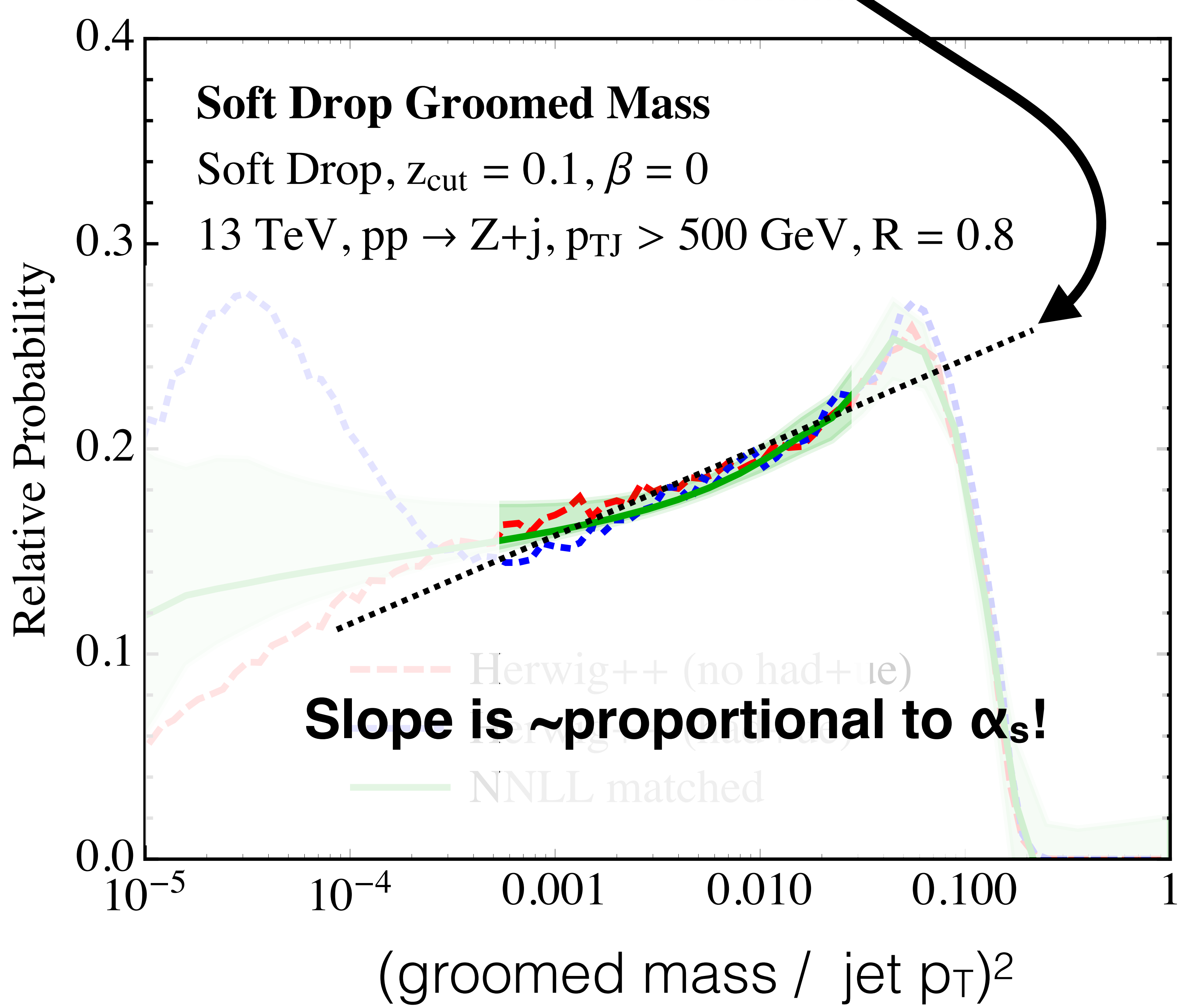
Groomed jet mass



C. Frye, A. Larkoski, M. Schwartz, K. Yan, JHEP 07 (2016) 064

Groomed jet mass for α_s

$$\frac{e_2^{(2)}}{\sigma}\frac{d\sigma}{de_2^{(2)}} \approx -\frac{\alpha_s C_i}{\pi}\left[\log(z_{\mathrm{cut}}) - B_i\right]\exp\left[\boxed{-\frac{\alpha_s C_i}{\pi}}\left[\log(z_{\mathrm{cut}}) - B_i\right]\log(e_2^{(2)})\right]$$



C. Frye, A. Larkoski, M. Schwartz, K. Yan, JHEP 07 (2016) 064

Measurement of the groomed jet mass

Highlights

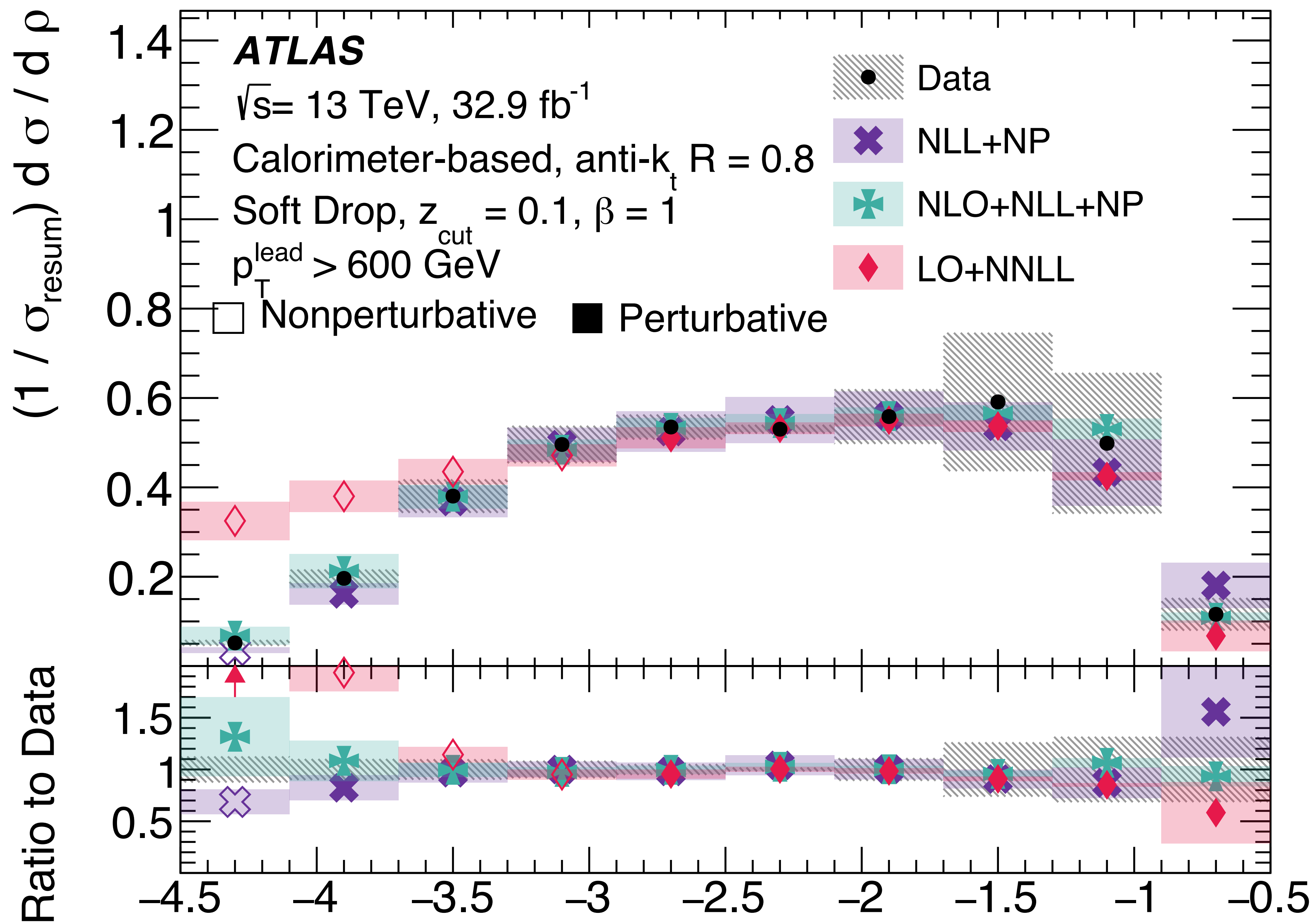
Per-particle calibration

Charged-only & all-particles (shown)

Separate for forward/central to expose q/g

Fragmentation modeling dominates uncertainty

Phys. Rev. D 101 (2020) 052007
Phys. Rev. Lett. 121 (2018) 092001
(see also JHEP 11 (2018) 113 from CMS)



ρ = log [(groomed mass / jet p_T)^2]

The path to α_s

- Non-Perturbative control [see e.g. A. Hoang et al., JHEP12 (2019) 002]

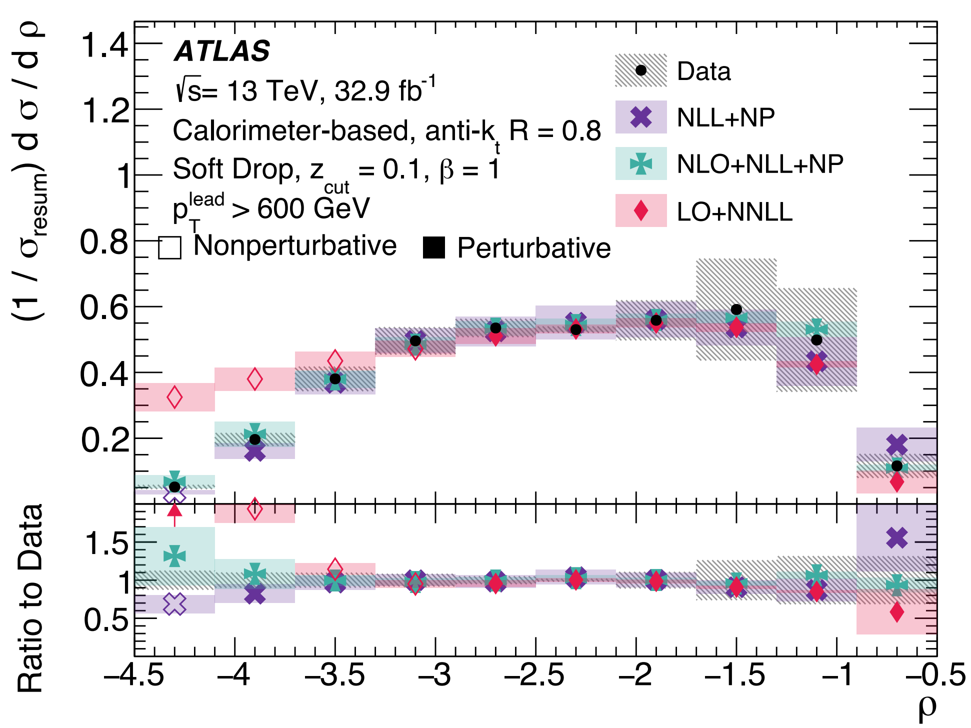
- Higher-order fixed-order (PDG requires NNLO) [Gehrmann et al., in preparation]

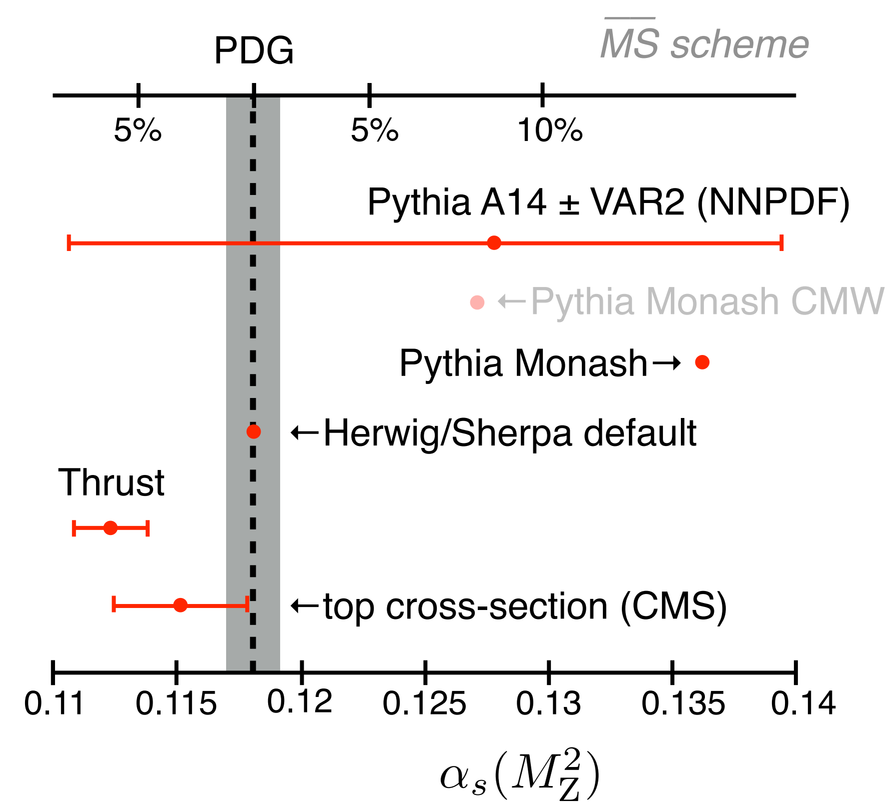
- Sensitivity to q/g fractions [P. Komiske, E. Metodiev, BPN, in preparation]

$$\frac{e_2^{(2)}}{\sigma}\frac{d\sigma}{de_2^{(2)}} = -\frac{\alpha_s C_i}{\pi}\left[\log(z_{\text{cut}}) - B_i\right] \exp\left[-\frac{\alpha_s C_i}{\pi}\left[\log(z_{\text{cut}}) - B_i\right] \log(e_2^{(2)})\right]$$

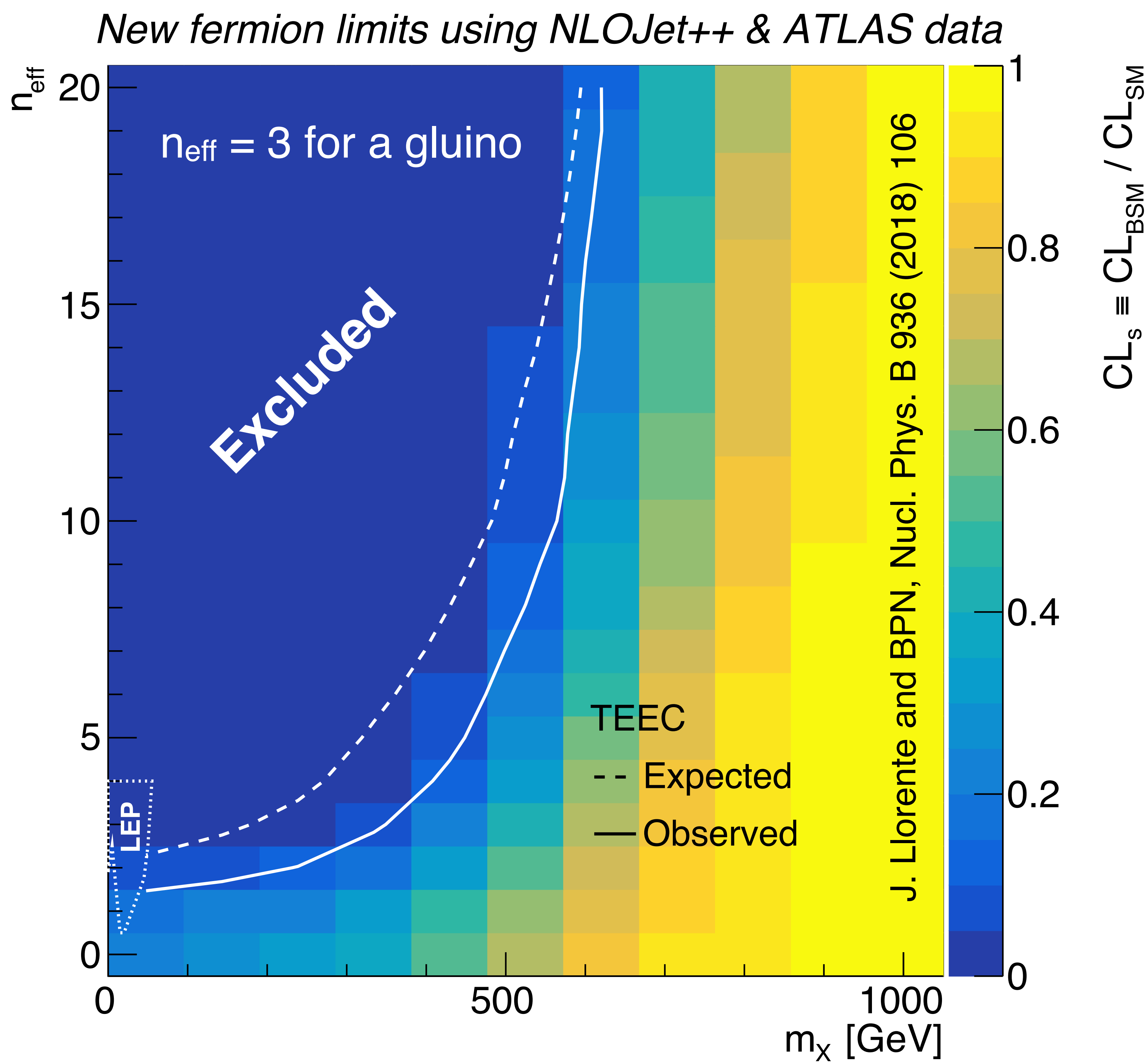
- Experimental precision

N.B. thrust and lattice disagree at ~3σ; even a 10% measurement would be interesting





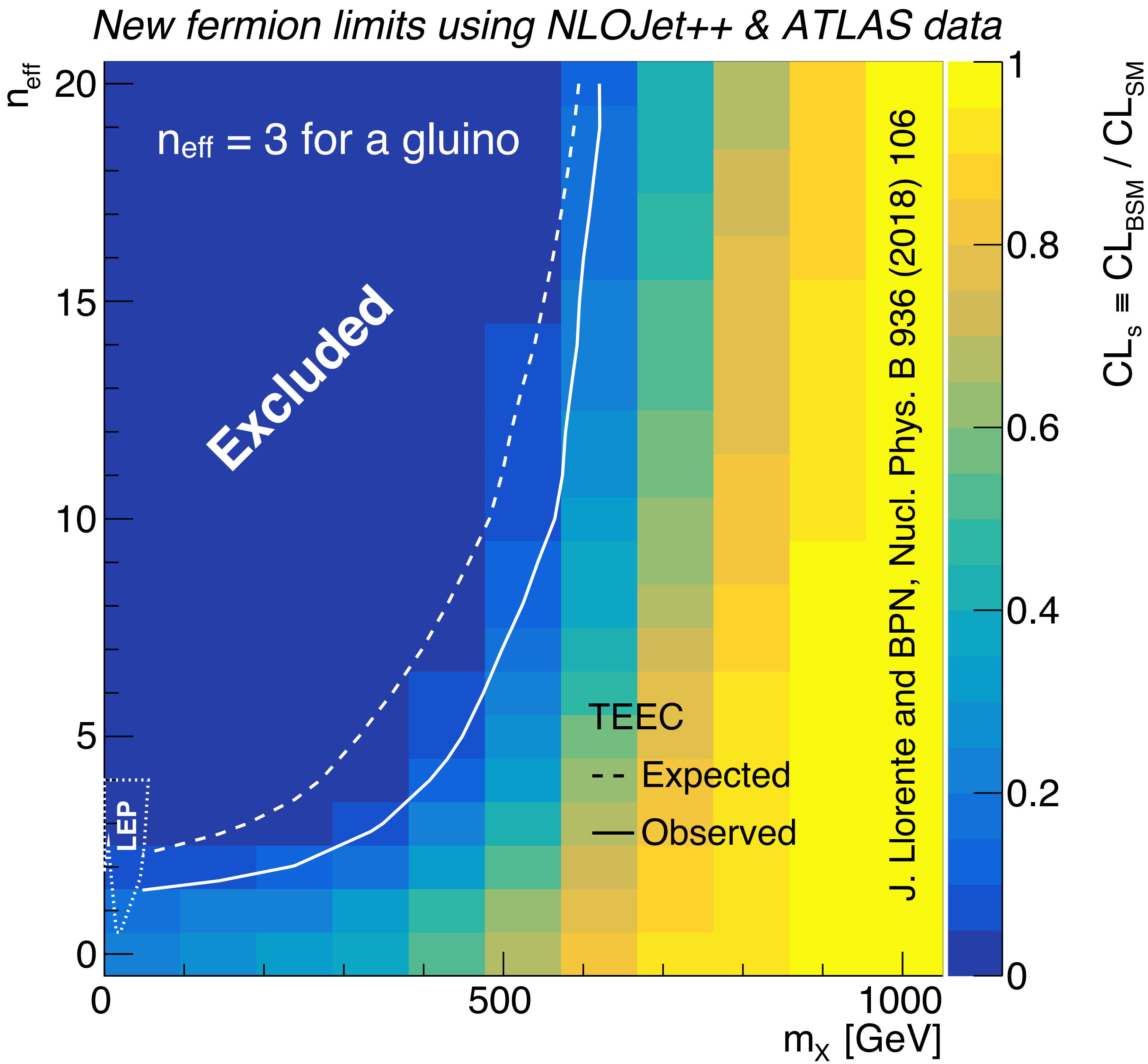
More from α_s: BSM from the running



Precision measurements can also allow us to look for subtle deviations and are complementary to direct searches

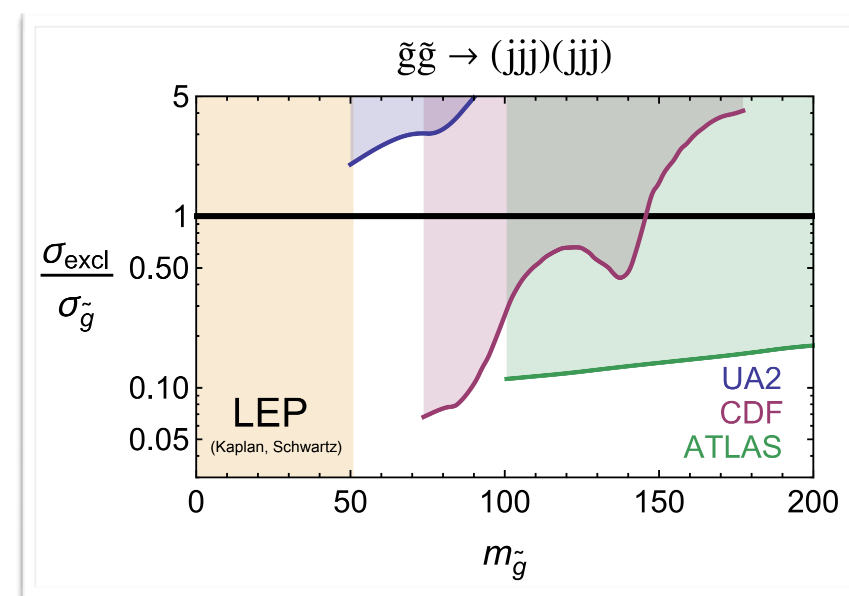
N.B. not based on soft drop mass, but transverse energy energy correlation functions

More from α_s: BSM from the running



Precision measurements can also allow us to look for subtle deviations and are complementary to direct searches

There are also gaps!



J. Evans and D. McKeen, 1803.01880

N.B. not based on soft drop mass, but transverse energy energy correlation functions

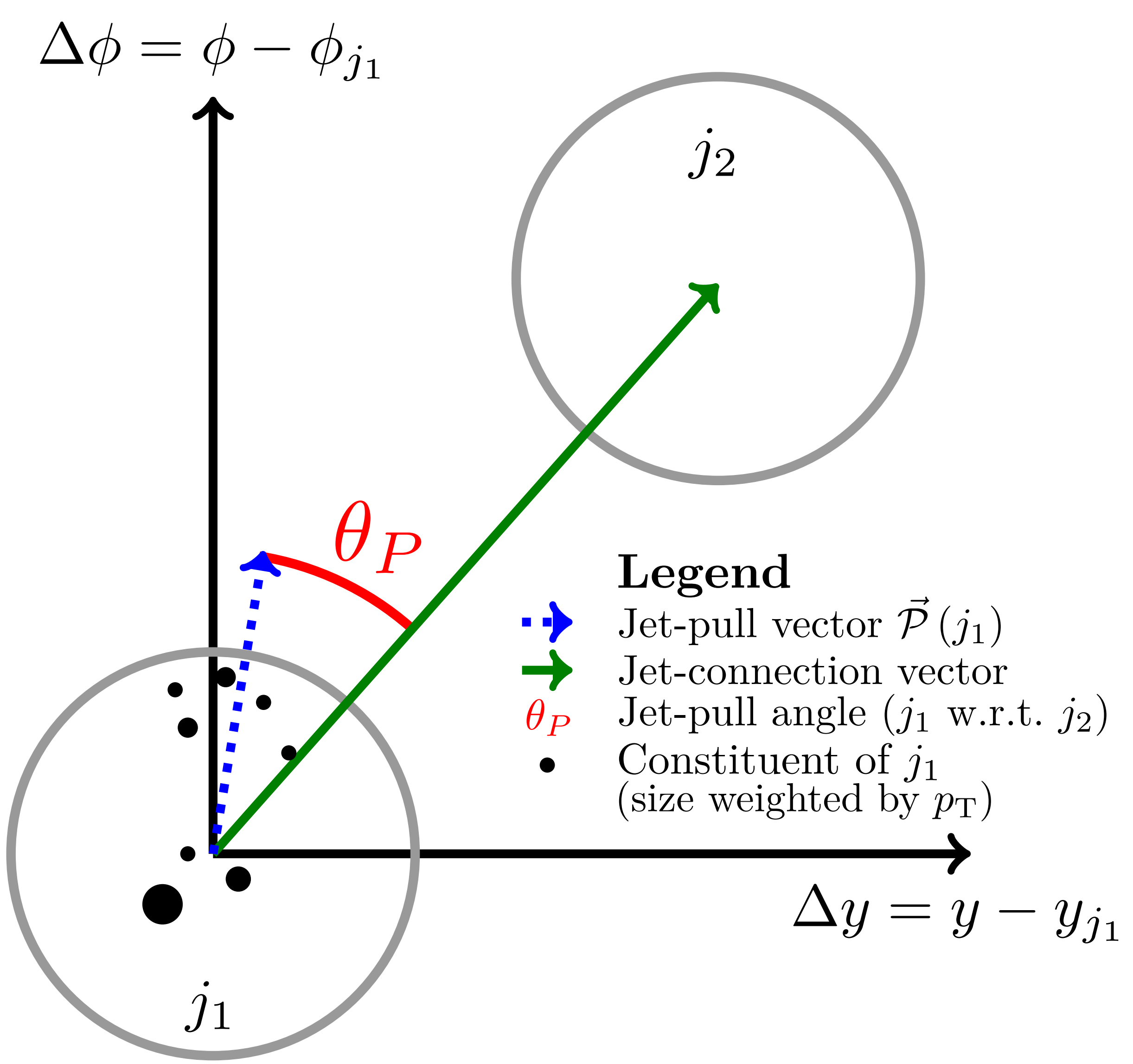
(2) Probe of fundamental physics

Theme: Use correlations between jets as a way to expose quantum properties

(2) Probe of fundamental physics

Theme: Use correlations between jets as a way to expose quantum properties

Example 1: Jet pull



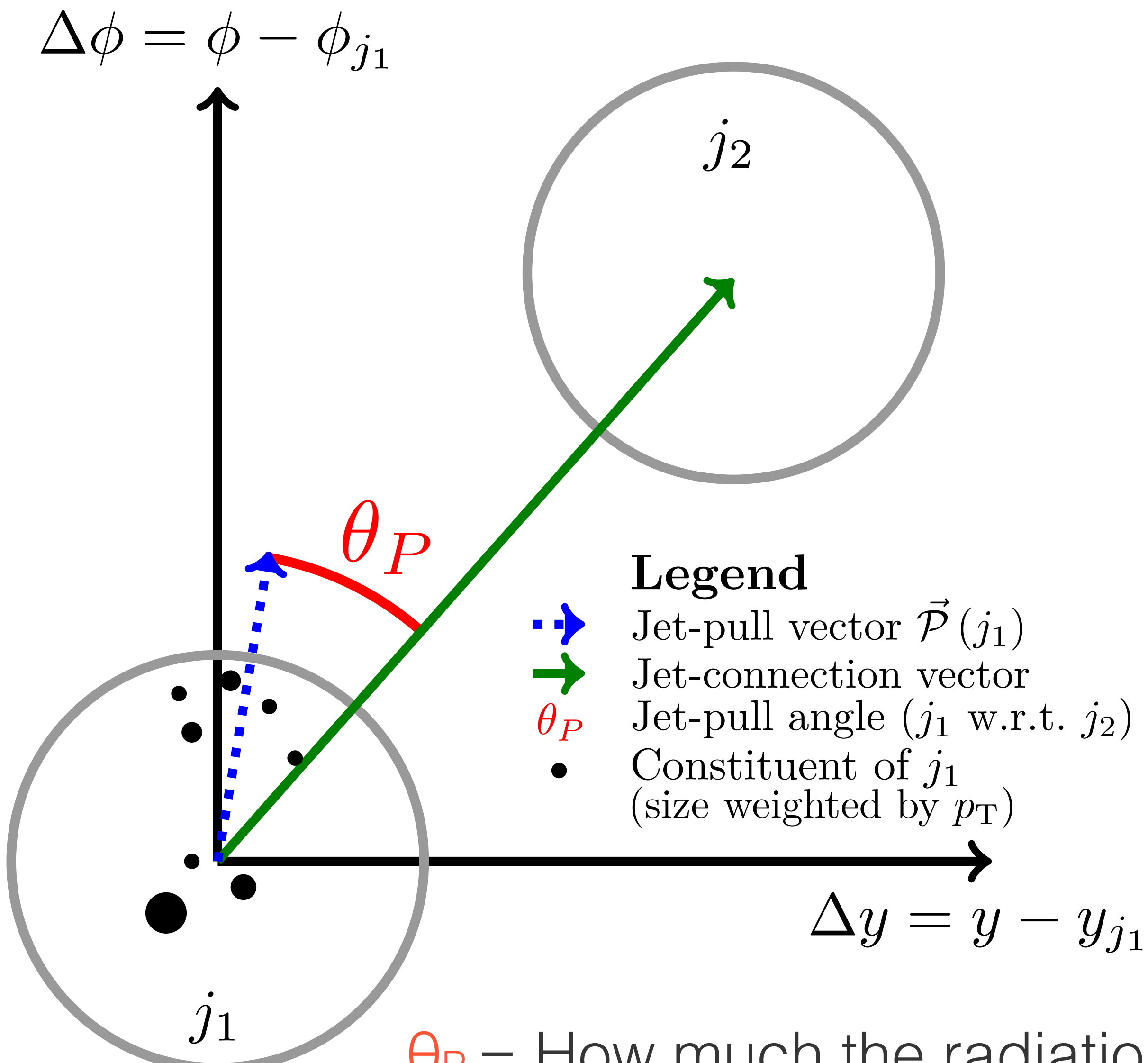
We can study QCD **entanglement** from correlations in the radiation patterns of pairs of jets.

An exciting laboratory for this work is boosted W bosons, a copious source of **singlet** → jets.

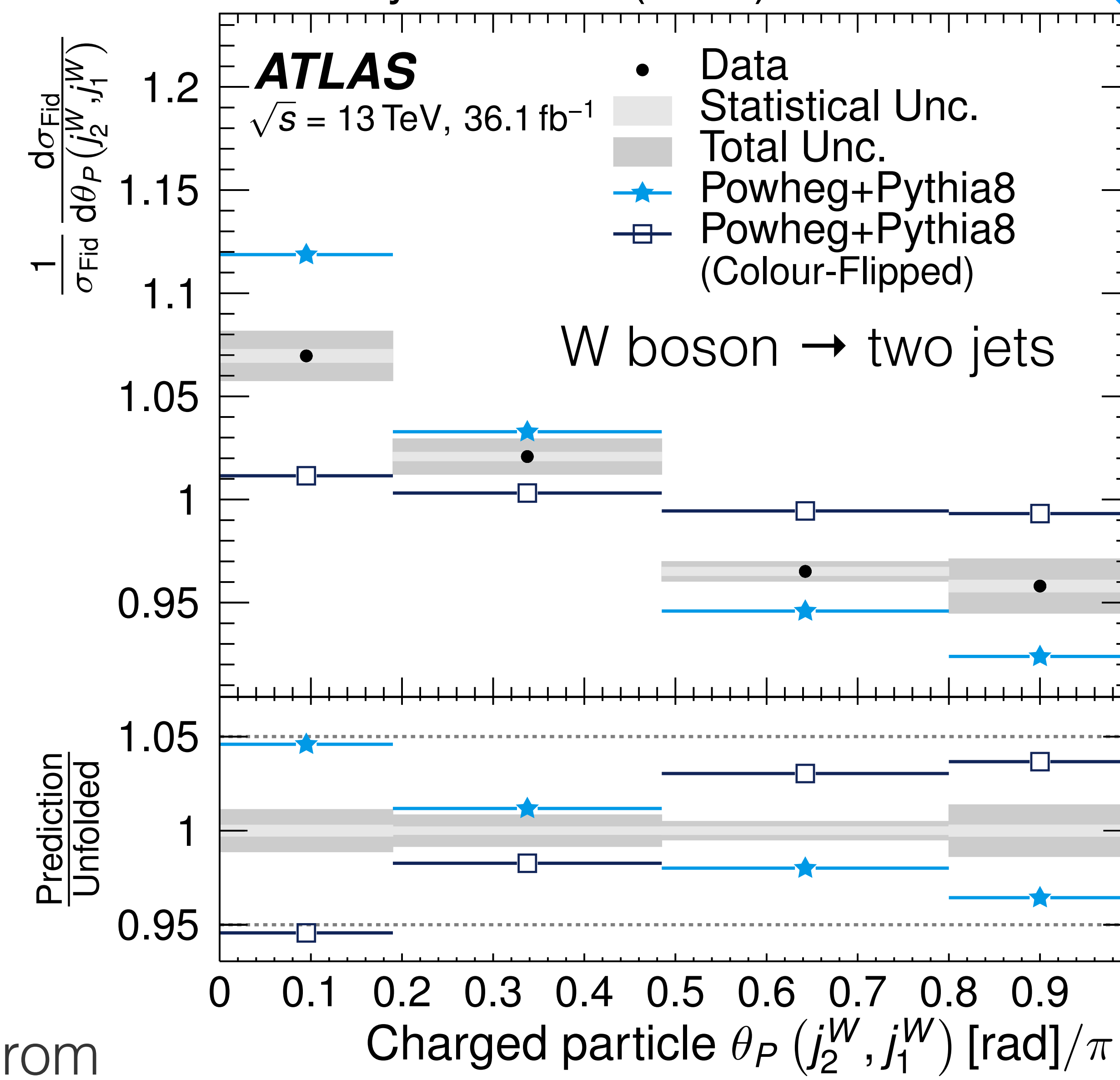
(2) Probe of fundamental physics

Theme: Use correlations between jets as a way to expose quantum properties

Example 1: Jet pull



Eur. Phys. J. C 78 (2018) 847



θ_P = How much the radiation from one jet "leans" toward the other.

(2) Probe of fundamental physics

Theme: Use correlations between jets as a way to expose quantum properties

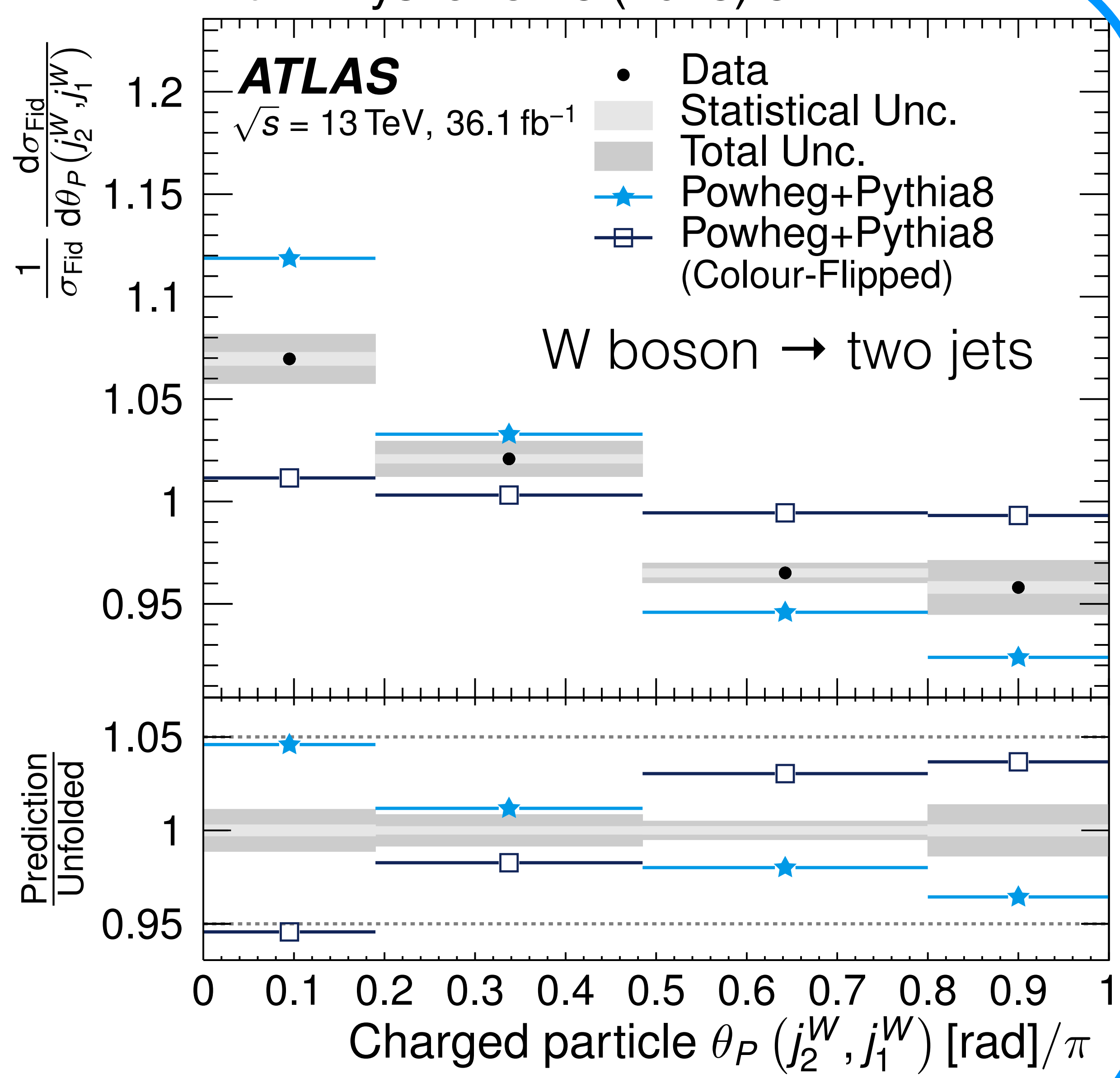
Example 1: Jet pull

Here is an observable where we can't distinguish between "entanglement" turned **"on"** and **"off"** !

Theory predictions are challenging, but in development

(see A. Larkoski, S. Marzani, C. Wu, PRD 99 (2019) 091502)

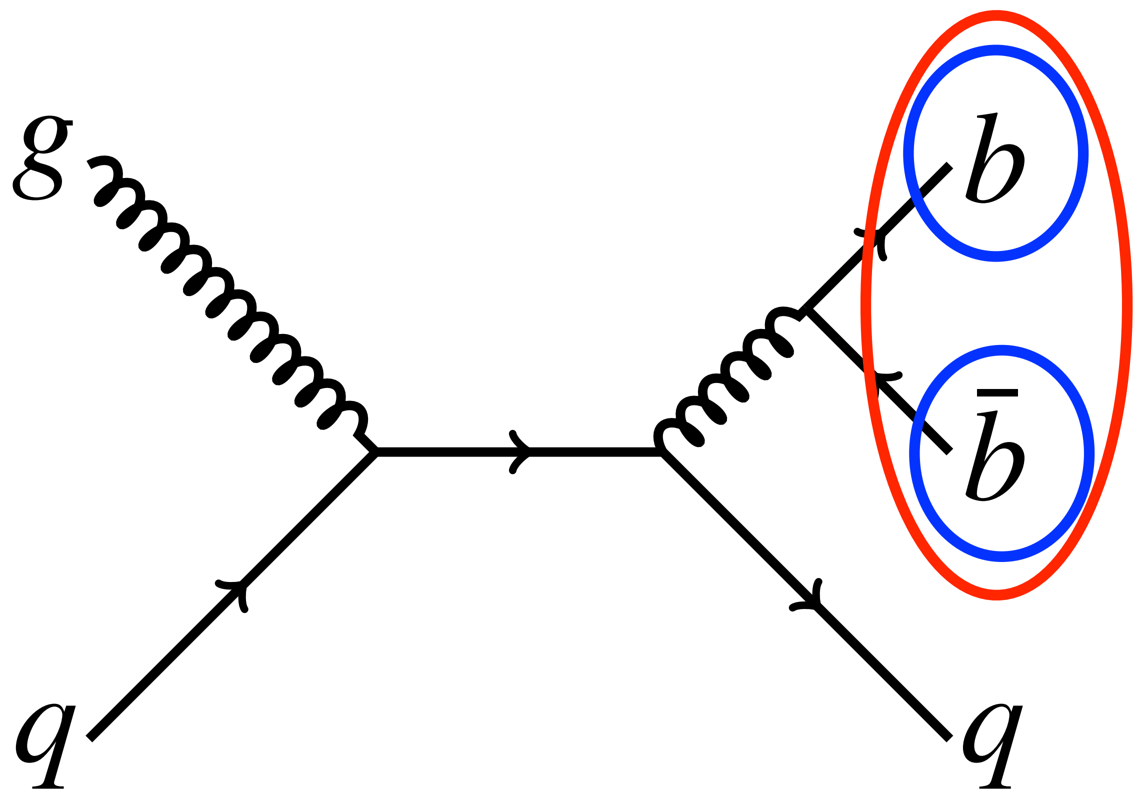
Eur. Phys. J. C 78 (2018) 847



(2) Probe of fundamental physics

Theme: Use correlations between jets as a way to expose quantum properties

Example 2: g → bb



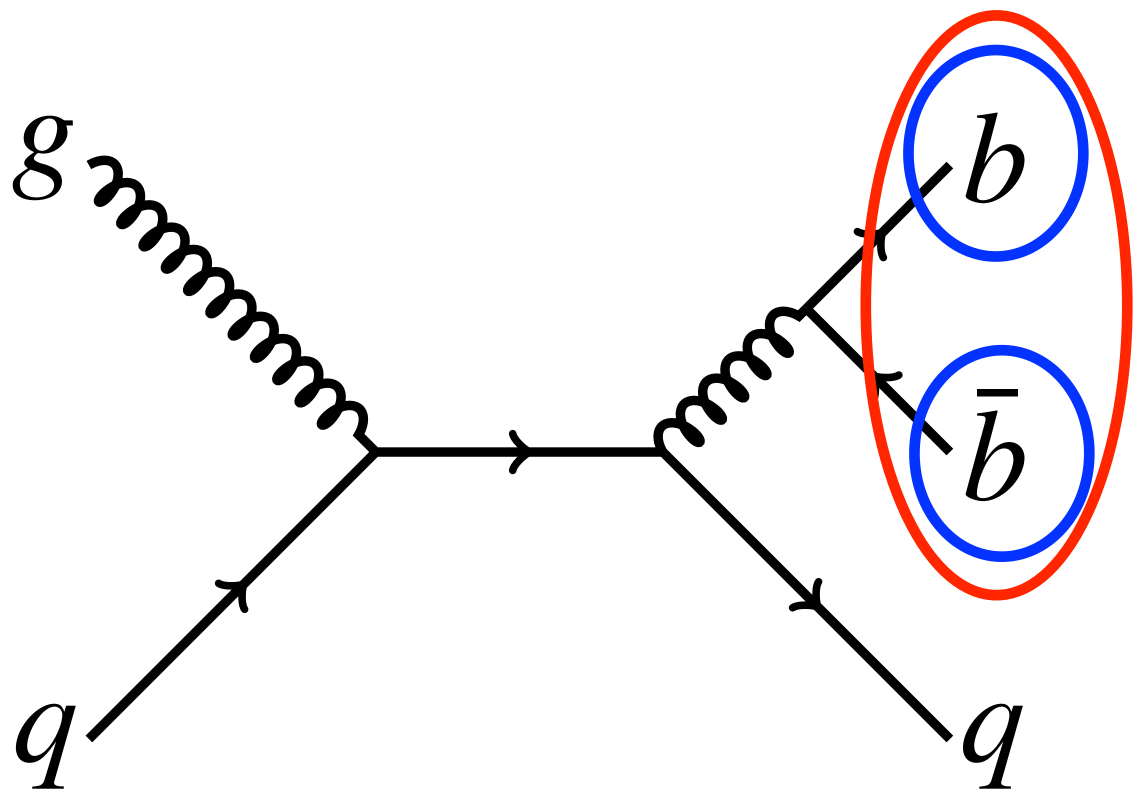
Gluon splitting to bottom quarks gives us the only ~pure access to QCD splitting functions.

(and of course, this is a very important process for Higgs)

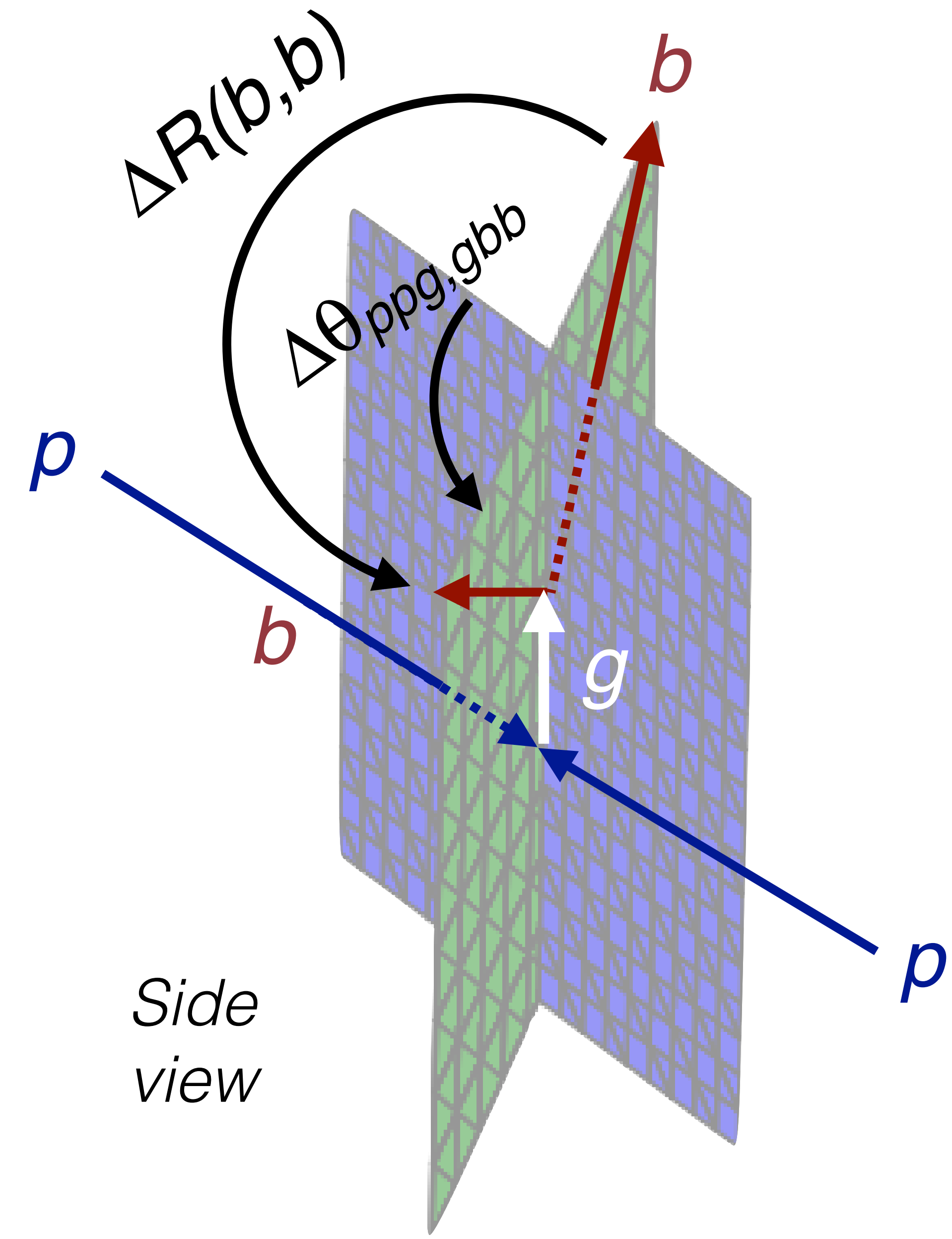
(2) Probe of fundamental physics

Theme: Use correlations between jets as a way to expose quantum properties

Example 2: g → bb



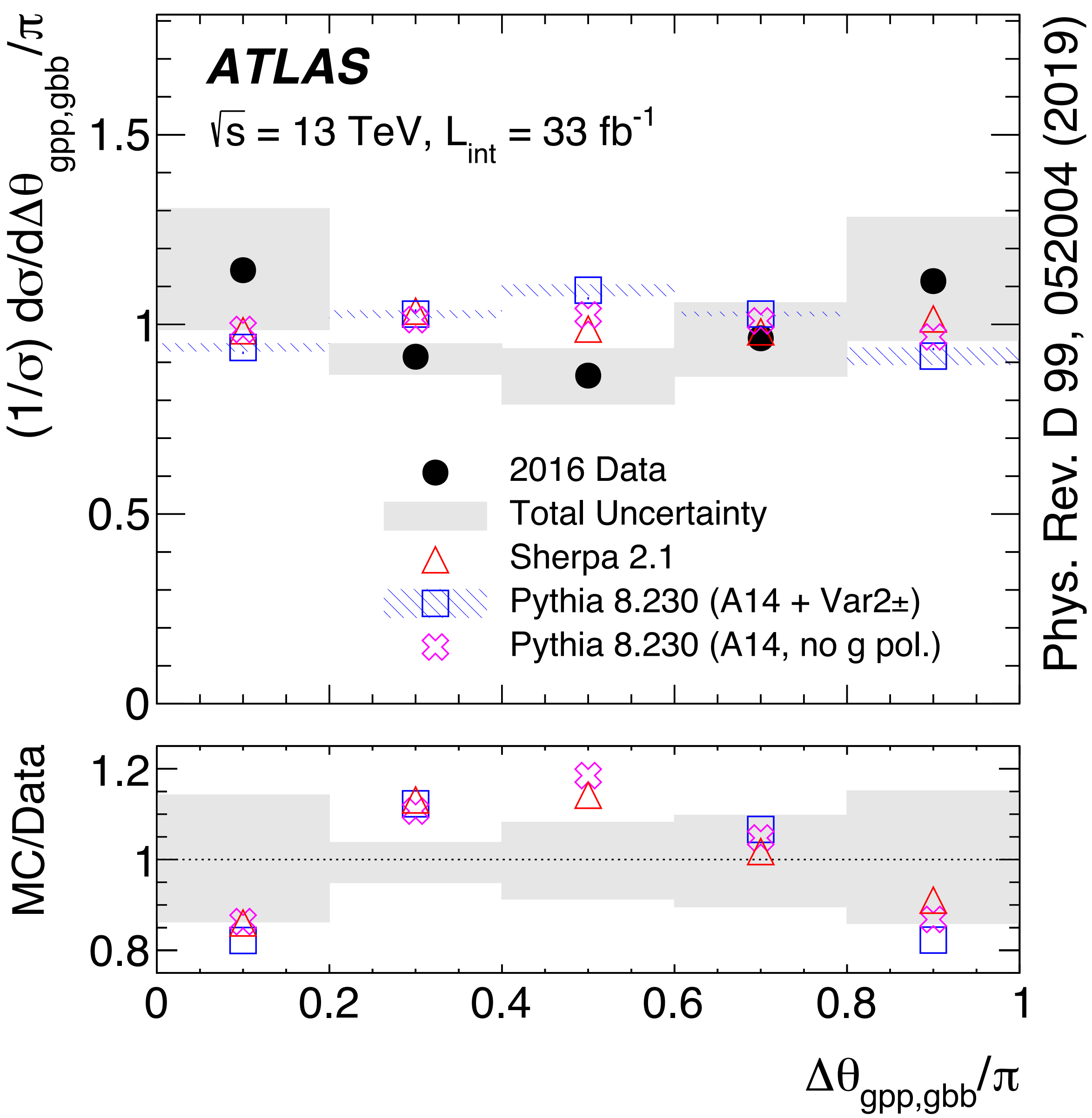
Relative angles to probe polarization

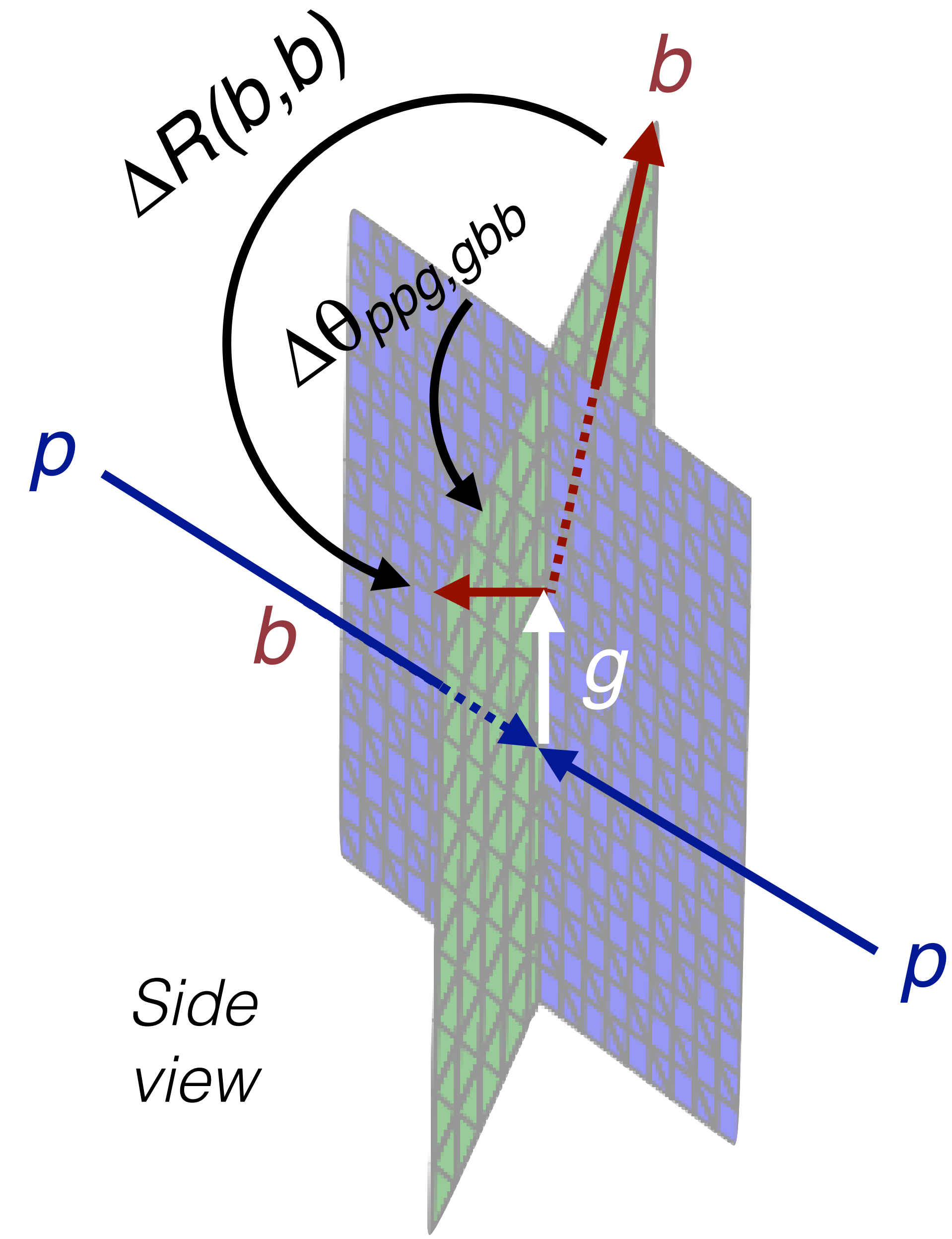


(2) Probe of fundamental physics

Theme: Use correlations between jets as a way to expose quantum properties

Example 2: g → bb

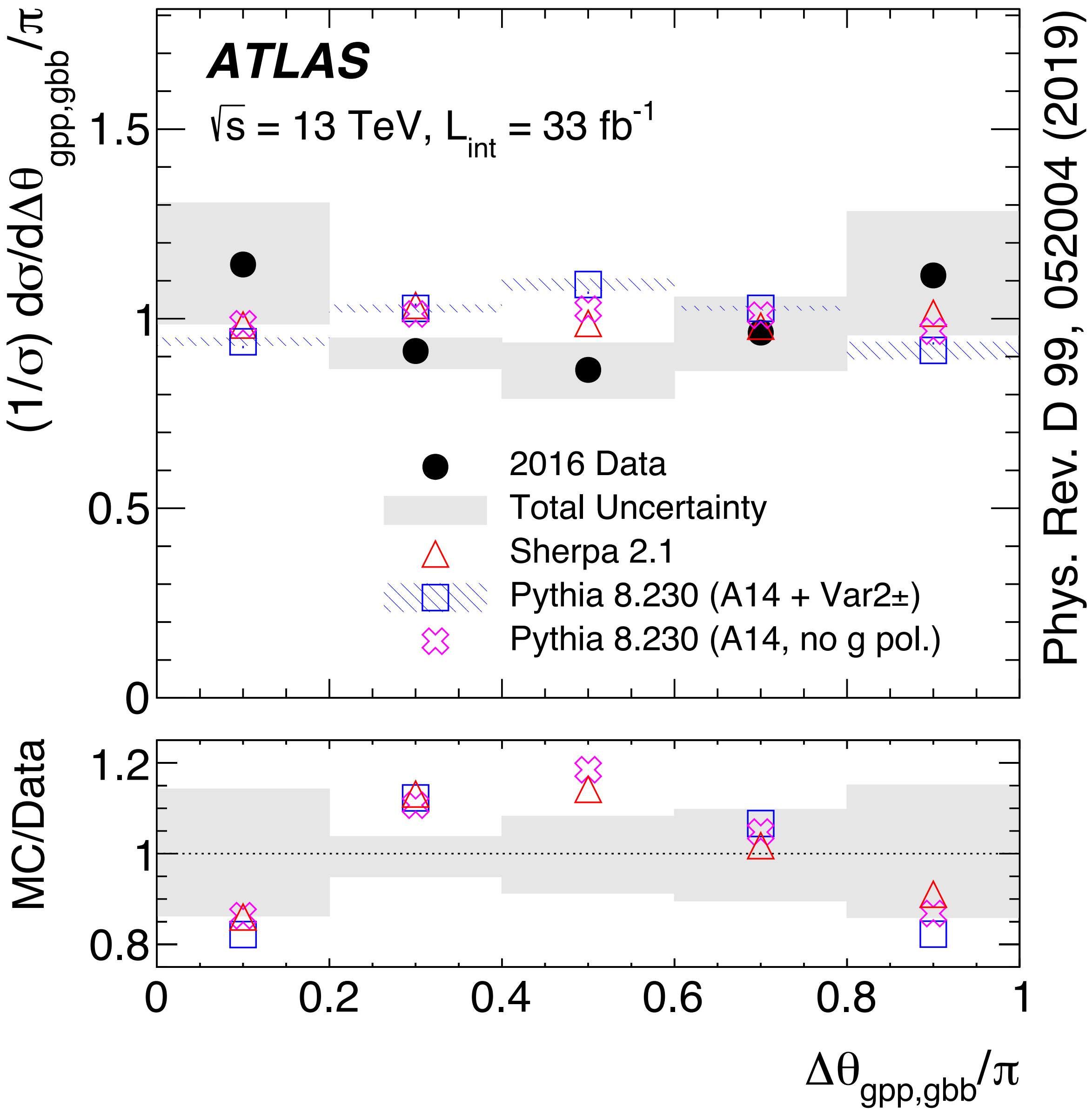




(2) Probe of fundamental physics

Theme: Use correlations between jets as a way to expose quantum properties

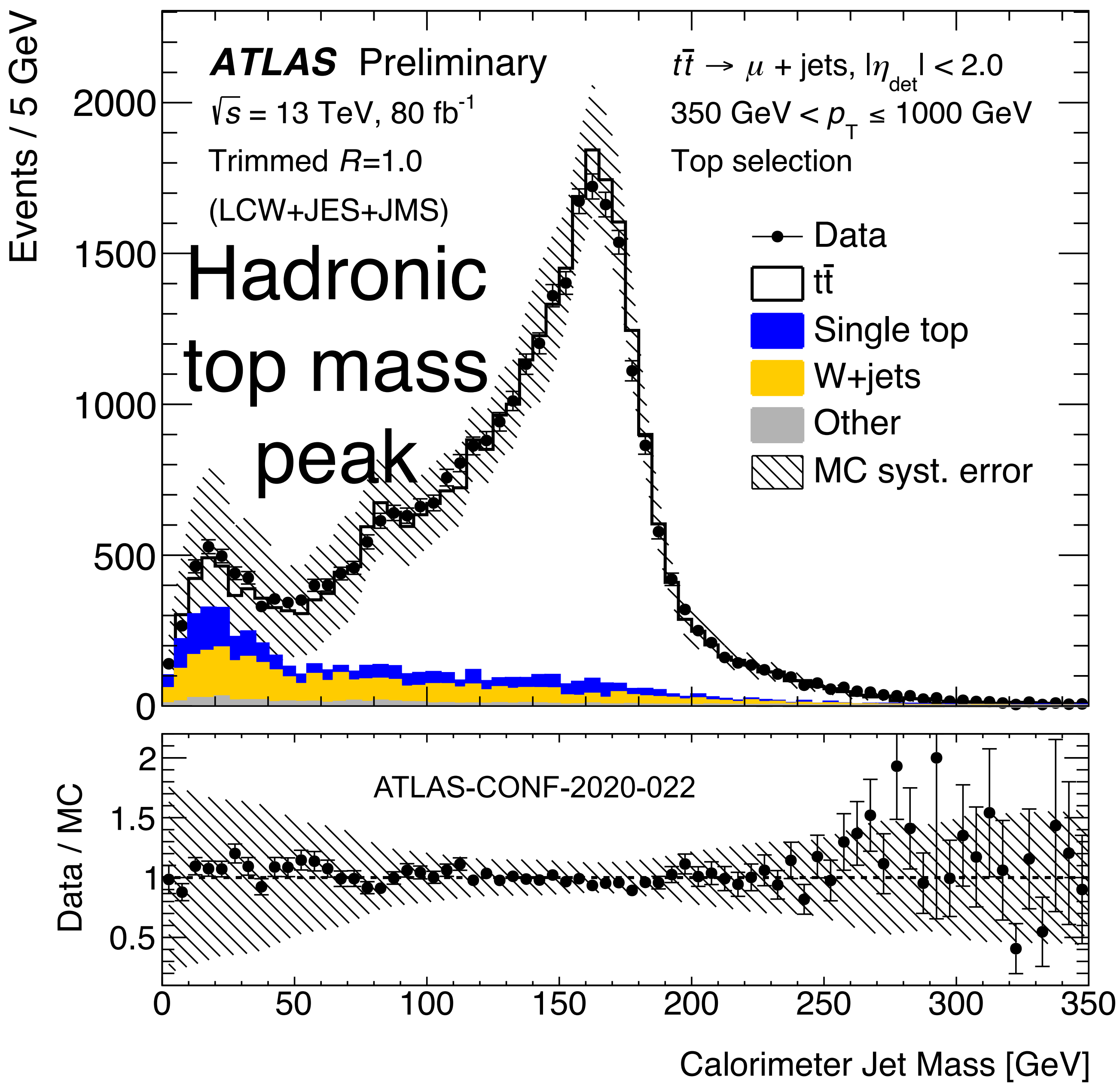
Example 2: g → bb



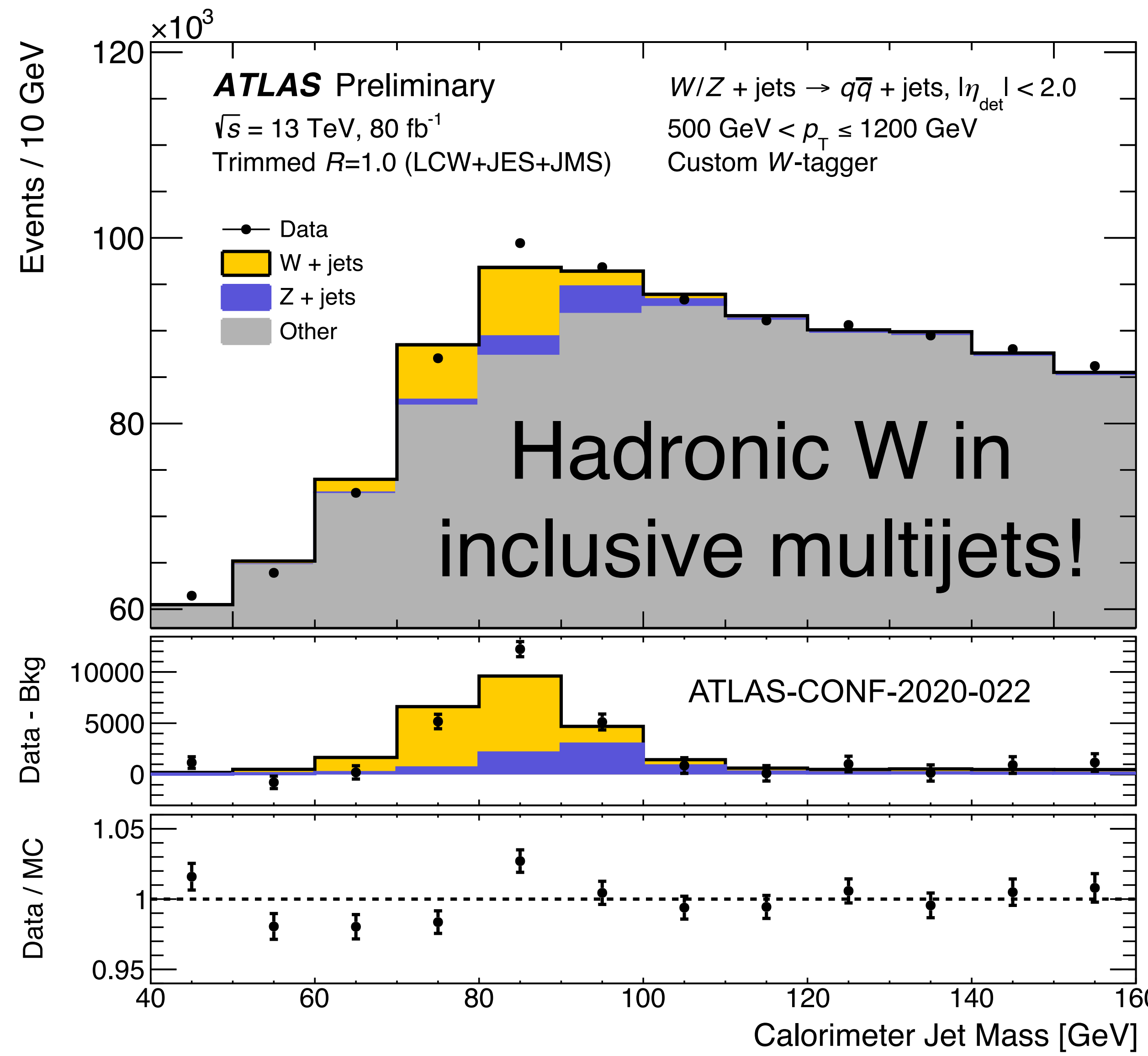
Gluons seems “more polarized” in data than in our predictions. Slight improvement from matrix element corrections (Sherpa 2 → 3).

See also Fischer, Lifson, Skands, EPJC 77 (2017) 719

(3) Direct Searches for BSM



When a massive particle is boosted, its decay products can be contained inside a single jet.

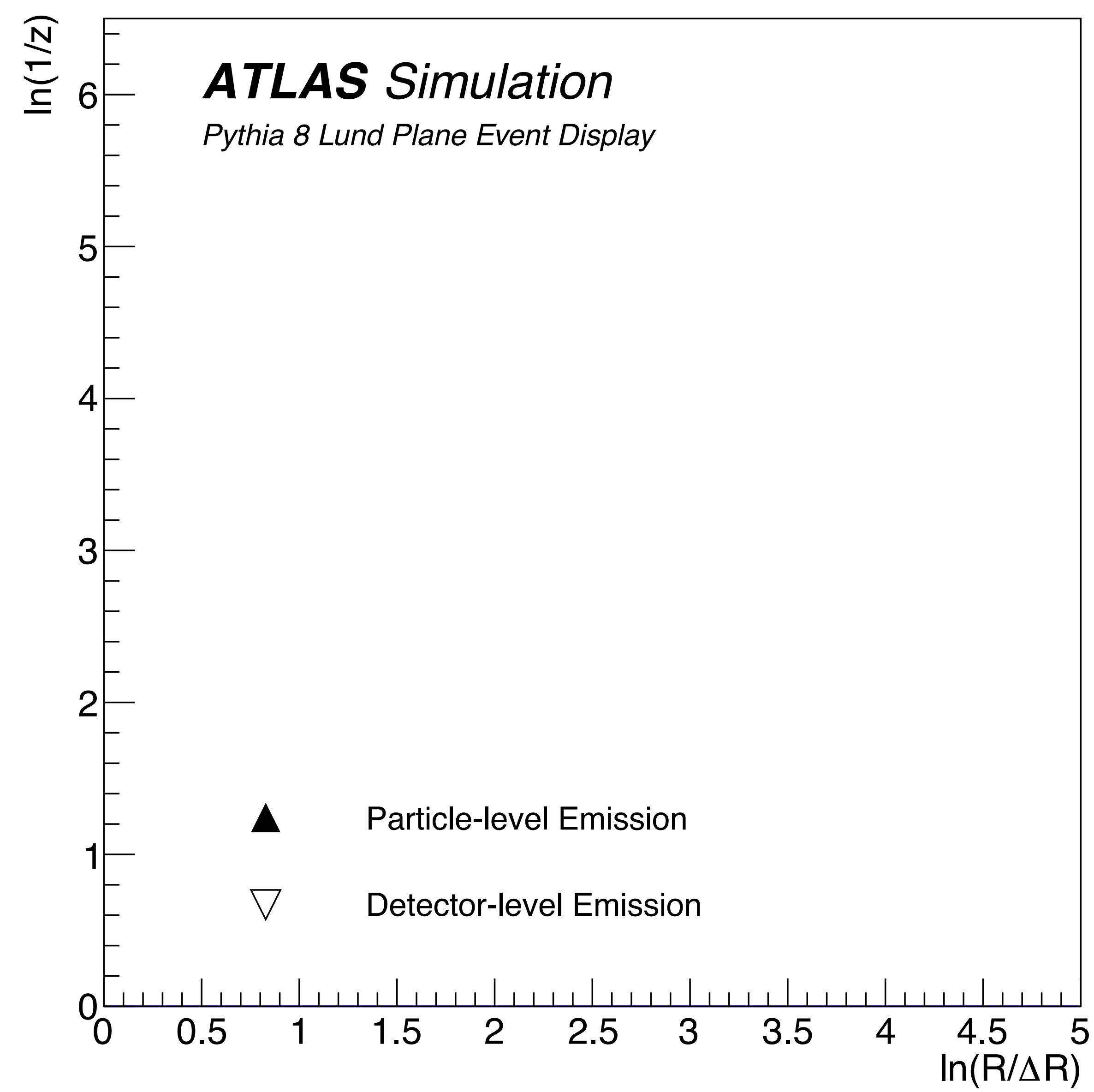


Capturing boosted top quarks, W/Z bosons, and Higgs bosons is now very standard !

(4) Towards improving Monte Carlo

Important: isolate effects with different physical origin

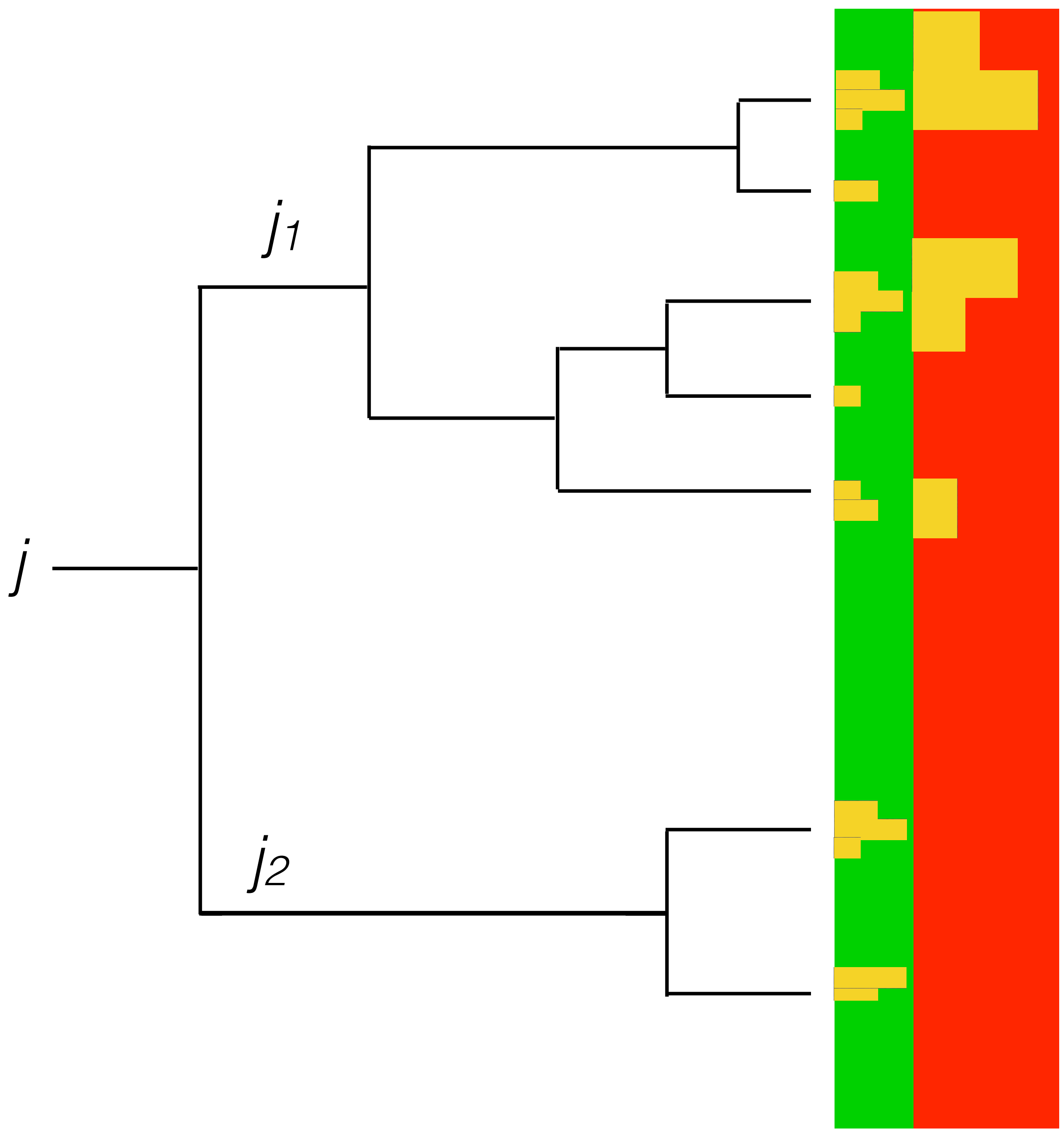
Tool: Lund plane to categorize all hard splittings at once

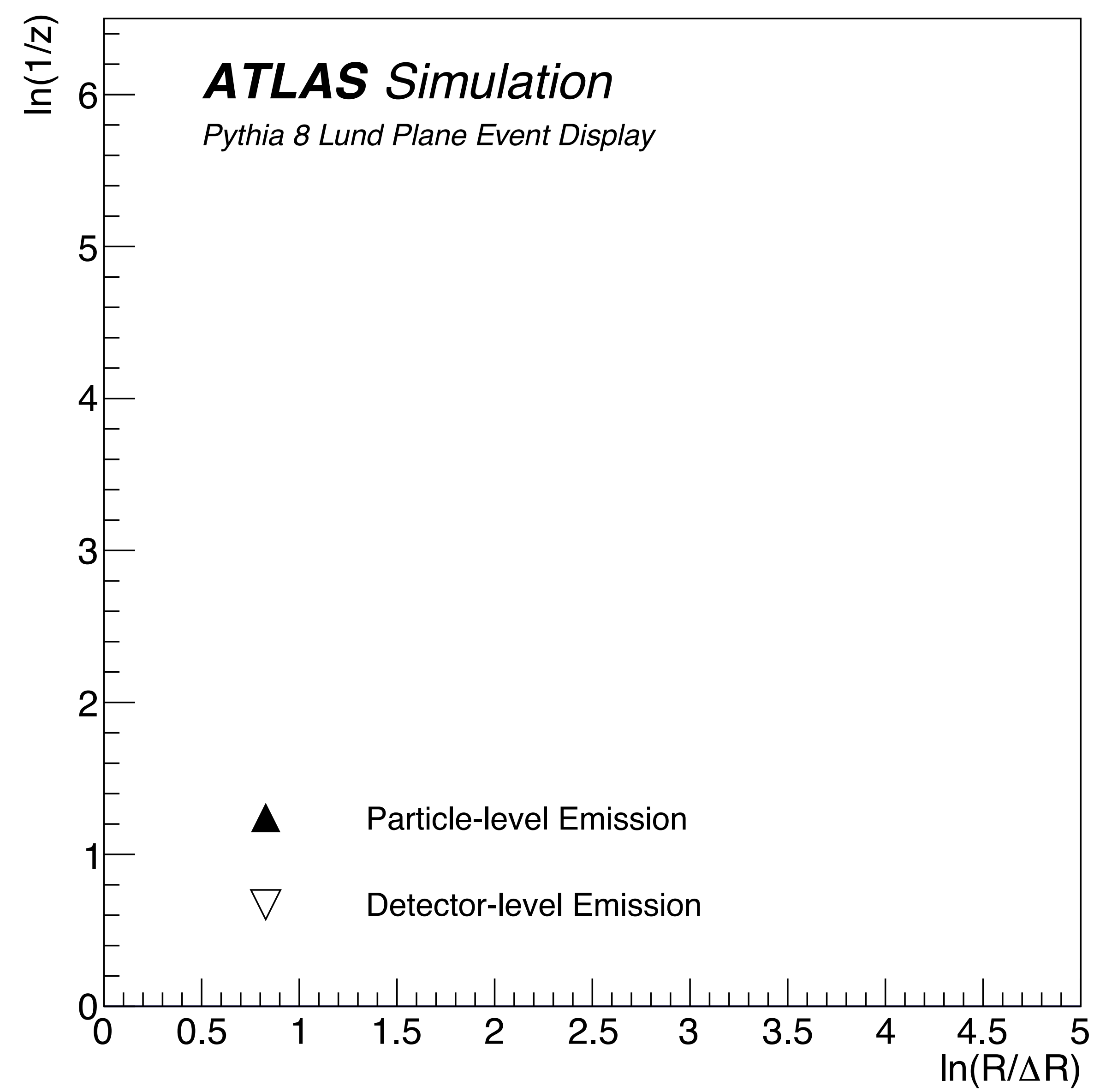


[Dreyer, Salam, Soyez, JHEP 12 (2018) 064]

Phys. Rev. Lett. 124, 222002 (2020)

Categorizing all hard splittings at once



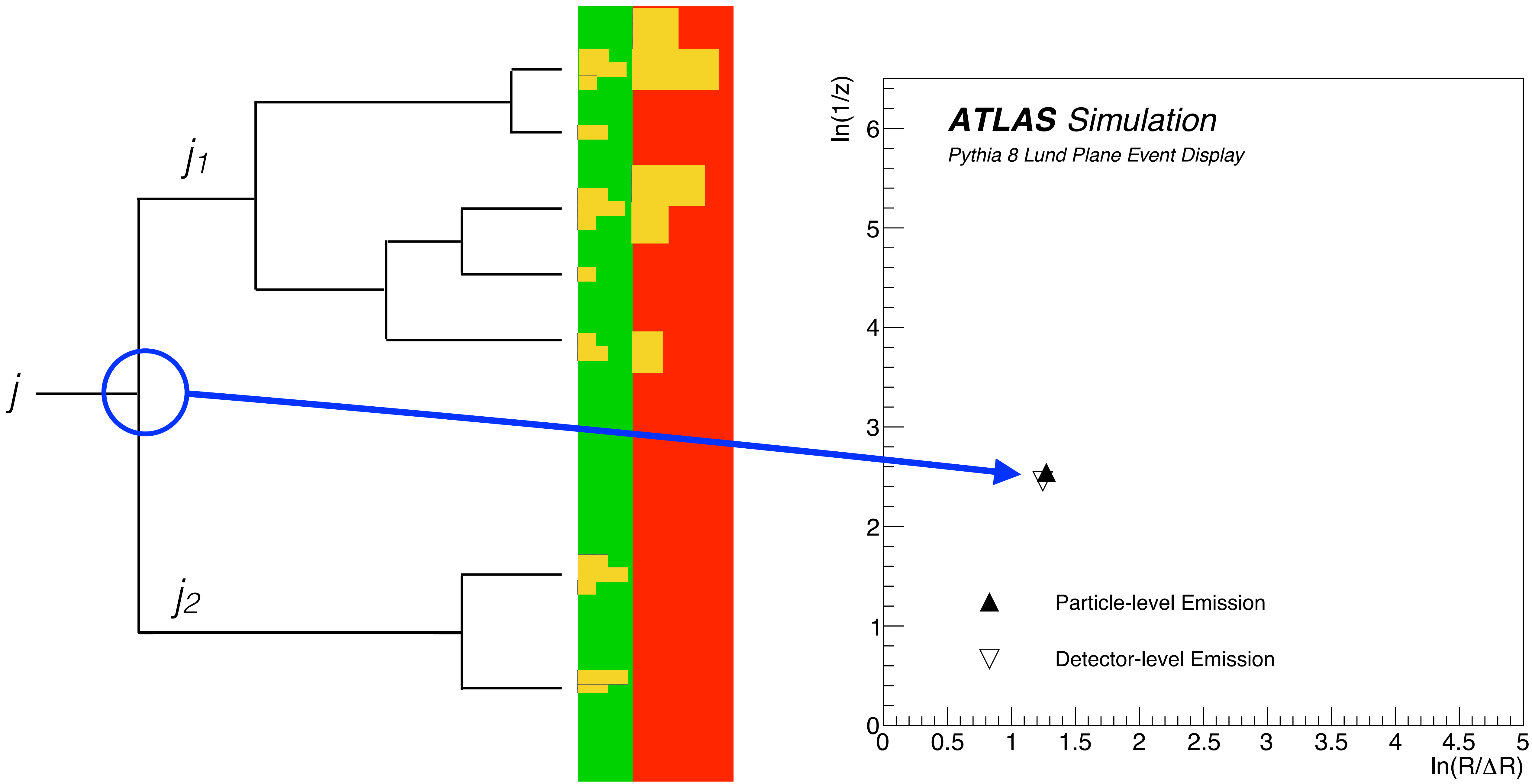


z = j_1 momentum fraction of j

ΔR = angle between j_1 and j_2

Phys. Rev. Lett. 124, 222002 (2020)

Categorizing all hard splittings at once

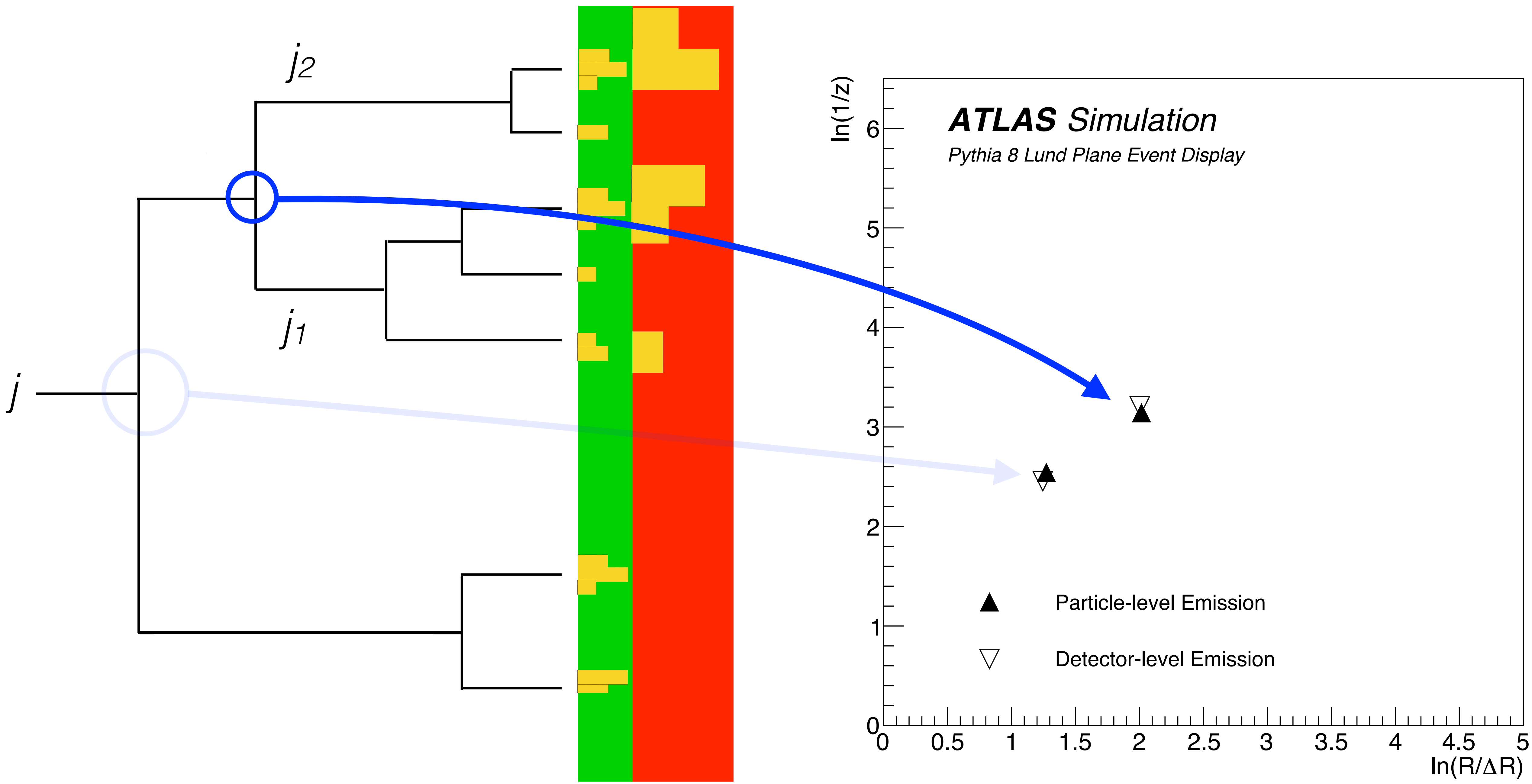


$z = j_1$ momentum fraction of j

ΔR = angle between j_1 and j_2

Phys. Rev. Lett. 124, 222002 (2020)

Categorizing all hard splittings at once

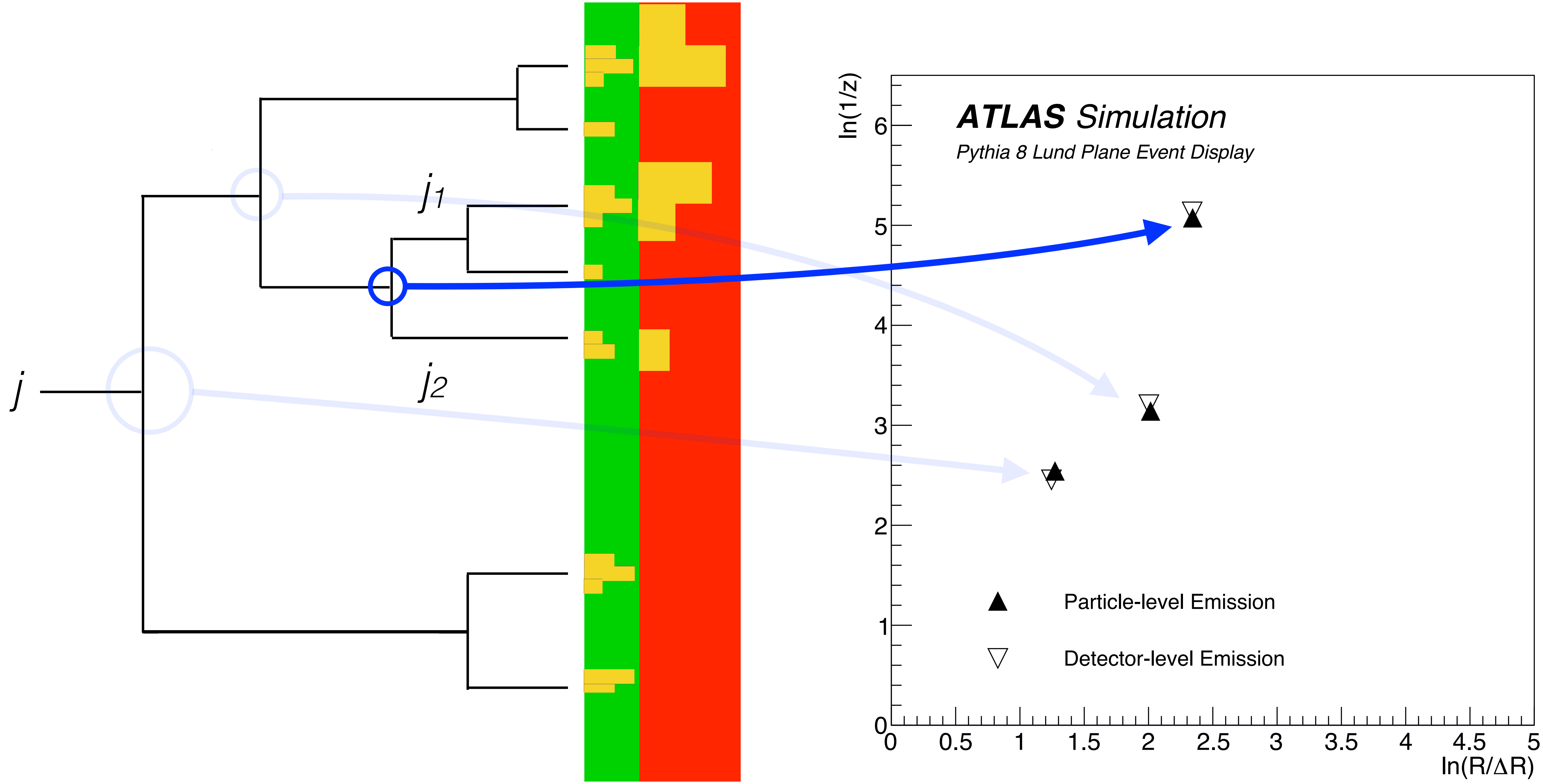


z = j_1 momentum fraction of j

ΔR = angle between j_1 and j_2

Phys. Rev. Lett. 124, 222002 (2020)

Categorizing all hard splittings at once

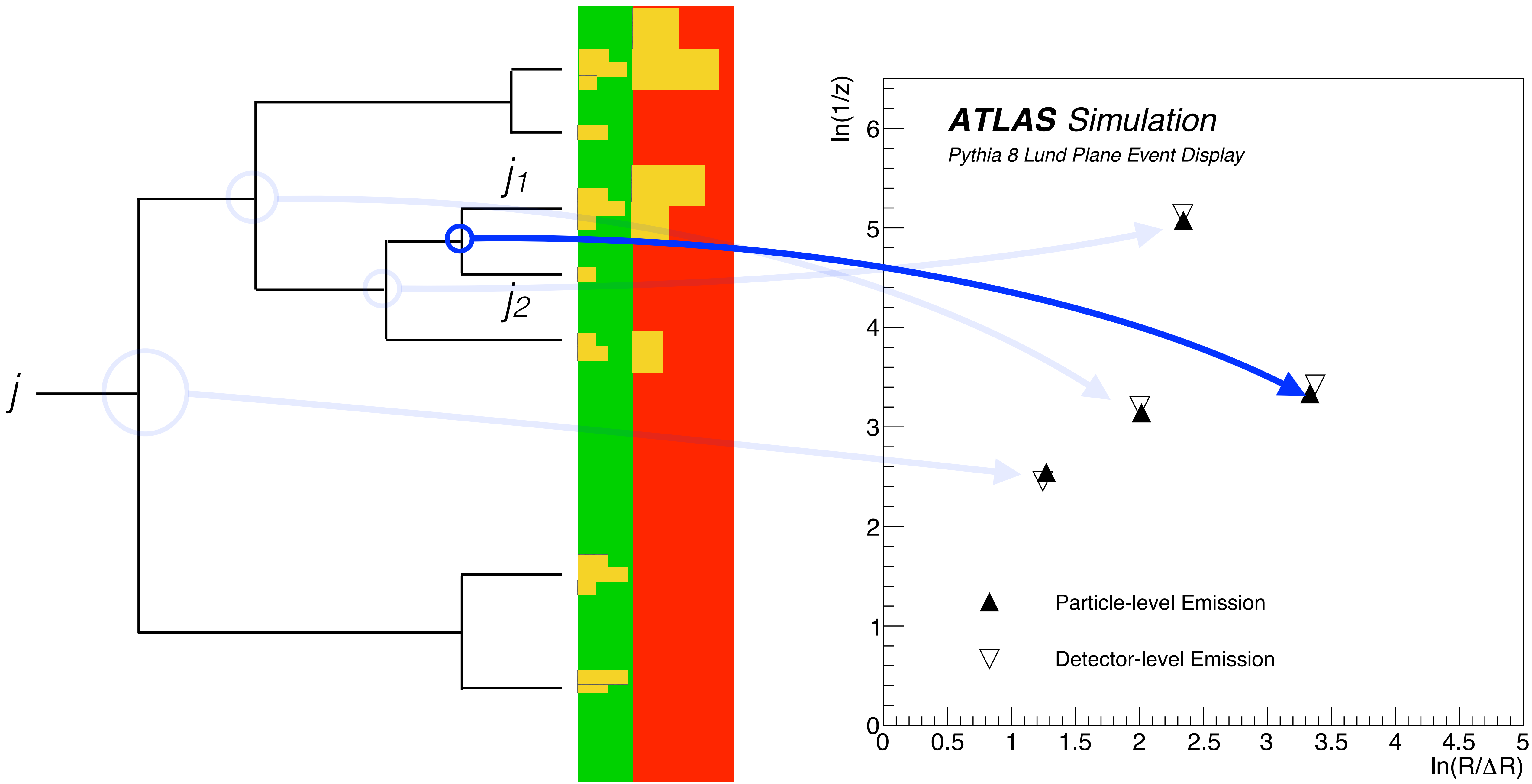


z = j_1 momentum fraction of j

ΔR = angle between j_1 and j_2

Phys. Rev. Lett. 124, 222002 (2020)

Categorizing all hard splittings at once

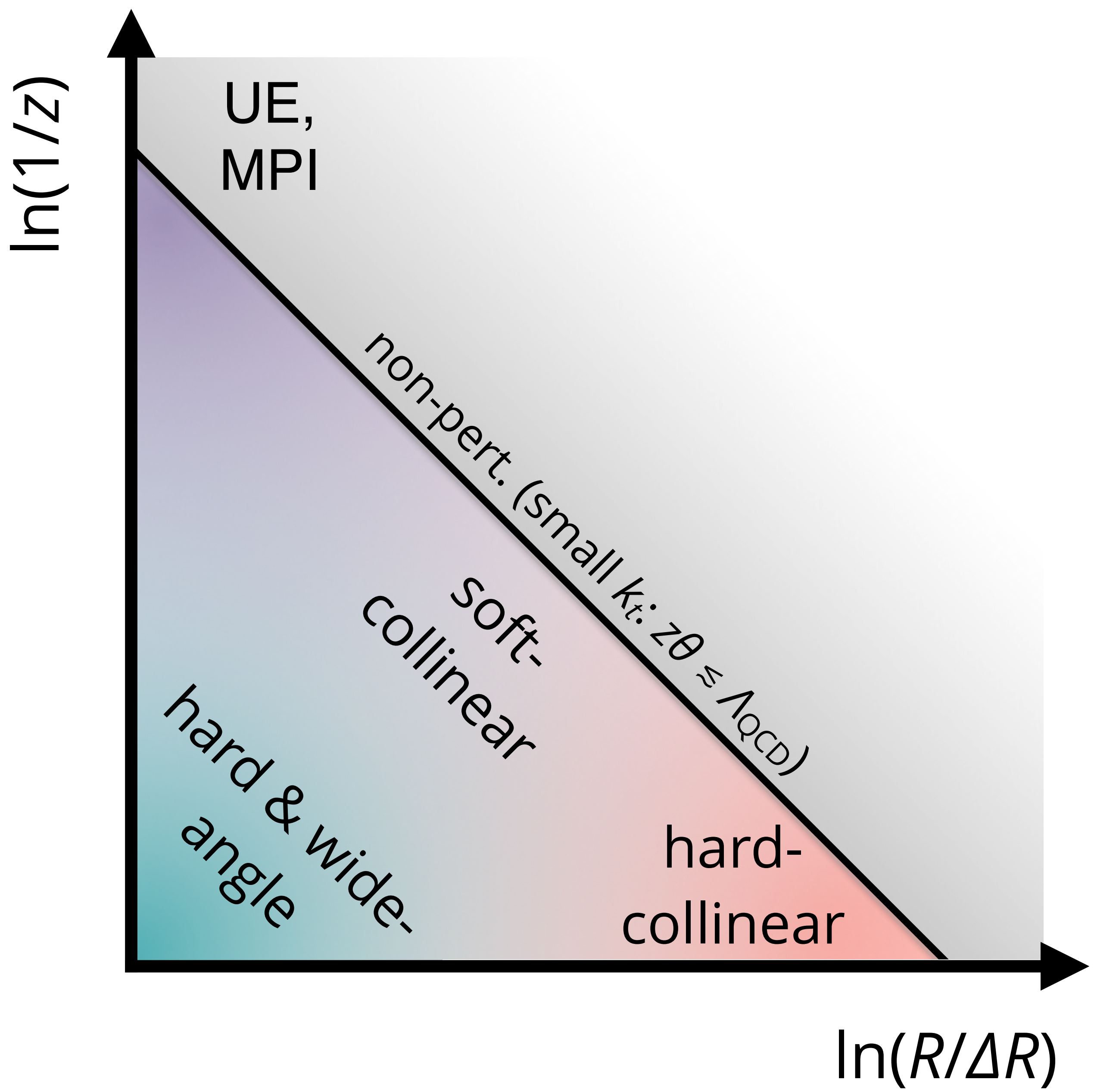


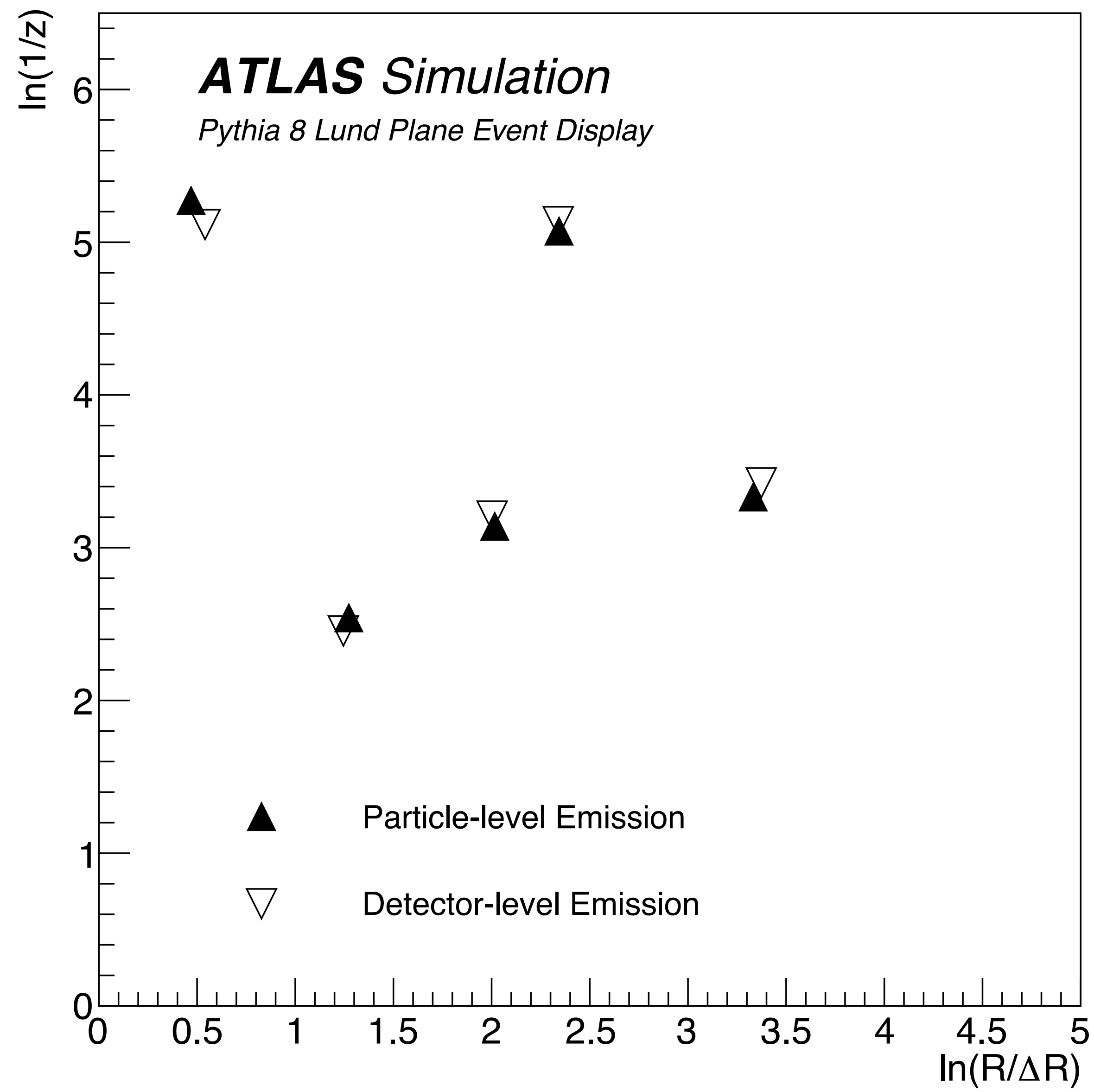
z = j_1 momentum fraction of j

ΔR = angle between j_1 and j_2

Phys. Rev. Lett. 124, 222002 (2020)

Categorizing all hard splittings at once

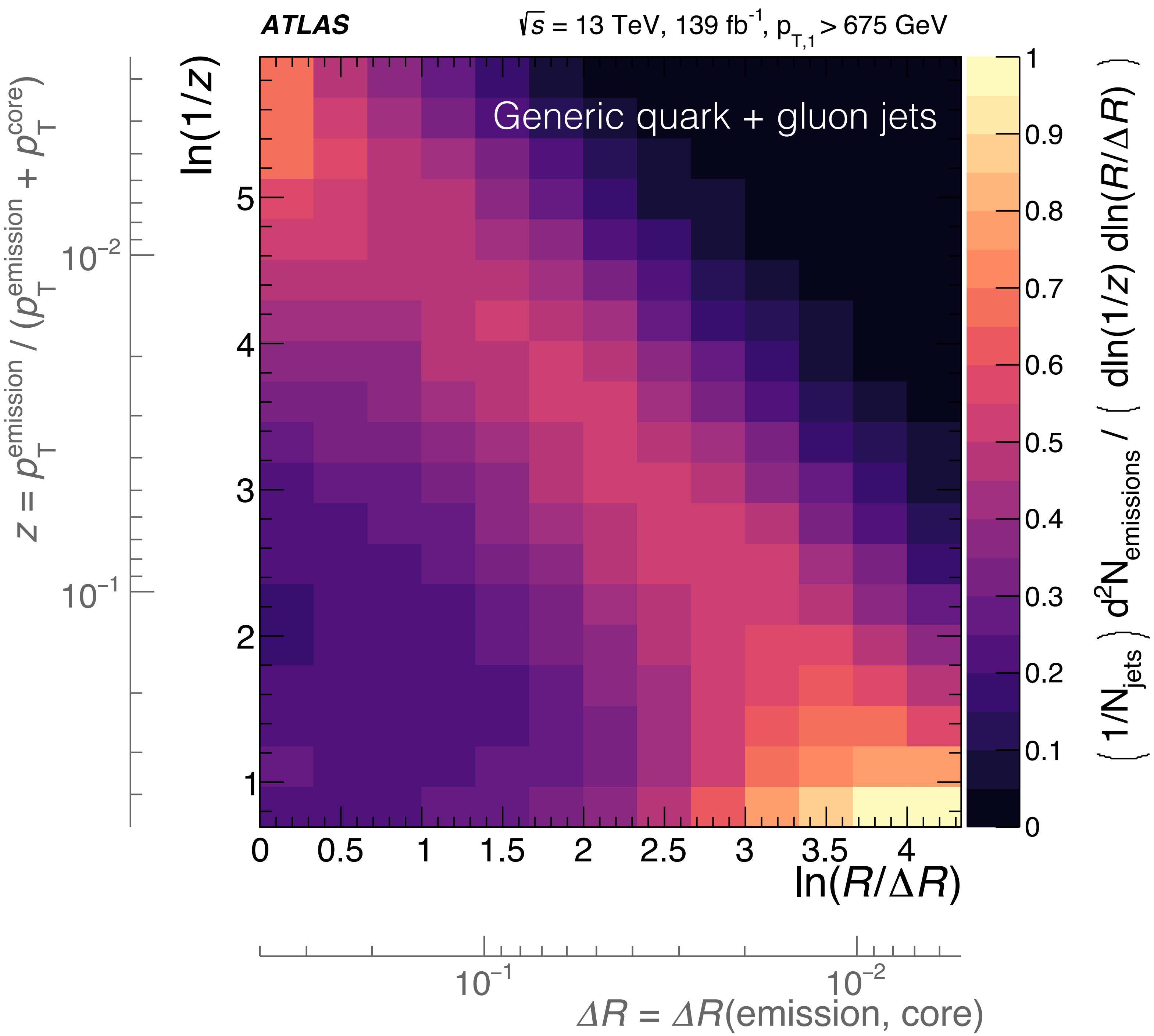




Factorize physical processes!

Phys. Rev. Lett. 124, 222002 (2020)

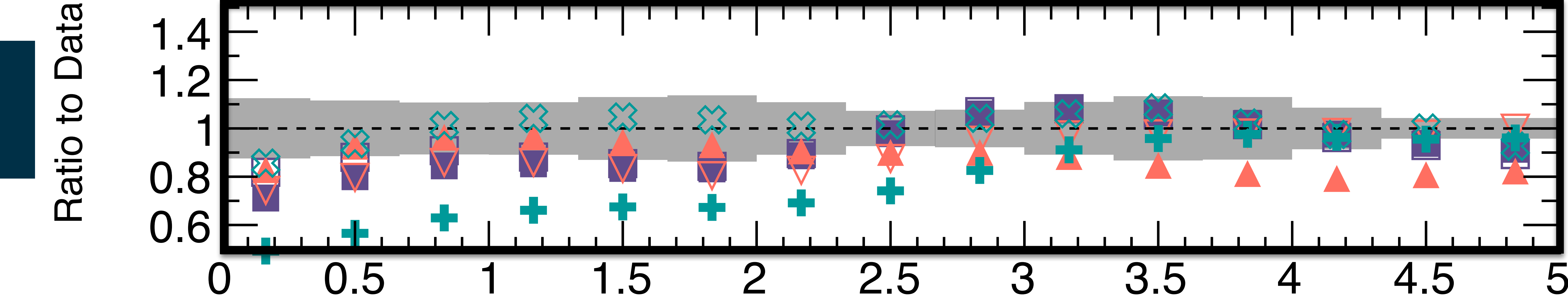
Categorizing all hard splittings at once

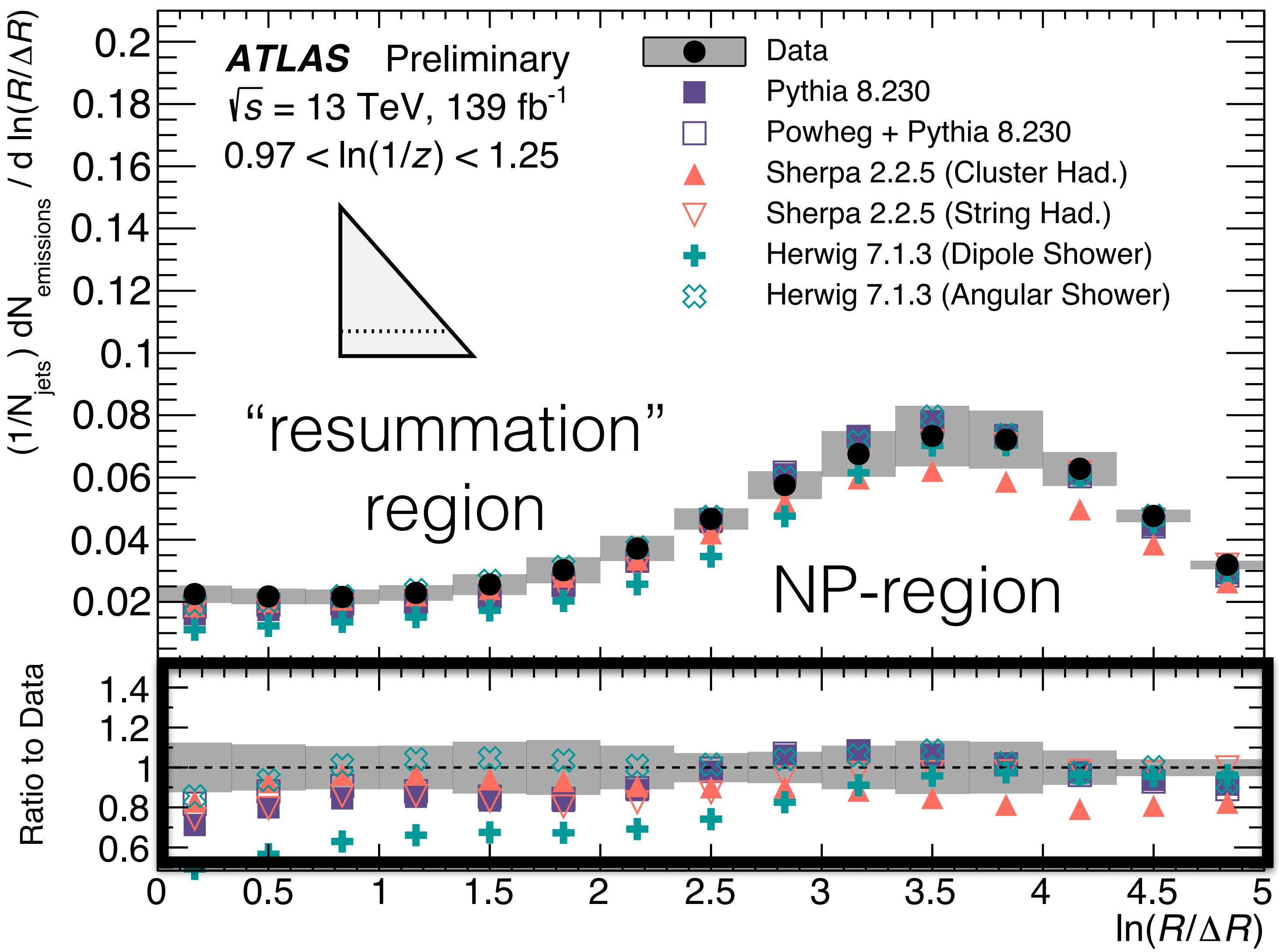


First measurement of the Lund jet plane!

…powerful tool for isolating hadronization, parton shower effects, and fixed-order effects

Key experimental challenge: **tracking inside dense environments**





▲ vs. ▽ hadronization

✚ vs. ✖ parton shower

First measurement of the Lund jet plane!

…powerful tool for isolating ***hadronization****,* ***parton shower effects****, and fixed-order effects*

Key experimental challenge:

tracking inside dense environments

The Future: Machine Learning

In addition to new theoretical and experimental insights, **machine learning** holds great potential for jet substructure

Theory of everything

↓

Physics simulators

↓

Detector-level observables

↓

Pattern recognition ⟷ Pattern recognition

Nature

↓

Experiment

↓

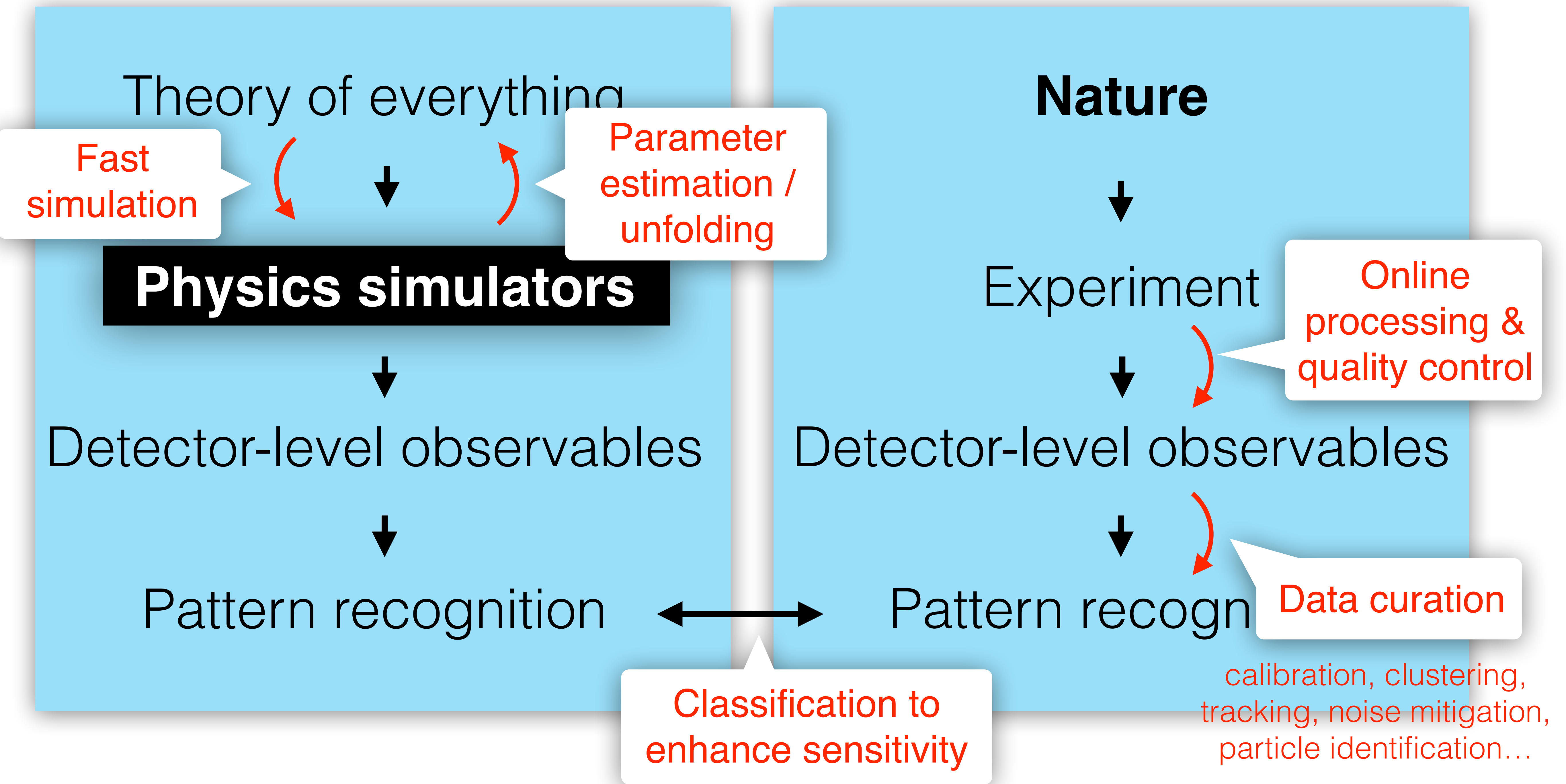
Detector-level observables

↓

Pattern recognition

The Future: Machine Learning

In addition to new theoretical and experimental insights, **machine learning** holds great potential for jet substructure



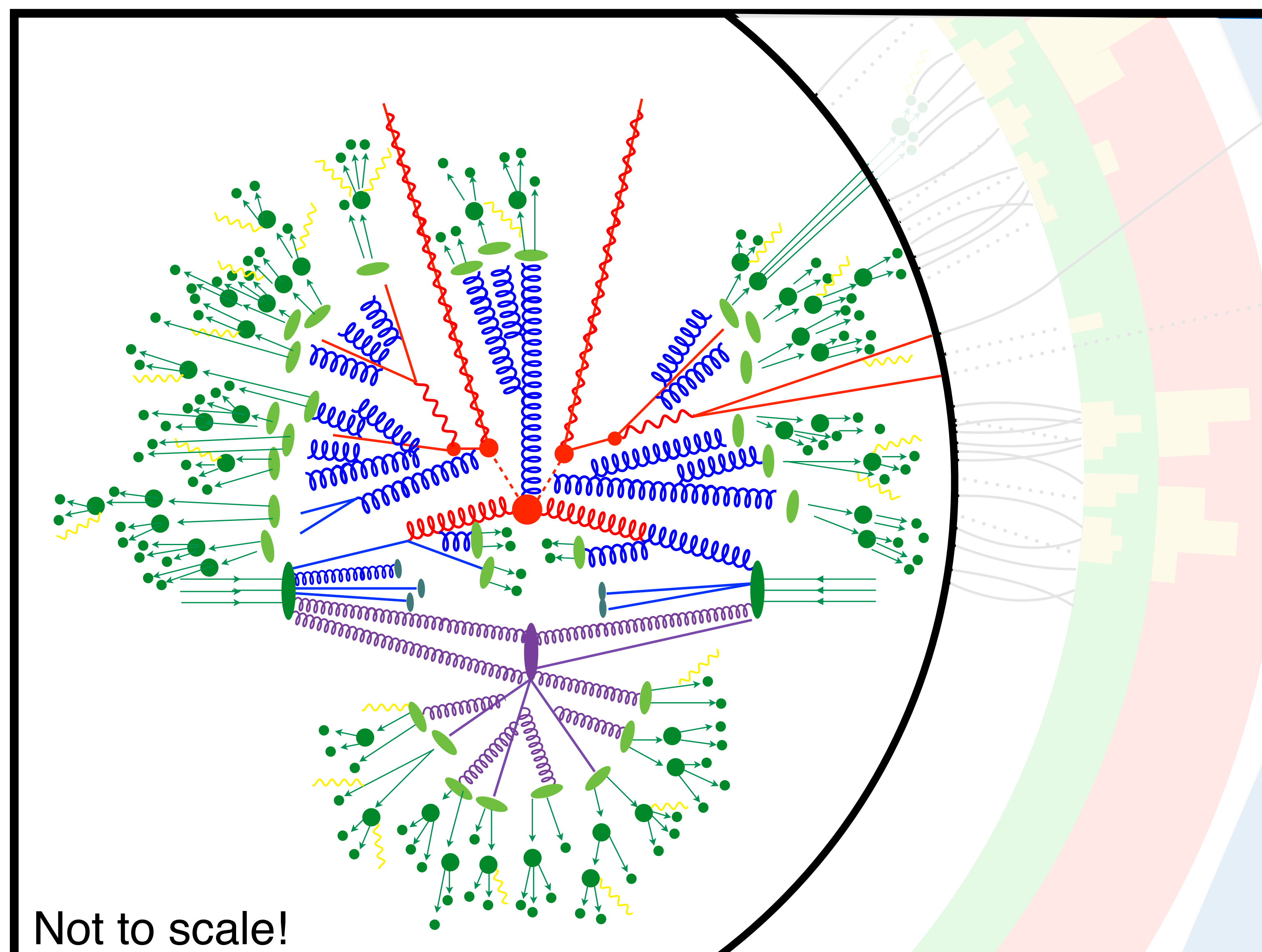
Hadronic Final States: A *hyper* challenge

Key **challenge** and **opportunity**: *hypervariate phase space* & *hyper spectral data*

Typical collision events at the LHC produce **O(1000+)** particles

We detect these particles with **O(100 M)** readout channels

Image inspired by JHEP 02 (2009) 007



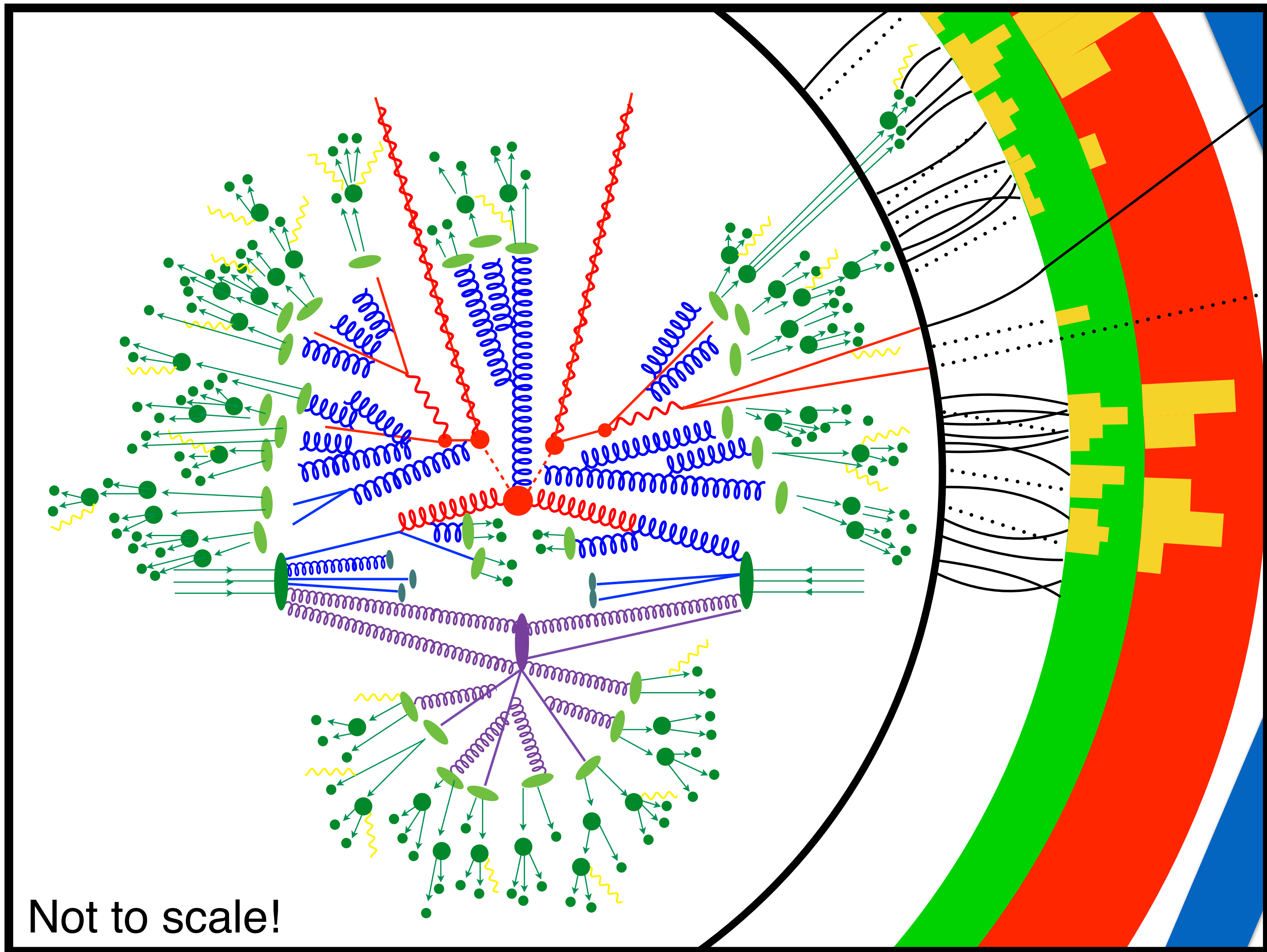
Hadronic Final States: A *hyper* challenge

Key **challenge** and **opportunity**: *hypervariate phase space*
& *hyper spectral data*

Typical collision events at the LHC produce **O(1000+)** particles

We detect these particles with **O(100 M)** readout channels

Image inspired by JHEP 02 (2009) 007



ML for Hadronic Final States

The rest of this talk will be about new ideas for ML for (1) measurements and (2) searches in hadronic final states.

ML for Hadronic Final States

The rest of this talk will be about new ideas for ML for (1) measurements and (2) searches in hadronic final states.

For (1), I will use the example of unfolding.

What if we could measure a full event instead of projecting it into a single event/jet shape?

ML for Hadronic Final States

The rest of this talk will be about new ideas for ML for (1) measurements and (2) searches in hadronic final states.

For (1), I will use the example of unfolding.

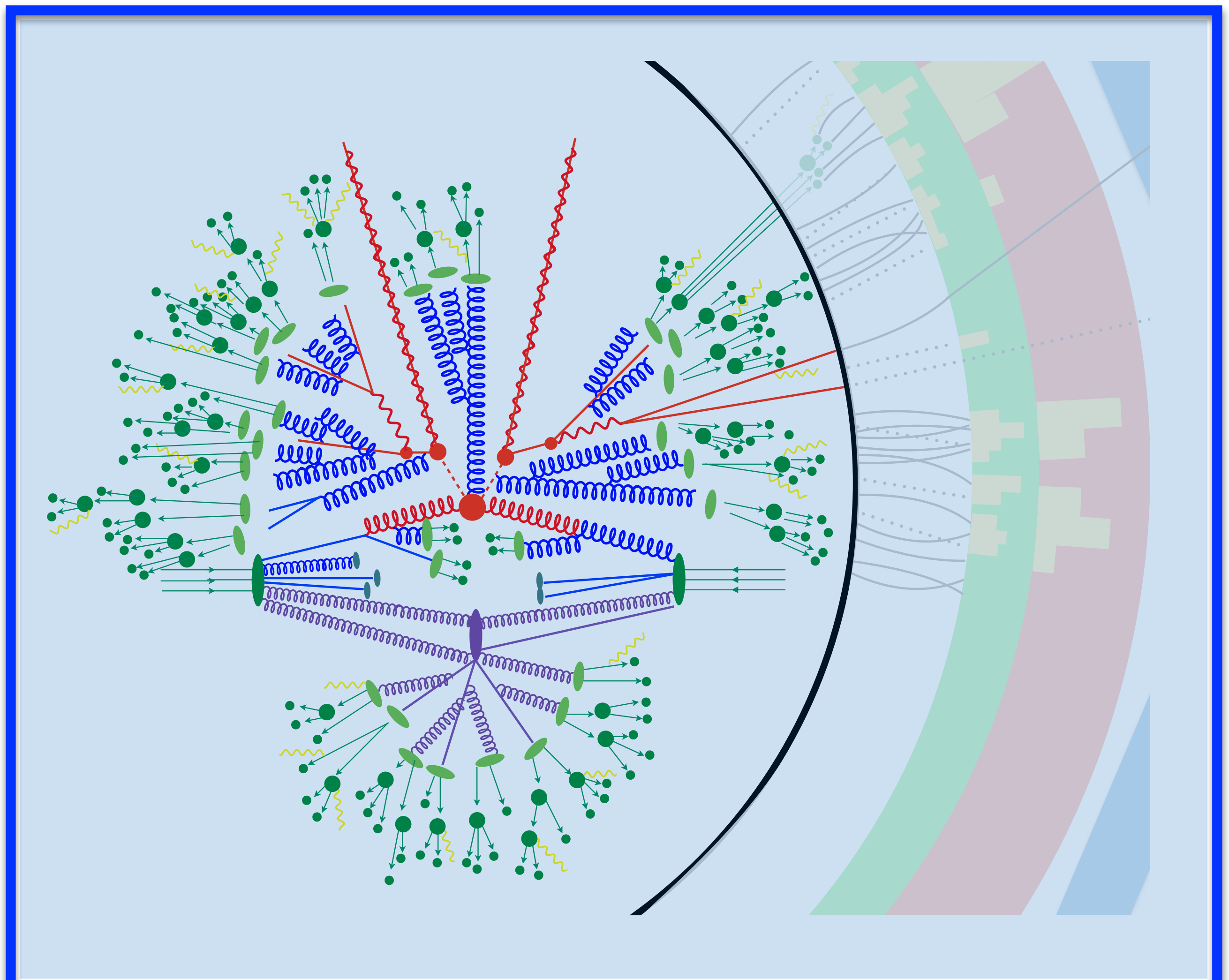
What if we could measure a full event instead of projecting it into a single event/jet shape?

For (2), I will use the example of anomaly detection.

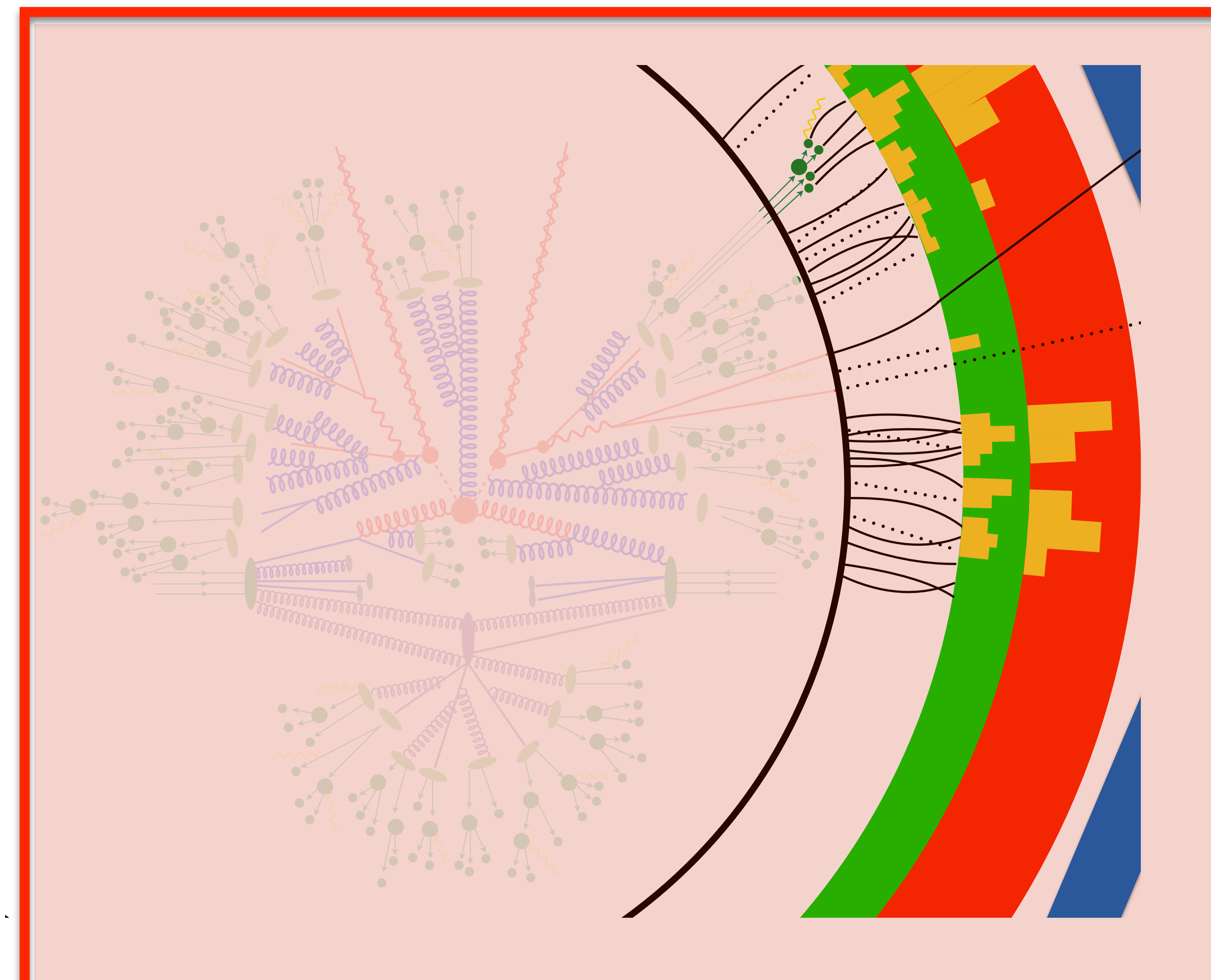
What if we could look for new physics without having a particular model in mind?

(1) Unfolding (Deconvolution)

Want this



Measure this

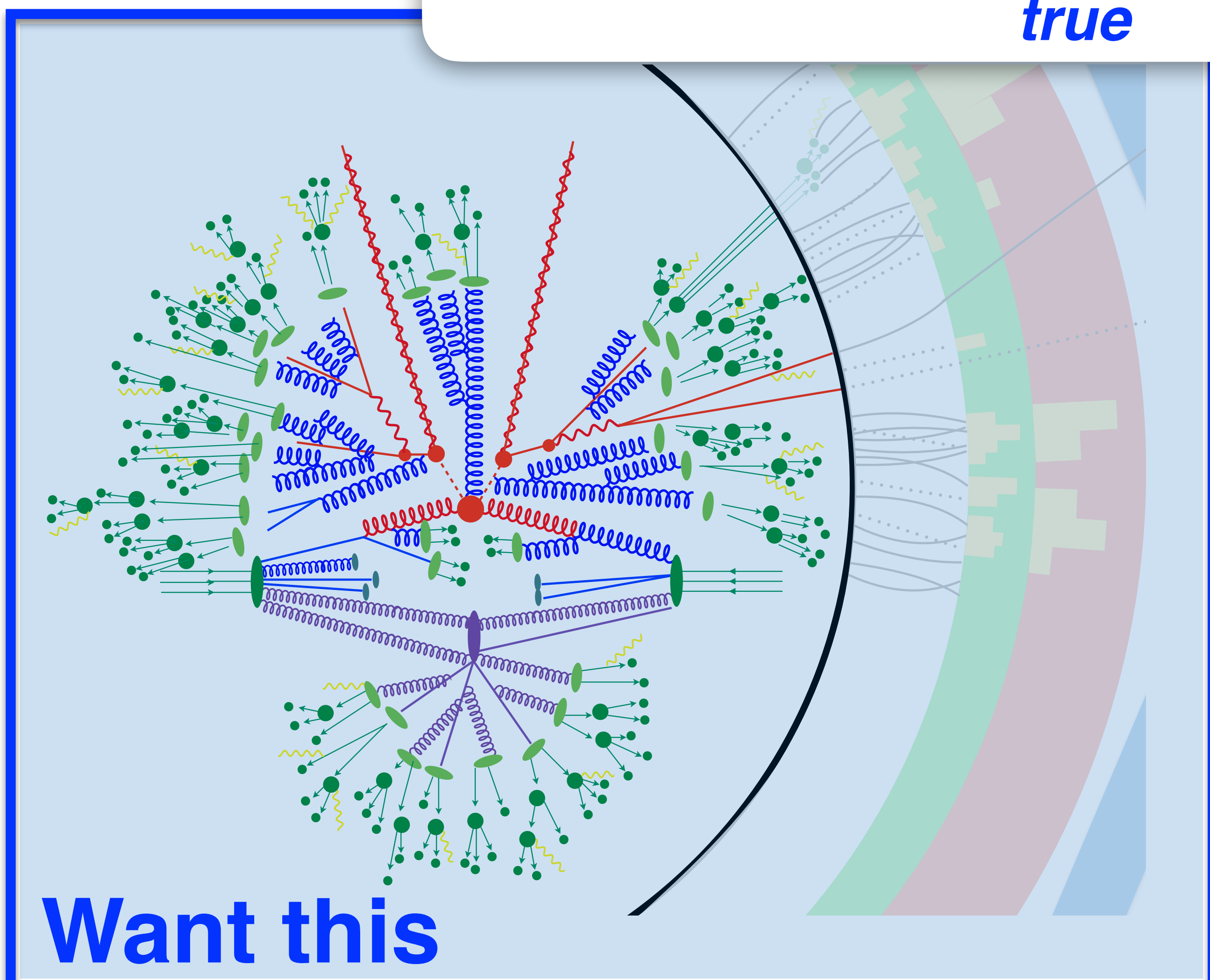


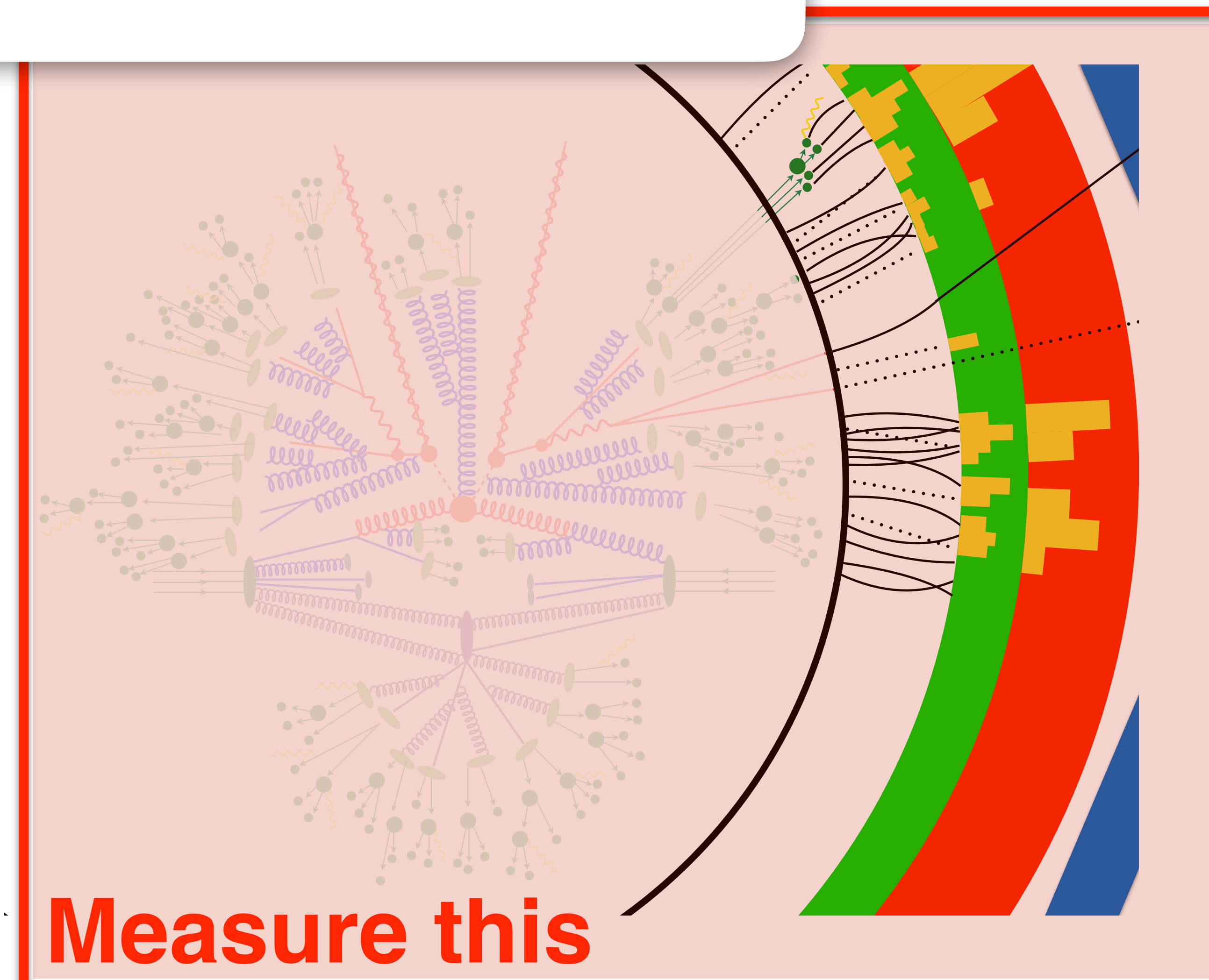
i.e. remove detector distortions

(1) Unfolding (Deconvolution)

If you know *p(meas. | true)*, could do maximum likelihood, i.e.

$$\textit{unfolded} = \underset{\textit{true}}{\operatorname{argmax}}\ p(\textit{measured} \mid \textit{true})$$





p(meas. | true) = "response matrix" or "point spread function"

(1) Unfolding (Deconvolution)

If you know *p(**meas.** / **true**)*, could do maximum likelihood, i.e.

$$\text{unfolded} = \underset{\text{true}}{\operatorname{argmax}}\ p(\text{measured} \,/\, \text{true})$$

Challenge: **measured** is hyperspectral and **true** is hypervariate … *p(**meas.** | **true**) is **intractable** !*

*p(**meas.** / **true**)* = "response matrix" or "point spread function"

(1) Unfolding (Deconvolution)

If you know *p(**meas.** / **true**)*, could do maximum likelihood, i.e.

$$\textit{unfolded} = \underset{\textit{true}}{\operatorname{argmax}}\; p(\textit{measured} \,/\, \textit{true})$$

Challenge: **measured** is hyperspectral and **true** is hypervariate … *p(**meas.** | **true**) is **intractable** !*

However: we have **simulators** that we can use to sample from *p(**meas.** | **true**)*

→ **Simulation-based (likelihood-free) inference**

*p(**meas.** / **true**)* = “response matrix” or “point spread function”

I'll briefly show you one solution to give you a sense of the power of likelihood-free inference.

Reweighting

I'll briefly show you one solution to give you a sense of the power of likelihood-free inference.

The solution will be built on ***reweighting***

dataset 1: sampled from ***p(x)***
dataset 2: sampled from ***q(x)***

Create weights ***w(x) = q(x)/p(x)*** so that when dataset 1 is weighted by ***w***, it is statistically identical to dataset 2.

Reweighting

I'll briefly show you one solution to give you a sense of the power of likelihood-free inference.

The solution will be built on ***reweighting***

dataset 1: sampled from ***p(x)***
dataset 2: sampled from ***q(x)***

Create weights ***w(x) = q(x)/p(x)*** so that when dataset 1 is weighted by ***w***, it is statistically identical to dataset 2.

What if we don't (and can't easily) know ***q*** and ***p***?

Classification for reweighting

Fact: Neutral networks learn to approximate the likelihood ratio = $q(x)/p(x)$

(or something monotonically related to it in a known way)

Solution: train a neural network to distinguish the two datasets!

This turns the problem of **density estimation** (**hard**) into a problem of **classification** (**easy**)

Classification for reweighting

Particularly useful for particle physics, where collisions may produce a variable # of particles which are interchangeable

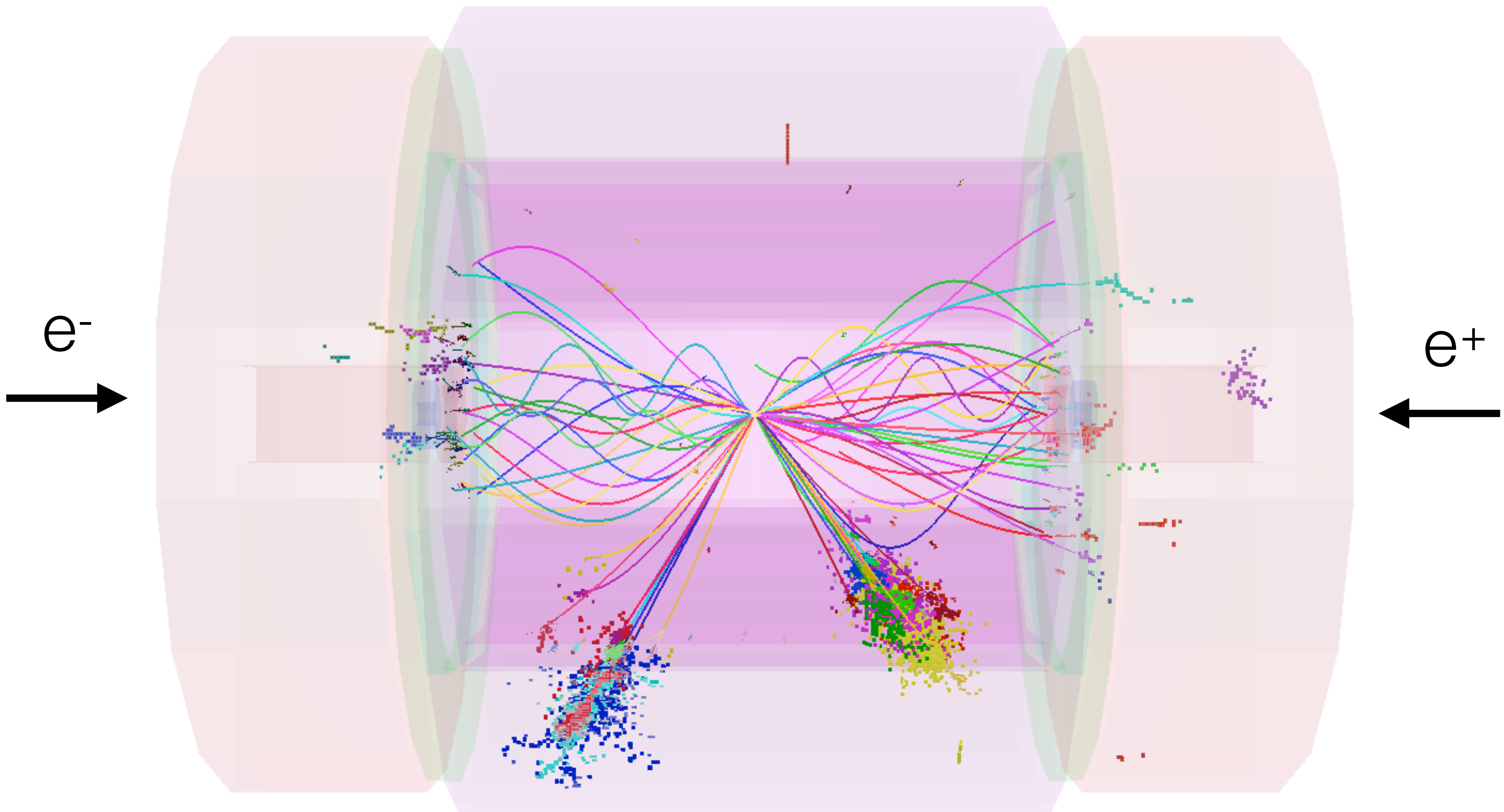


Image: Linear Collider Detector Project

Example: electron-positron collisions

Learn a classifier on the full observable phase space (momenta + particle flavor) and then check with some standard observables.

Our events have a variable number of particles & due to quantum mechanics, are permutation invariant. Thus, we use a deep-sets variant called **particle flow networks**.

PFNs: Komiske, Metodiev, Thaler, JHEP 01 (2019) 121

Deep sets: Zaheer et al., NIPS 2017

Example: electron-positron collisions

Learn a classifier on the full observable phase space (momenta + particle flavor) and then check with some standard obs rvables.

Our events have a variable number of rticles & due to quantum mechanics, a e use a deep-sets varia .

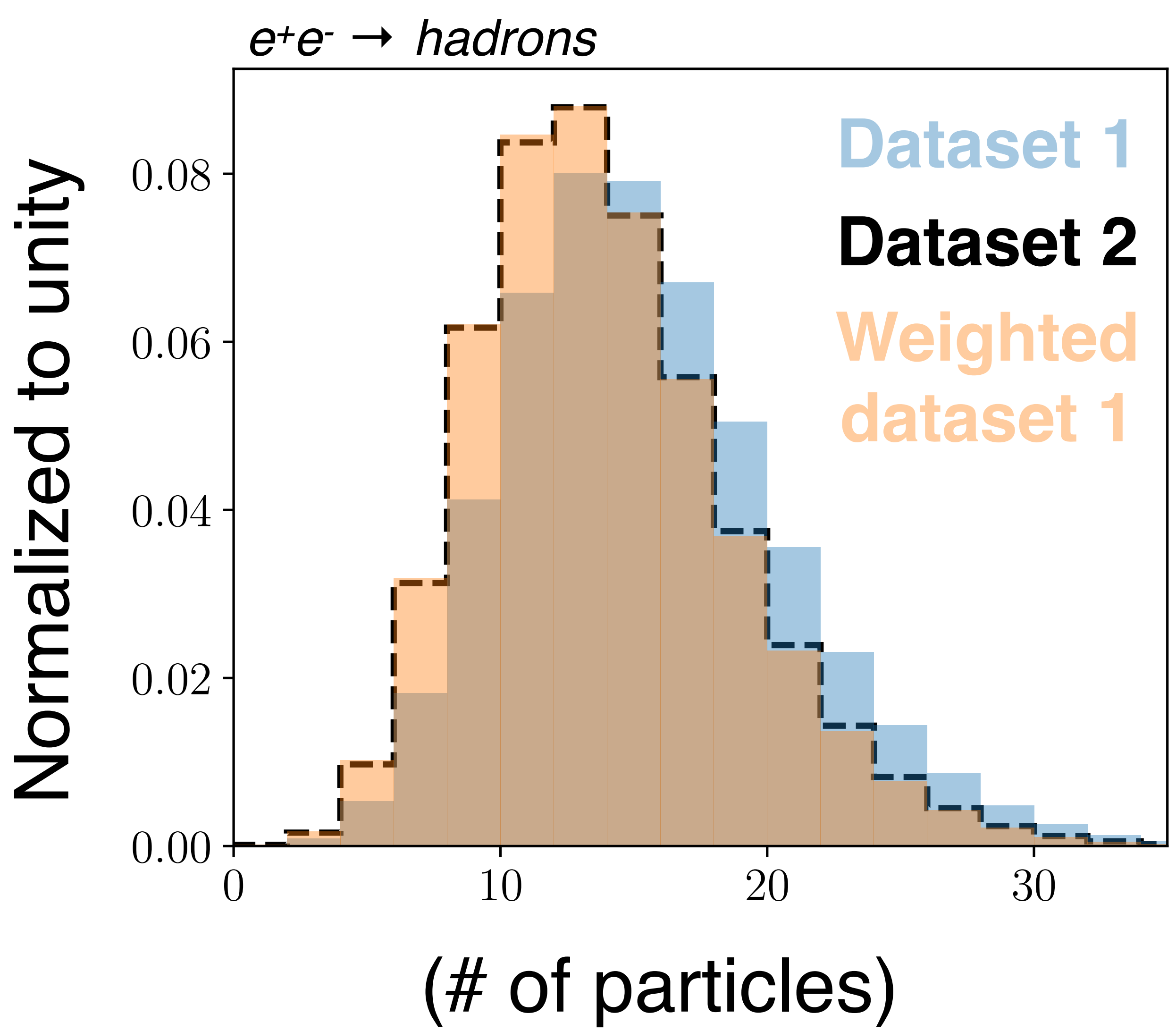
Just to stress: this gives you a new simulation with all the 4-vectors that is statistically indistinguishable.

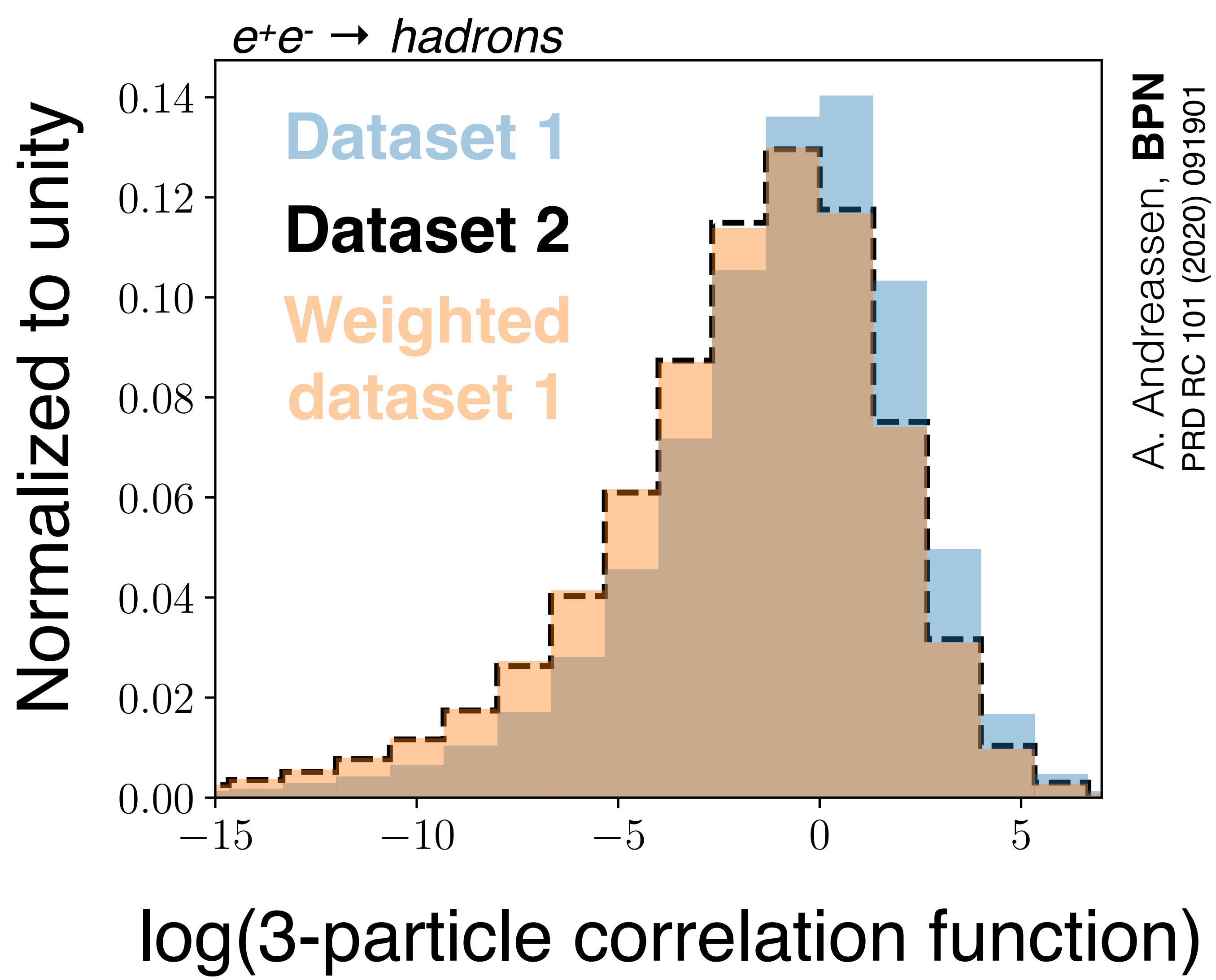
PFNs: Komiske, Metodiev, Thaler, JHEP 01 (2019) 121

Deep sets: Zaheer et al., NIPS 2017

Classification for reweighting

Reweight the **full phase space** and then check for various binned 1D observables.

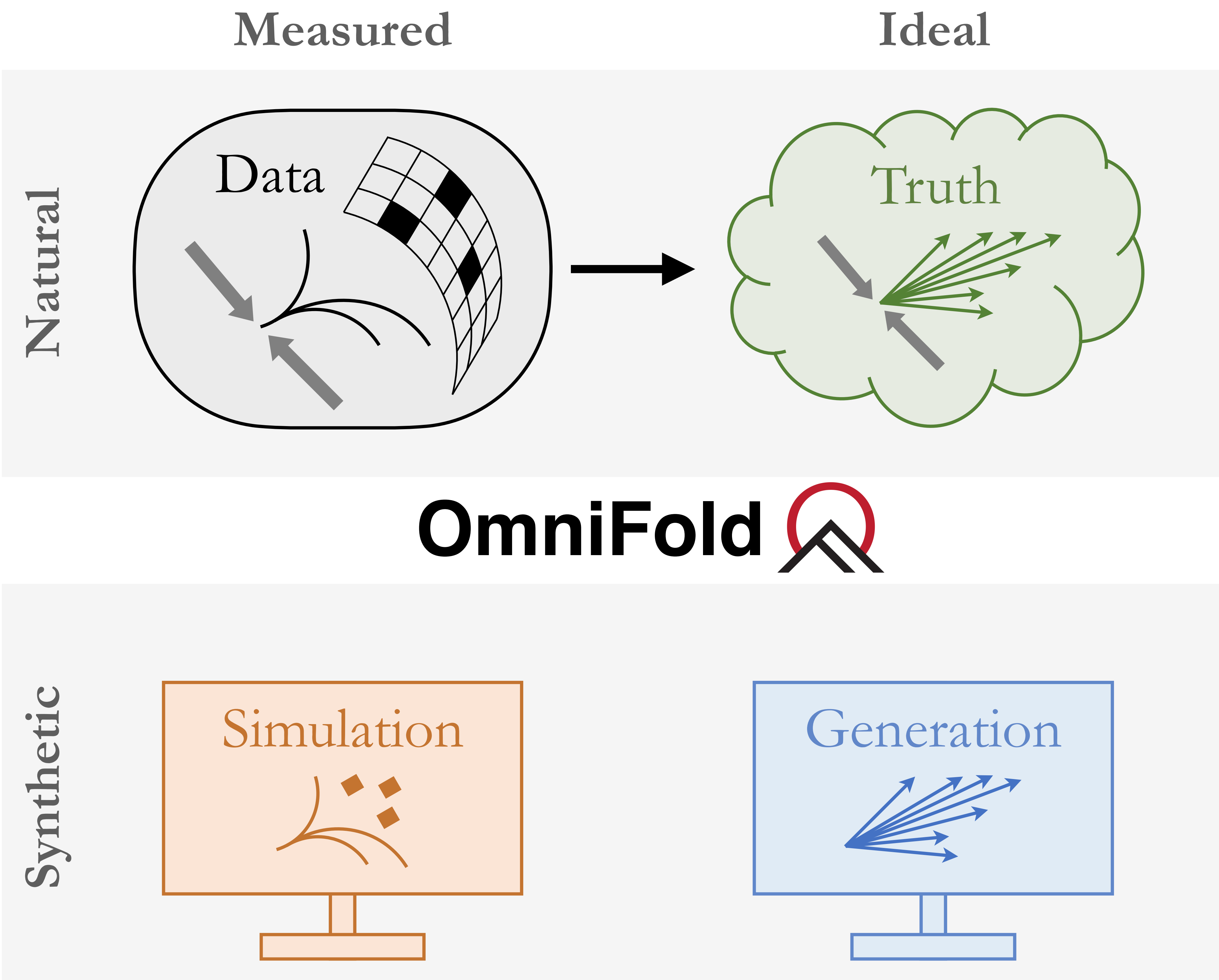




A. Andreassen, **BPN**
PRD RC 101 (2020) 091901

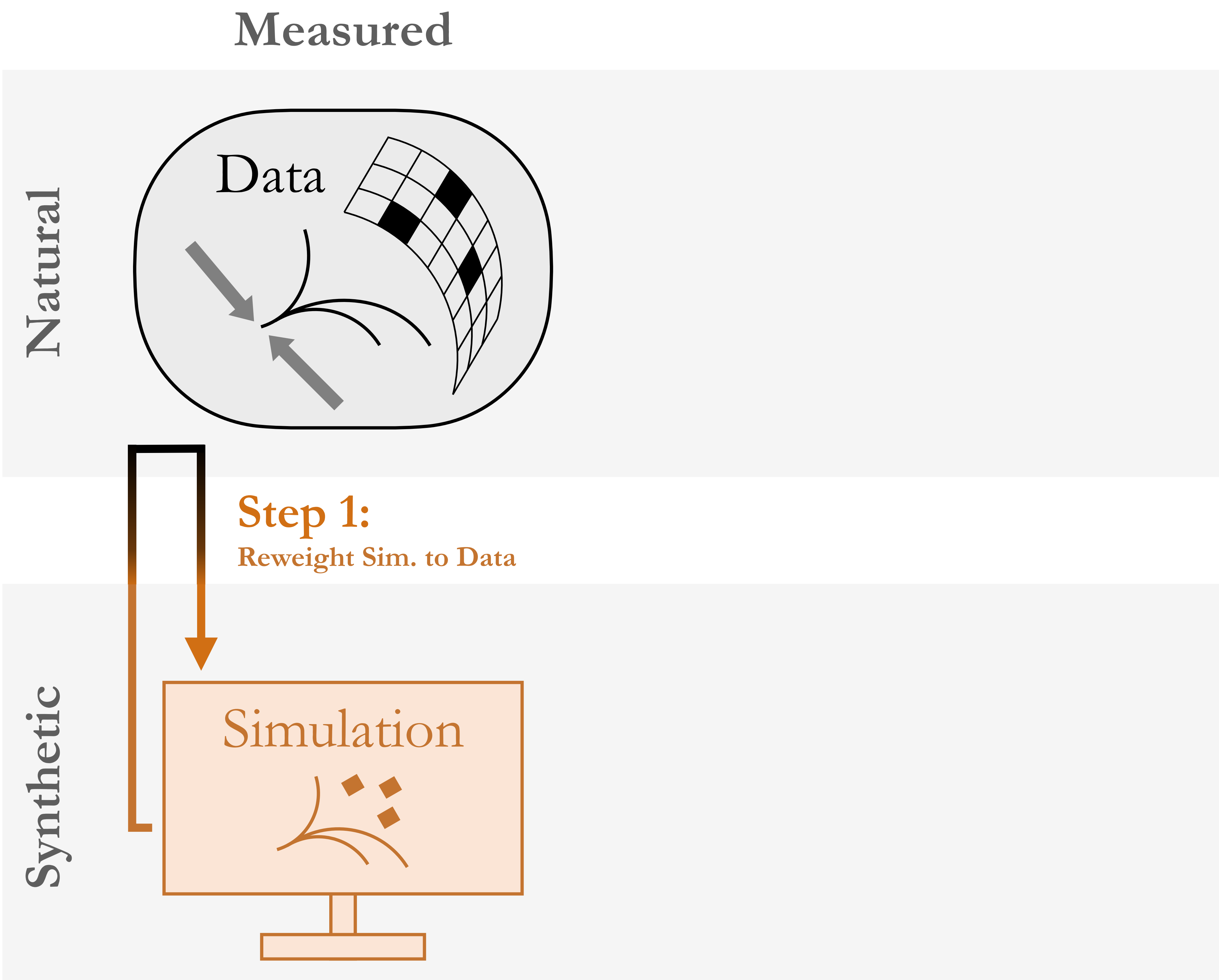
We call this Deep neural networks using Classification for Tuning and Reweighting or "**DCTR**"

Unfolding by iterated reweighting



A. Andreassen, P. Komiske, E. Metodiev, **BPN**, J. Thaler, PRL 124 (2020) 182001

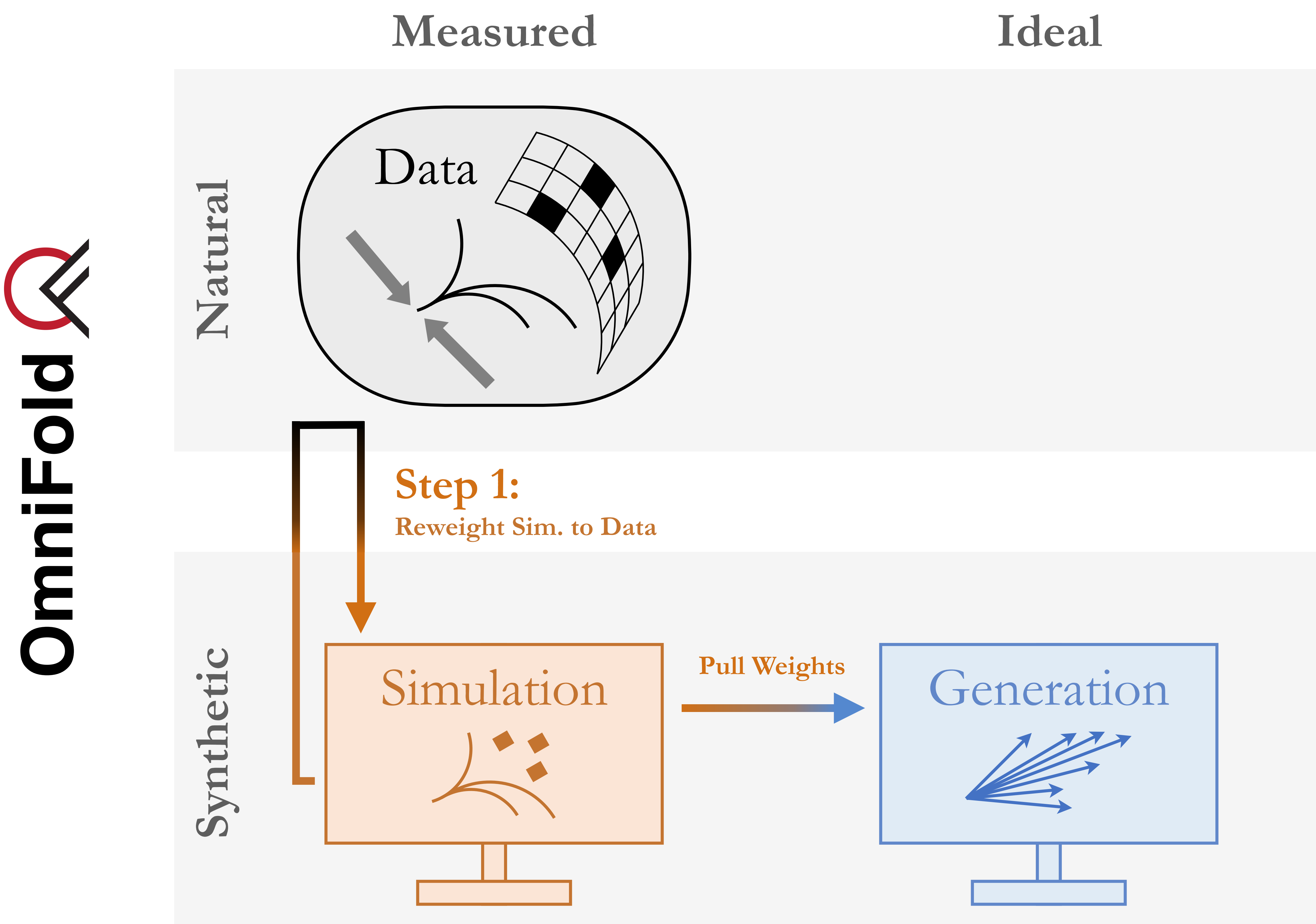
Unfold by iterating: OmniFold



OmniFold

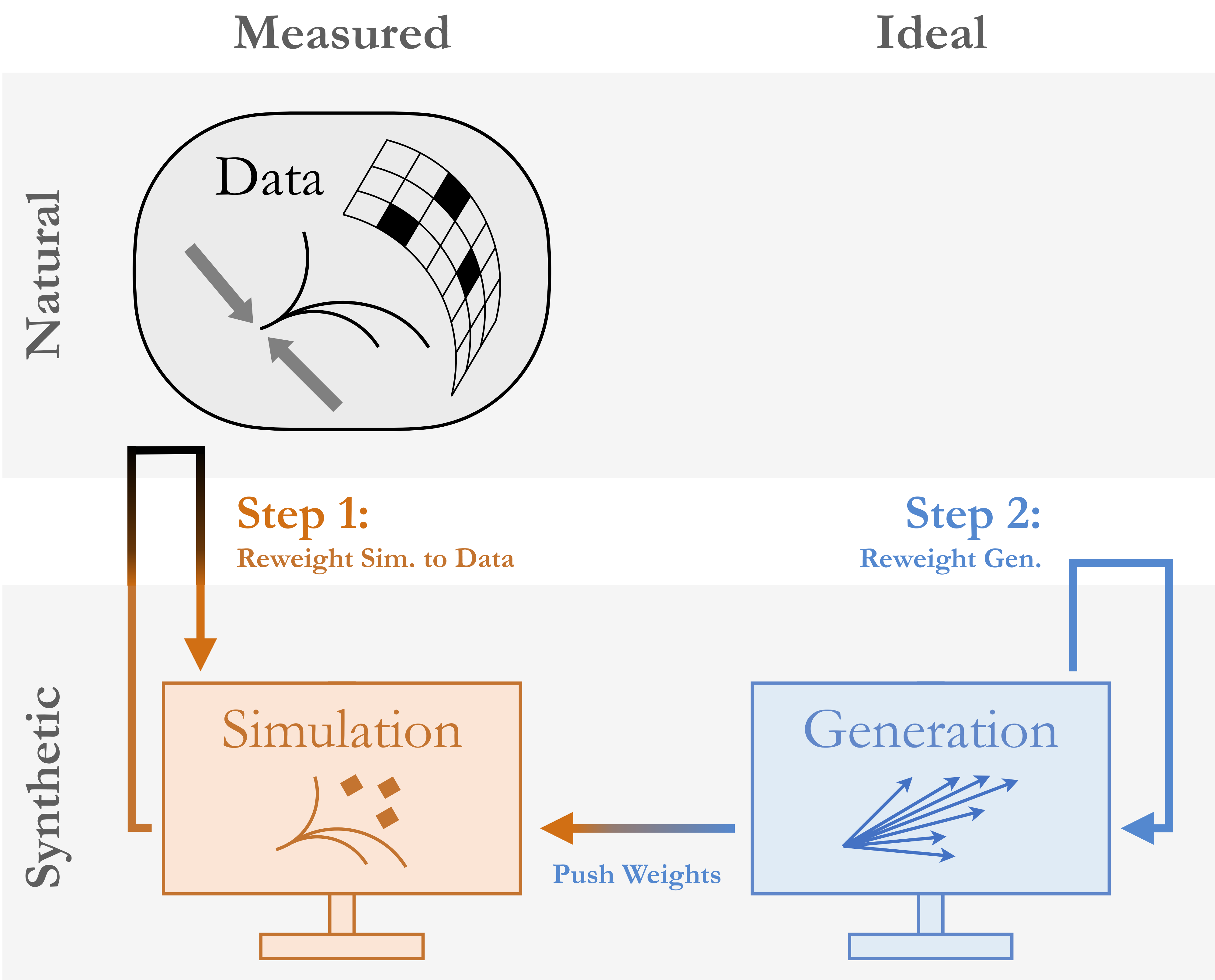
A. Andreassen, P. Komiske, E. Metodiev, **BPN**, J. Thaler, PRL 124 (2020) 182001

Unfold by iterating: OmniFold



A. Andreassen, P. Komiske, E. Metodiev, **BPN**, J. Thaler, PRL 124 (2020) 182001

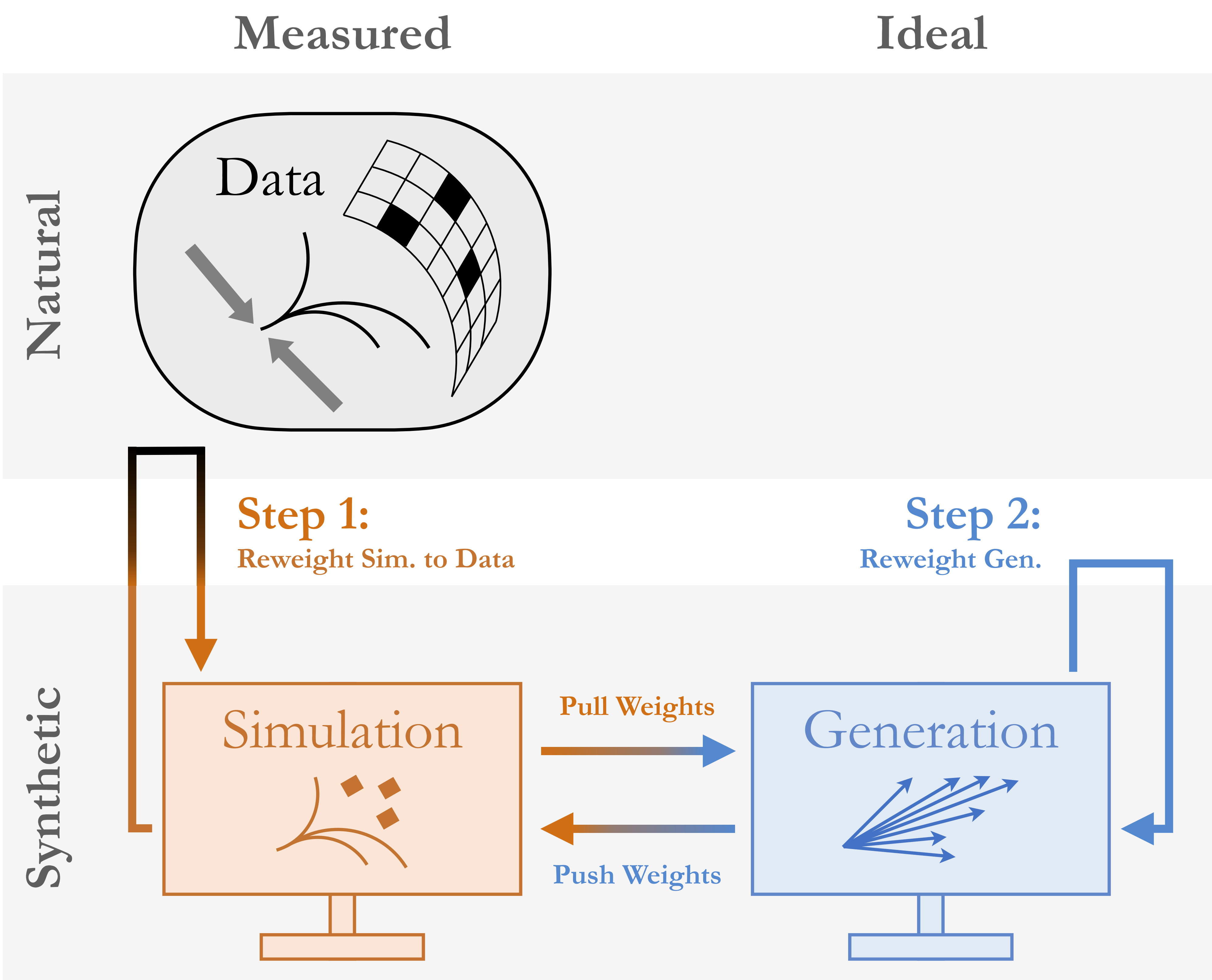
Unfold by iterating: OmniFold



OmniFold

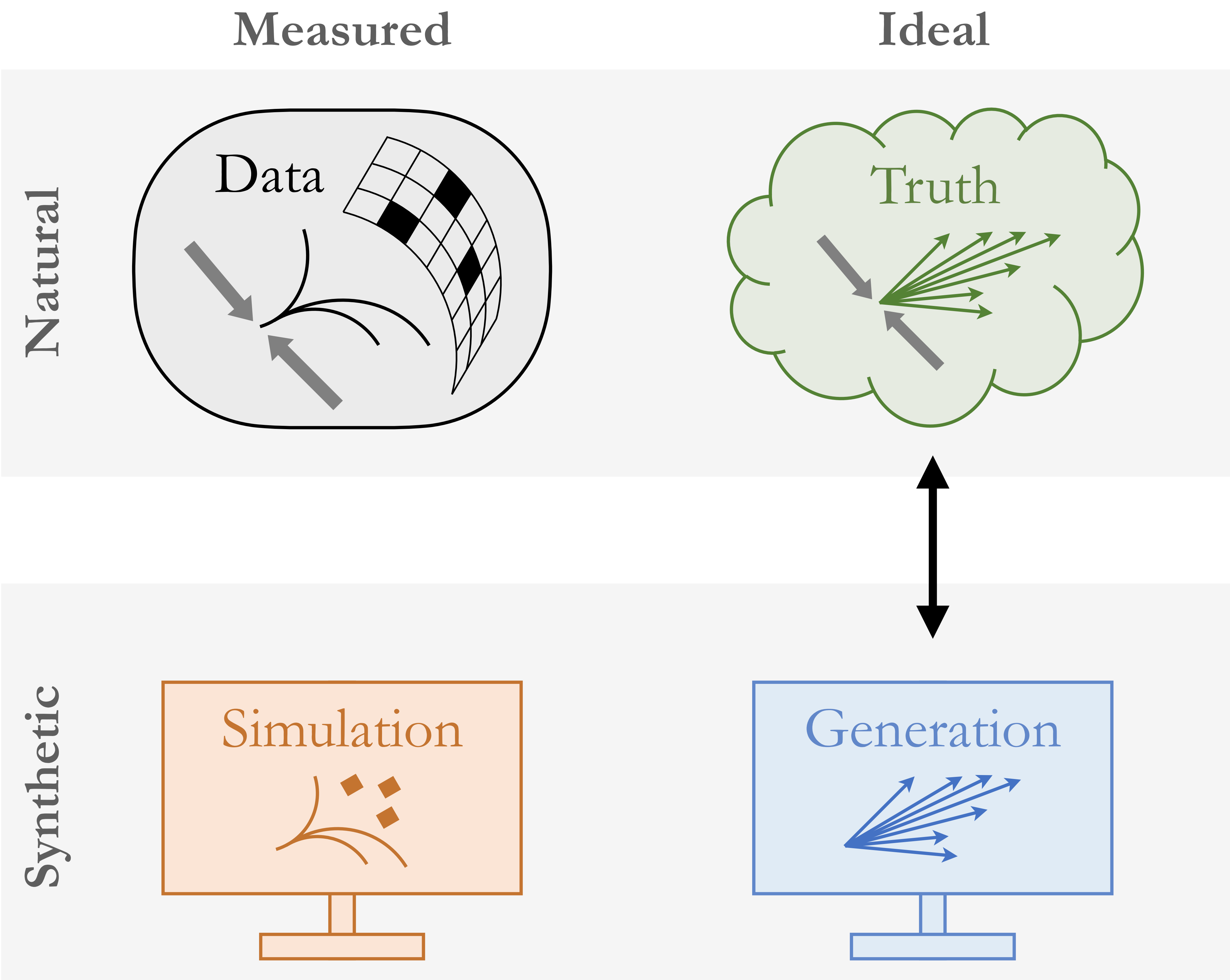
A. Andreassen, P. Komiske, E. Metodiev, **BPN**, J. Thaler, PRL 124 (2020) 182001

Unfold by iterating: OmniFold



A. Andreassen, P. Komiske, E. Metodiev, **BPN**, J. Thaler, PRL 124 (2020) 182001

Unfold by iterating: OmniFold

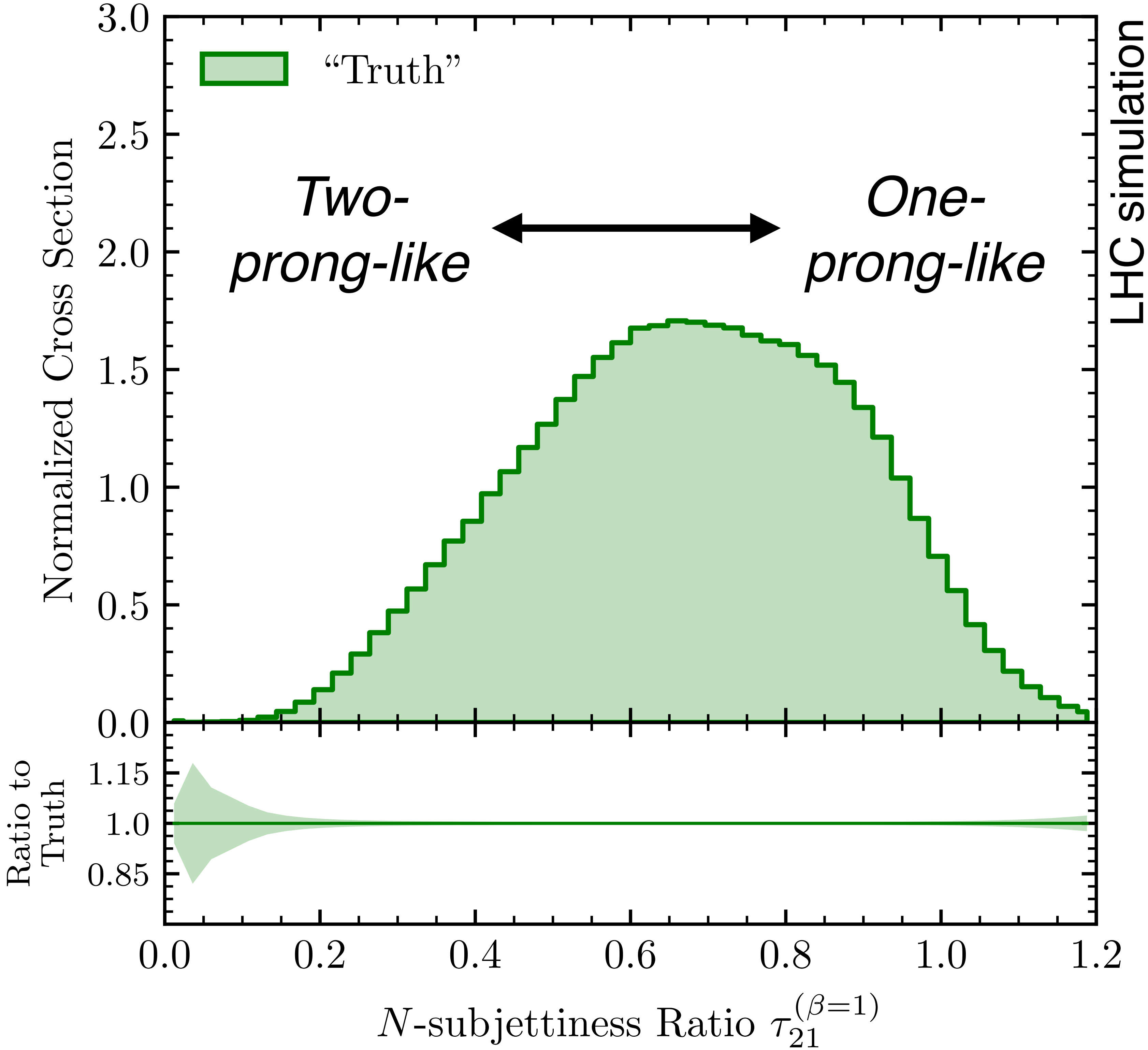


OmniFold

A. Andreassen, P. Komiske, E. Metodiev, **BPN**, J. Thaler, PRL 124 (2020) 182001

Results

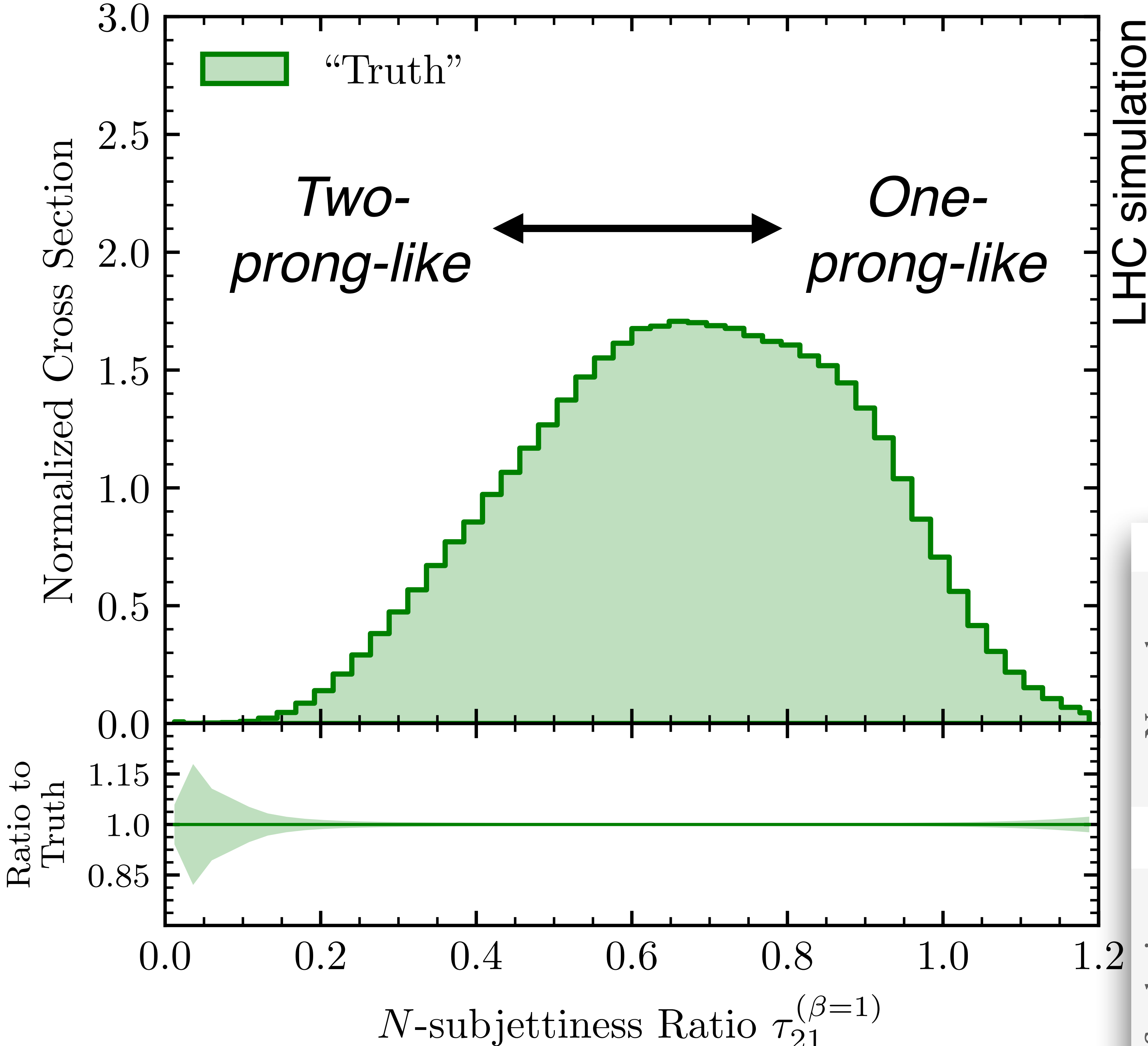
A. Andreassen, P. Komiske, E. Metodiev, **BPN**, J. Thaler, PRL 124 (2020) 182001



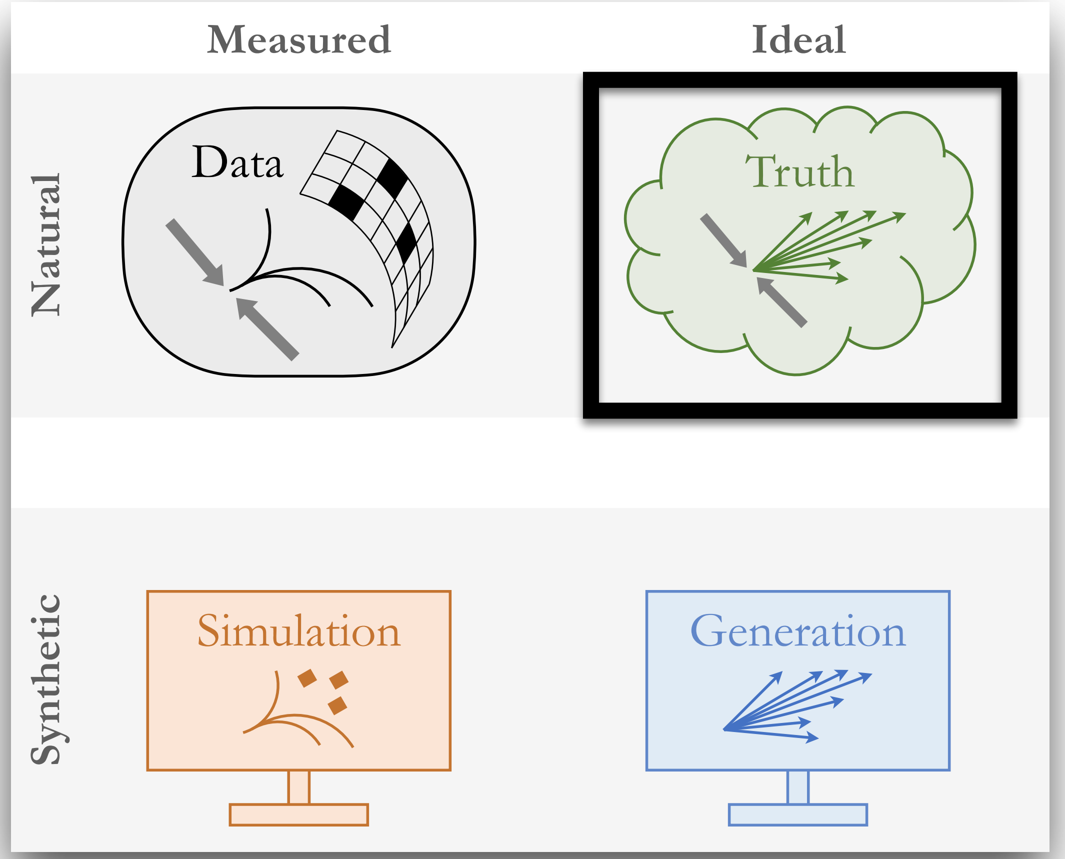
Consider this observable, which characterizes the substructure

Results

A. Andreassen, P. Komiske, E. Metodiev, **BPN**, J. Thaler, PRL 124 (2020) 182001

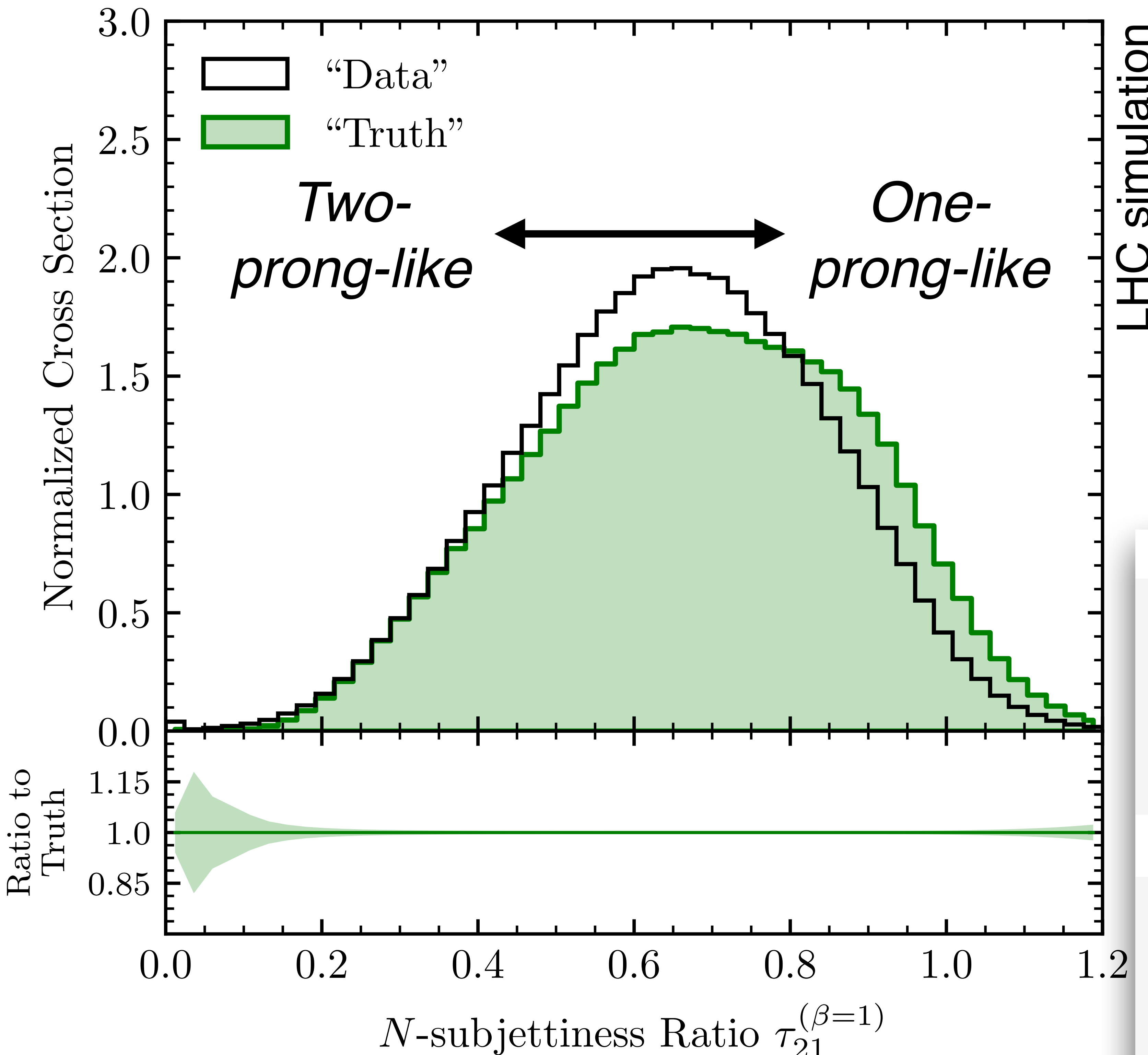


Consider this observable, which characterizes the substructure

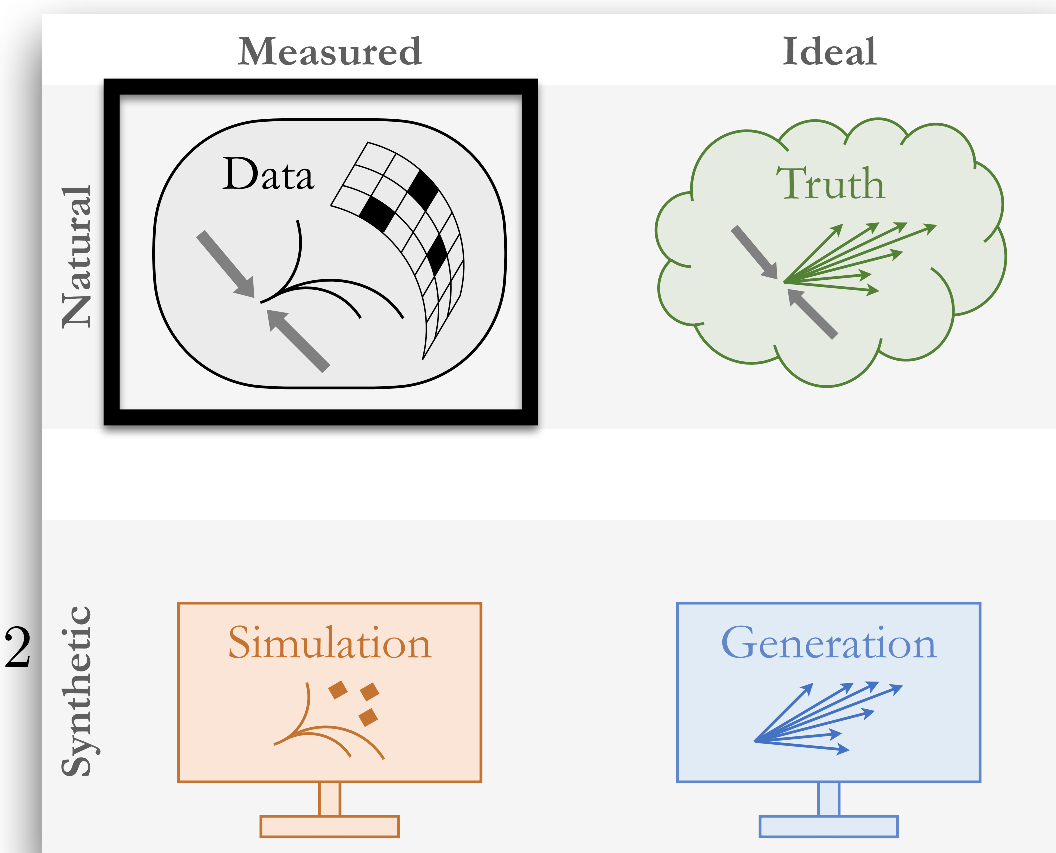


Results

A. Andreassen, P. Komiske, E. Metodiev, **BPN**, J. Thaler, PRL 124 (2020) 182001

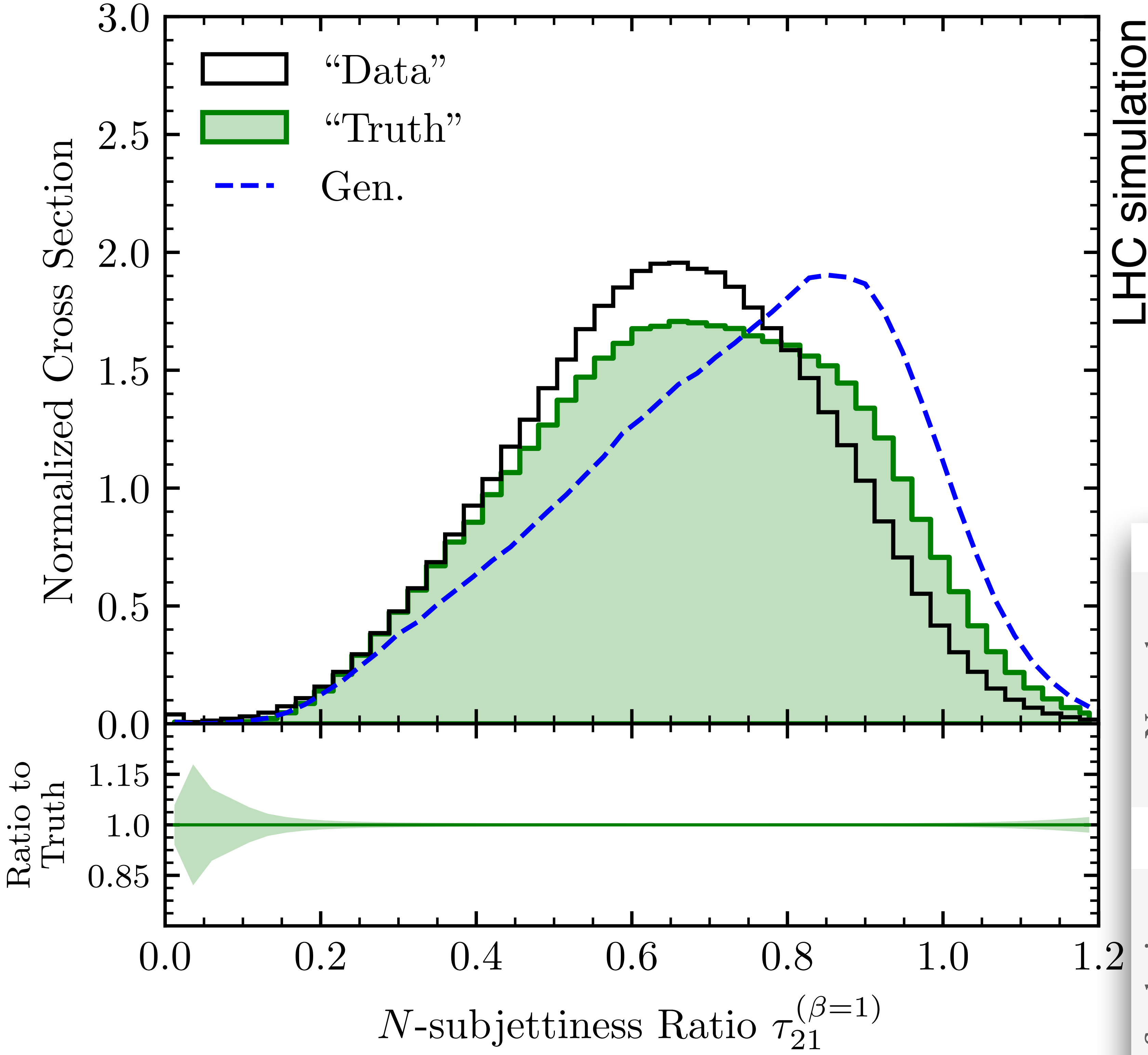


Consider this observable, which characterizes the substructure

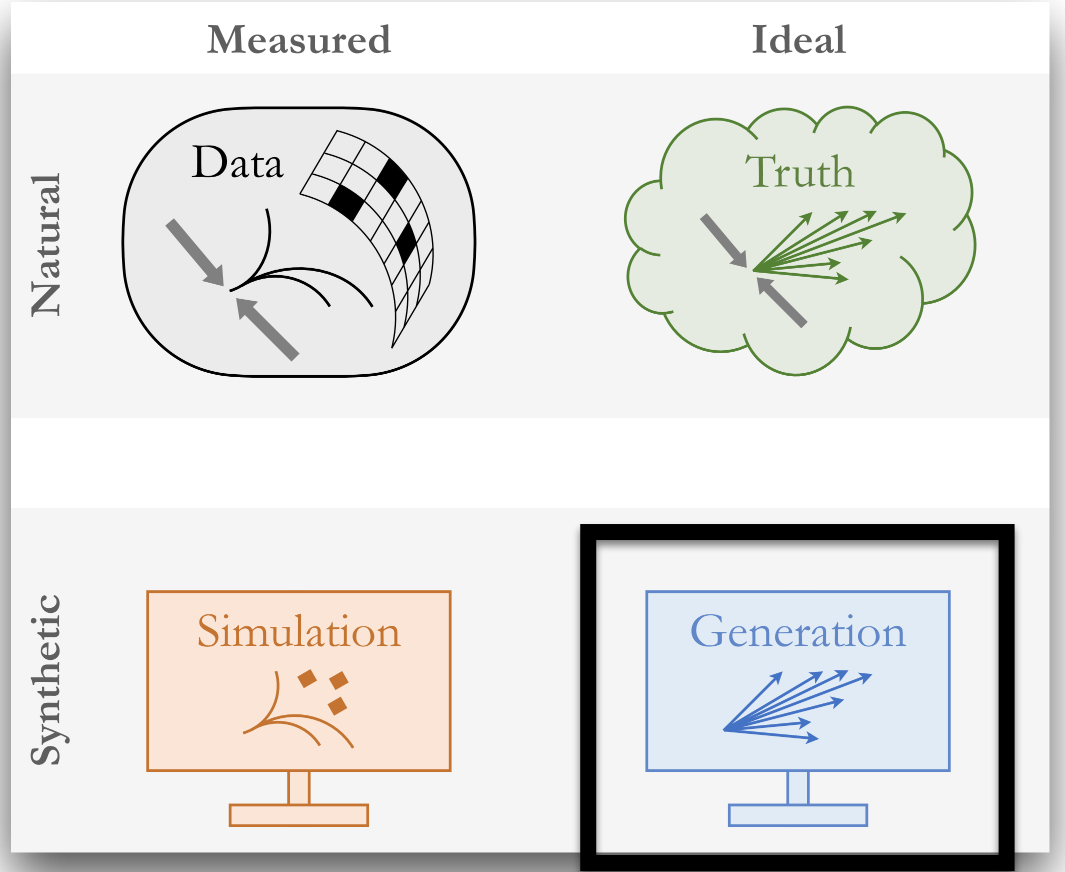


Results

A. Andreassen, P. Komiske, E. Metodiev, **BPN**, J. Thaler, PRL 124 (2020) 182001

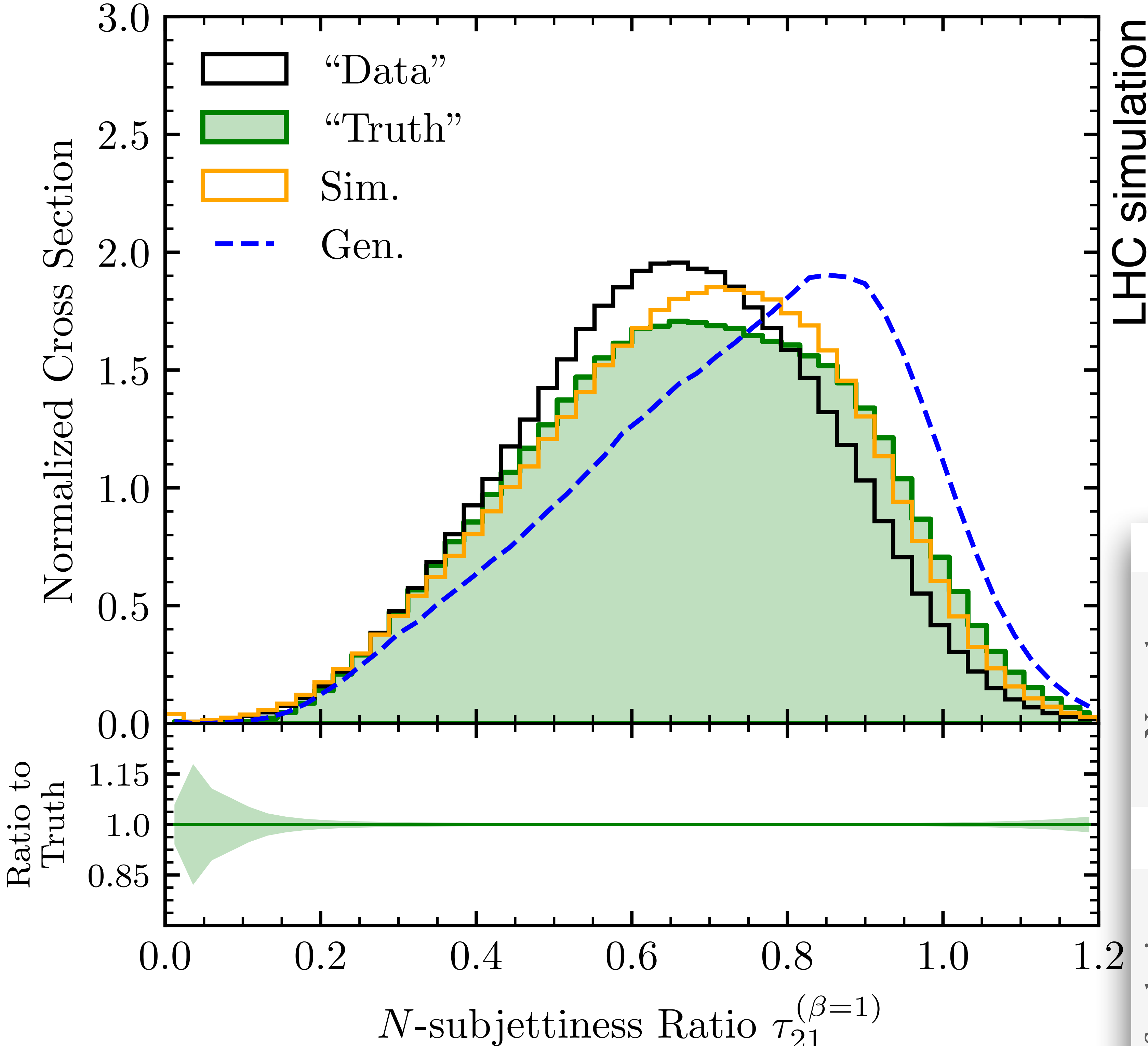


Consider this observable, which characterizes the substructure

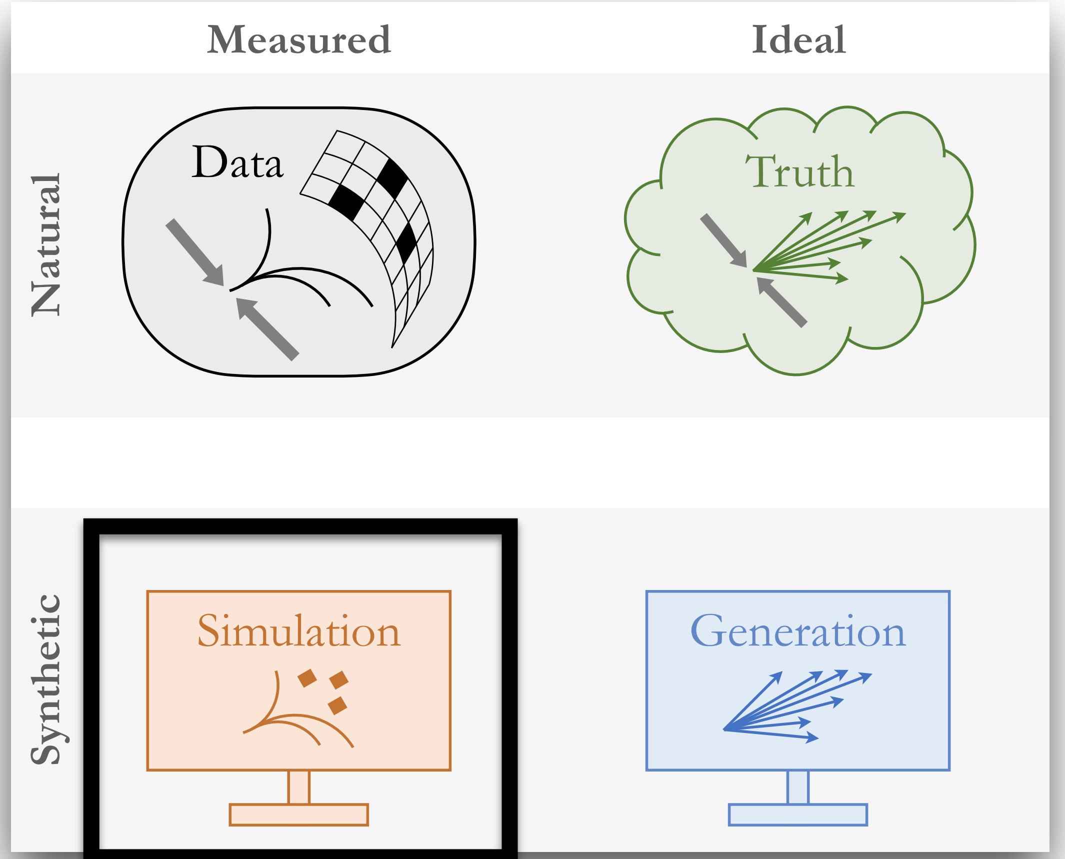


Results

A. Andreassen, P. Komiske, E. Metodiev, **BPN**, J. Thaler, PRL 124 (2020) 182001

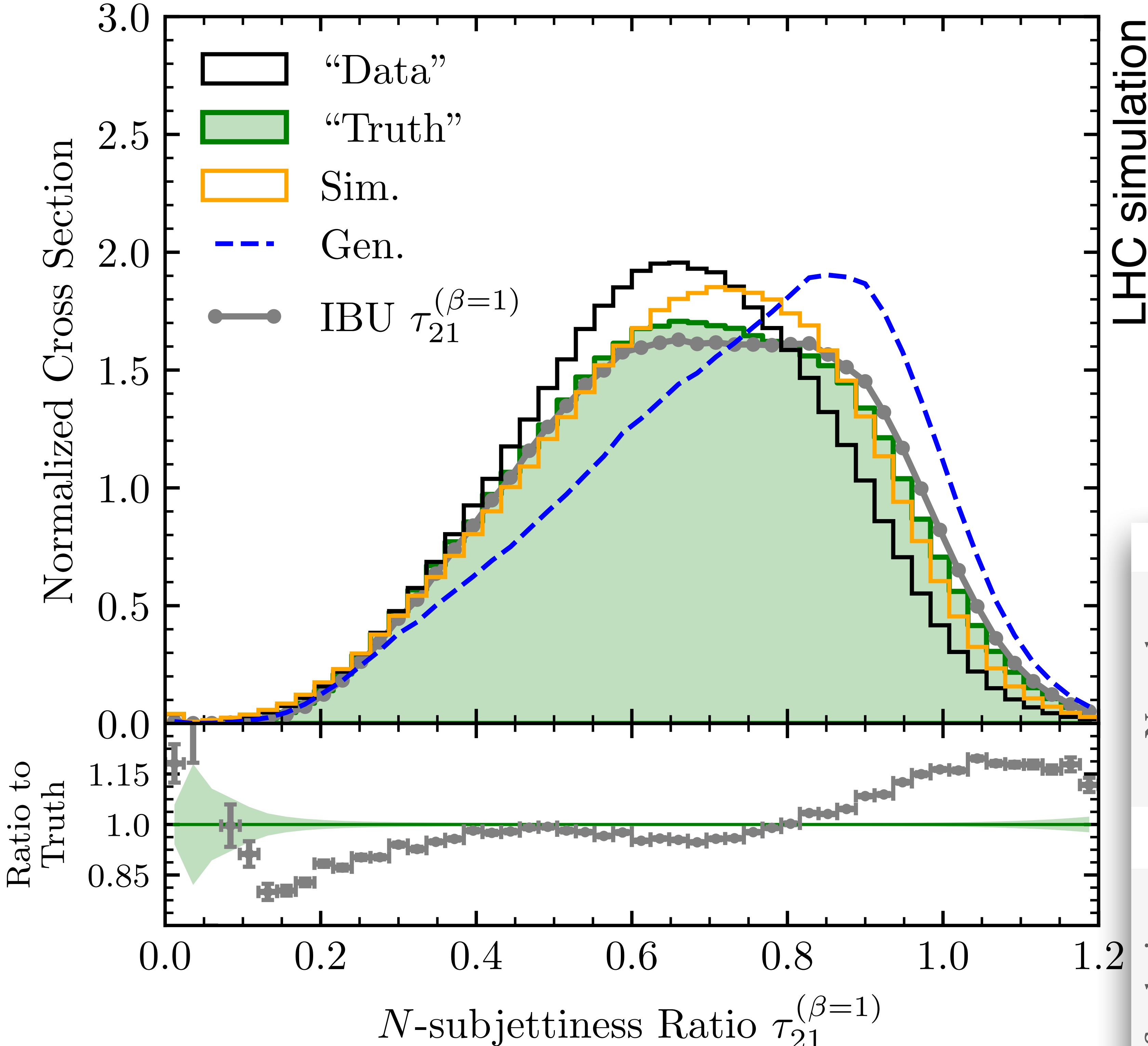


Consider this observable, which characterizes the substructure

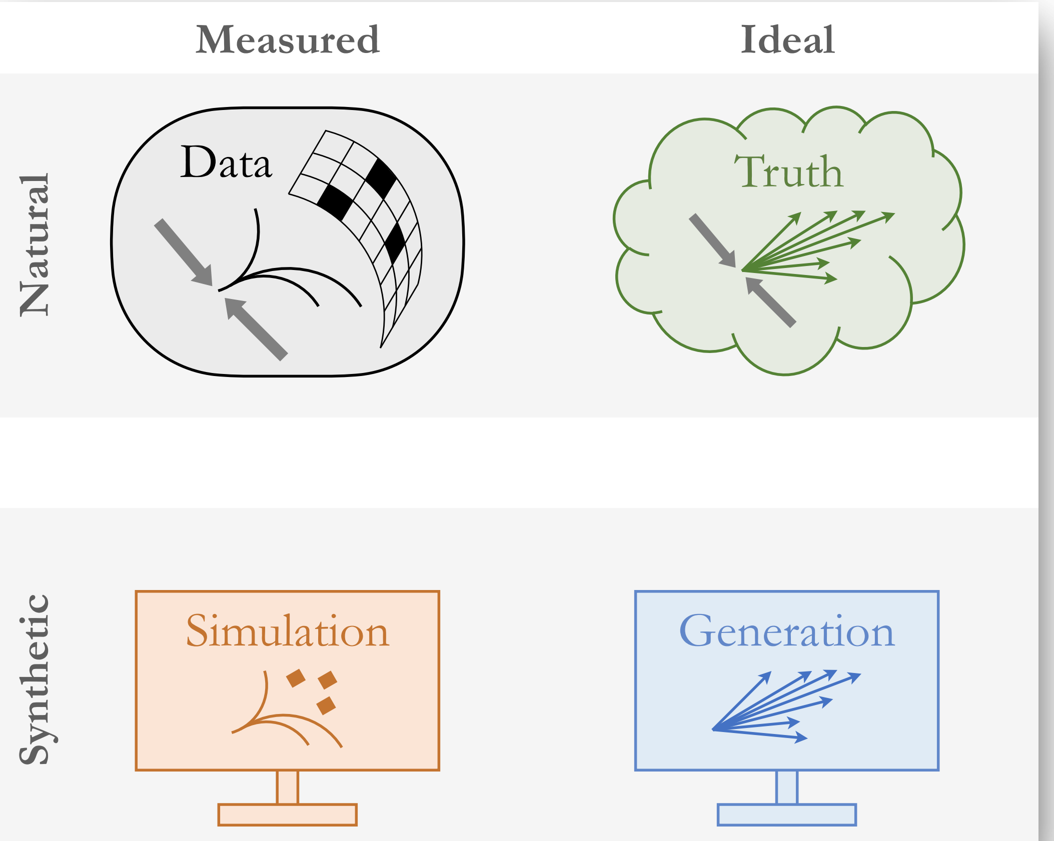


Results

A. Andreassen, P. Komiske, E. Metodiev, **BPN**, J. Thaler, PRL 124 (2020) 182001

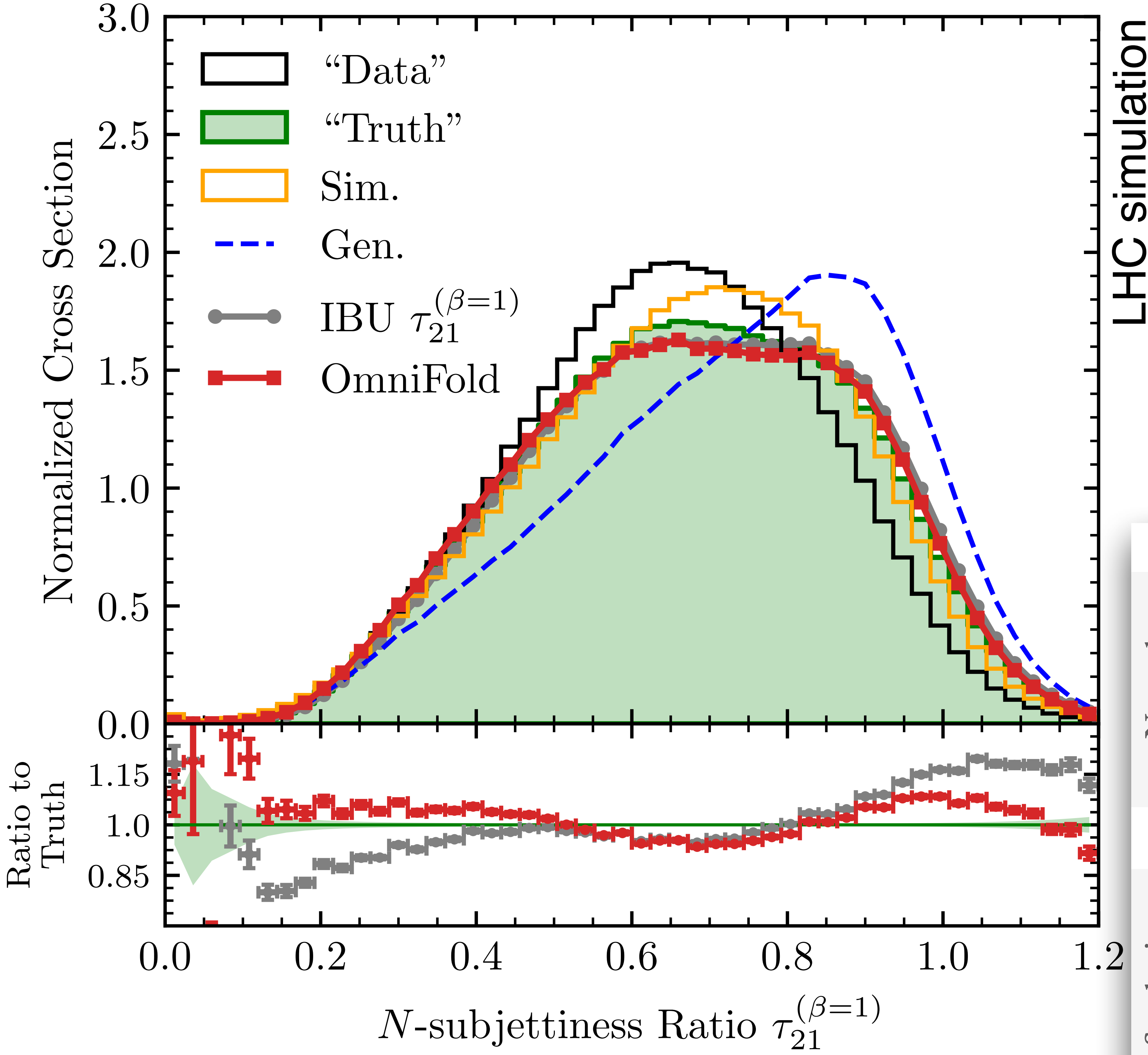


IBU is the current standard. It is a1D binned and iterative approach.

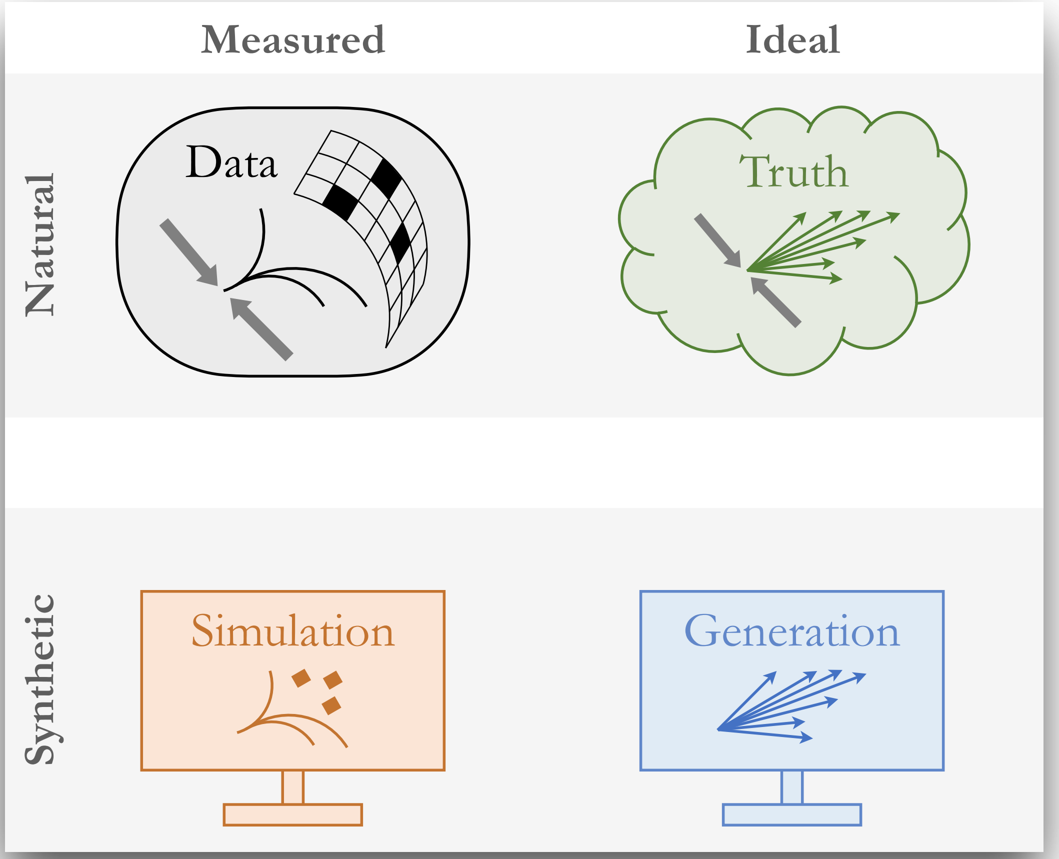


Results

A. Andreassen, P. Komiske, E. Metodiev, **BPN**, J. Thaler, PRL 124 (2020) 182001

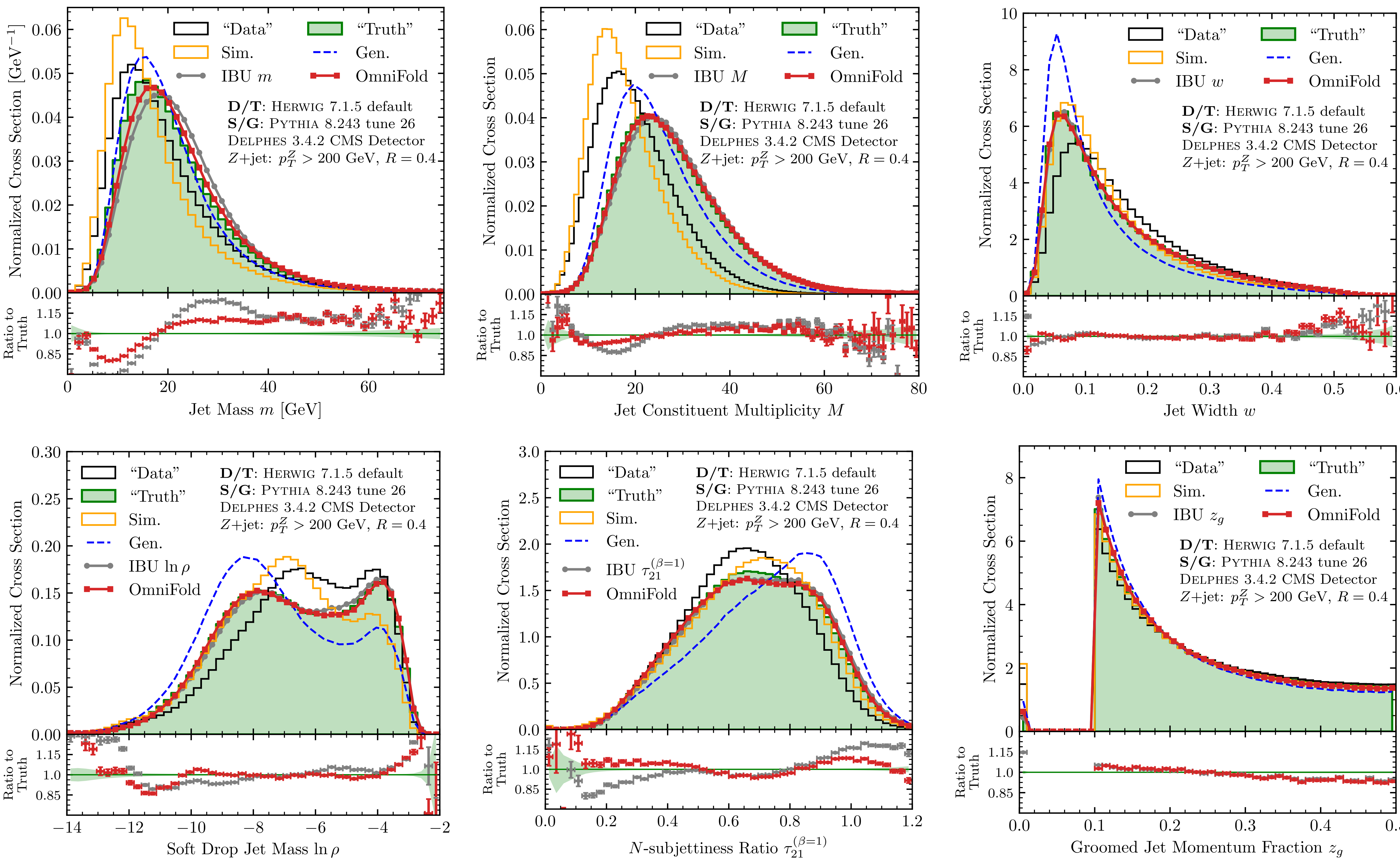


OmniFold outperforms IBU even though it is not tailored to this observable



Results

A. Andreassen, P. Komiske, E. Metodiev, **BPN**, J. Thaler, PRL 124 (2020) 182001

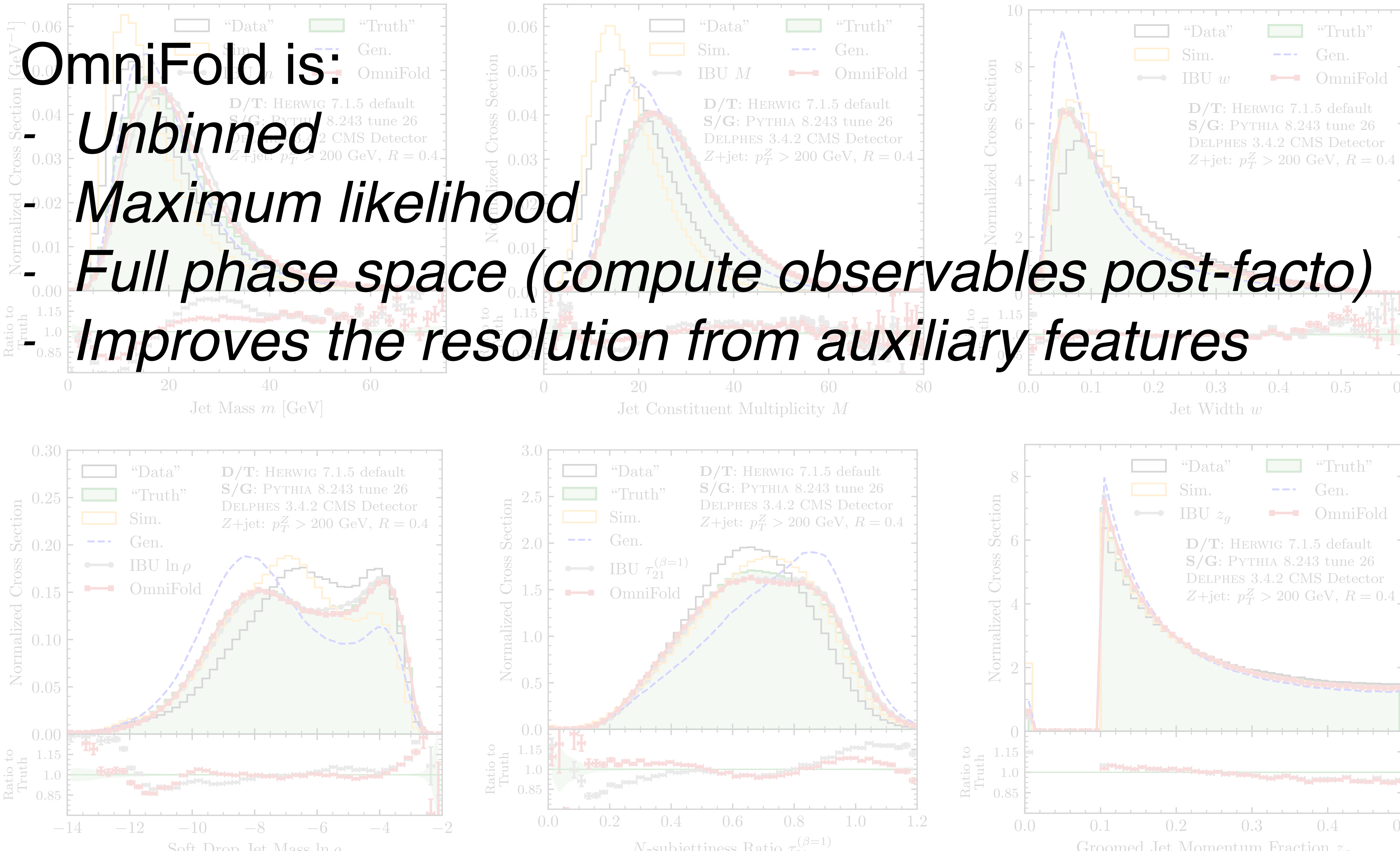


Results

[A. Andreassen, P. Komiske, E. Metodiev, BPN, J. Thaler, 1907.08209]

OmniFold is:

- *Unbinned*
- *Maximum likelihood*
- *Full phase space (compute observables post-facto)*
- *Improves the resolution from auxiliary features*



(2) Model agnostic searches

One strategy is to learn directly from data without training on a particular signal model.

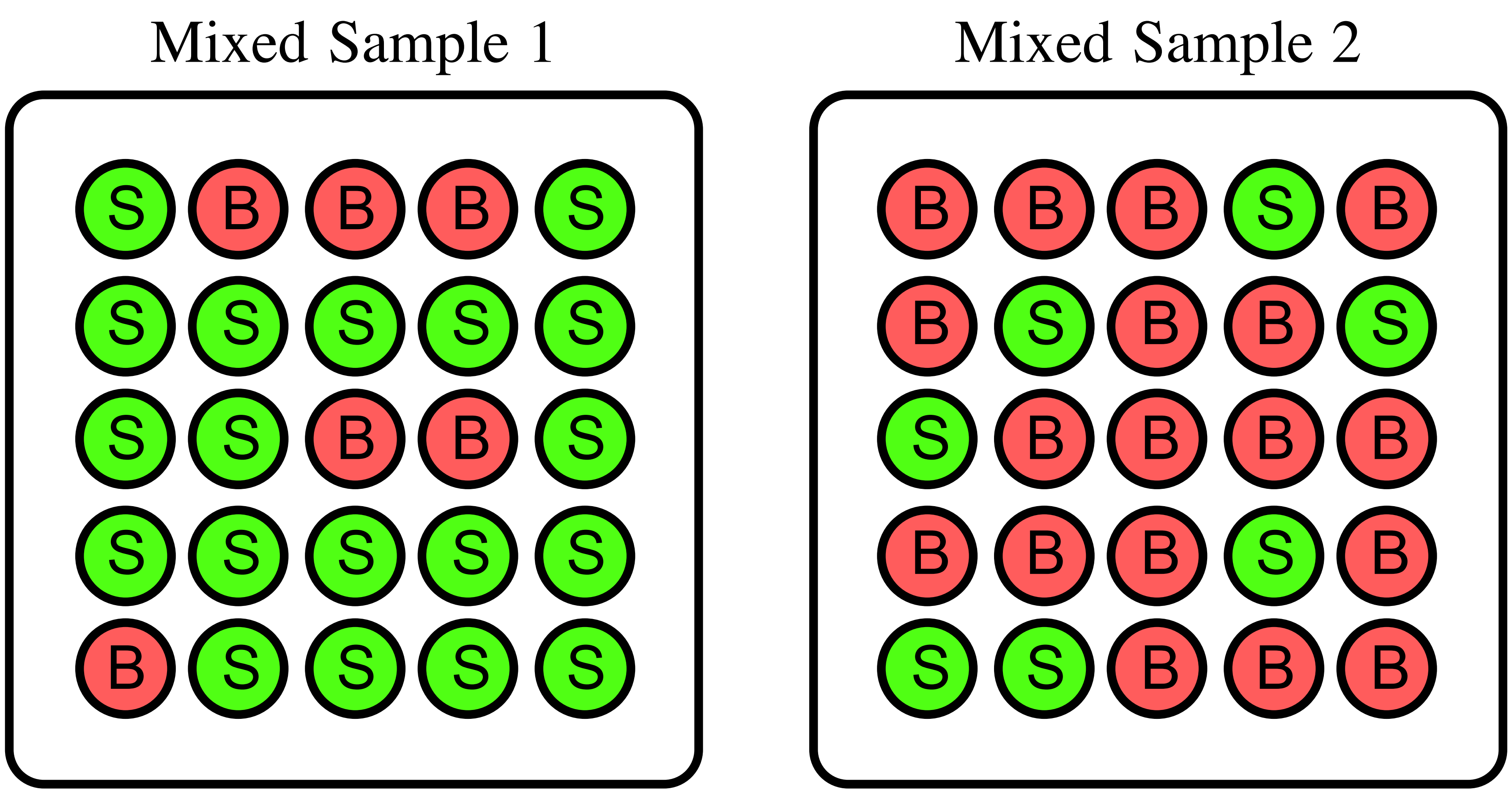
(2) Model agnostic searches

One strategy is to learn directly from data without training on a particular signal model.

However, the data do not have labels. How can we train a classifier without labels??

Learning from unlabeled data

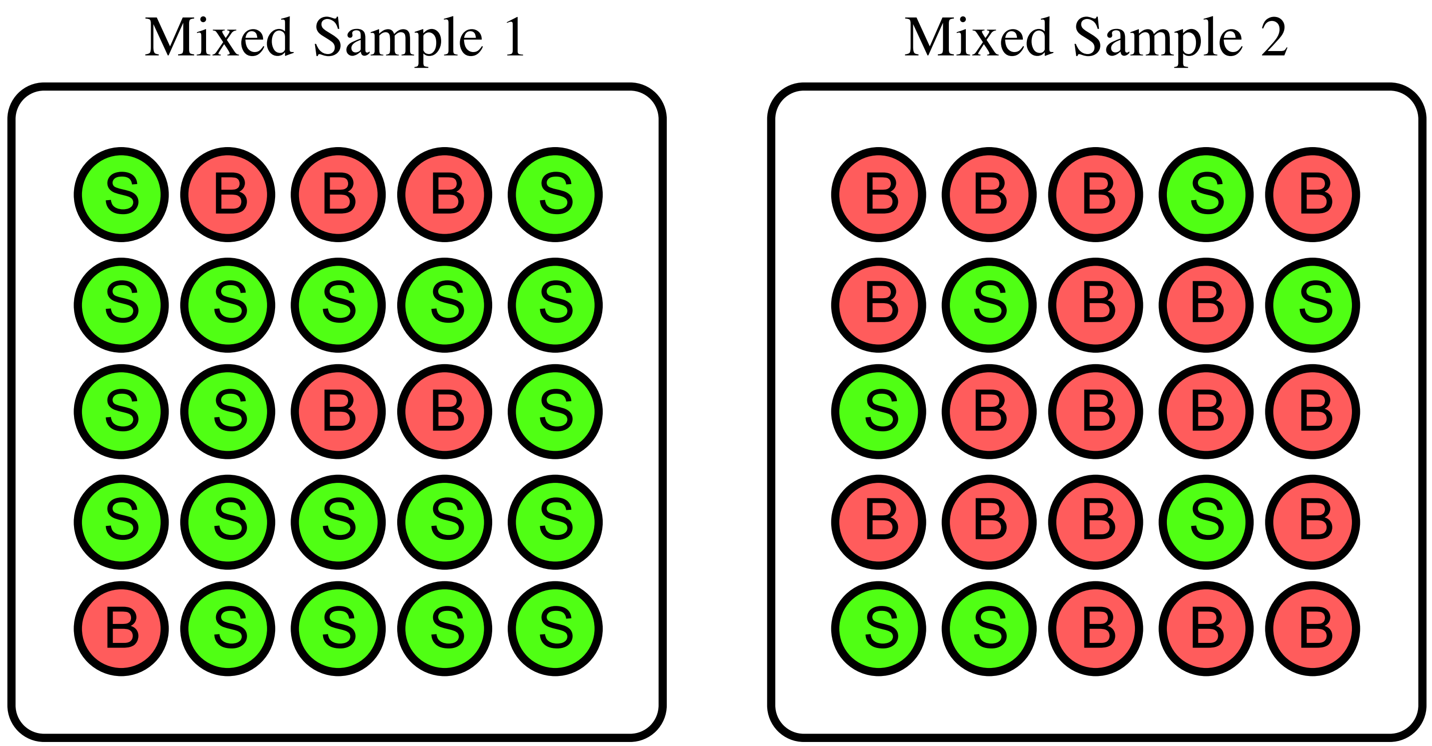
The data are unlabeled and in the best case, come to us as mixtures of two classes (“signal” and “background”).



(we don’t get to observe the color of the circles)

Weak supervision: *Classification Without Labels*

Can we learn without any label information?



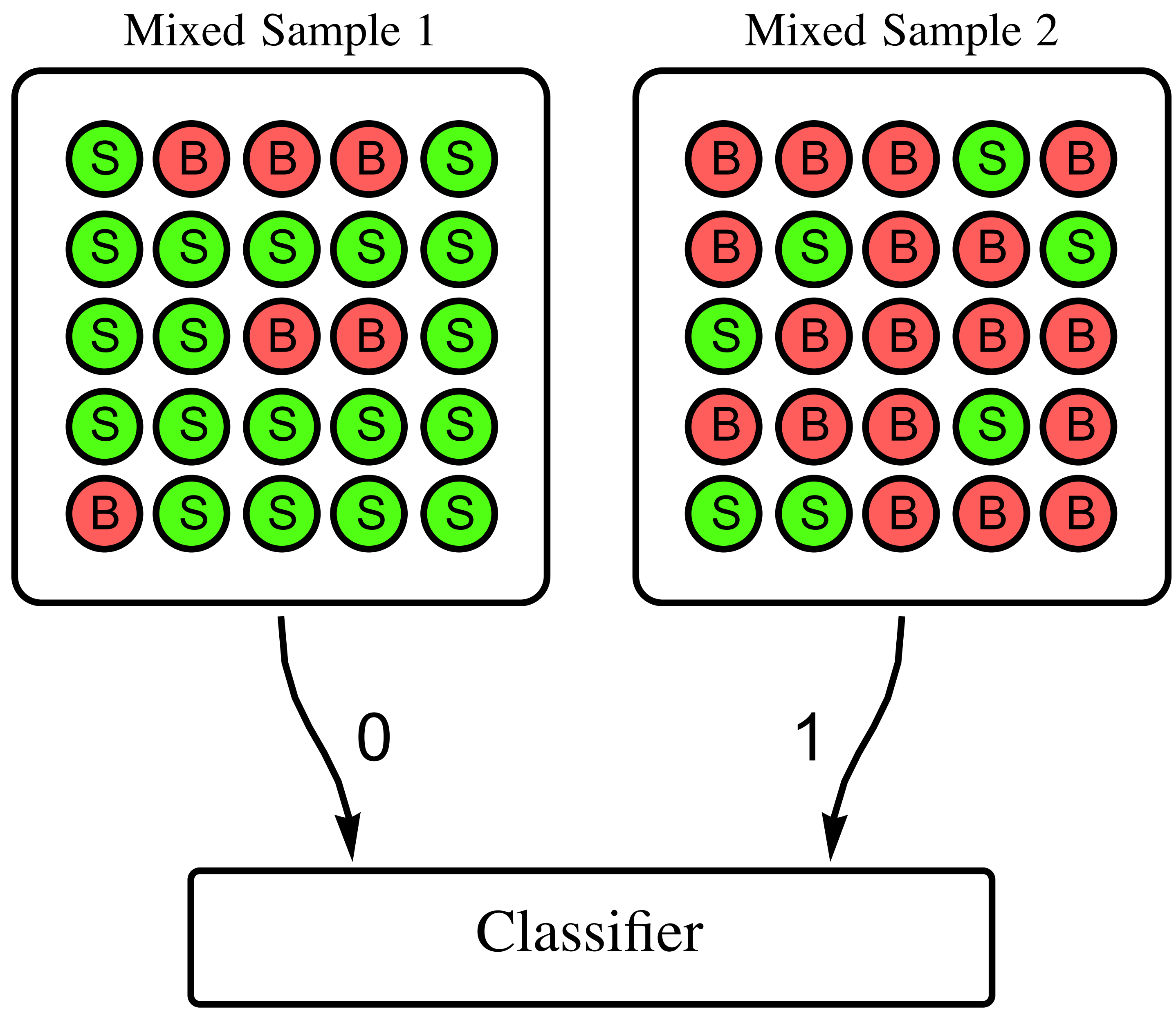
E. Metodiev, **BPN**, J. Thaler, JHEP 10 (2017) 51

Weak supervision: *Classification Without Labels*

Can we learn without any label information?

Yes !

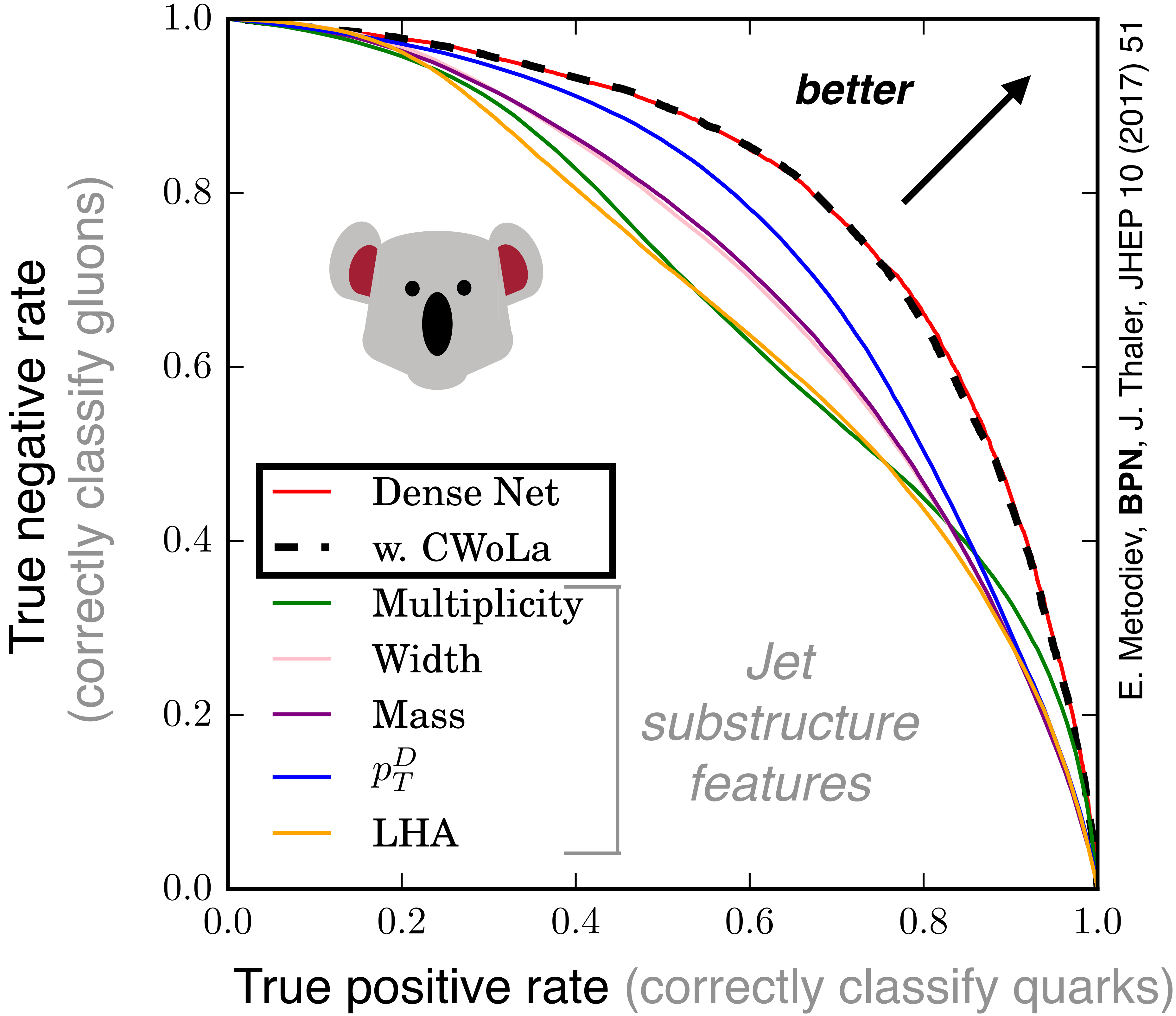
Training on impure samples is (asymptotically) equivalent to training on pure samples



E. Metodiev, **BPN**, J. Thaler, JHEP 10 (2017) 51

Works!

Example: classify quarks versus gluons



Anomaly detection

How can we use CWoLa to **hunt for new** particles?

CWoLa Hunting

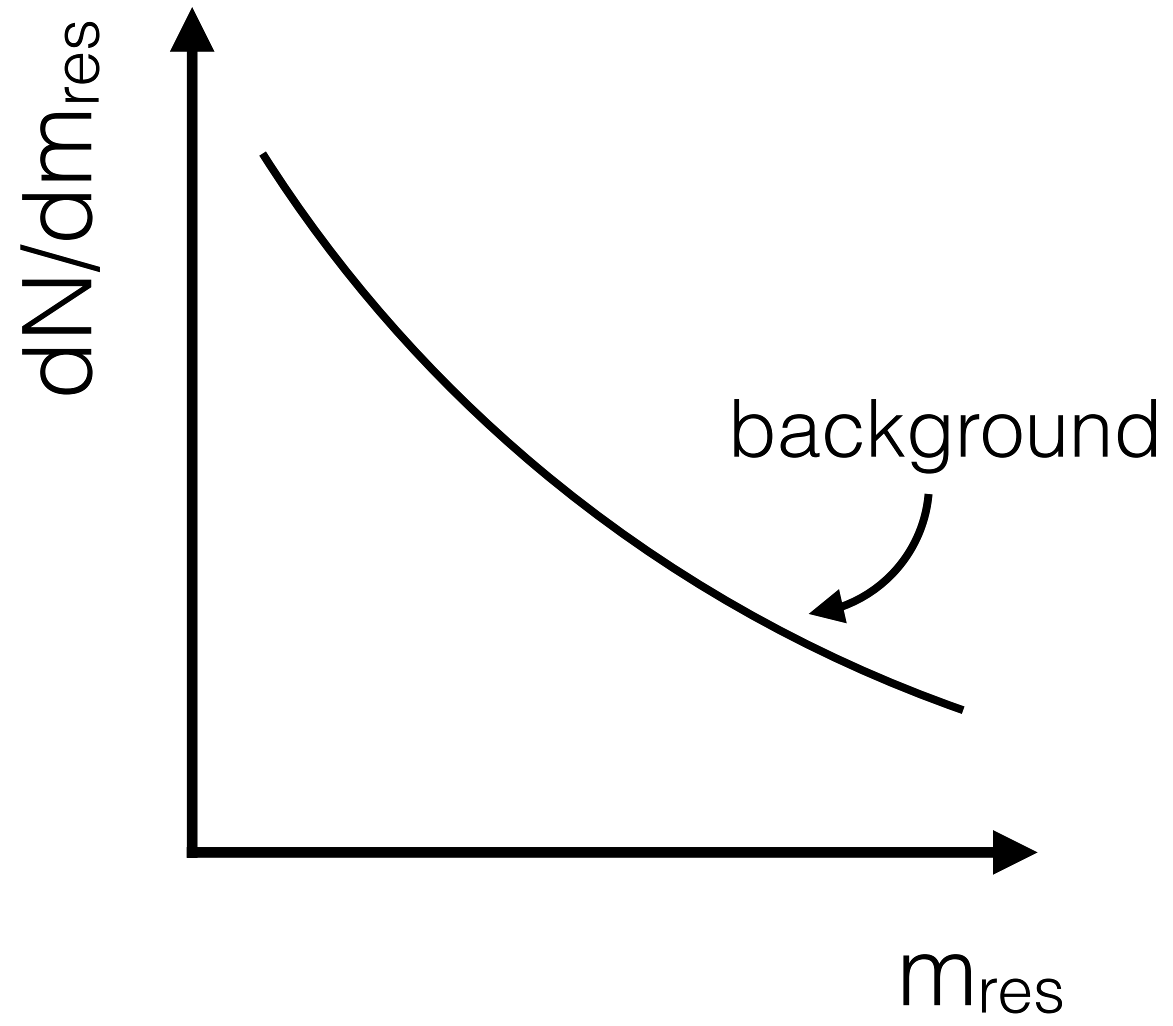
How can we use CWoLa to **hunt for new** particles?

*Image from *The Courier Mail*. Koala is actually being freed - I do not condone violence against these animals!

CWoLa Hunting

Phys. Rev. Lett. 121 (2018) 241803

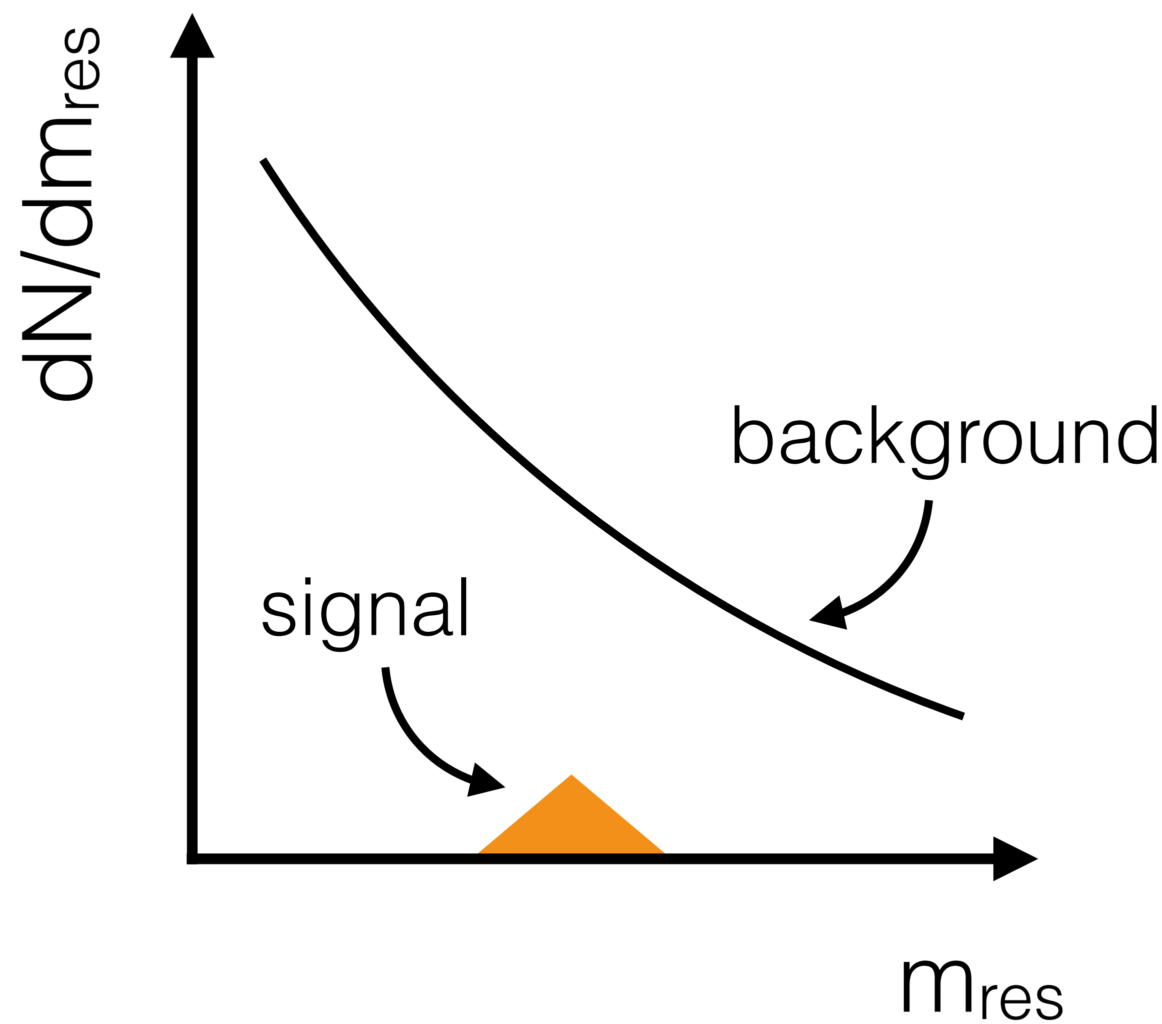
J. Collins, K. Howe, **BPN**



CWoLa Hunting

[*Phys. Rev. Lett. 121 (2018) 241803*]

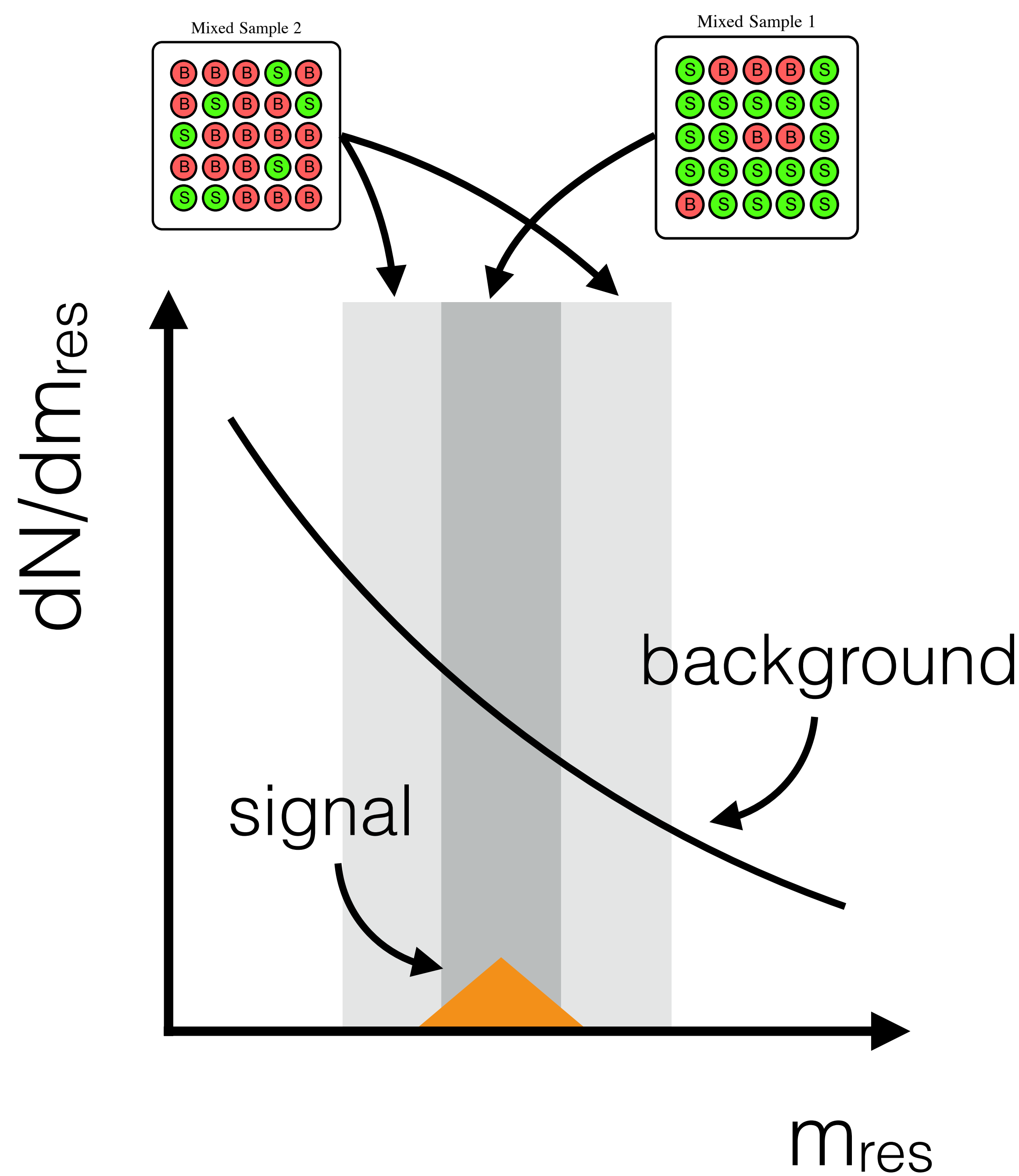
J. Collins, K. Howe, **BPN**



CWoLa Hunting

[*Phys. Rev. Lett. 121 (2018) 241803*]

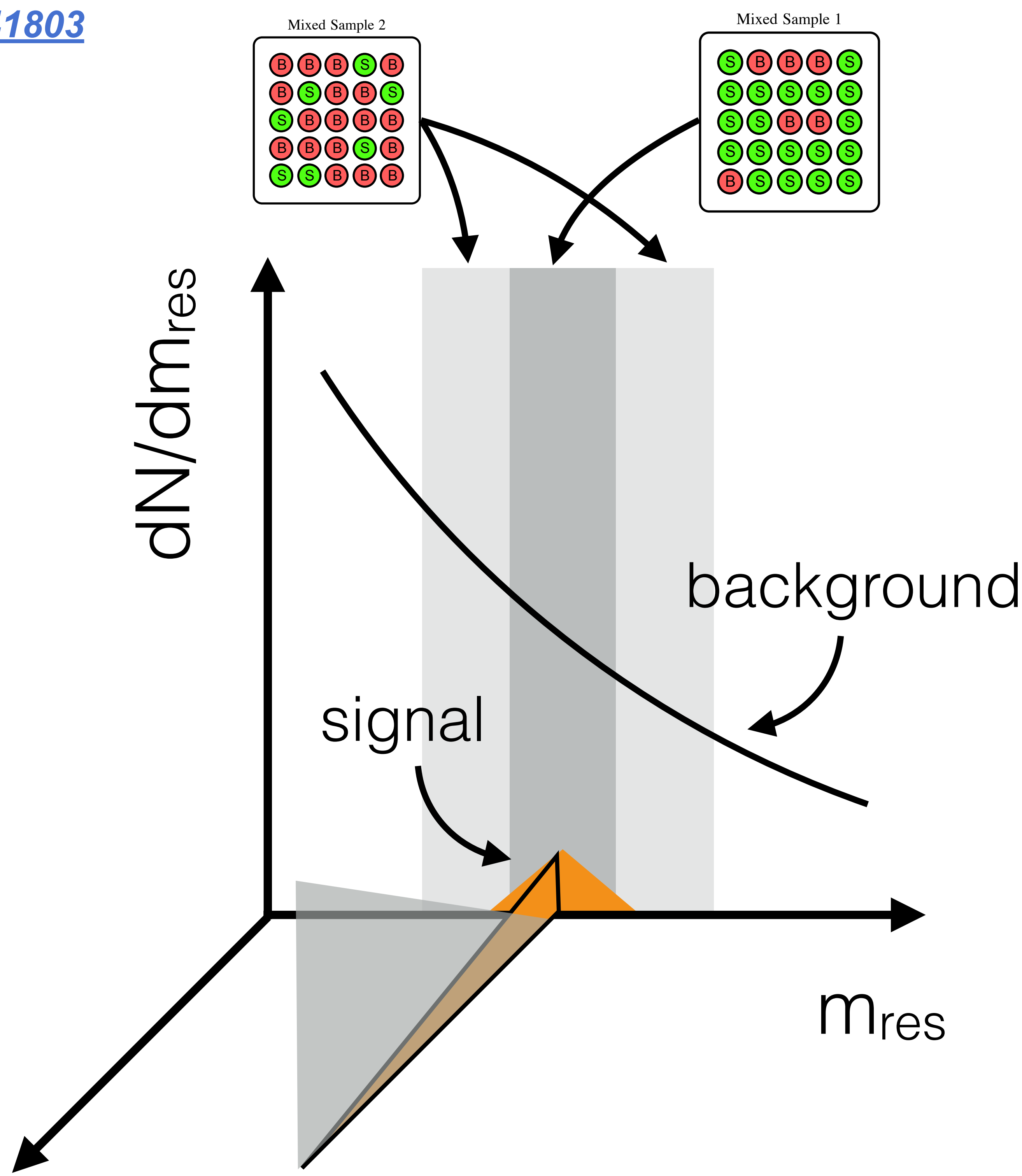
J. Collins, K. Howe, **BPN**



CWoLa Hunting

[*Phys. Rev. Lett. 121 (2018) 241803*]

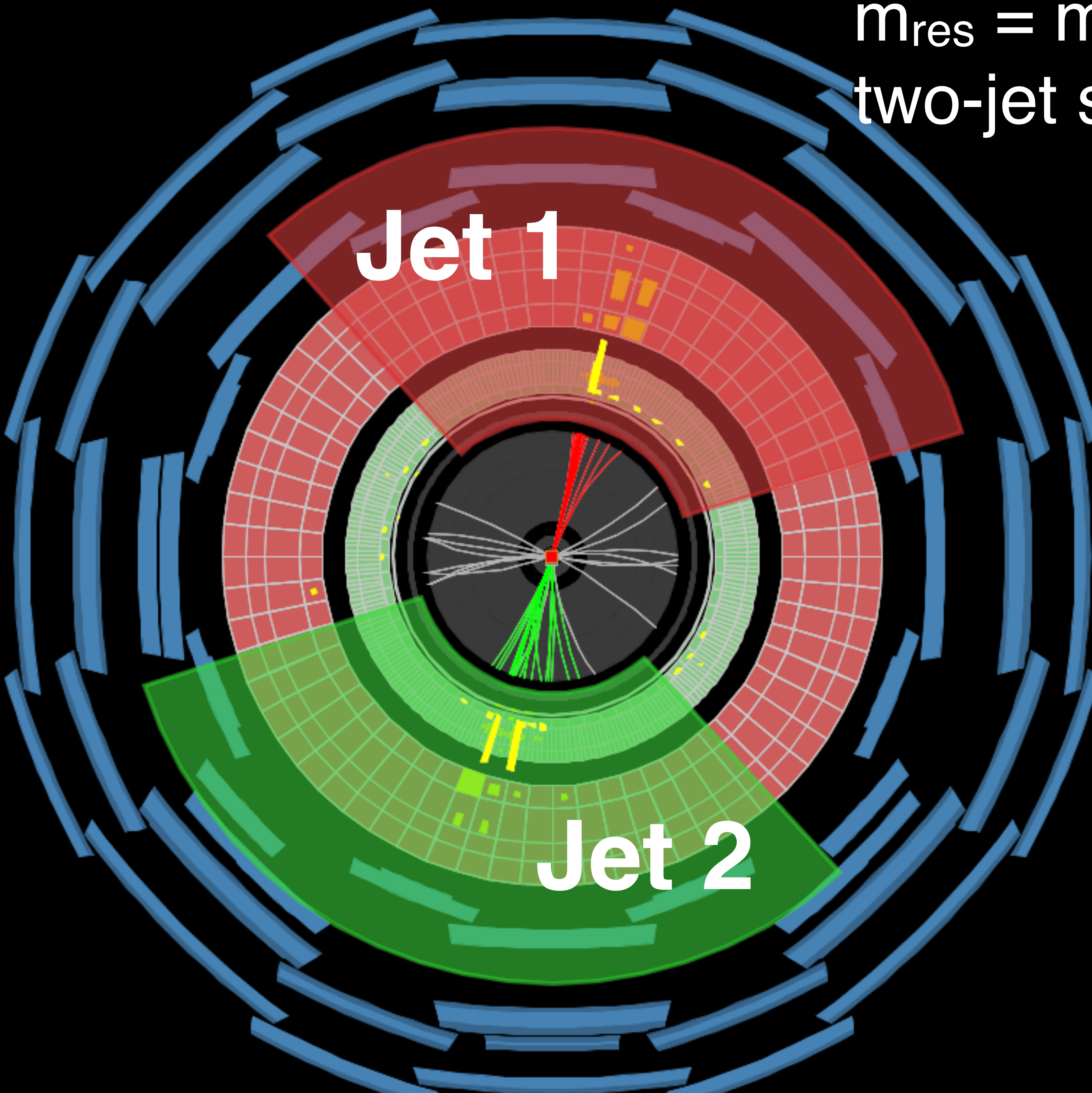
J. Collins, K. Howe, **BPN**

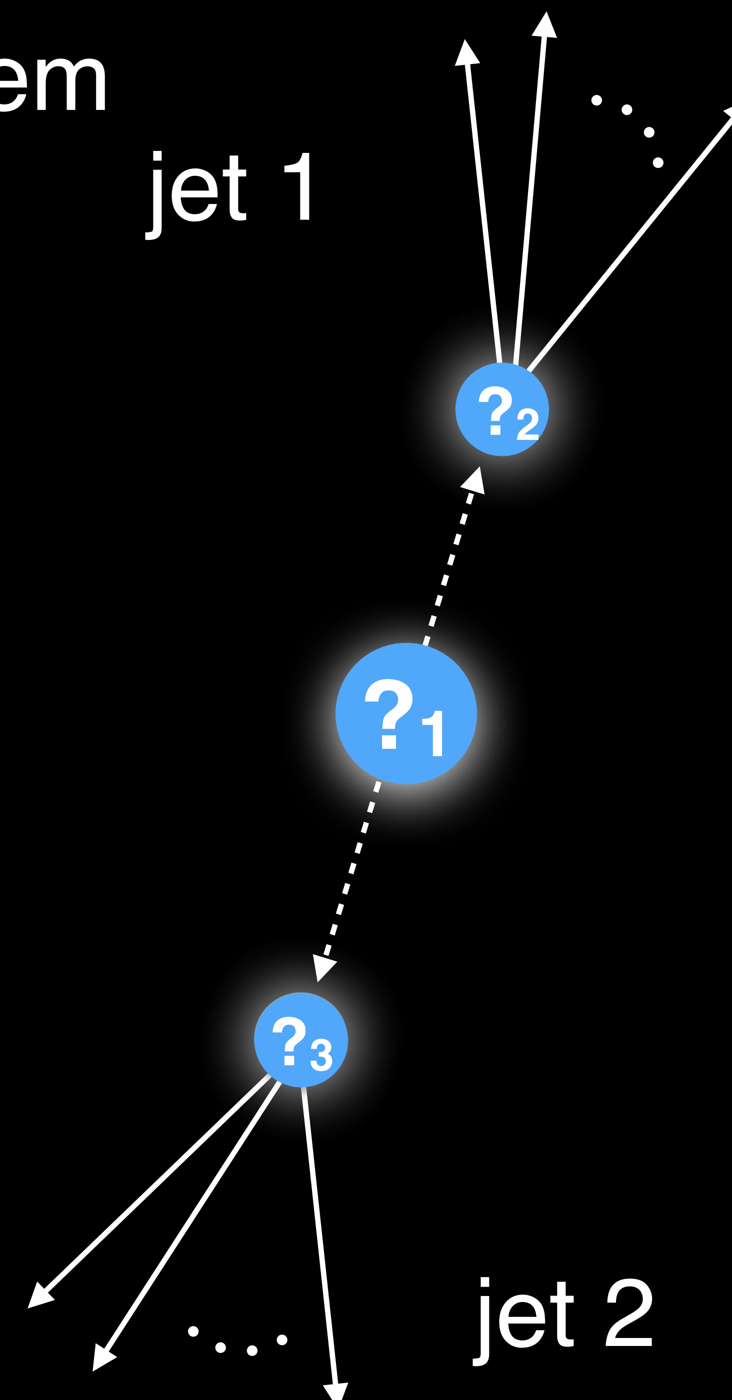


+ be careful to not pay a big trails factor
(ask if interested)

Example: two-“jet” search

m_{res} = mass of two-jet system

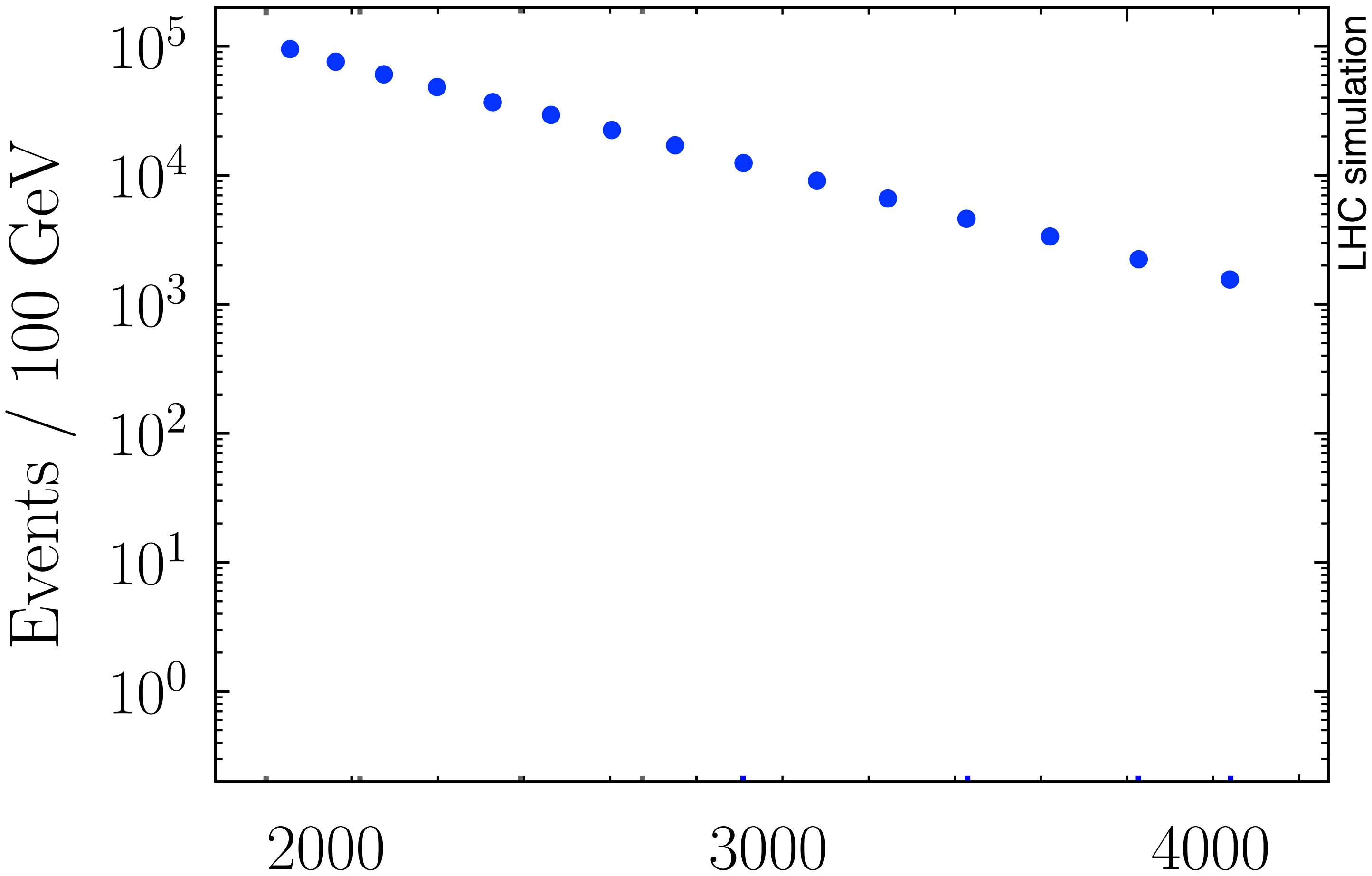




collisions in/out of page

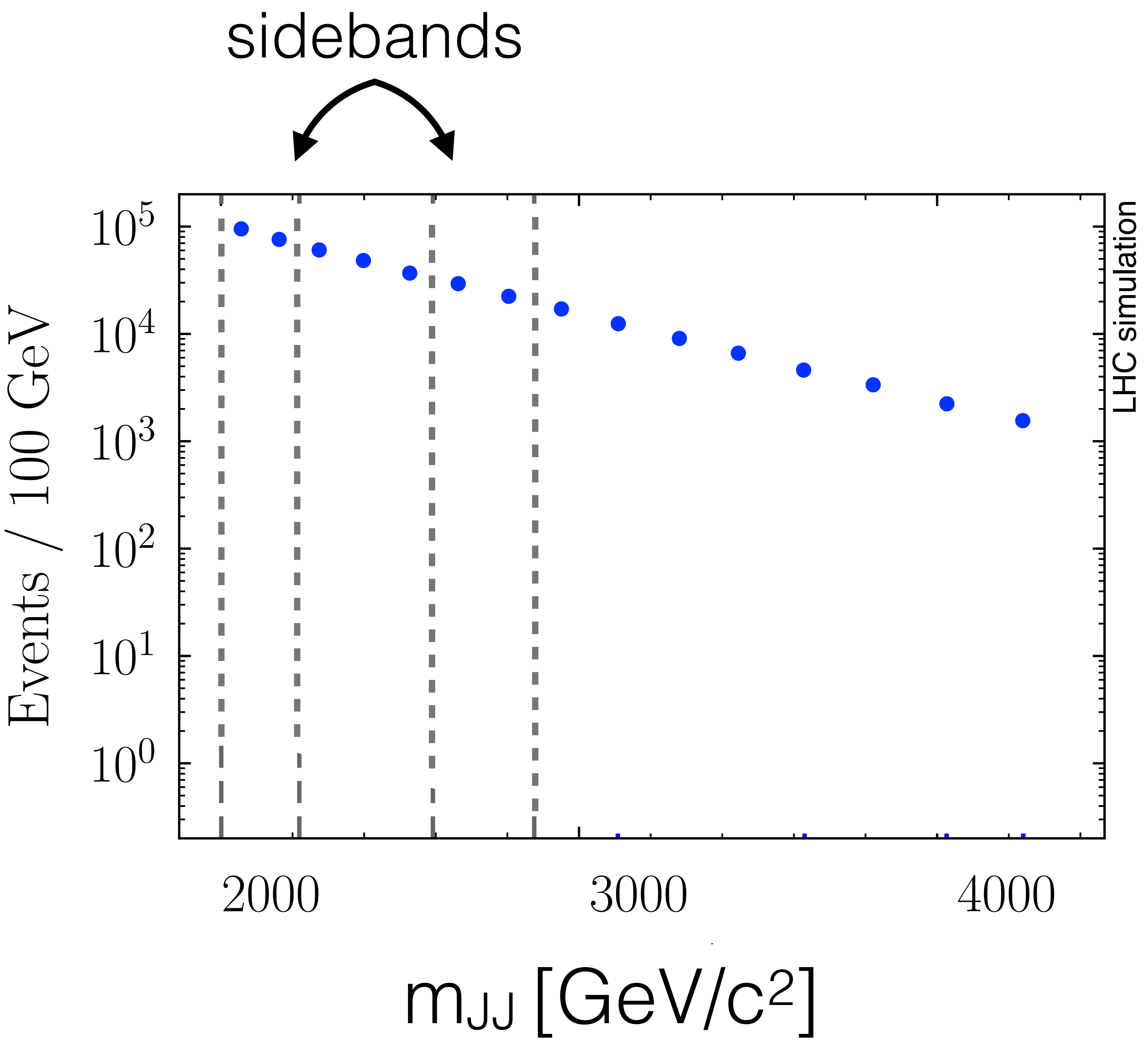
y = substructure of the two jets

Example: two-jet search

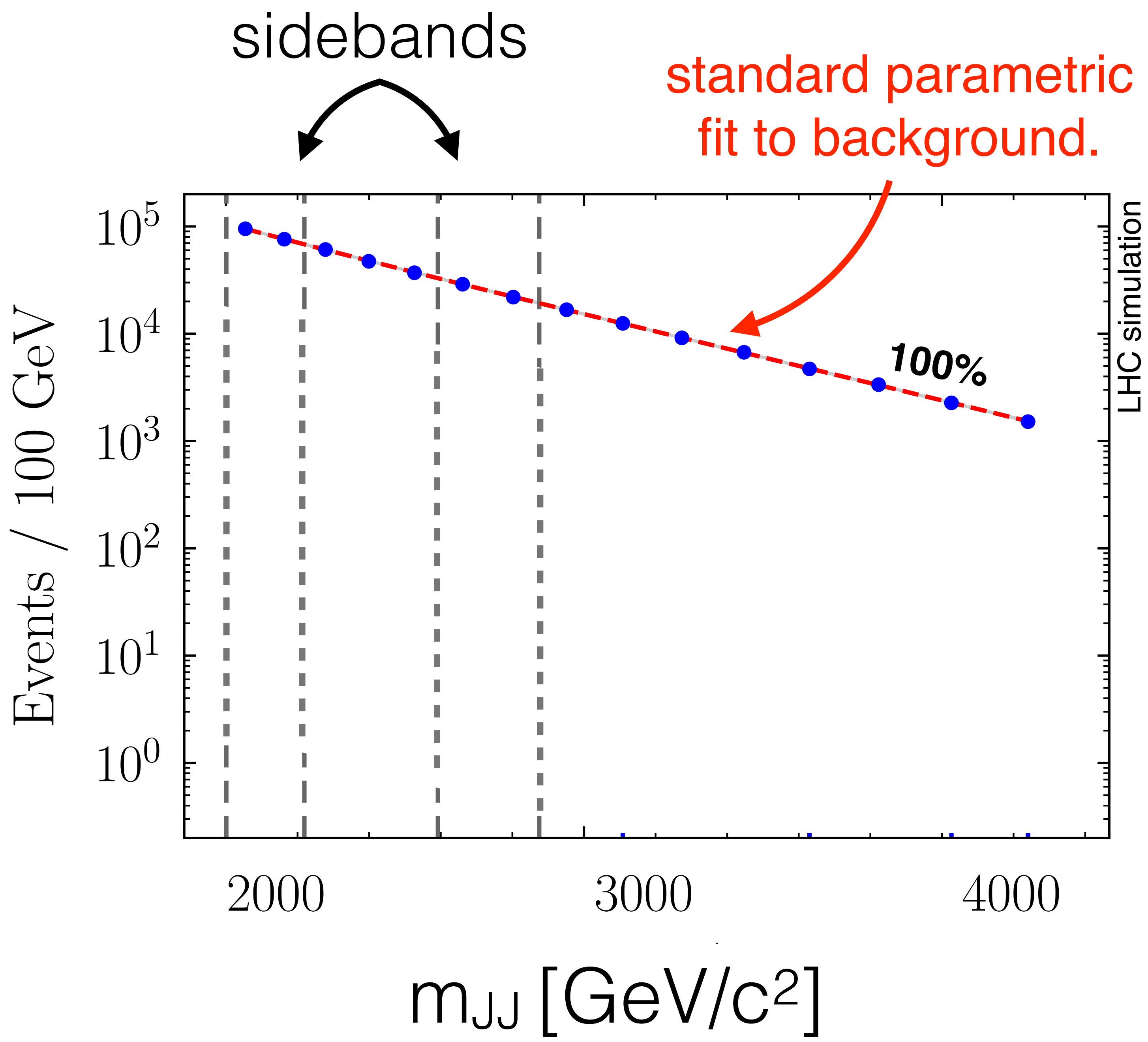


m_{JJ} [GeV/c^2]

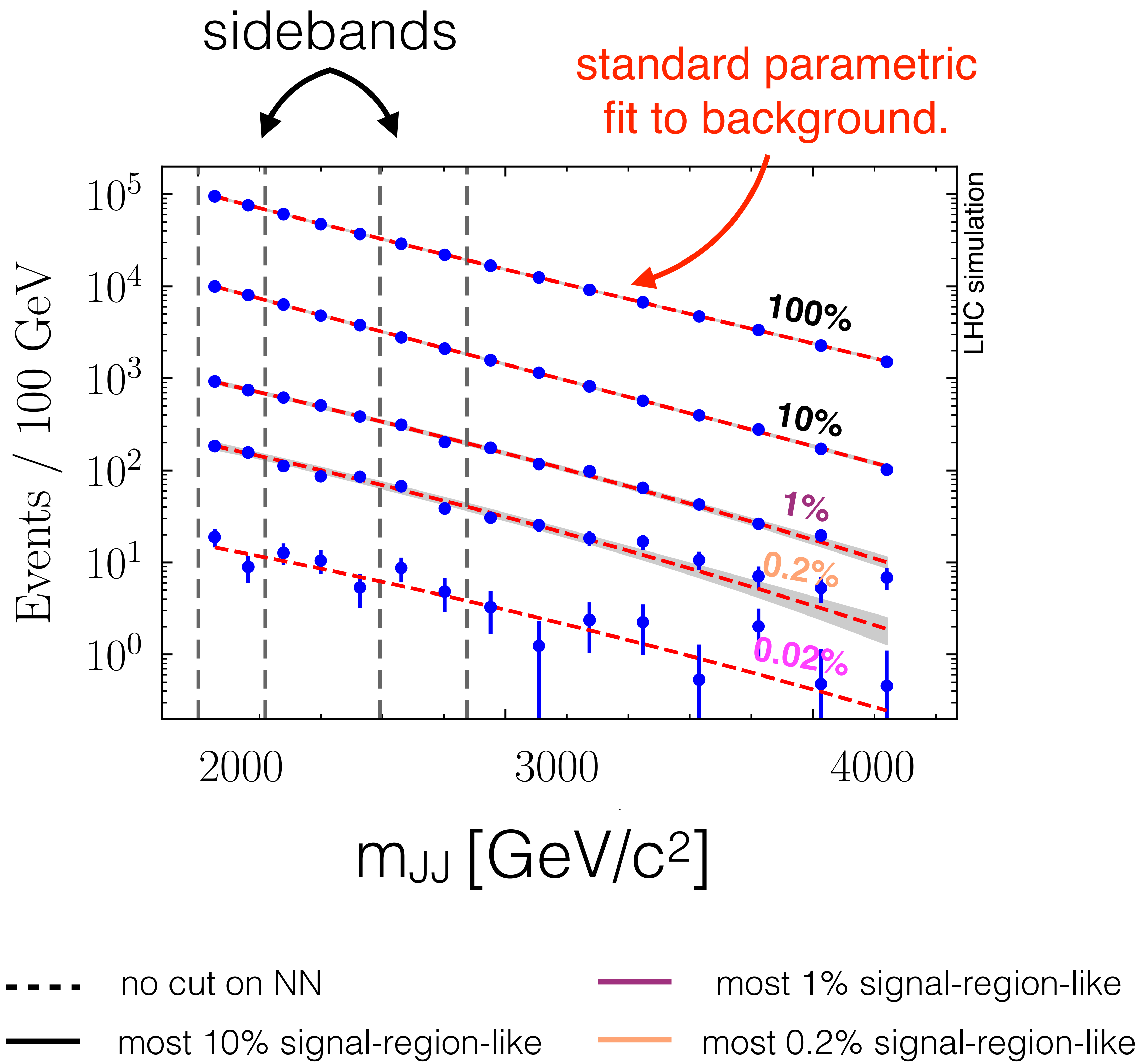
Example: two-jet search



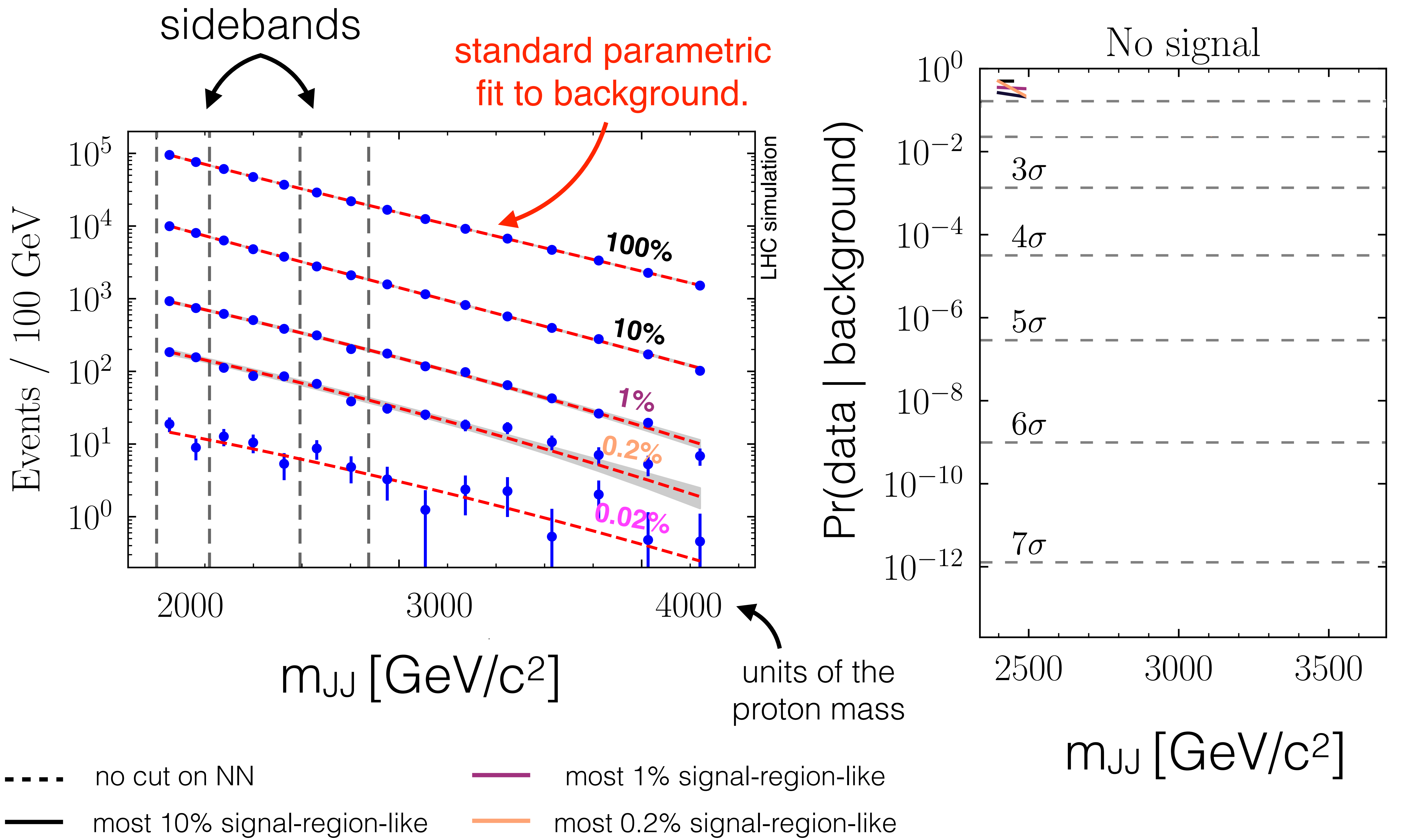
Example: two-jet search



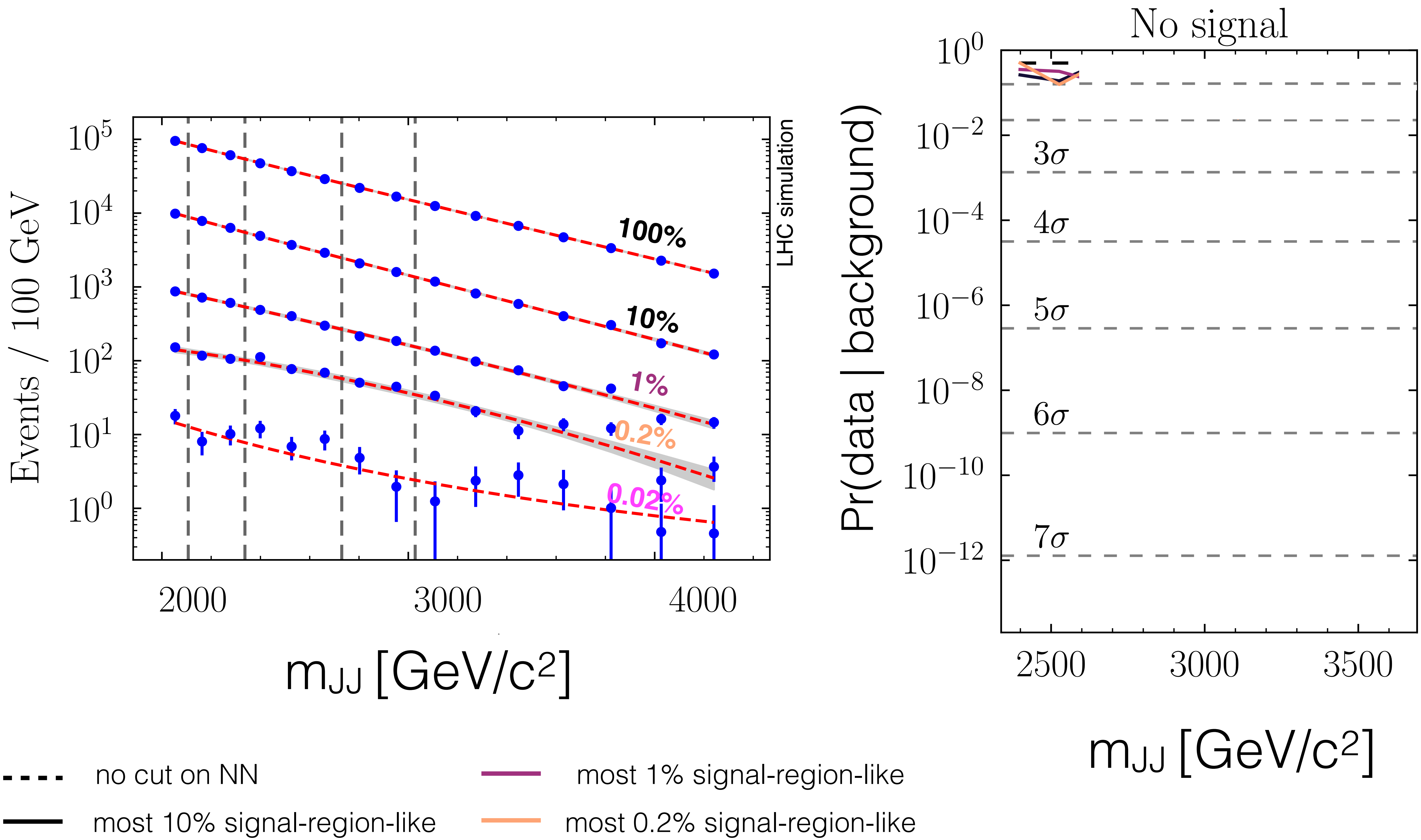
Example: two-jet search



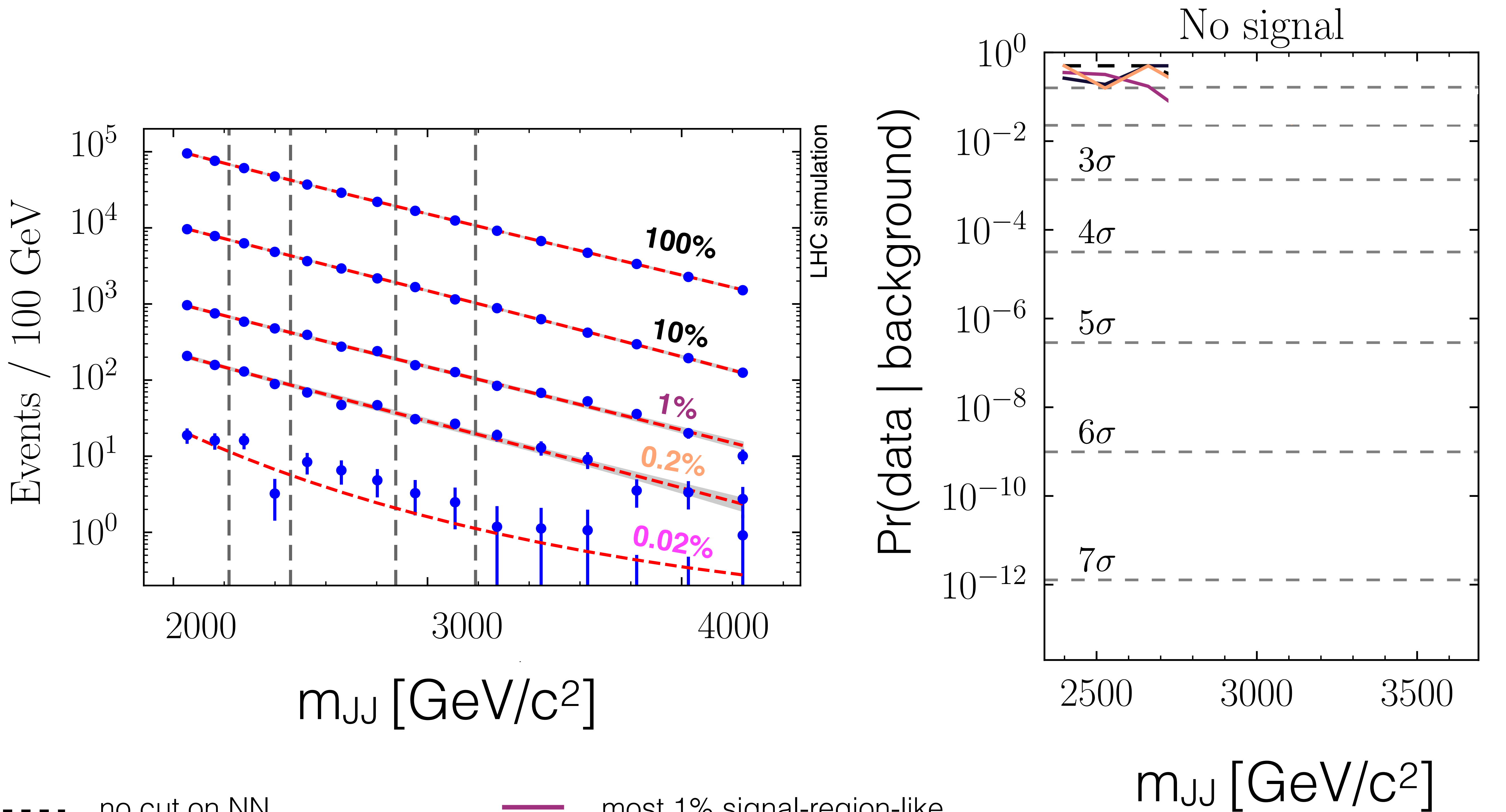
Example: two-jet search



Example: two-jet search

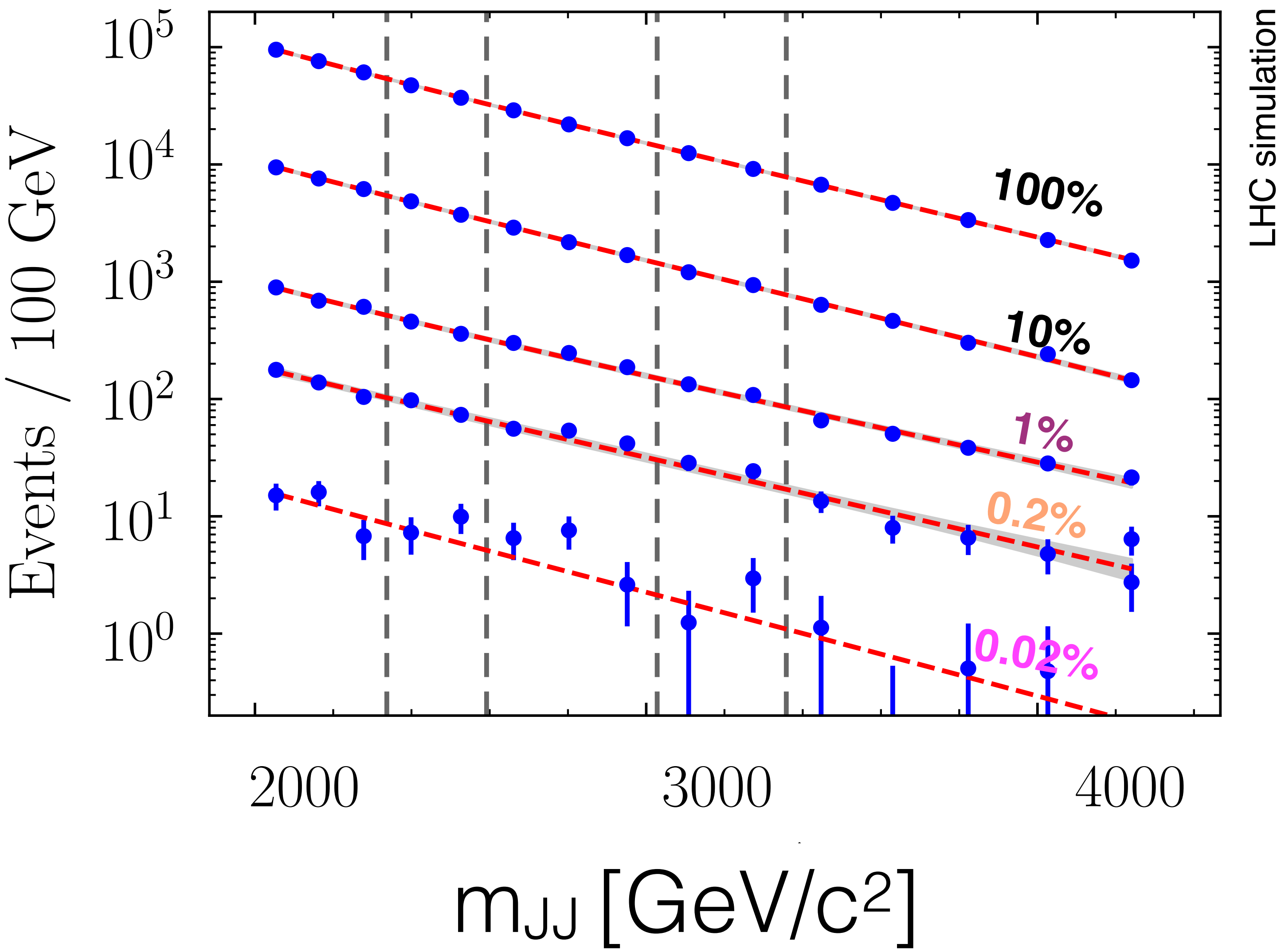


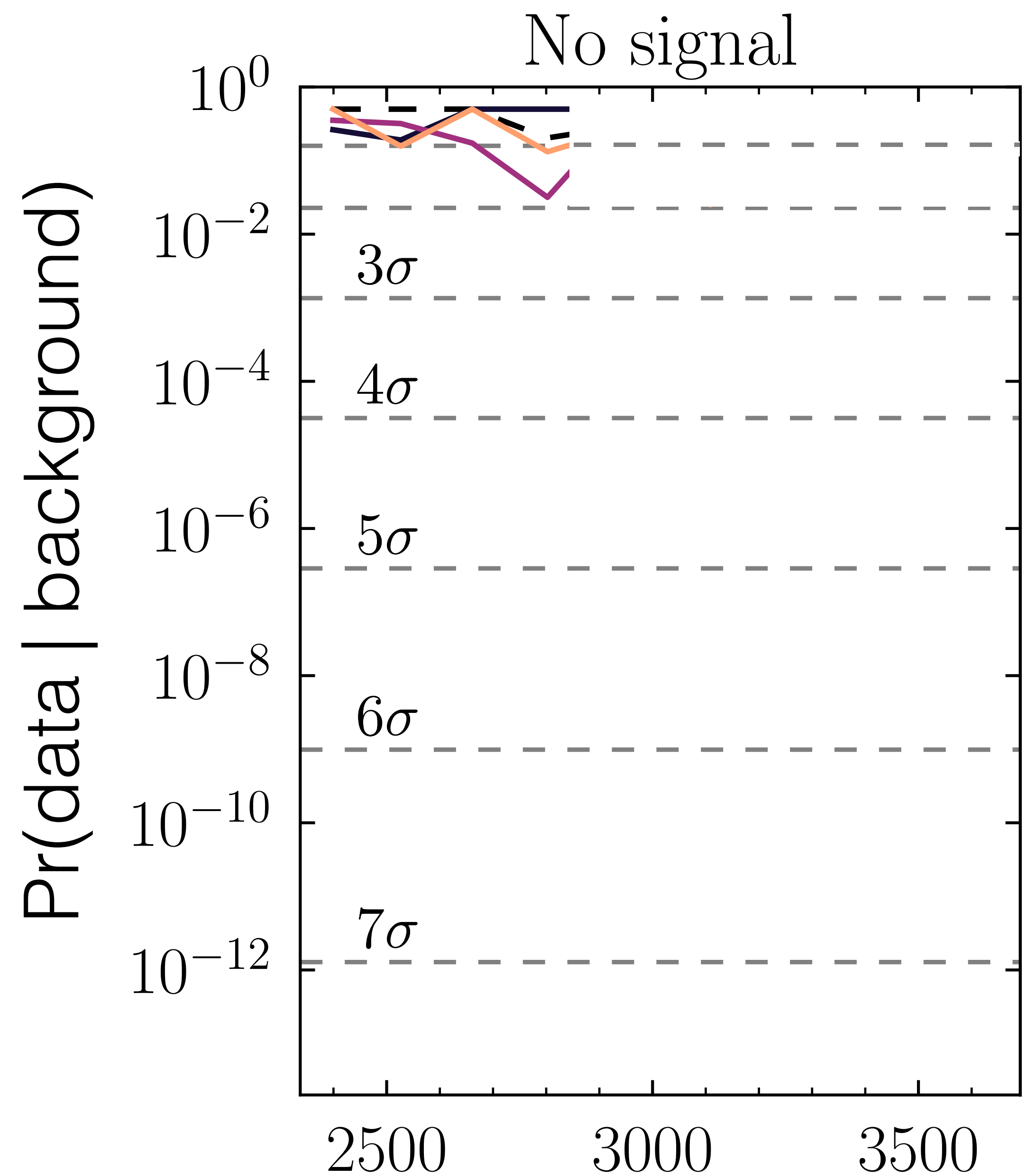
Example: two-jet search



no cut on NN
most 10% signal-region-like
most 1% signal-region-like
most 0.2% signal-region-like

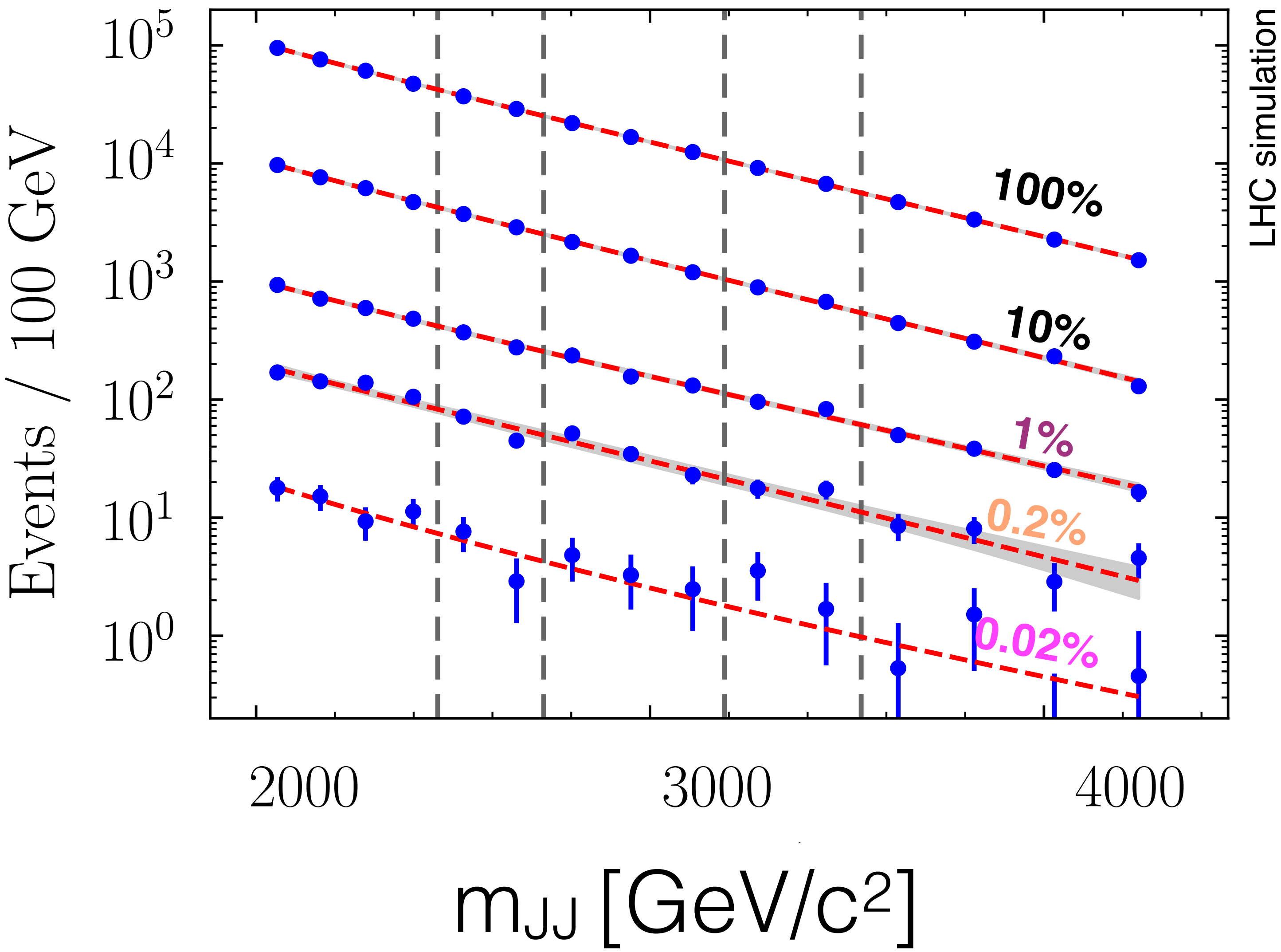
Example: two-jet search

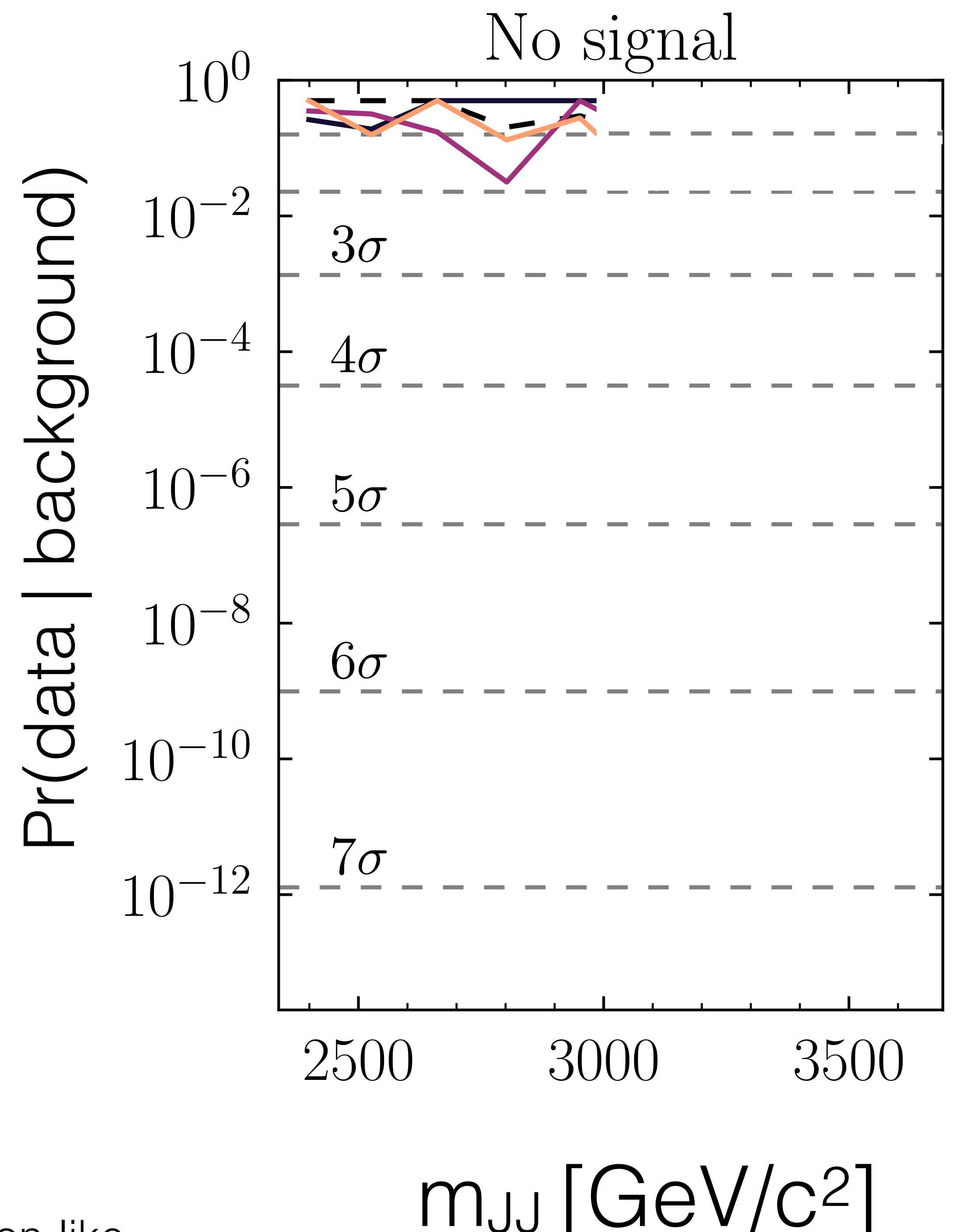




no cut on NN
most 10% signal-region-like
most 1% signal-region-like
most 0.2% signal-region-like

Example: two-jet search





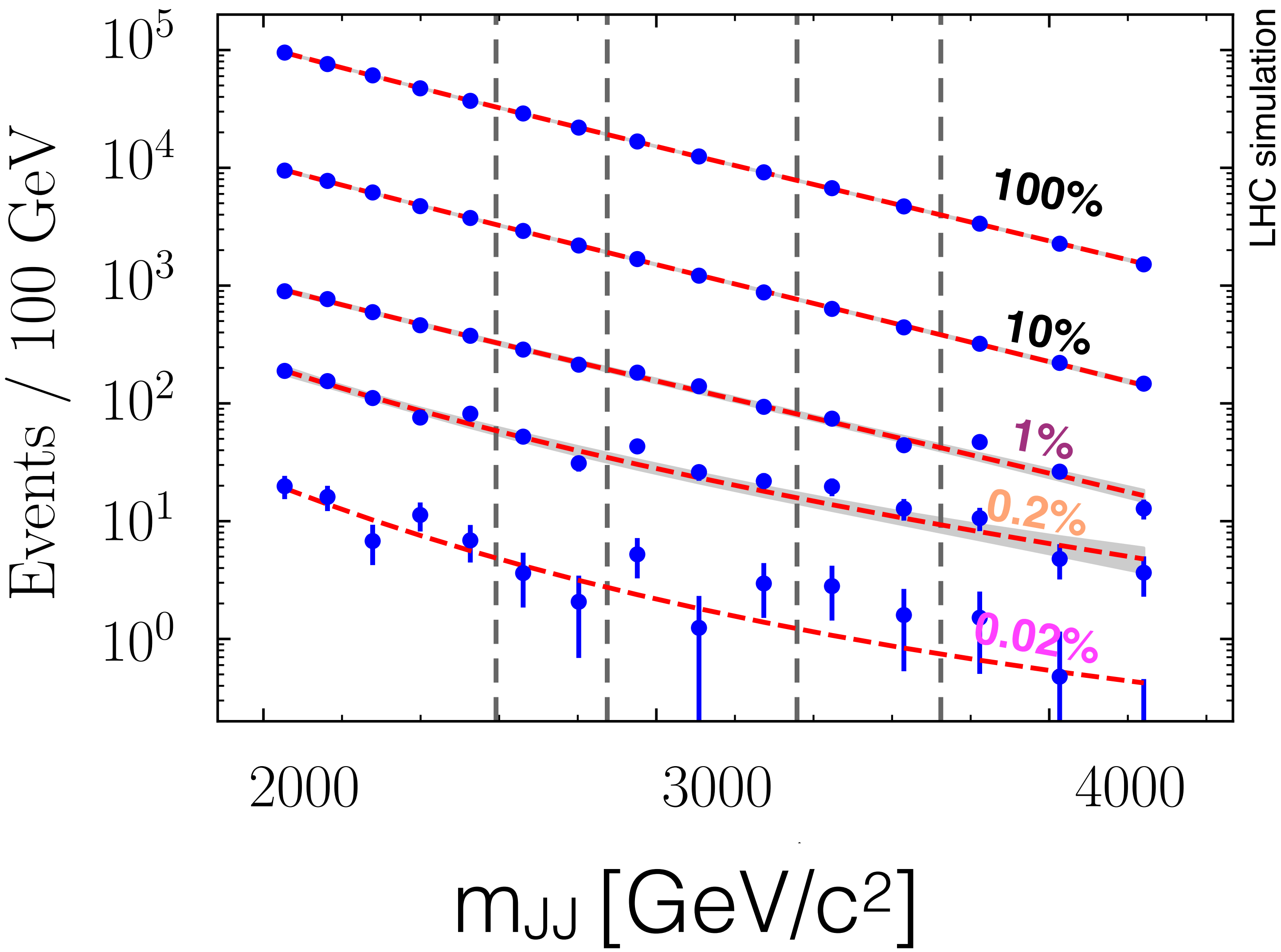
no cut on NN

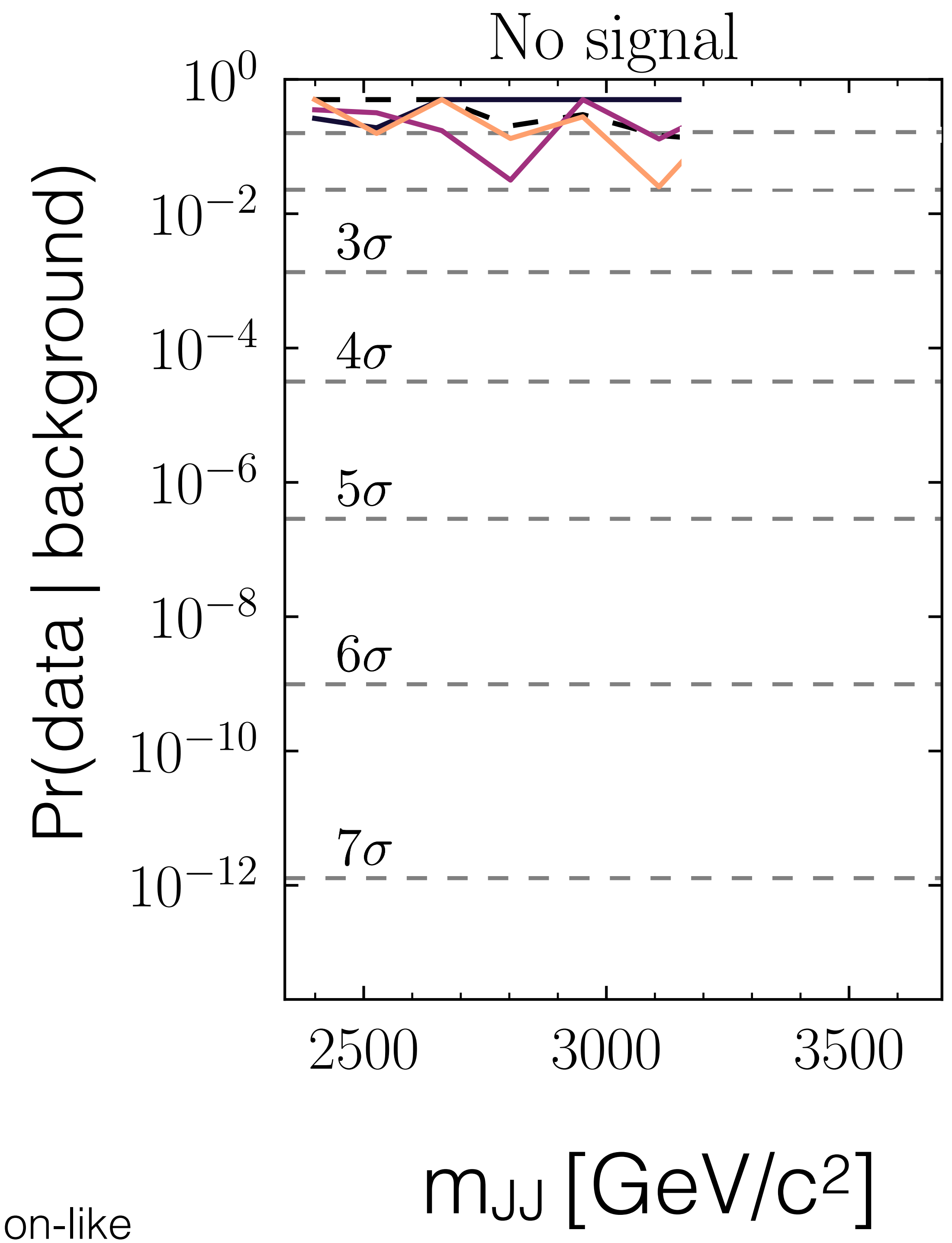
most 10% signal-region-like

most 1% signal-region-like

most 0.2% signal-region-like

Example: two-jet search





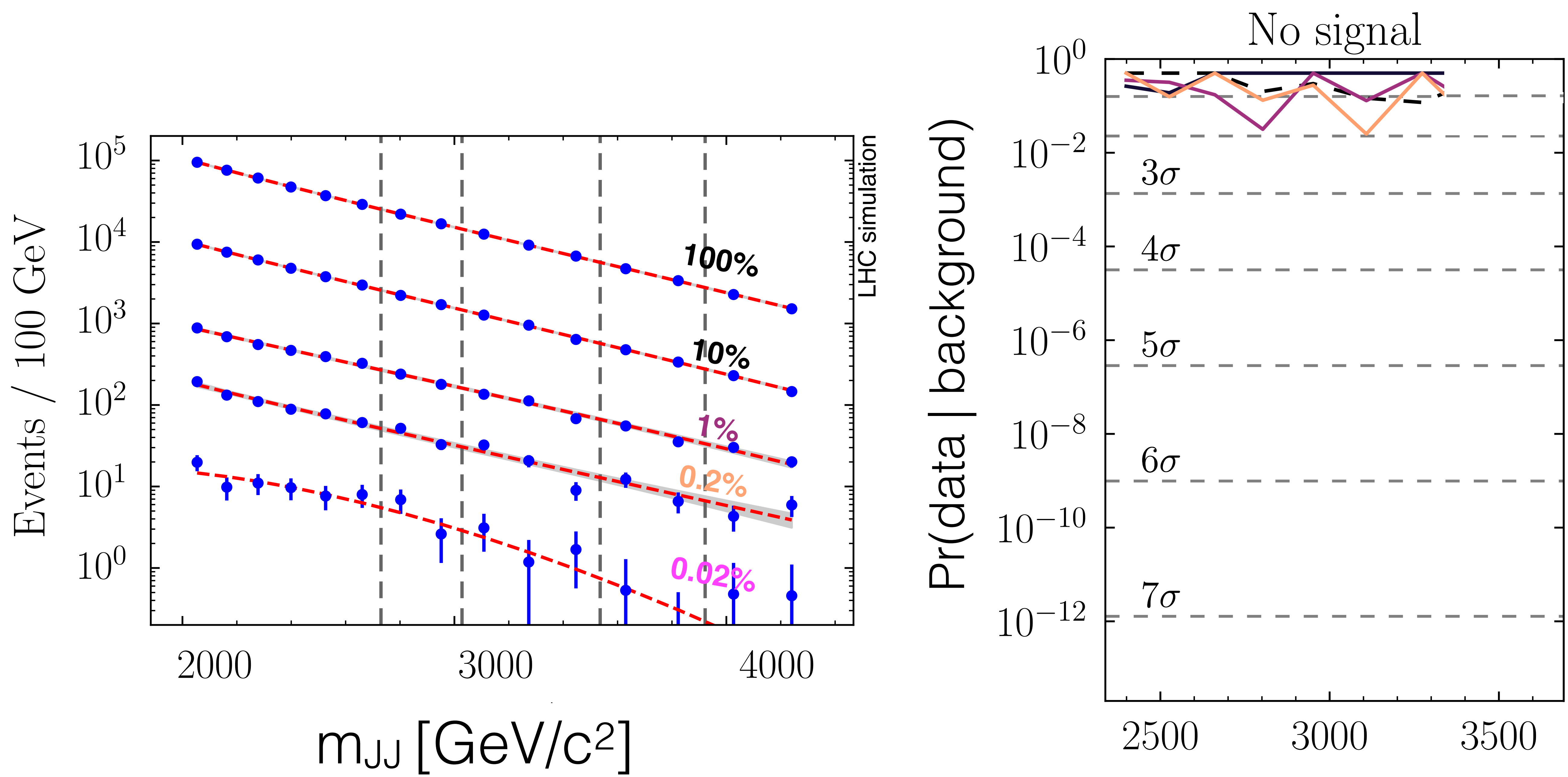
no cut on NN

most 10% signal-region-like

most 1% signal-region-like

most 0.2% signal-region-like

Example: two-jet search



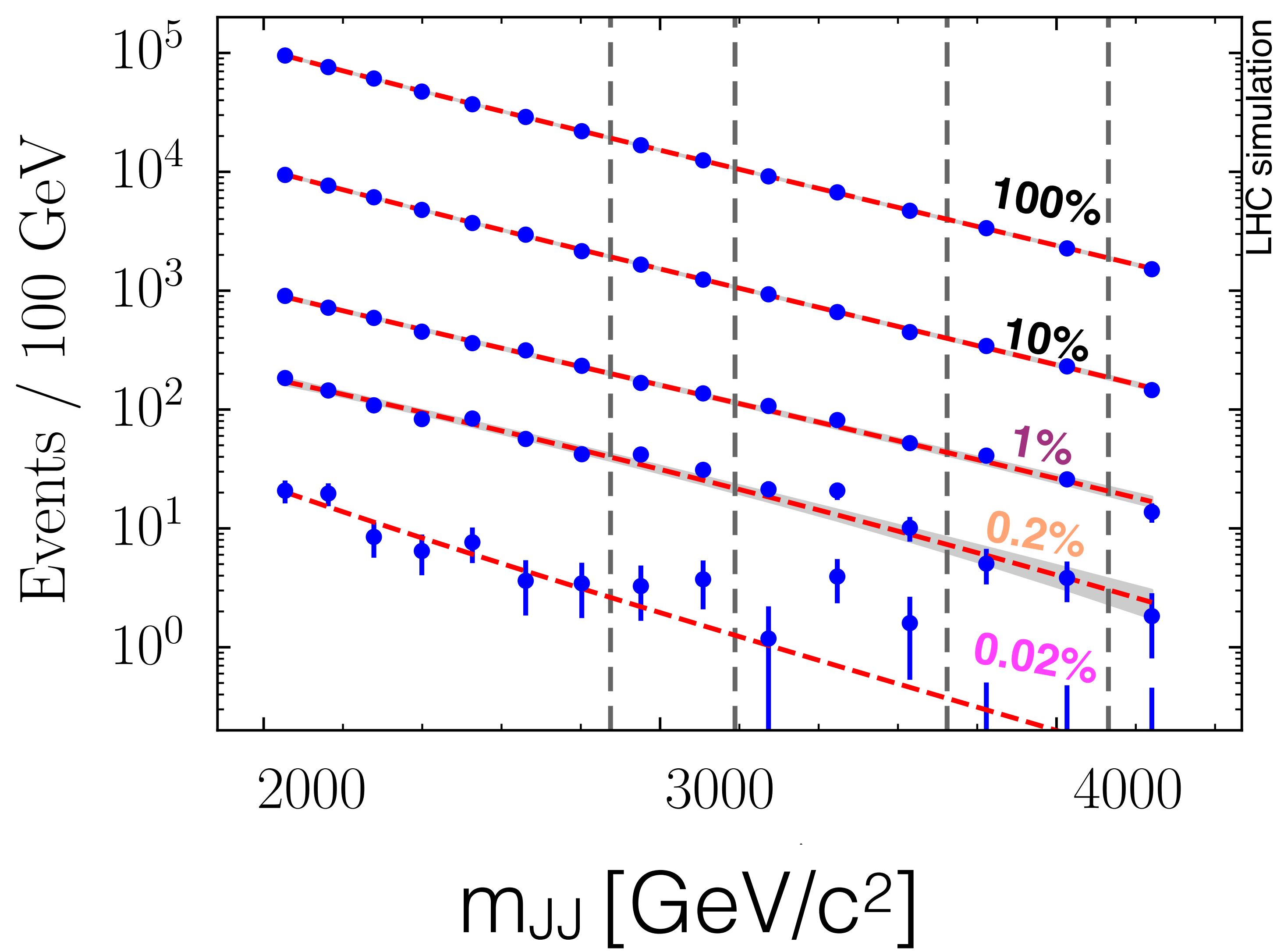
no cut on NN

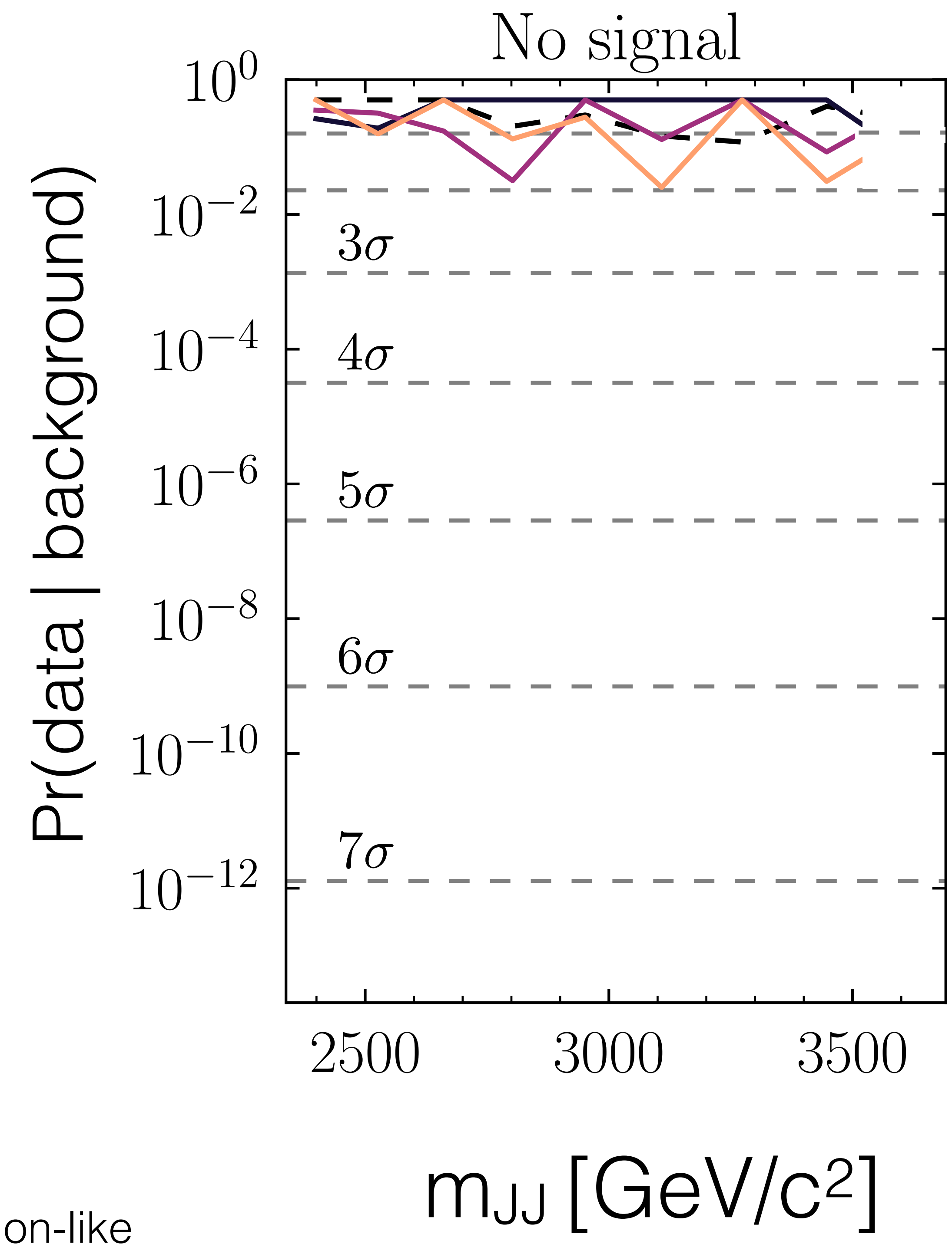
most 10% signal-region-like

most 1% signal-region-like

most 0.2% signal-region-like

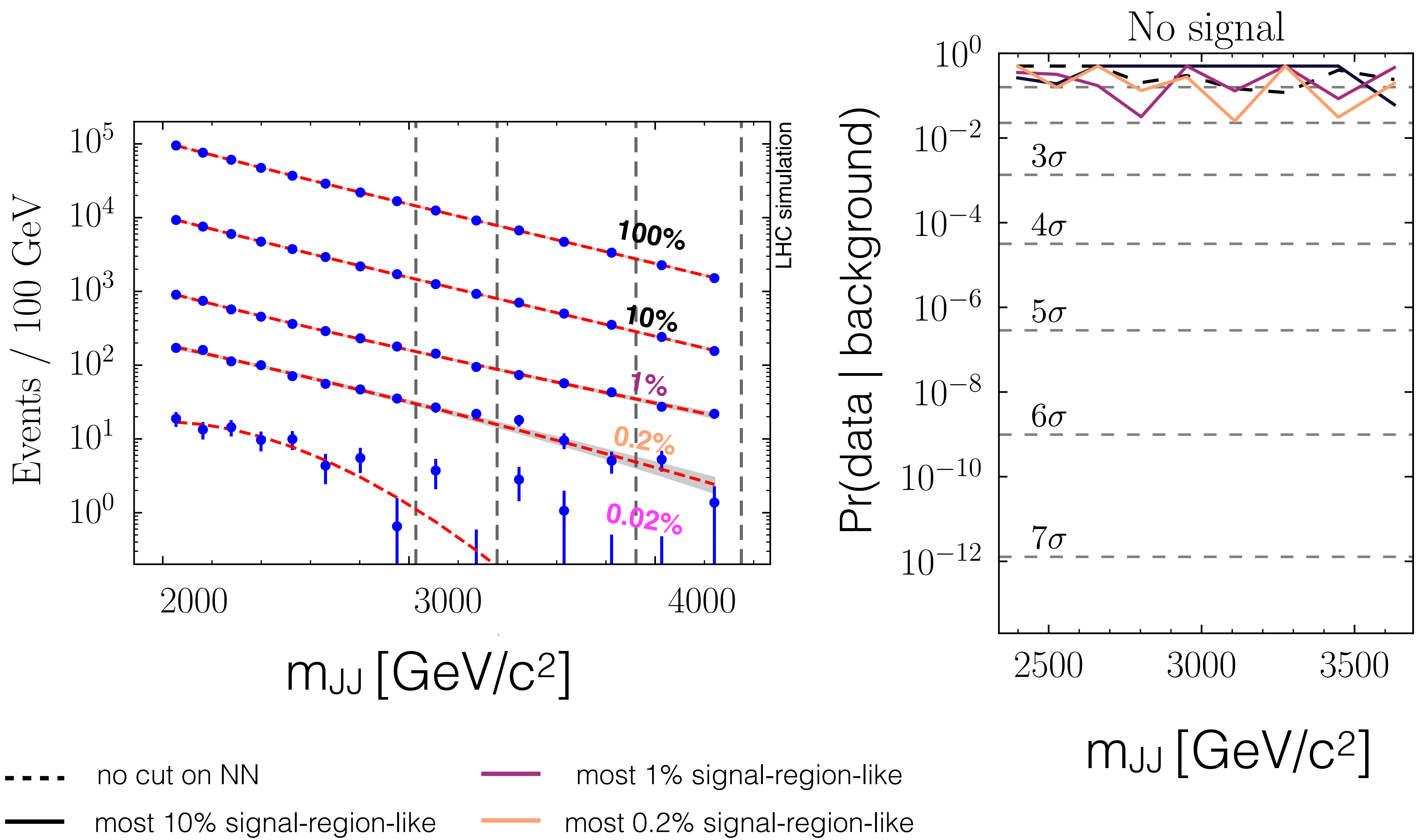
Example: two-jet search



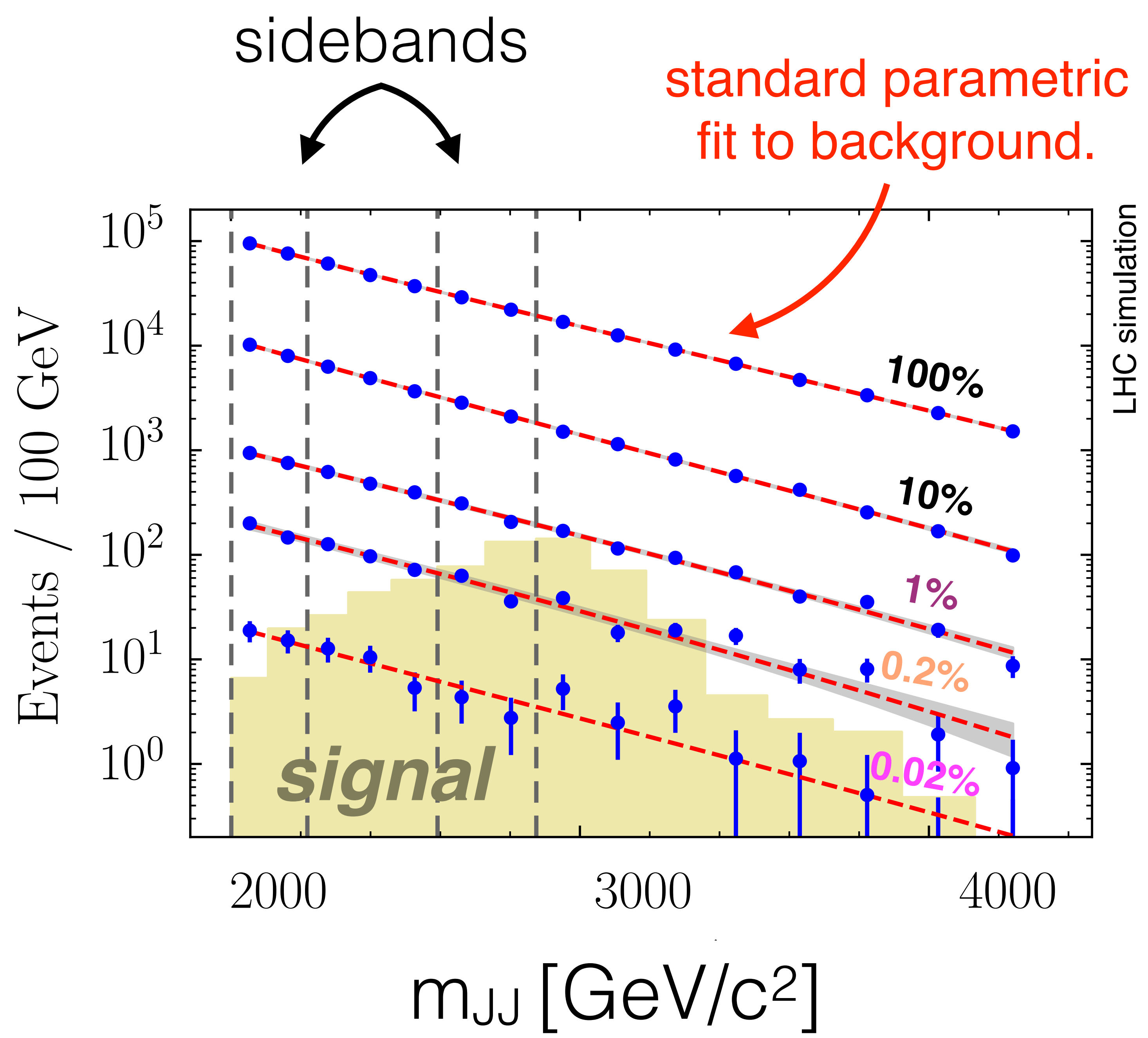


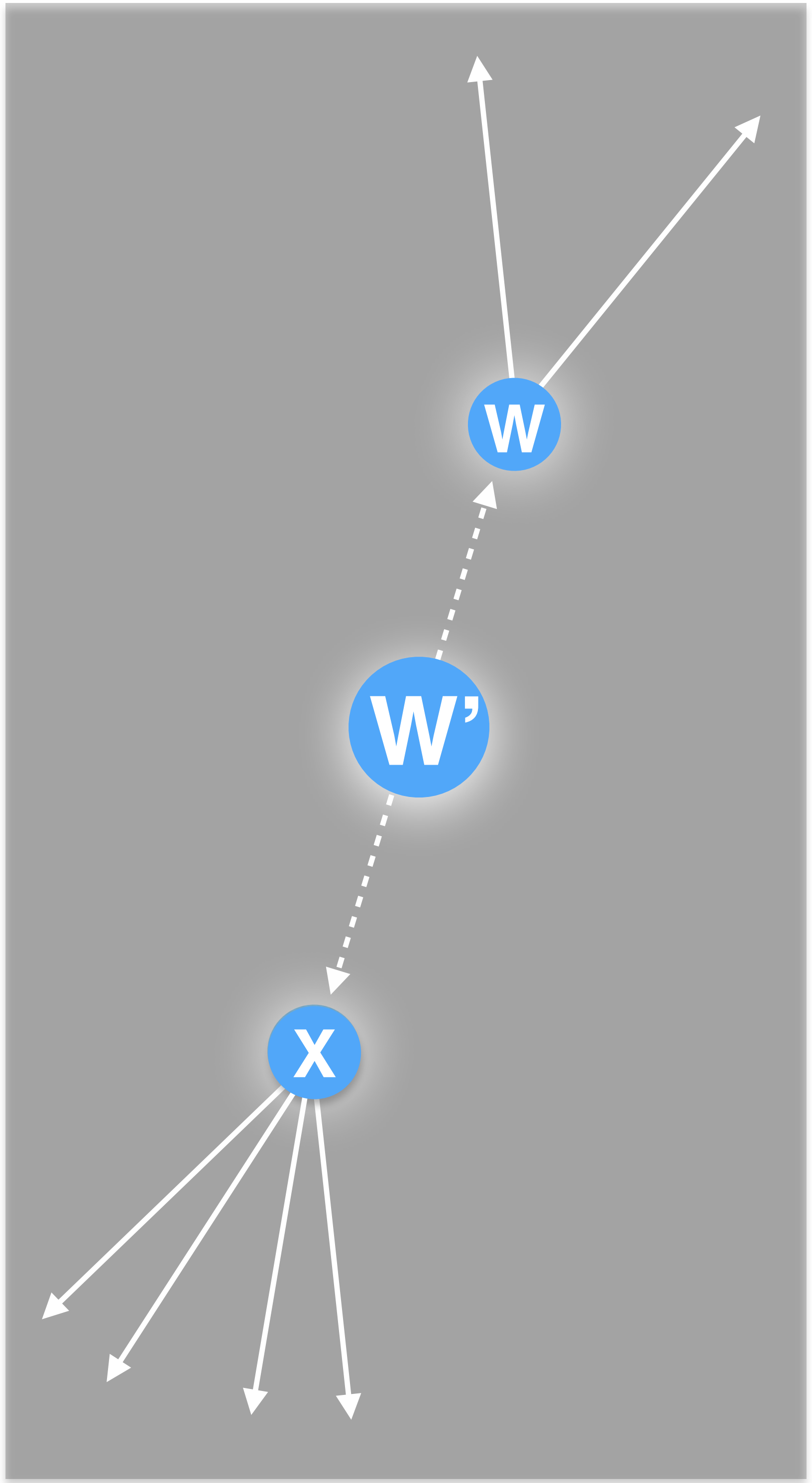
- - - no cut on NN
— most 10% signal-region-like
— most 1% signal-region-like
— most 0.2% signal-region-like

Example: two-jet search

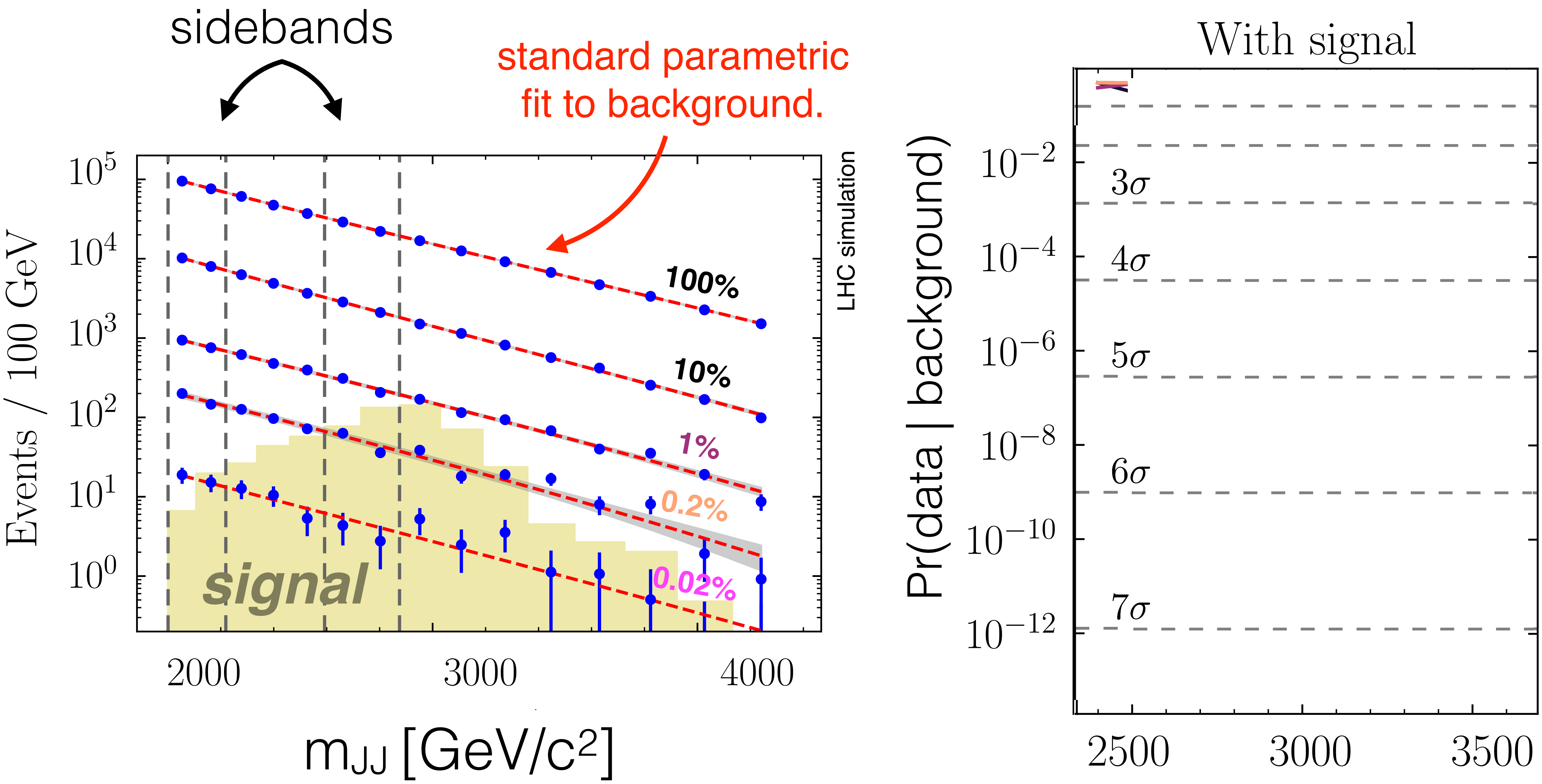


…and when there is a signal?





…and when there is a signal?



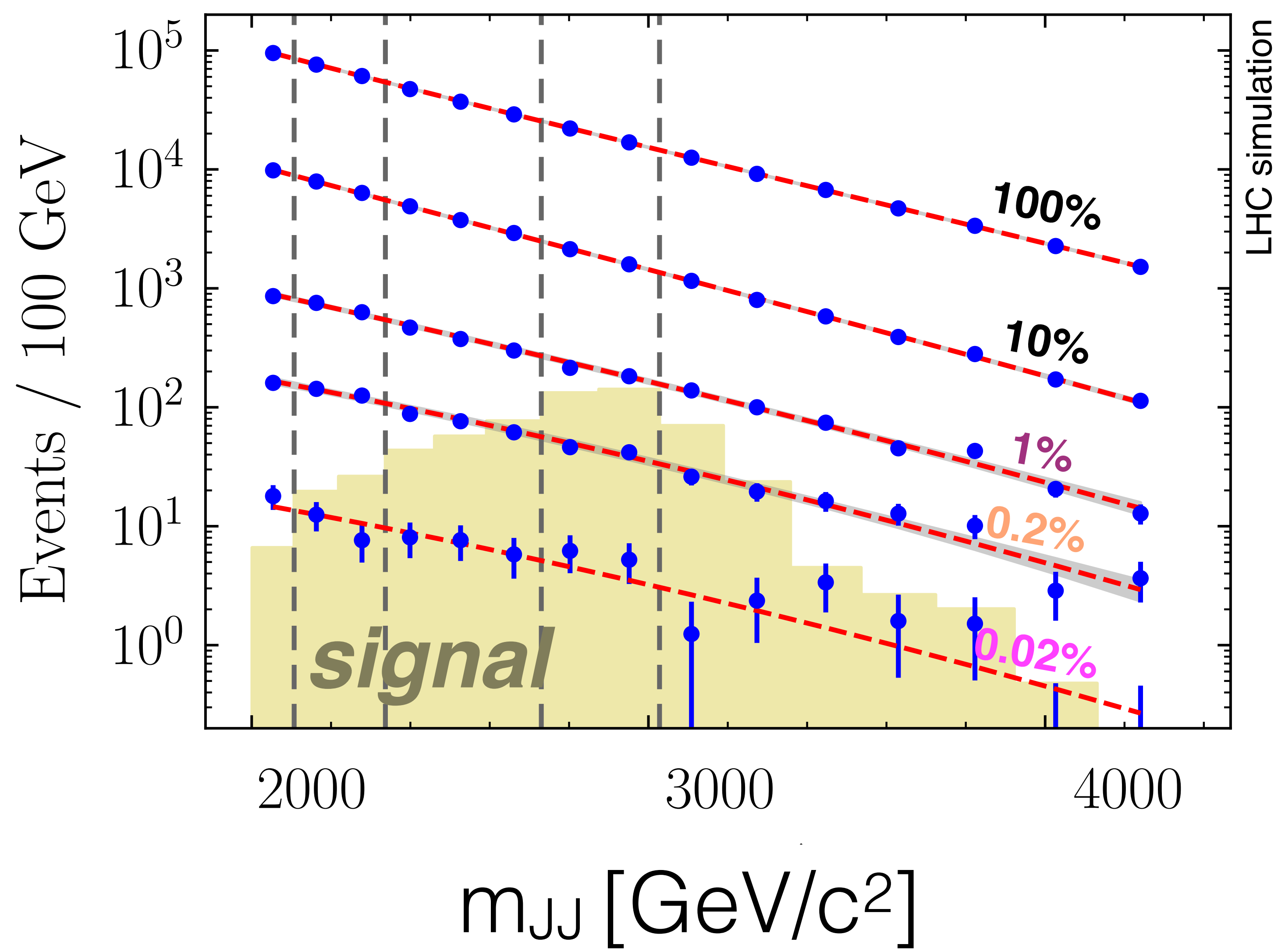
no cut on NN

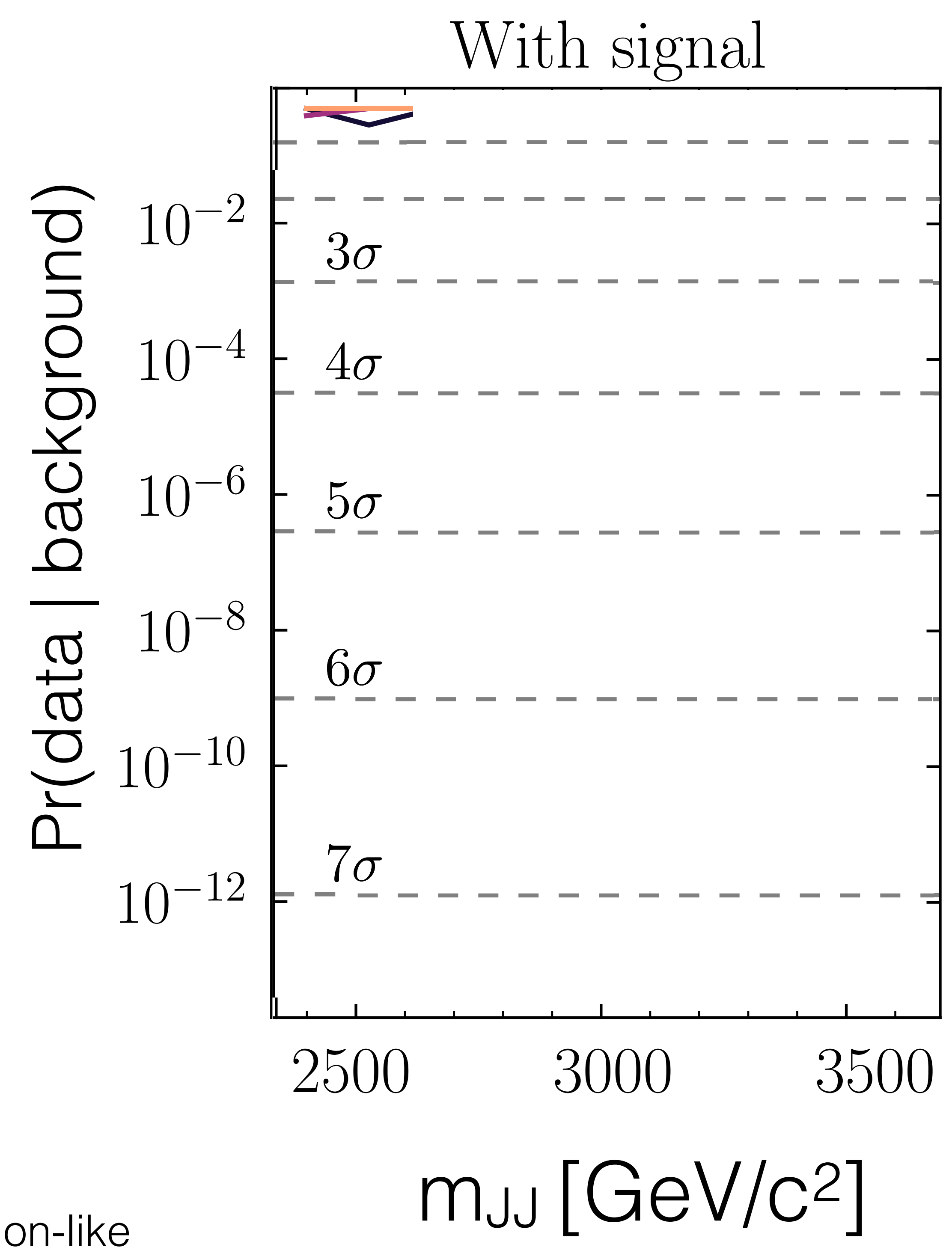
most 10% signal-region-like

most 1% signal-region-like

most 0.2% signal-region-like

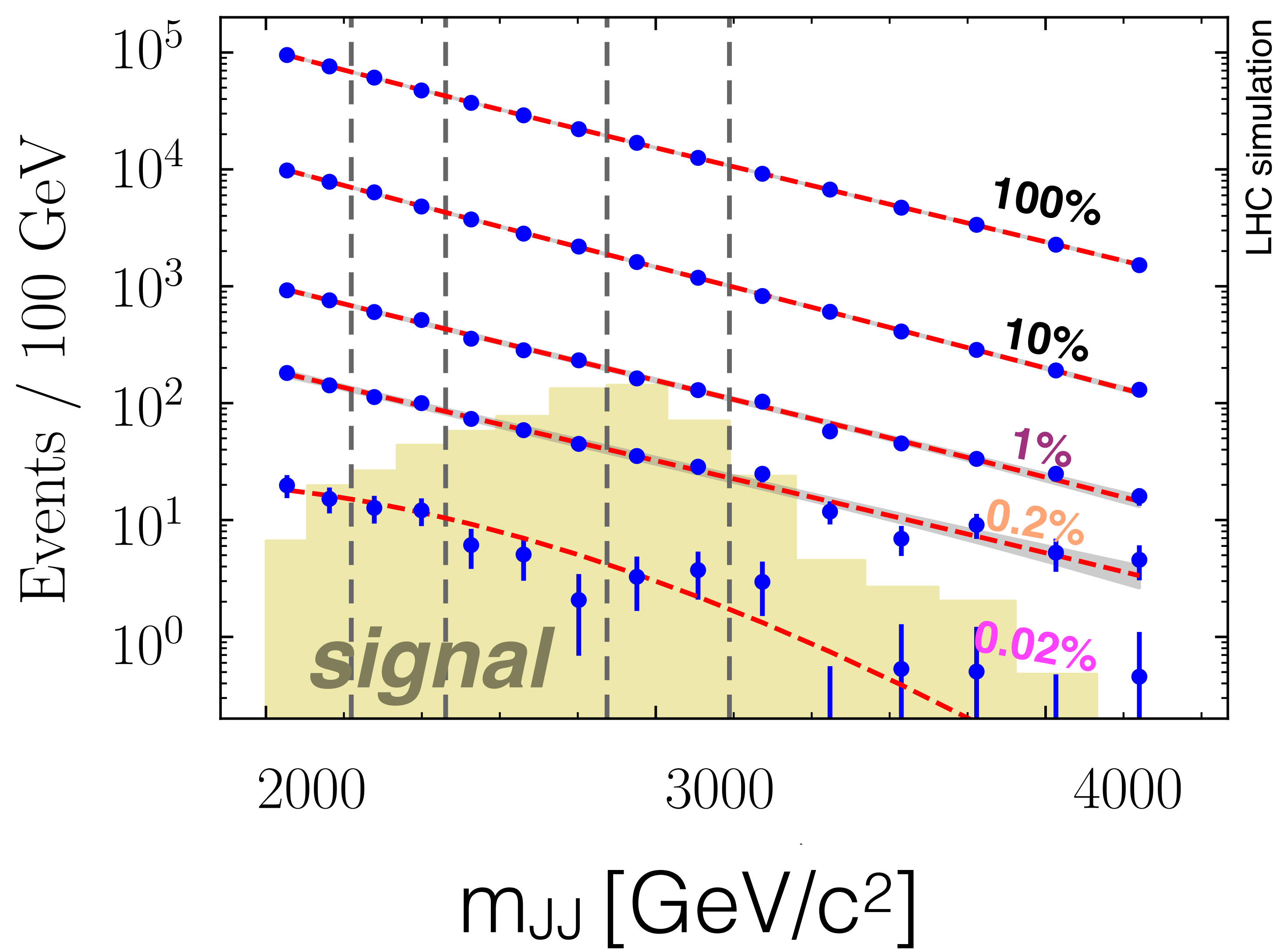
…and when there is a signal?

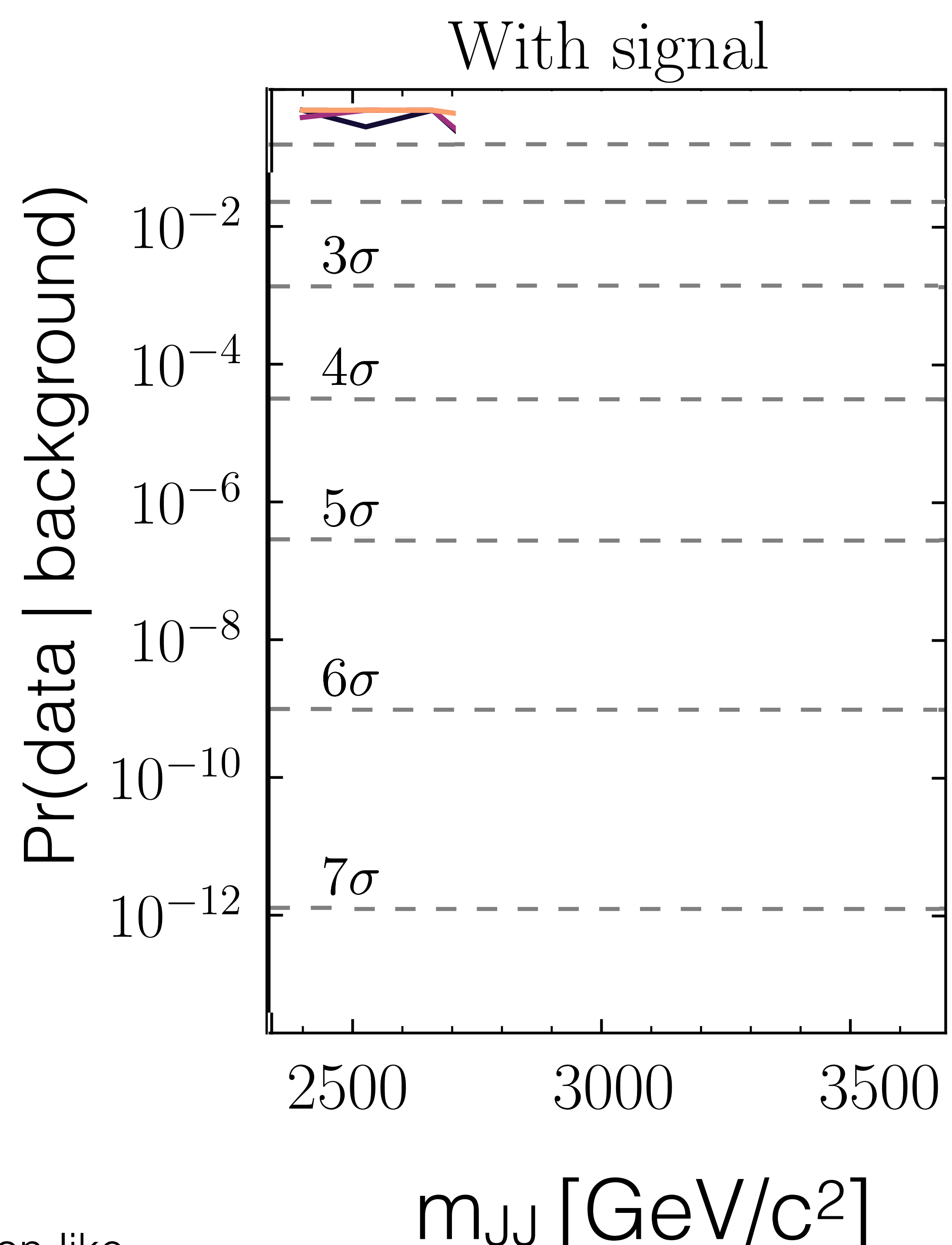




no cut on NN
most 10% signal-region-like
most 1% signal-region-like
most 0.2% signal-region-like

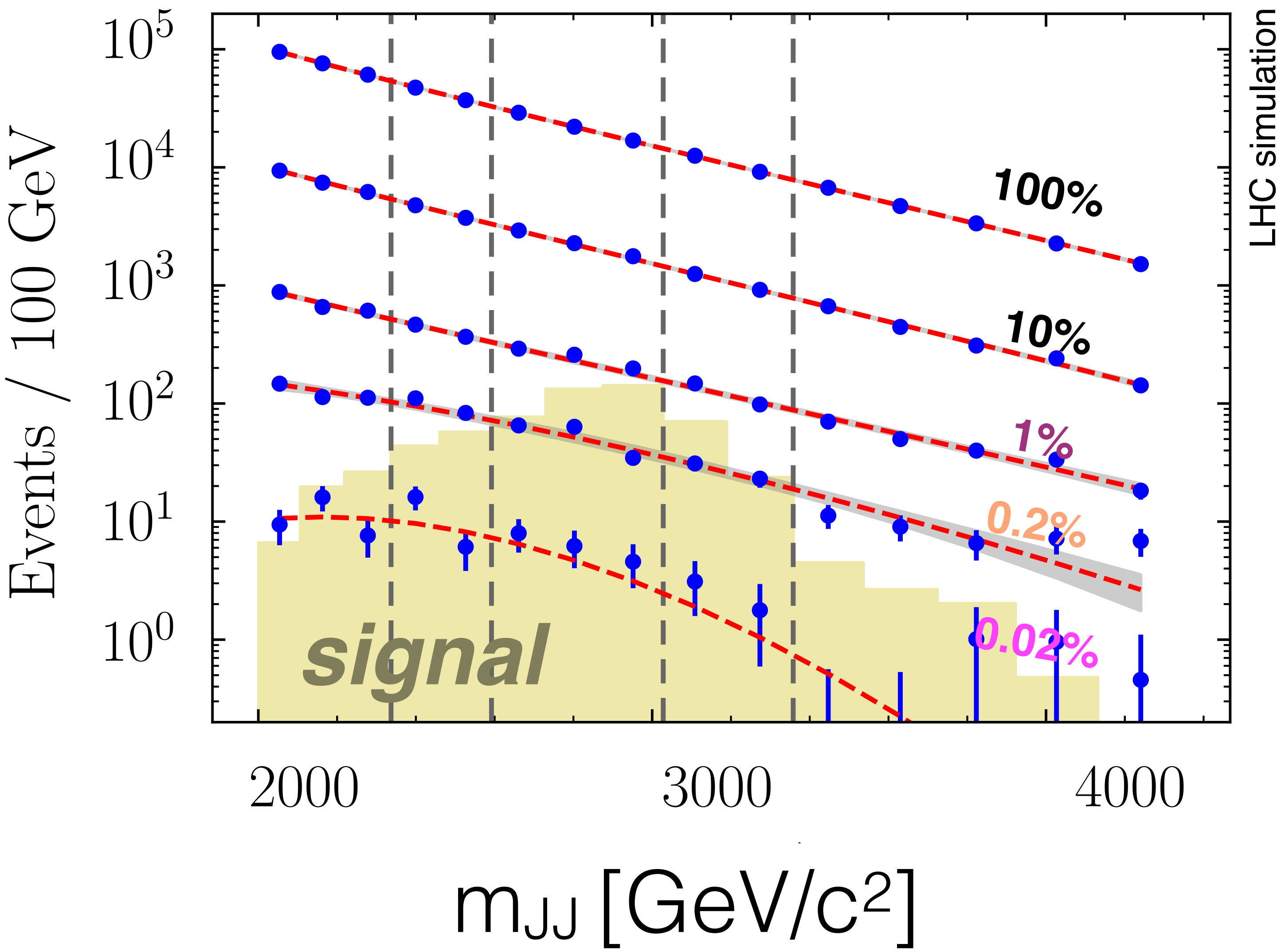
…and when there is a signal?

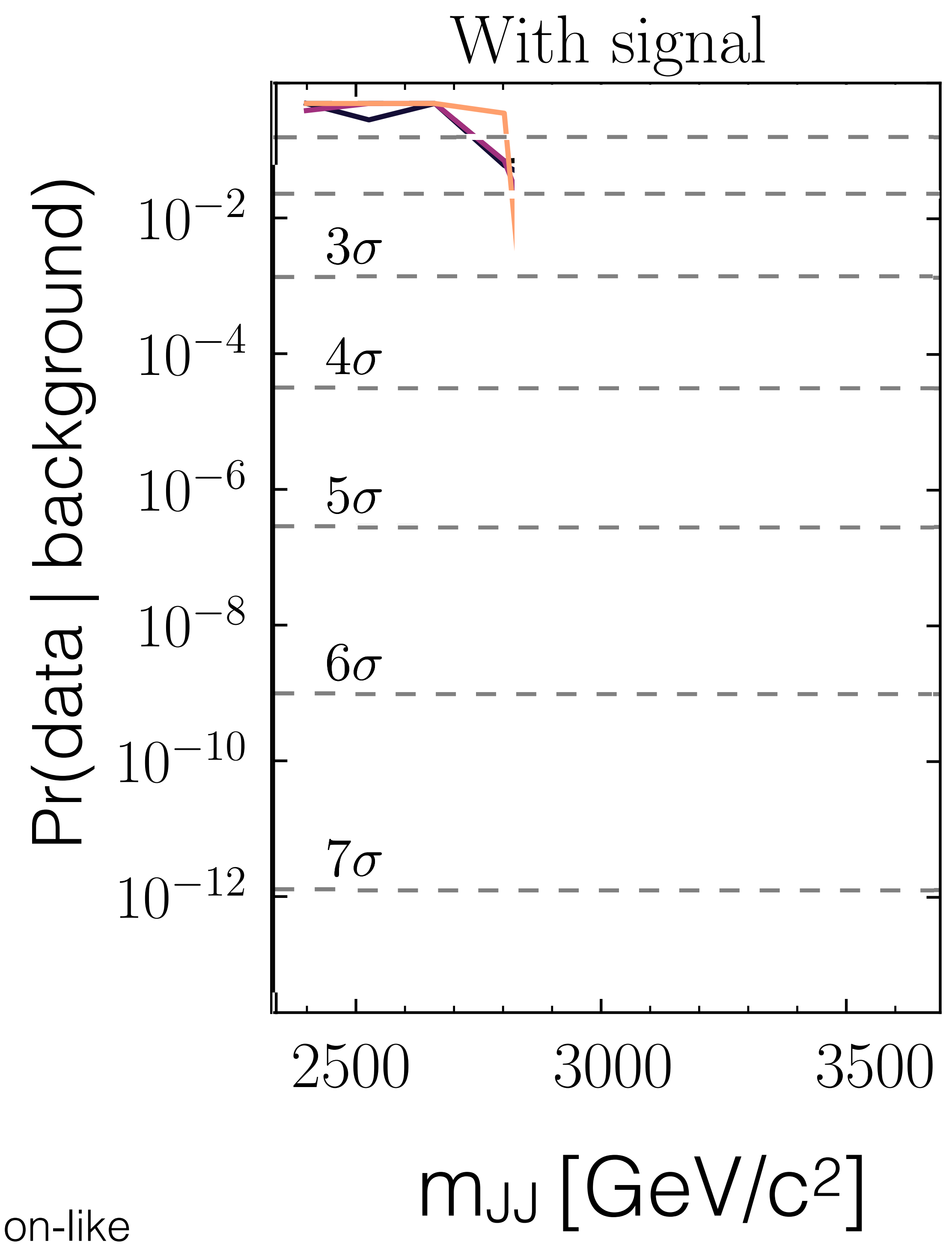




no cut on NN
most 10% signal-region-like
most 1% signal-region-like
most 0.2% signal-region-like

…and when there is a signal?





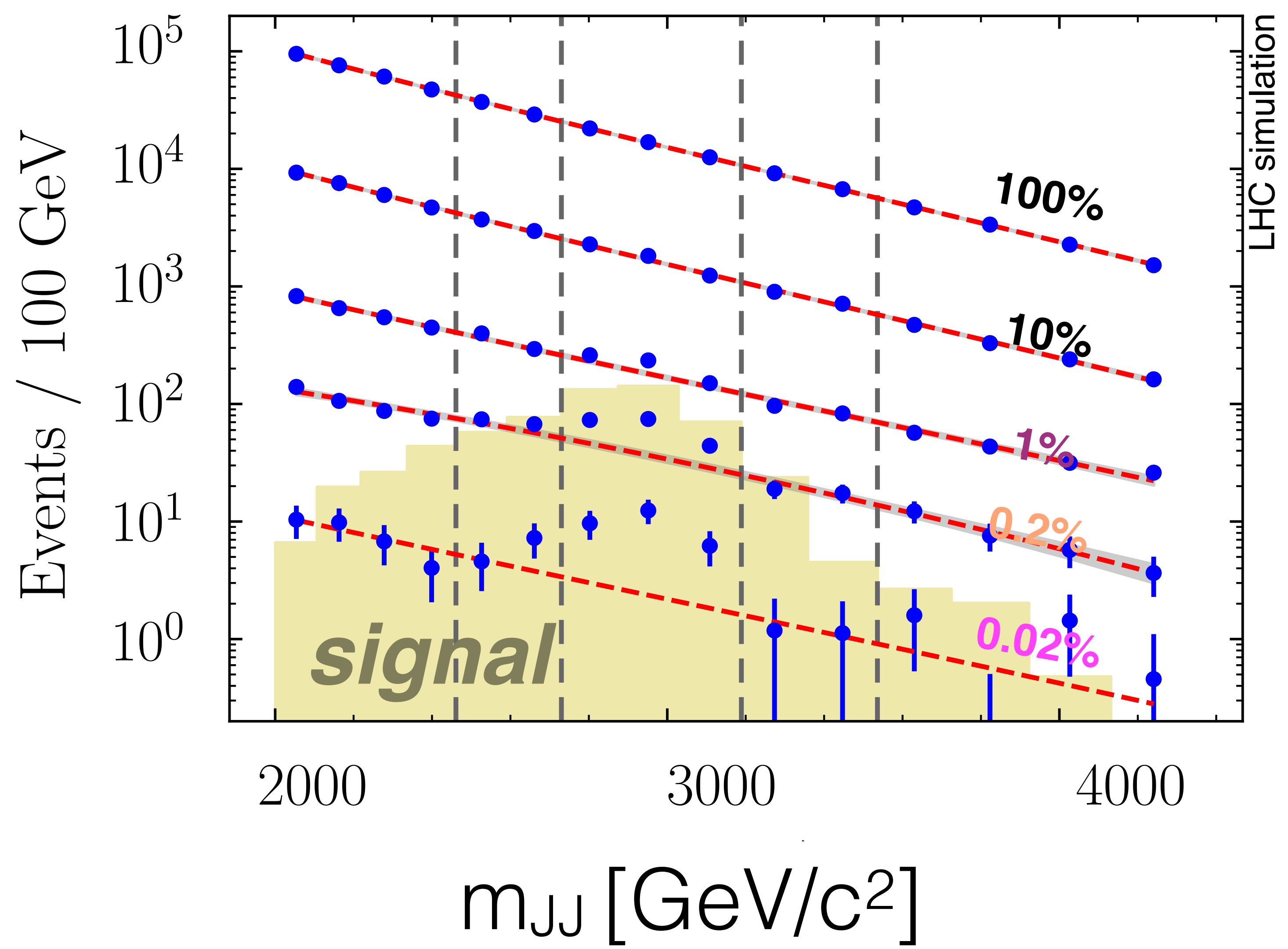
no cut on NN

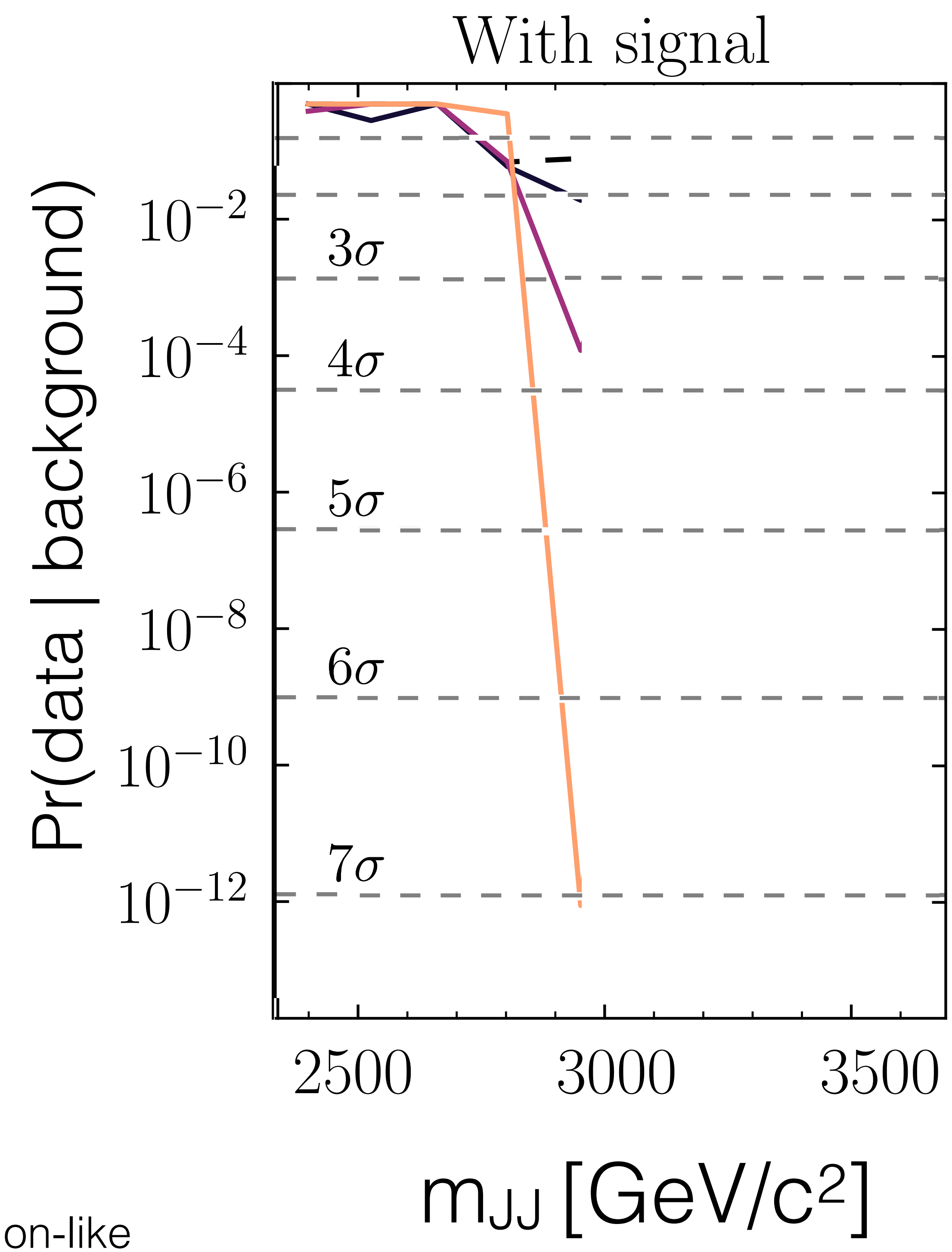
most 10% signal-region-like

most 1% signal-region-like

most 0.2% signal-region-like

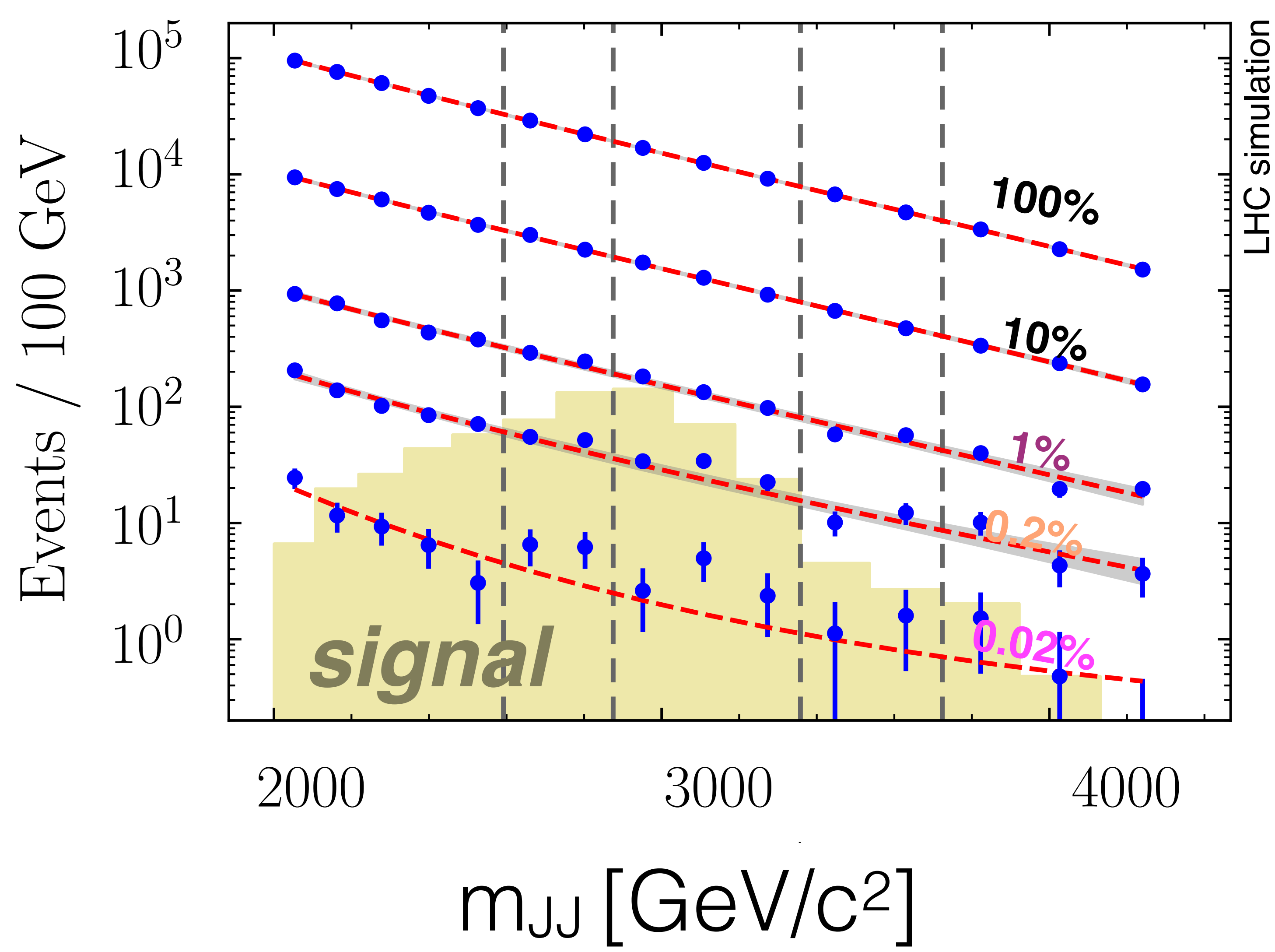
…and when there is a signal?

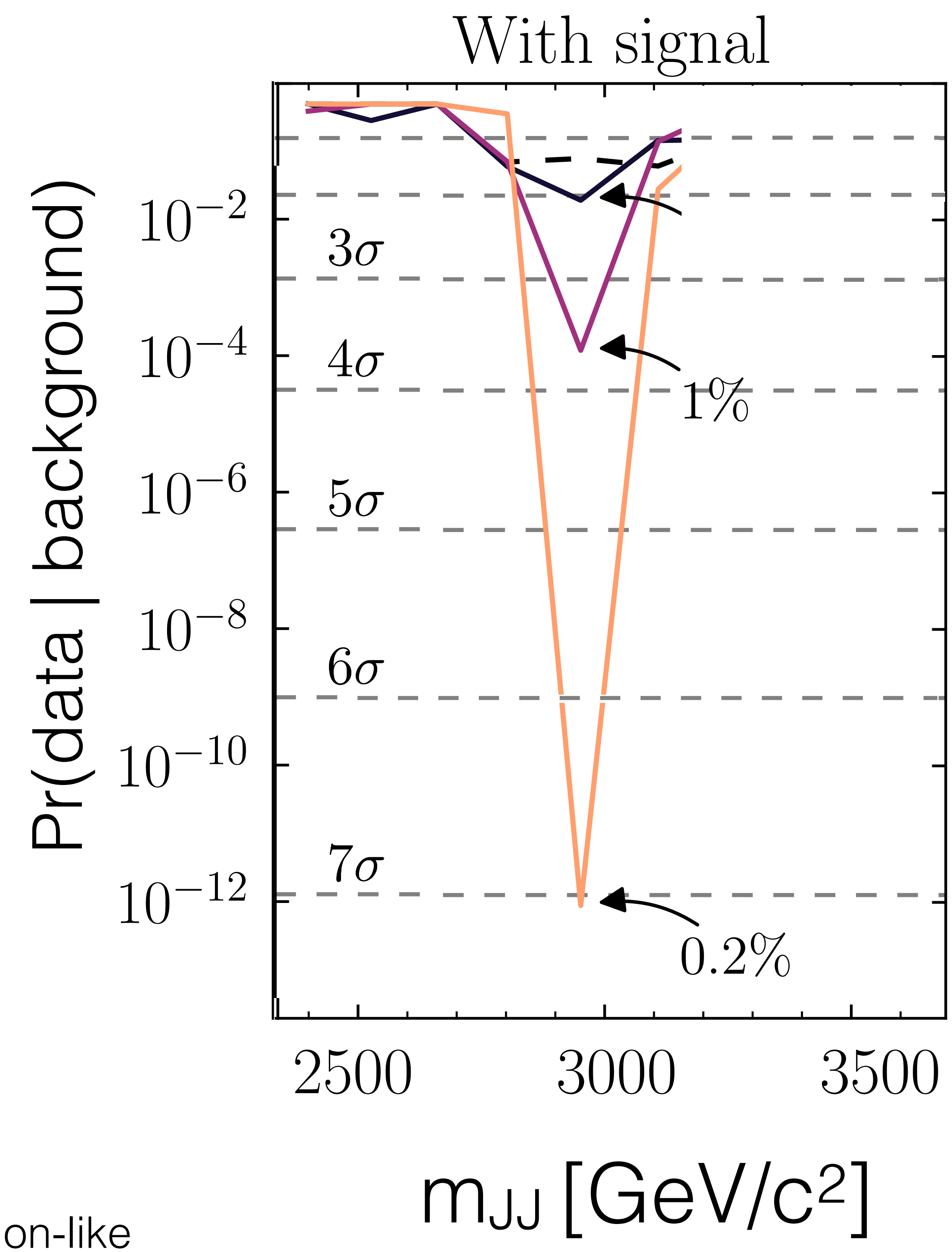




no cut on NN
most 10% signal-region-like
most 1% signal-region-like
most 0.2% signal-region-like

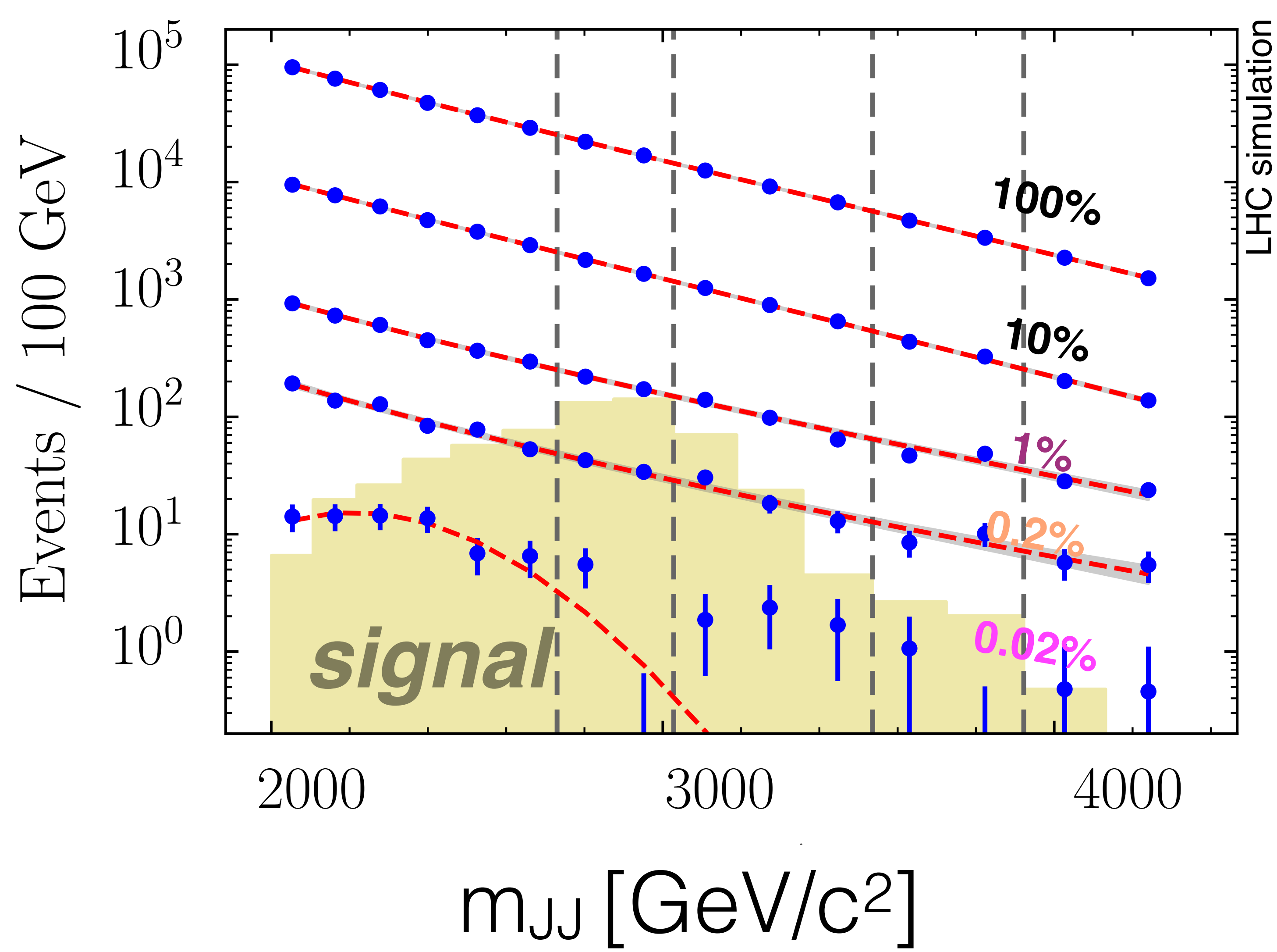
…and when there is a signal?

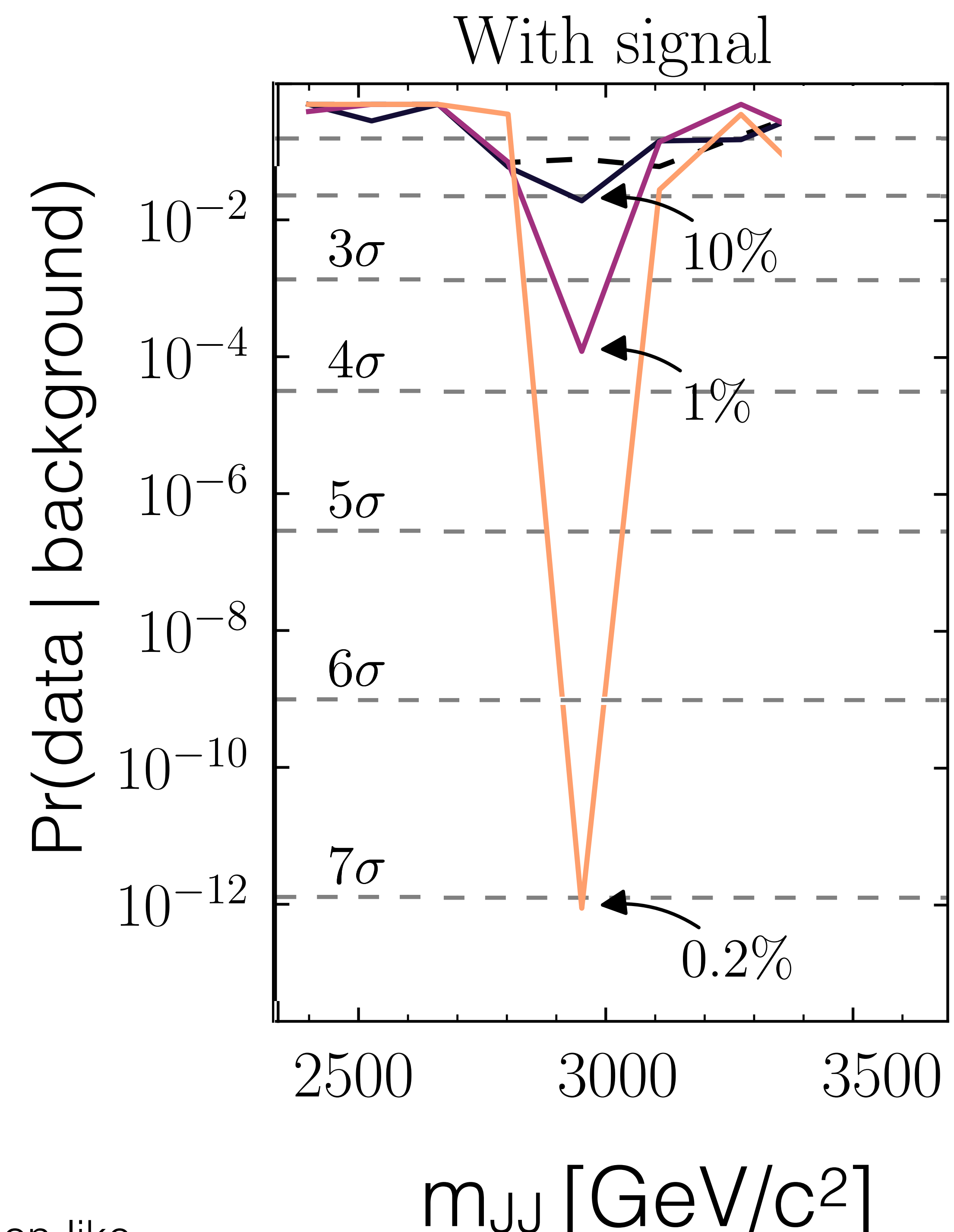




no cut on NN
most 10% signal-region-like
most 1% signal-region-like
most 0.2% signal-region-like

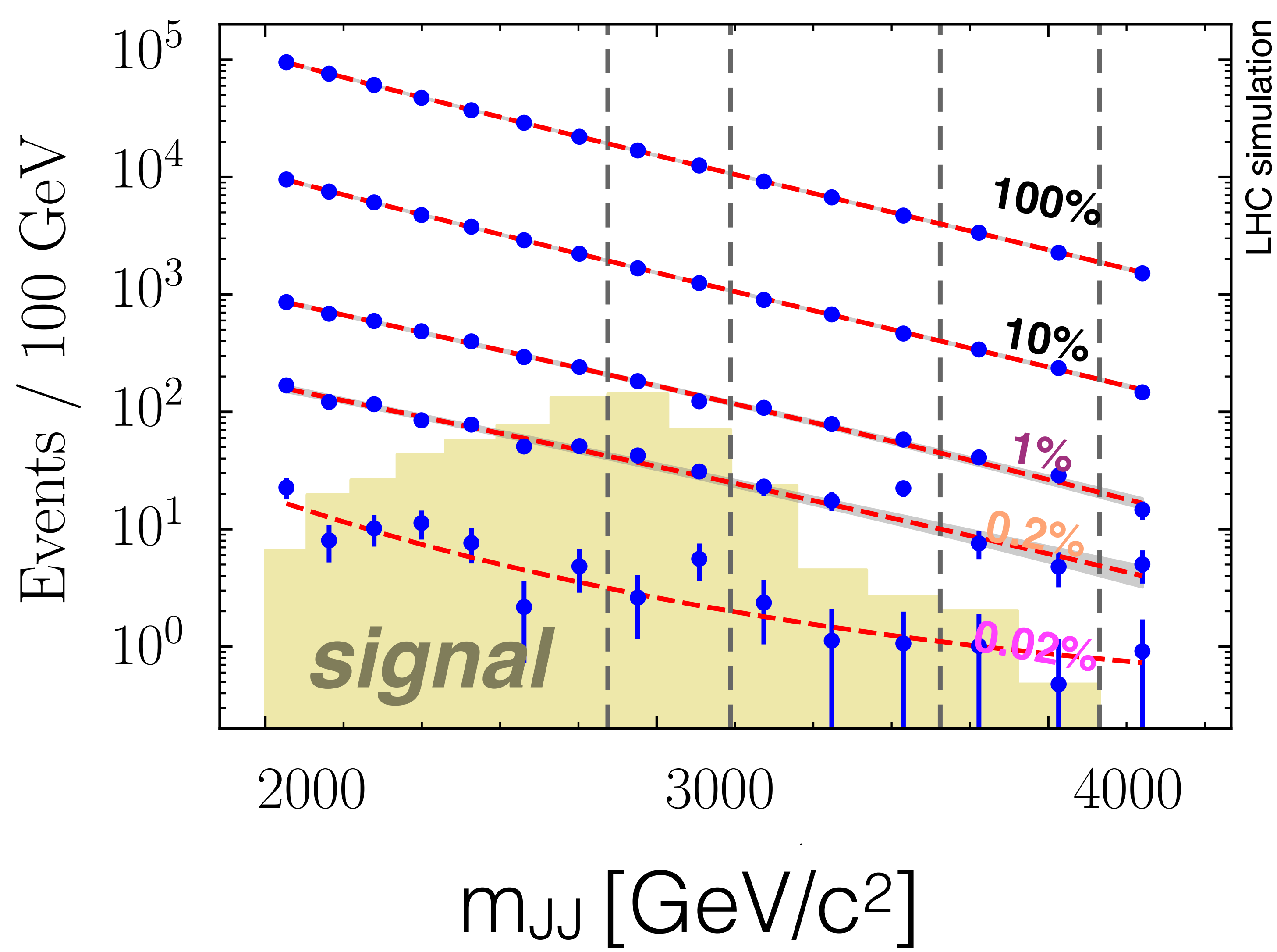
…and when there is a signal?

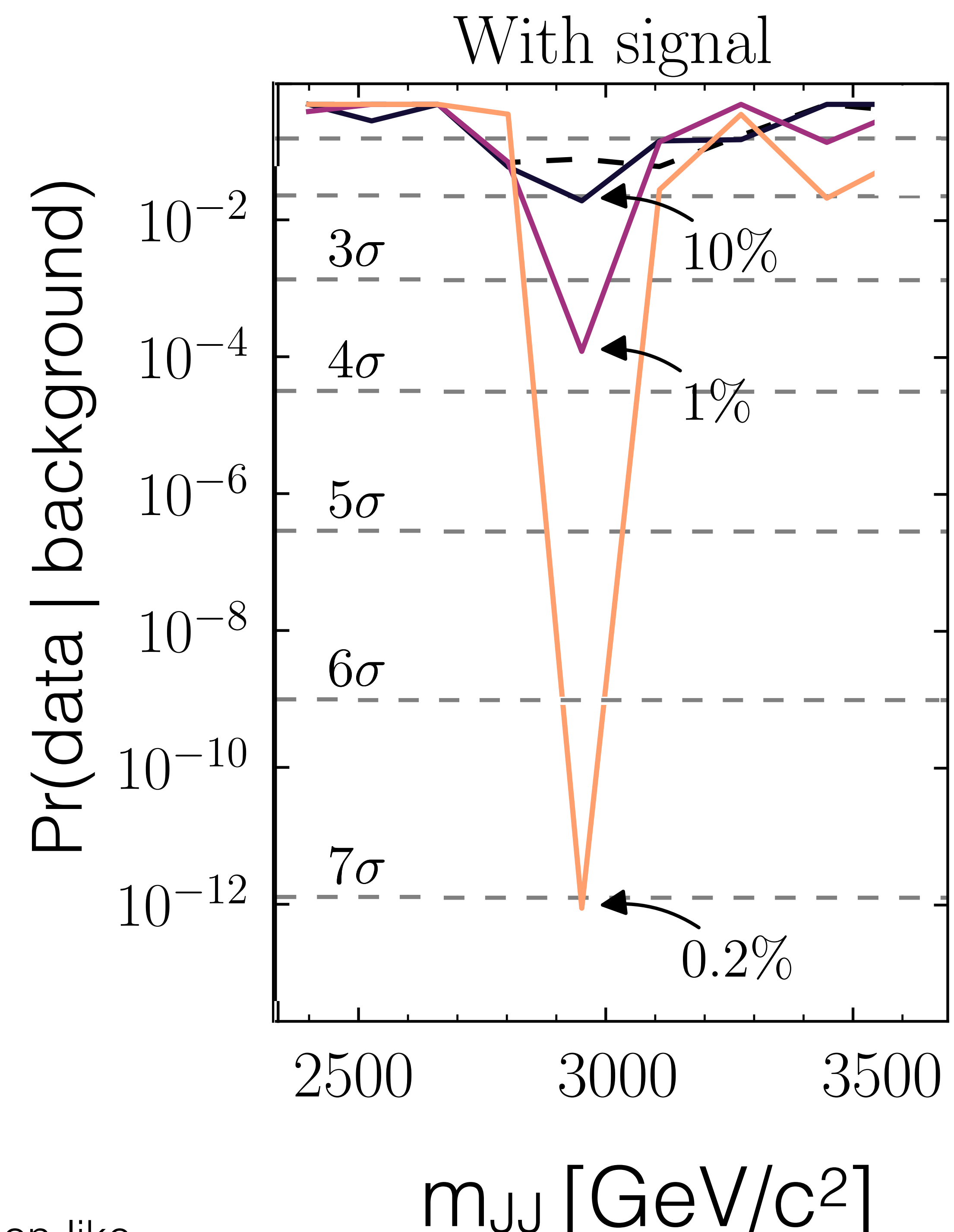




no cut on NN
most 10% signal-region-like
most 1% signal-region-like
most 0.2% signal-region-like

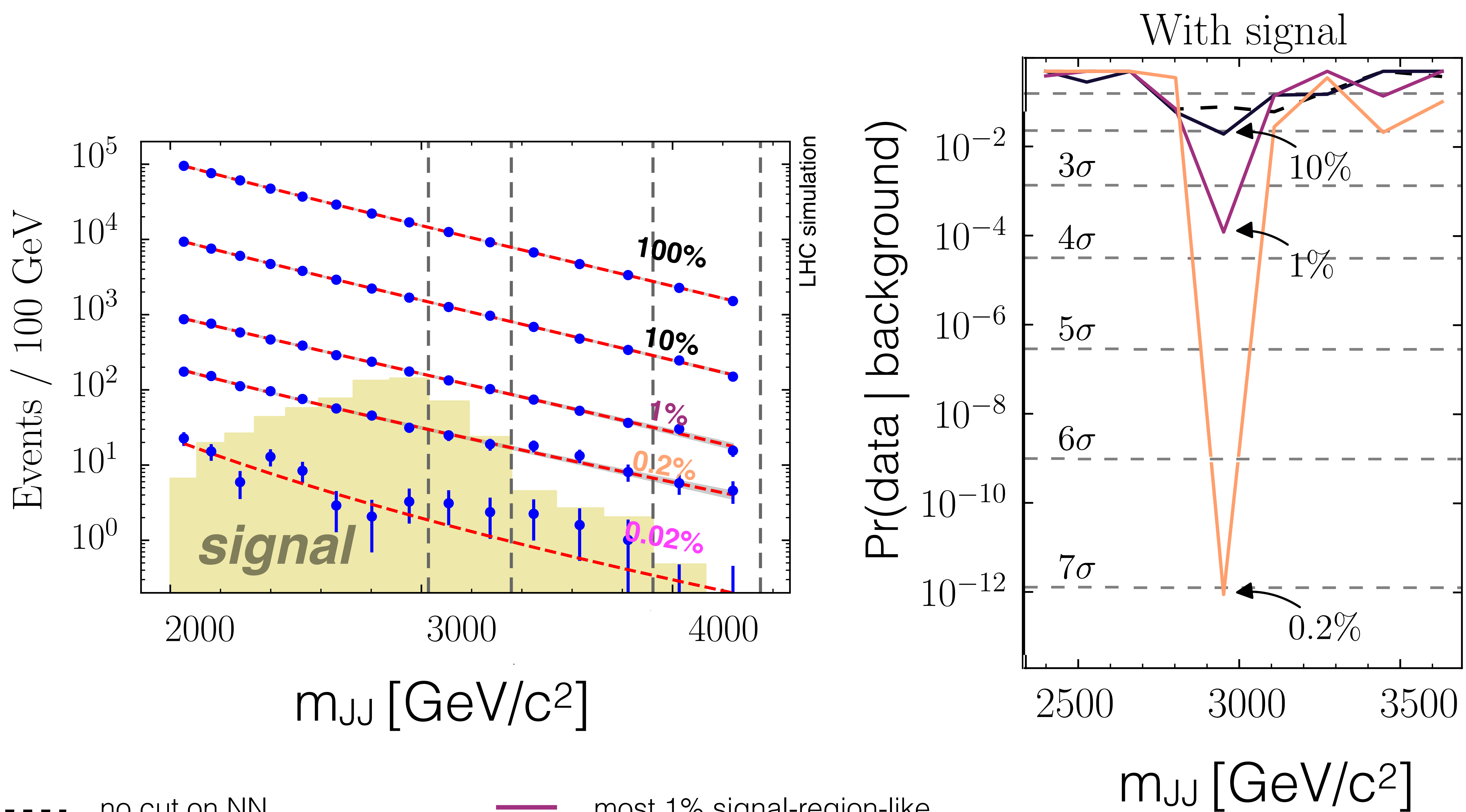
…and when there is a signal?





no cut on NN
most 10% signal-region-like
most 1% signal-region-like
most 0.2% signal-region-like

…and when there is a signal?



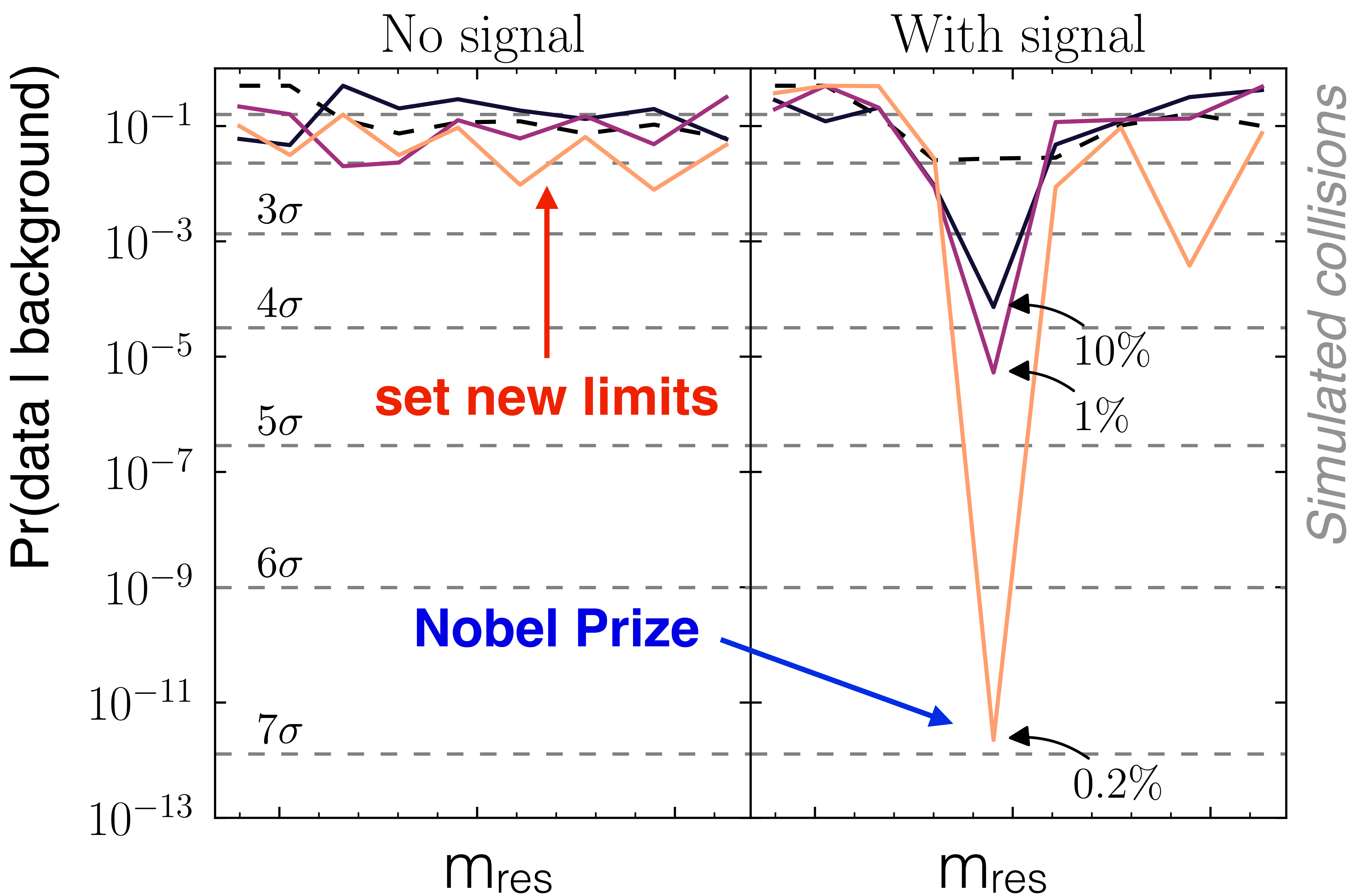
no cut on NN — most 1% signal-region-like

most 10% signal-region-like — most 0.2% signal-region-like

CWoLa hunting: overview

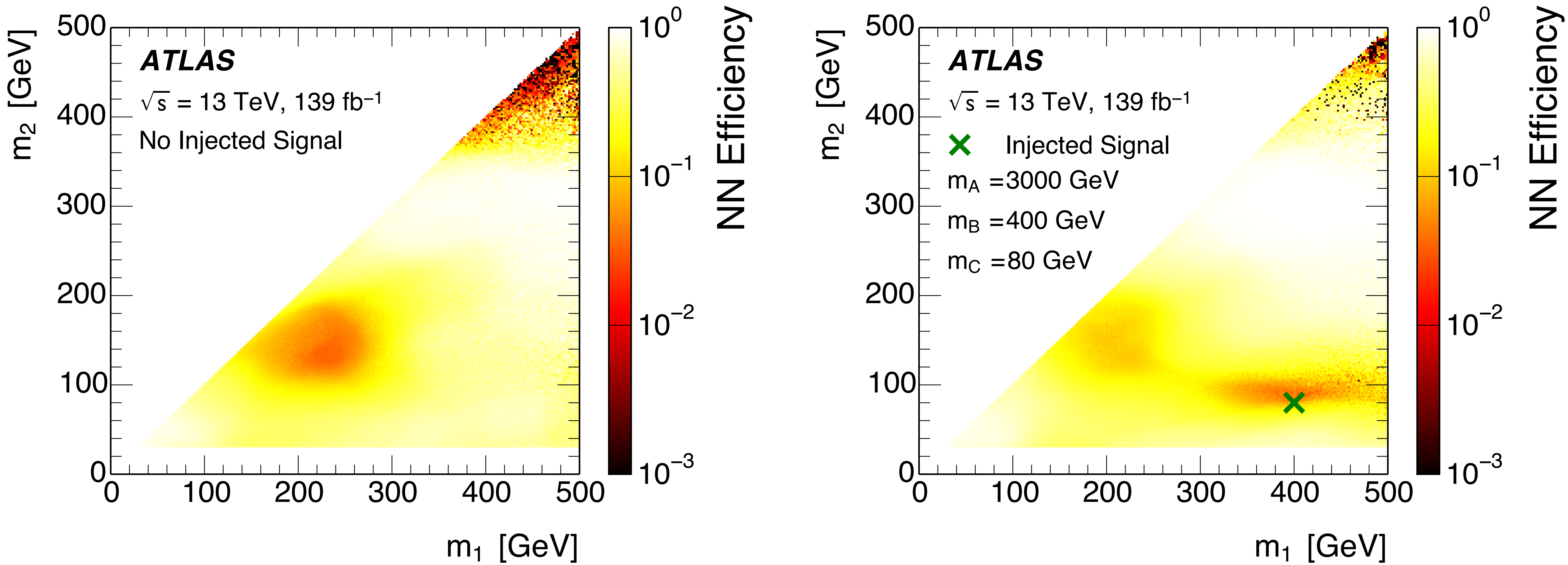
[Phys. Rev. Lett. 121 (2018) 241803]

J. Collins, K. Howe, **BPN**



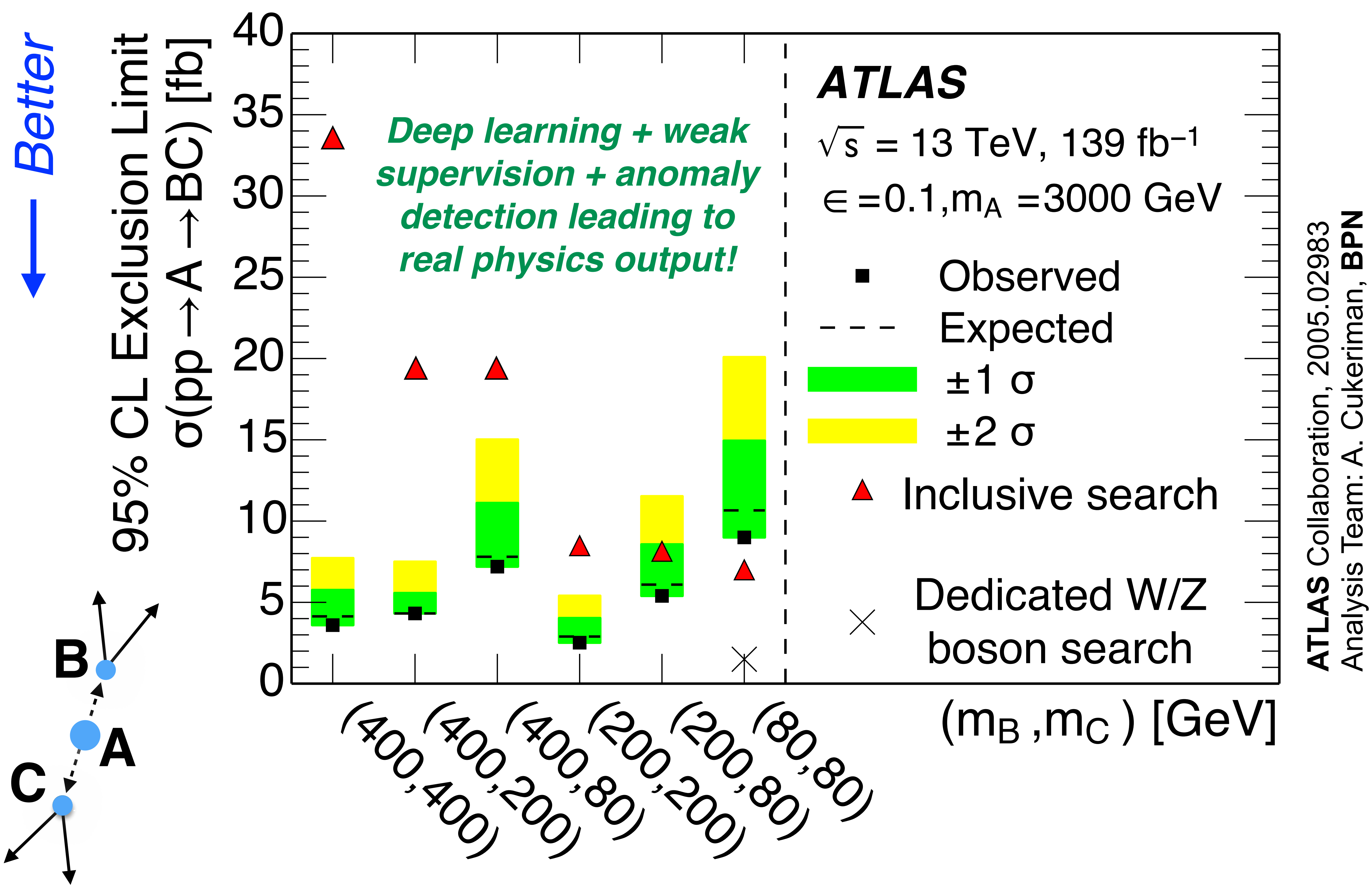
Collision data results New

ATLAS Collaboration, 2005.02983
Analysis Team: A. Cukeriman, **BPN**



First round, keep it simple: feature space is 2D (jet masses)

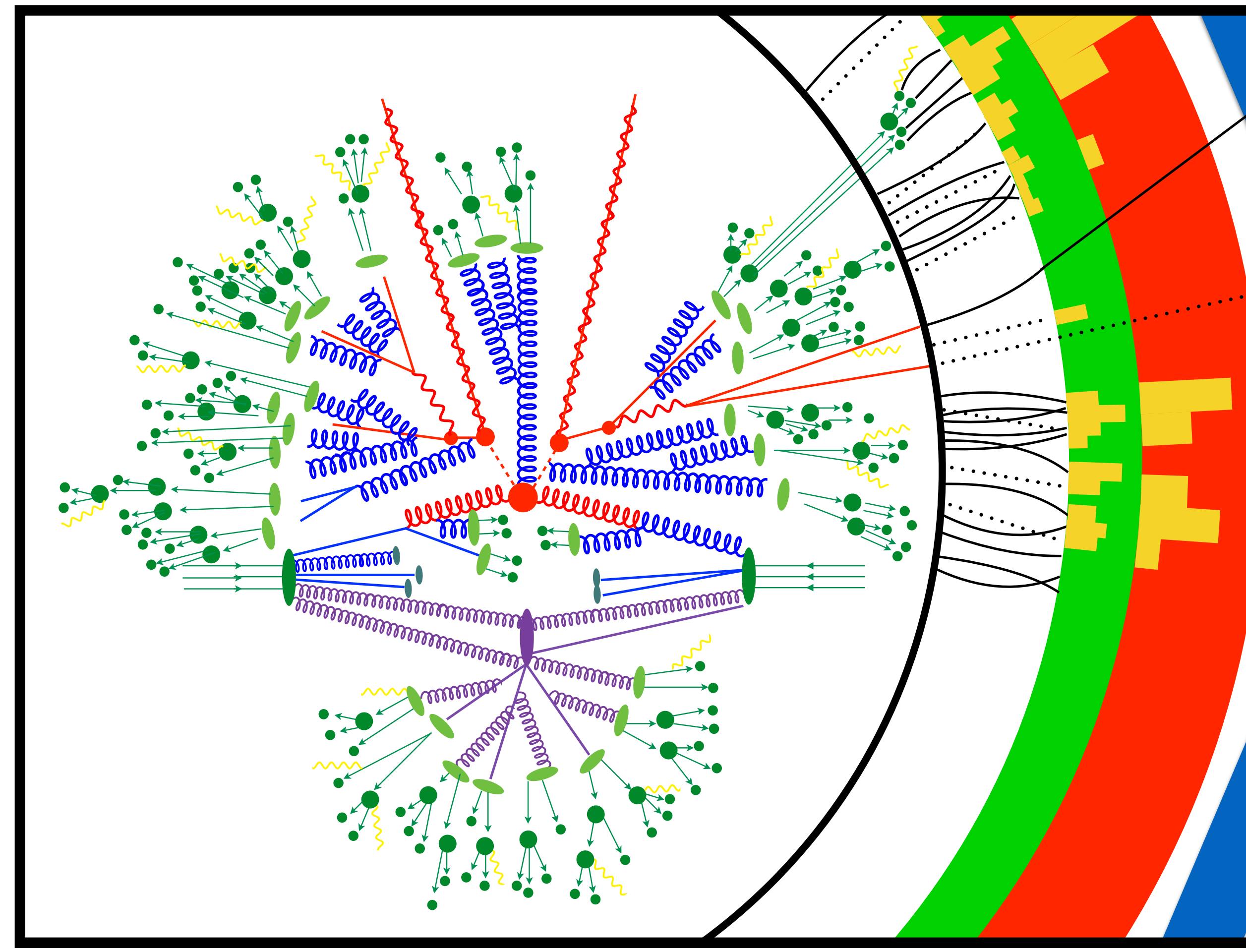
Collision data results New



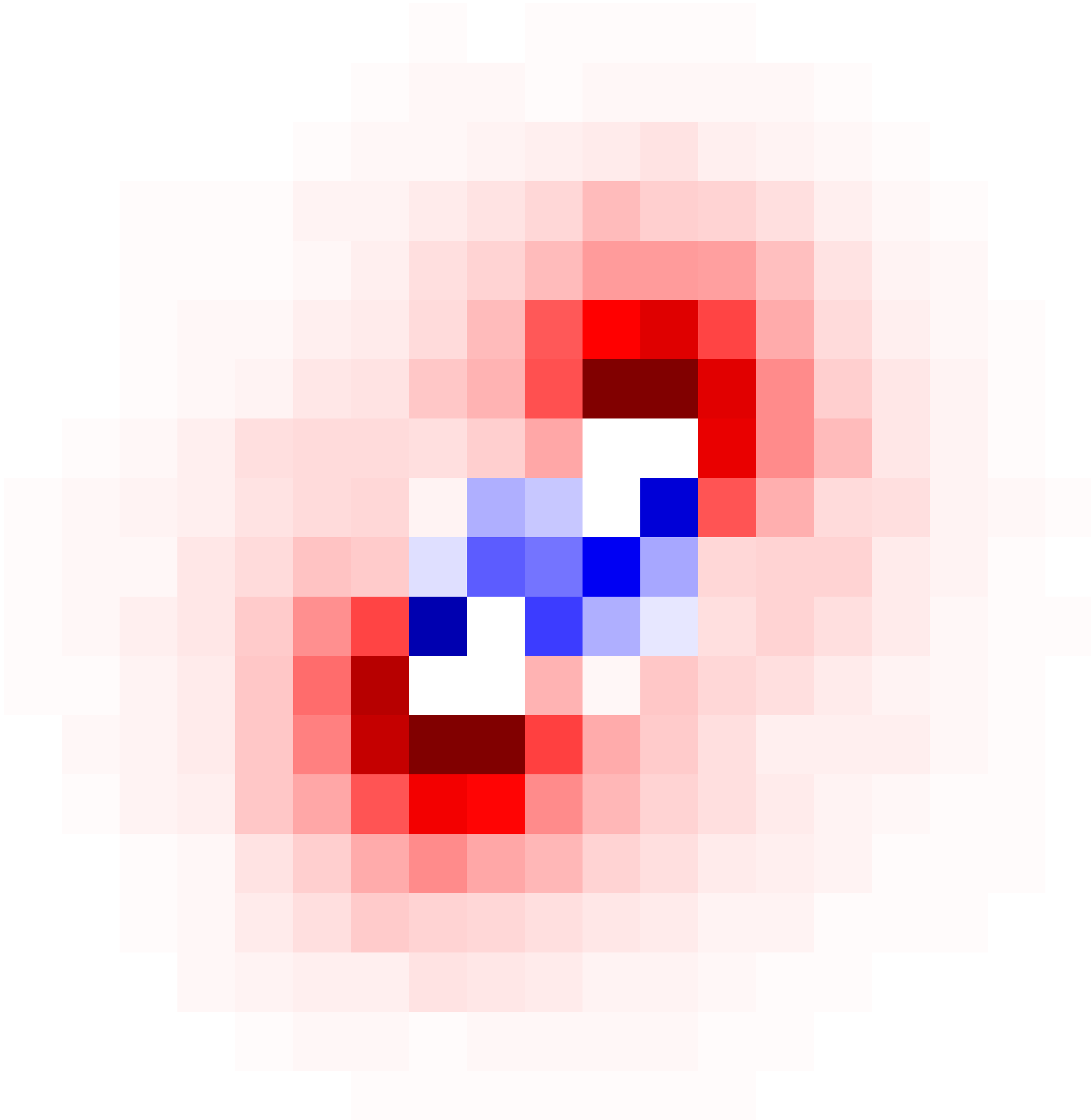
Conclusions and outlook

Jet substructure offers a rich physics program at the LHC, and elsewhere

Deep learning has a great potential to **enhance**, **accelerate**, and **empower** HEP analyses



The **full phase space** of our experiments is now explorable & combined with deep physics insight, we will be able to learn something new and fundamental about nature !



Fin.